A167000

(A)

PHYSICS
SPACE ENVIRONMENT

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFGL-TR-85-0315 SR, No. 253	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Air Force Geophysics Laboratory 6b. OFFICE SYMBOL (If applicable) CA	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Hanscom AFB Massachusetts 01731	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62101F; PROJECT NO. 9993; TASK NO. GE; WORK UNIT ACCESSION NO. OH

11. TITLE (Include Security Classification)
Handbook of Geophysics and the Space Environment

12. PERSONAL AUTHOR(S)
Adolph S. Jursa, Editor

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Scientific Final	FROM _____ TO _____	1985 December 5	1062

16. SUPPLEMENTARY NOTATION Additional copies of this document are available from the Air Force Geophysics Laboratory (SULR), Hanscom AFB, Massachusetts 01731.

17 COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Geophysics; Solar wind; Ionosphere, Space environment; Solar irradiance; Aurora; Sun; Radiation belts; Air glow.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The fourth edition of the Air Force Handbook of Geophysics and the Space Environment has been completely revised. It was conceived as a commemorative issue in recognition of the twenty-fifth anniversary of the International Geophysical Year. It begins with chapters on the sun and its emissions, then treats the earth's magnetic field and the radiation belts, and follows with chapters on the ionosphere and the aurora. The subject of electrical charging of space vehicles has been of special concern to the Air Force and has been included to aid designers interested in that problem. The next group of chapters deals with properties of the atmosphere, and the Handbook concludes with chapters on the earth sciences and infrared astronomy.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Adolph S. Jursa	22b. TELEPHONE (Include Area Code) (617) 377-3010 22c. OFFICE SYMBOL AFGL/CA

Cont. of Block 18:

Density	Winds	Geodesy
Temperature	Atmosphere	IR astronomy
Pressure	Gravity	

HANDBOOK OF GEOPHYSICS AND THE SPACE ENVIRONMENT

SCIENTIFIC EDITOR

Adolph S. Jursa

AIR FORCE GEOPHYSICS LABORATORY

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
1985

ASSISTANT EDITOR

Marylou Tschirch

CONTRIBUTING AUTHORS

	Chapter
Aarons, J.	10
Altrock, R.C.	1
Anderson, G.P.	21
Barnes, A.A.	16
Barron, W.R.	11
Basu, S.	10
Basu, Su	9
Battis, J.C.	23
Bertoni, E A.	16
Brown, H.A.	16
Buchau, J.	10
Burke, H.K.	20
Burke, W.J.	8
Cabaniss, G.H.	23
Champion, K.S.W.	14
Cipar, J.J.	23
Cliver, E.W.	10, 11
Clough, S.A.	18
Cohen, I.D.	16
Cole, A.E.	14, 15, 17
Cronin, J P.	11
Dandekar, B.S.	10
DeMastus, H.L.	1
Dyer, R.M.	16, 19
Evans, J.W.	1
Falcone, V.J., Jr.	19
Fenn, R.W.	18
Feynman, J.	3
Field, E.C.	10
Fitzgerald, D.R.	20
Gallery, W.O	18
Garrett, H.B.	7
Glass, M.	16
Good, R.E.	18, 21
Grantham, D.D.	16, 17
Gringorten, I.I.	15, 16, 17
Guidice, D.A.	11
Hall, L.A.	2
Hardy, D.A.	8
Hardy, K.R.	16
Hecksher, J.L.	10
Heroux, L.J.	2
Hinteregger, H.E.	2
Huffman, R.E.	22
Izumi, Y.	15, 16
Jekeli, C.	24
Johnston, J.C.	23
Kantor, A.J.	14, 15, 16, 17
Keil, S.L. Capt. USAF	1
Klobuchar, J.A.	10
Knecht, D.J.	4
Kneizys, F.X.	18
Kossey, P.A.	10
Kunkel, B.A.	16
Lewis, E.A.	10
Marcos, F.A.	17
McNamara, L.F.	10
Mendillo, M.F.	10
Metcalf, J.I.	16
Mill, J.D. Lt. Col. USAF	18
Millman, G.H.	10
Muench, H.S.	21
Murdock, T.L.	25
Neidig, D.F.	1
Novak, J.M. Capt. USAF	23
O'Neil, R.R.	12
Philbrick, C.R.	21
Picard, R.H.	12
Price, S.D.	25
Radick, R.R.	1
Rich, F.J.	9, 10
Rothman, L.S.	18
Rothwell, P.L.	5
Sagalyn, R.C.	20
Sharma, R.D.	13
Shea, M.A.	6
Shearer, J.A.	23
Shettle, E.P.	18
Simon, G.W.	1
Smart, D.F.	6
Spjeldvik, W.N.	5
Shuman, B.M.	4
Swider, W.	21
Tattelman, P.	15, 16, 17
Vancour, R.P.	8
Volz, F.E.	18
Weber, E.J.	10
Whalen, J.A.	12

FOREWORD

The space, atmospheric, and terrestrial environments influence the functioning of all Air Force systems. As technology advances, the role of the enviroment becomes more important to system performance. In many cases, the environment determines the limit in technical capability. This Handbook reflects the world of geophysics as honed and shaped by the special needs of the Air Force.

Today's operational systems were designed with the environmental knowledge made available by the Air Force Geophysics Laboratory and its predecessor, the Air Force Cambridge Research Laboratories. Tomorrow's Air Force depends on the quality and breadth of today's research as reflected in this Handbook. Although written primarily by Air Force Geophysics Laboratory scientists, this work reflects the state-of-knowledge of thousands of collaborative investigators, both U.S. and foreign. It is a mixture of basic research and exploratory development, and as such, represents the entire range of our technology based efforts. The very breadth of Air Force operations from space, through the atmosphere, to the earth makes it challenging, exciting, and rewarding to do front line research across this broad spectrum and to present a useable summary of the results to a wide community via this publication. This Handbook is one of our deliverable end products.

The recognition by the Air Force of the value of geophysics research to its mission is clearly illustrated in this work. The advances made by in-house scientists, the unique world class facilities created to do this work, and the excellent support of space based experiments, all attest to the long term dedication required to advance our understanding of the environment. Many significant players, military and civilian, have come and gone since the last issue of this Handbook. It is in the context of pride in these people and pride in our accomplishments for our country that we take great pleasure in dedicating this book to the twenty-fifth anniversary of the International Geophysical Year.

J. R. JOHNSON, Colonel, USAF
Commander

A. T. STAIR, JR.
Chief Scientist

PREFACE

This fourth edition of the Air Force Handbook of Geophysics and the Space Environment has been completely revised. It was conceived as a commemorative issue in recognition of the twenty-fifth anniversary of the International Geophysical Year. This was particularly appropriate since the Air Force Geophysics Laboratory—at that time the Air Force Cambridge Research Center—had been an active participant in the IGY in 1957–58 and the first edition of the Handbook had been published in 1957. The complex task of preparing this edition involved several years of effort during which time the twenty-fifth anniversary of the IGY occurred.

The purpose of the Handbook remains the same, that is, to provide Air Force designers, engineers, and systems operators with facts and data about the environment in which the Air Force operates. Extraordinary progress has occurred in this scientific field since the first edition was published due both to rapid advances in computer capability and the development of large rockets and satellites which have made available to experimenters platforms for upper atmosphere and space research that did not exist twenty-seven years ago.

The first edition of the Handbook was published the same year the first artificial earth satellite was orbited. At the time of the third edition, there was considerable activity in space but man had not yet gone to the moon. Since the publication of the third edition, research in and use of space has grown at an astonishing pace and its importance in man's future is an accepted fact. The Air Force has been active in space since the development of intercontinental ballistic missiles and early reconnaissance satellites, and is preparing for an even more active role as the space shuttle becomes operational.

In keeping with the Air Force interest in space, the format of the Handbook has been arranged to give those chapters dealing with the space environment increased emphasis. The fourth edition begins with chapters on the sun and its emissions, then treats the earth's magnetic field and the radiation belts, and follows with chapters on the ionosphere and the aurora. The subject of electrical charging of space vehicles has been of special concern to the Air Force and has been included to aid designers interested in that problem. Of no less importance are Air Force operations in the atmosphere and on the earth itself. The next group of chapters deals with properties of the atmosphere, and the Handbook concludes with chapters on the earth sciences and infrared astronomy.

Readers familiar with earlier editions will note that some chapters of the older versions have been deleted, while new chapters have been added. The choice of subject matter has been determined primarily by changes in today's Air Force requirements and operational activities. The contents also are related to the technical programs at the Air Force Geophysics Laboratory and reflect the expertise of the scientific staff, but should not be considered a technical report on such programs. In fact, the Handbook is only an introduction to several areas of geophysics. Subjects such as galactic x ray astronomy, astrophysics, planetary physics, meteors, some aspects of atmospheric pollution, oceanography, and others not of current concern to the Air Force do not appear here. Such subjects are covered very adequately elsewhere.

Numerous references have been included in each chapter so that the reader may pursue a specific subject area to whatever depth desired. The authors were not constrained by chapter format or length but were allowed to express themselves freely. As a result the chapters range from brief statements on a subject to textbook treatment of others. The Handbook has been brought as up to date as possible. Other government organizations such as NASA and NOAA have also published surveys and technical memoranda that deal with aspects of the atmosphere and space environment. Together with this Handbook they can provide the reader with an up-to-date picture of our understanding of the atmosphere and the near-earth space environment.

A. S. J.

ACKNOWLEDGMENTS

Many people have contributed to make this edition of the Handbook of Geophysics and Space Environment possible. The decision to proceed with the fourth edition of the Handbook was made by Colonels James Baker and Gerald D'Arcy, Commander and Vice Commander, respectively, of the Air Force Geophysics Laboratory in 1981. We are indebted to the authors of the twenty-five chapters; without them there would be no Handbook. Recognition is due three co-editors who undertook the necessary preliminary planning steps to organize the Handbook before I assumed responsibility as scientific editor. They are Dr. John N. Howard, Dr. Morton Barad, and Dr. Jules Aarons, who also contributed as co-author and scientific editor of Chapter 10. I would like especially to acknowledge the very important contribution of my assistant editor, Marylou Tschirch, who has been primarily responsible for the detailed editing, coordination, and final preparation of the manuscript for publication. Her tireless effort has been a major factor in the success of the Handbook. Mr. John Dempsey assisted with the editing of several chapters as well as contributing his expertise to discussions of format, typesetting, and printing.

Thanks are due to many secretaries in the Laboratory who typed the draft manuscripts of the chapters for the authors and, in particular, to Mrs. Ann Turner for her competent and extensive typing, copying, and accomplishment of many other administrative tasks that were necessary in the preparation of this Handbook.

A.S.J.

TABLE OF CONTENTS

CHAPTER 1. THE SUN

CHAPTER 2. SOLAR ULTRAVIOLET IRRADIANCE

TABLE OF CONTENTS

CHAPTER 3. SOLAR WIND

CHAPTER 4. THE GEOMAGNETIC FIELD

CHAPTER 6. GALACTIC COSMIC RADIATION AND SOLAR ENERGETIC PARTICLES

CHAPTER 7. THE CHARGING OF SPACECRAFT SURFACES

CHAPTER 8. MAGNETOSPHERIC AND HIGH LATITUDE IONOSPHERIC ELECTRODYNAMICS

TABLE OF CONTENTS

CHAPTER 9. IONOSPHERIC PHYSICS

CHAPTER 10. IONOSPHERIC RADIO WAVE PROPAGATION

CHAPTER 11. SOLAR RADIO EMISSION

CHAPTER 12. THE AURORA

CHAPTER 15. ATMOSPHERIC TEMPERATURES, DENSITY AND PRESSURE

CHAPTER 16. WATER VAPOR, PRECIPITATION, CLOUDS, AND FOG

TABLE OF CONTENTS

CHAPTER 17. WINDS

CHAPTER 19. ELECTROMAGNETIC WAVE PROPAGATION IN THE LOWER ATMOSPHERE

CHAPTER 20. ATMOSPHERIC ELECTRICITY

CHAPTER 21. ATMOSPHERIC COMPOSITION

CHAPTER 22. ATMOSPHERIC EMISSION AND ABSORPTION OF ULTRAVIOLET RADIATION

CHAPTER 23. GEOKINETICS

CHAPTER 24. GEODESY AND GRAVITY

LE OF CONTENTS

CHAPTER 25. INFRARED ASTRONOMY

Chapter 1

THE SUN

R.C. Altrock
H.L. DeMastus
J.W. Evans
S.L. Keil, Capt., USAF
D.F. Neidig
R.R. Radick
G.W. Simon

Radiant energy from the sun determines the surface temperature of the earth and supplies virtually all the energy for natural processes on the earth's surface and in its atmosphere. The sun is about 5×10^9 years old and geological and paleontological evidence indicates it has shone with its present intensity for about a billion years. The solar radiation is, however, not absolutely constant. There are short-period fluctuations in ultraviolet and x-ray radiation and in corpuscular emission that induce complex responses in the earth's atmosphere. In this chapter, solar characteristics and activity are identified; other chapters deal with the geophysical responses themselves. We shall discuss first the *quiet sun*, then solar activity.

1.1 THE SUN'S "STEADY-STATE" BEHAVIOR (THE QUIET SUN)

Although all observable solar features change on timescales of seconds to days, the statistical description of many such features taken together remains unchanged over substantial periods of time. These "steady-state" features of the so-called quiet sun are described here.

1.1.1 Basic Characteristics

The sun is a relatively small, faint, cool star at an average distance of 1.49×10^8 km from the earth. Its mass is 1.99×10^{33} g. Composition of the outer layers by mass is approximately 73% H, 25% He, and 2% other elements [Allen, 1973]. Table 1-1 gives the run of temperature, density, mass concentration, energy generation, and pressure as functions of radius [Allen, 1973]. The total solar luminosity is 3.83×10^{33} erg/s.

A schematic cross-section through the sun is shown in Figure 1-1. Energy is generated in the solar core by nuclear reactions. The primary reaction is the fusion of hydrogen nuclei to form helium nuclei ($4H^1 \rightarrow He^4$). For every gram of hydrogen entering the reaction, 0.007 g is converted into energy. To produce the solar luminosity, about 4.3×10^{12} g/s must be converted to energy.

The energy generated in the core is transferred outward by radiation. Because of the high densities, the radiation is absorbed and re-emitted many times on its outward journey. Radiation that began in the core as high energy gamma rays is degraded by these successive absorptions and re-emissions until they finally emerge as visible radiation characteristics of the solar surface. However, the energy is not carried all the way to the surface as radiation. As it moves outward from the core, the temperature, density, and pressure drop. As the temperature drops, free electrons can be trapped by atoms into bound states, causing an increase in the opacity. Thus radiation becomes less effective in transporting energy. A large temperature gradient results and

Table 1-1. Solar structure as function of radius [Allen, 1973].

	r	T	ρ	M_r	L_r	log P
$R_\odot$	10^3 km	10^6 K	g/cm^3	$M_\odot$	$L_\odot$	dyn/cm^2
0.00	0	15.5	160	0.000	0.00	17.53
0.04	28	15.0	141	0.008	0.08	17.46
0.1	70	13.0	89	0.07	0.42	17.20
0.2	139	9.5	41	0 35	0.94	16.72
0.3	209	6.7	13.3	0.64	0.998	16.08
0.4	278	4.8	3.6	0.85	1.00	15.37
0.5	348	3.4	1.00	0.94	1.000	14.67
0.6	418	2.2	0.35	0.982	1.000	14.01
0.7	487	1.2	0.08	0.994	1.000	13.08
0.8	557	0.7	0.018	0.999	1.000	12.18
0.9	627	0.31	0.0020	1.000	1.000	10.94
0.95	661	0.16	0.0^34	1.000	1.000	9.82
0.99	689	0.052	0.0^45	1.000	1.000	8.32
0.995	692.5	0.031	0.0^42	1.000	1.000	7.68
1.000	696.0	0.006	0.0	1.000	1.000	—

T = temperature; ρ = density; P = pressure; M_r = mass within radius 4; L_r = energy generation within radius r; $M_\odot$, $R_\odot$, $L_\odot$ = mass, radius, energy generation of whole Sun, with $M_\odot = 1.99 \times 10^{33}$g and $L_\odot = 3.83 \times 10^{33}$ergs.

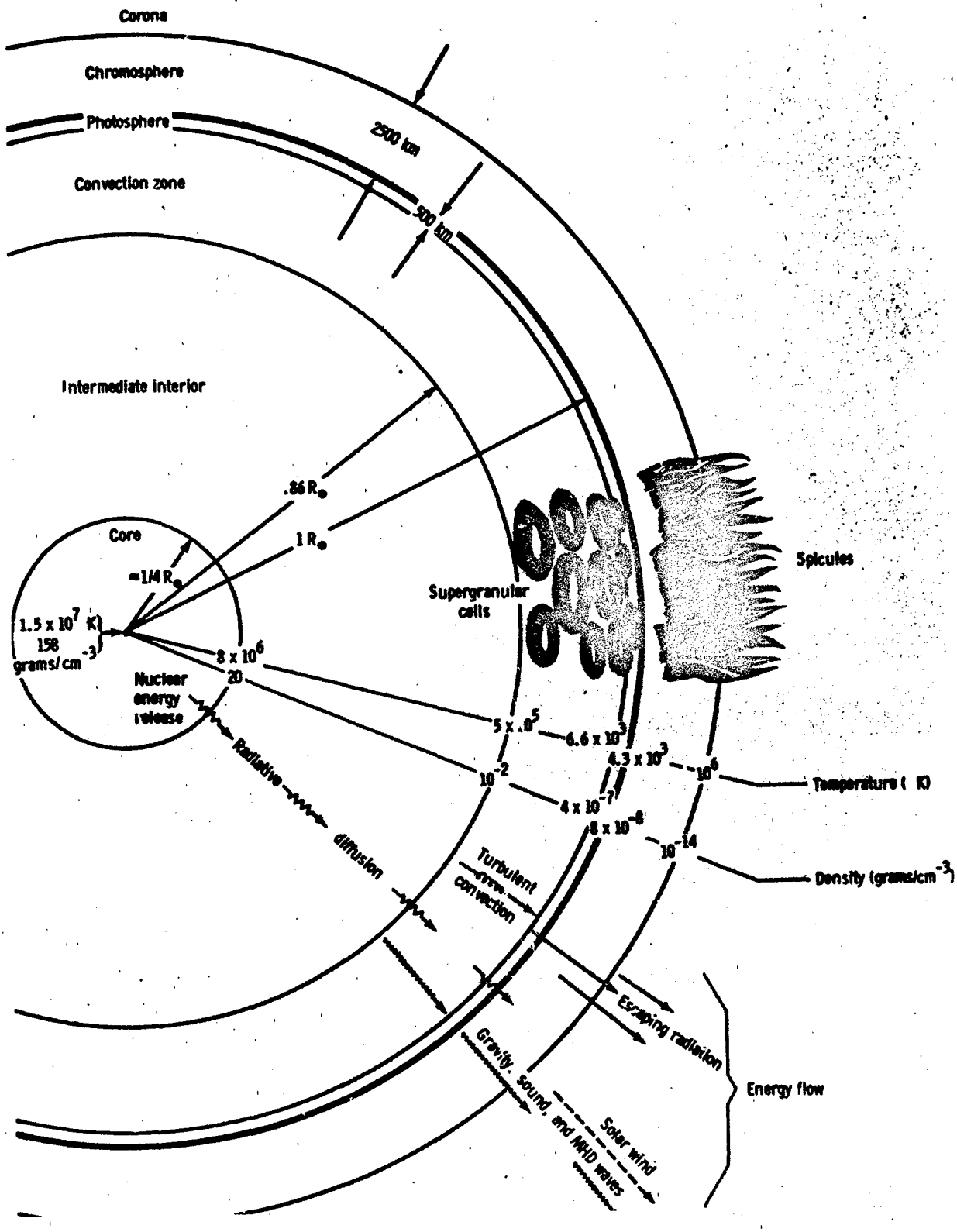


Figure 1-1. Idealized general solar properties, structure, and modes of outward energy flow. The features shown are not to scale and provide a qualitative picture only [Gibson, 1973].

convection becomes the primary energy transport mechanism. The exact depth at which this occurs is model dependent. However, observations clearly show that this *convection zone* reaches the solar surface and that at least 3 or 4 scales of convection are present. Near the solar surface, radiation can escape into space and it again becomes the primary energy transport mechanism.

Attempts to make accurate measurements of the solar diameter have led to the discovery of global solar oscillations. The observed periods agree approximately with theoretically predicted radial p-mode oscillations. Table 1-2 compares some observed periods of global oscillations with predictions based upon a theoretical solar model [Christensen-Dalsgaard and Gough, 1976]. A summary of attempts to measure global solar oscillations can be found in Hill [1979]. These modes of oscillation can be used in a manner

Table 1 2. Observed and theoretical modes of solar oscillations [Christensen-Dalsgaard and Gough, 1976].

Periods of Solar Oscillations (min)									
		l = 0		l = 2		l = 4		l = 2	
1973	1975	Mode	Period	Mode	Period	Mode	Period	Mode	Period
		p_1	62.7 (63.9)					g_1	55.1
52	47.9			f	46.0				
				p_1	42.2 (42.3)	f	39.6		
		p_2	43.8			p_1	38.0	g_2	61.5
33		p_3	32.6 (31.8)	p_2	34.3			g_3	70.9
	30.3					p_2	29.3		
		p_4	26.0	p_3	26.7			g_4	81.8
23.8						p_3	23.4		
	21.0	p_5	21.0 [21.5]	p_4	21.5			g_5	93.0
						p_4	19.5		
		p_6	18.2	p_5	17.9			g_6	105
16.7	17.1					p_5	16.3		
		p_7	15.8 [15.4]	p_6	16.0			g_7	117
	14.6					p_6	14.6		
		p_8	14.0	p_7	14.1			g_8	130
13.3						p_7	13.0		
		p_9	12.5	p_8	12.6			g_9	142
11.9	11.8					p_8	11.7		
		p_{10}	11.3	p_9	11.4			g_{10}	154 (147)
10.4	10.5	p_{11}	10.4	p_{10}	10.4			g_{11}	167 (159)
		p_{12}	9.55					g_{12}	180 (171)
9.2	8.8	p_{13}	8.84 [9.12]						
		p_{14}	8.52						
	7.9	p_{15}	7.99						
7.6		p_{16}	7.53						
7.0	7.2	p_{17}	7.12						
		p_{18}	6.75						
		p_{19}	6.41						
		p_{20}	6.10						

The periods headed 1973 and 1975 are from observations by Hill, Stebbins, and Brown [1975] and are accurate to ± 5%. Those periods not reported in the publication were kindly supplied by Profesor Hill. Periods in the other columns were obtained theoretically using a solar model with $Z = 0.02$, except those in parenthesis which are periods of oscillation of a model with $Z = 0.04$. The periods in square brackets were obtained by repeating the computation with $Z = 0.02$ using a different distribution of mesh points. (Reprinted with permission from Macmillan Journals Ltd © 1976.)

analogous to terrestrial seismology to probe the deep interior of the sun. This should eventually permit us to refine the model shown in Table 1-1 and place better constraints on theoretical models of the solar interior.

The total solar irradiance is quantitatively defined as the amount of radiant energy at all wavelengths received per unit time and area at the top of the earth's atmosphere, corrected to the mean earth-sun distance, and is customarily expressed in units of watts per square meter (W/m^2). The term *solar constant* is also used occasionally to denote this

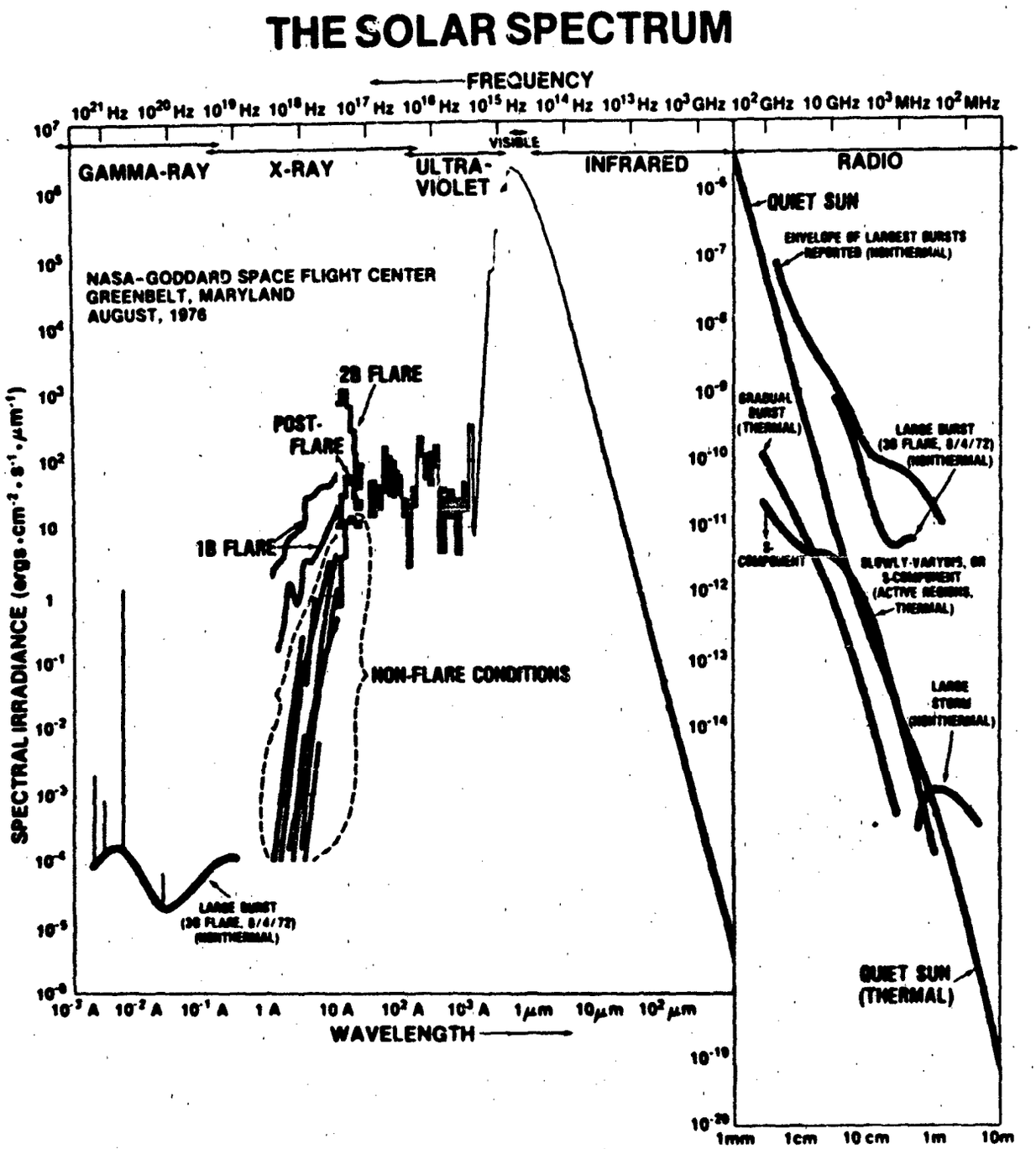


Figure 1-2. Spectral distribution of solar irradiance [White, 1977].

radiant flux but is somewhat misleading, since the total solar irradiance seems to fluctuate slightly. The best measurement to date of the total solar irradiance, 1368 ± 7 W/m^2, was obtained in 1980 using radiometers aboard the NASA Solar Maximum Mission (SMM) satellite.

More than 70% of the solar irradiance is concentrated in the near ultraviolet, visible, and near infrared portions of the spectrum lying between the atmospheric transmission cut-offs near 0.32 μm and 1.0 μm. Since most of this radiation reaches the earth's surface, a large fraction of its energy enters the lower atmosphere through the evaporation-precipitation cycle of water. About 2% of the solar irradiance appears at ultraviolet and x-ray wavelengths shorter than 0.32 μm. This radiation is all strongly absorbed in the upper atmosphere where it drives several important photochemical reactions, including ozone production. The remainder of the solar irradiance appears at infrared and radio wavelengths longer than 1.0 μm. Much of this radiation, especially in the infrared, is strongly absorbed by atmospheric water vapor and carbon dioxide and thus contributes to the energy budget of the lower atmosphere.

Figure 1-2 illustrates the spectral distribution of the solar irradiance. A complete description of the ultraviolet and radio portions of the solar irradiance will be found in Chapters 2 and 11.

1.1.2 Solar Rotation

The rotation period of the sun is a function of solar latitude, varying from about 24 days at the equator to 34 days at the pole. This variation is called *differential rotation*. Taking into account the earth's orbital motion, an effective period of 27 days is usually satisfactory for predicting rotation recurrences of solar-terrestrial disturbances. Because the sun's axis of rotation is inclined only 7° to the plane of the earth's orbit, most of the solar surface can be scanned in the course of one rotation.

When solar rotation is measured by following tracers (features on the solar surface such as sunspots) one observes faster rotation rates than from the Doppler shifts of solar spectral lines. Figure 1-3 shows the latitudinal dependence of rotation rate measured by tracers compared to the Doppler shift of lines [Wilcox and Howard, 1970]. Furthermore, Doppler measurements show that the differential rotation varies with time, and seems to have large eddies or cells imbedded in it. These temporal variations can be as large as the variation in latitude. Evidence has also been found for toroidal oscillations superposed on the solar rotation [Howard and LaBonte, 1980].

Differential rotation probably has its origin in the interaction between convection and rotation. This same interaction underlies most of the magnetic activity observed on the sun (Section 1.2.2).

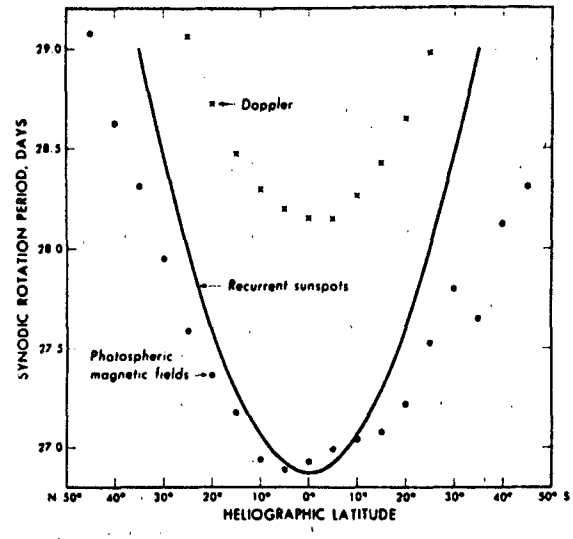


Figure 1-3. Solar differential rotation for sunspots (solid curve), large scale photospheric magnetic fields (circles) and quiet photosphere (crosses) [Wilcox and Howard, 1970].

1.1.3 The Solar Atmosphere

Because the sun is gaseous throughout, there is no rigid boundary marking the start of the atmosphere and end of the interior. However, a definite optical boundary, below which we cannot see, exists and is commonly referred to as the solar surface. As can be seen from Table 1-1, the solar temperature decreases outward from the core until we reach this optical surface. For a few hundred kilometers above this level the temperature continues to decrease. However, at about 400 km one reaches the *temperature minimum*, above which the temperature again begins to increase with increasing height. Although no rigid boundaries exist, it is convenient to divide the solar atmosphere into layers having characteristic temperatures and densities. Moving outward from the optical surface these layers are known as the *photosphere*, *chromosphere*, *transition zone*, and *corona*. Each of these regions is discussed below.

1.1.3.1 The Photosphere. The photosphere is the apparent solar surface, with a very sharply defined limb, seen by ordinary observations in white light. Its location with respect to other solar features is shown in Figure 1-1. Its diameter, usually considered to be the diameter of the sun, is 1.391 × 10^6 km. Practically all of the solar mass is contained within the radius of the photosphere. Figure 1-4 shows the variation of temperature, density, and pressure in a standard model of the lower solar atmosphere. Height is given in kilometers above a depth in the atmosphere for which the probability of a photon escaping is 1/e. The thickness of the photosphere is only about 100 km, or 0.014% of $R_\odot$, where $R_\odot$ is the solar radius. Below the photosphere

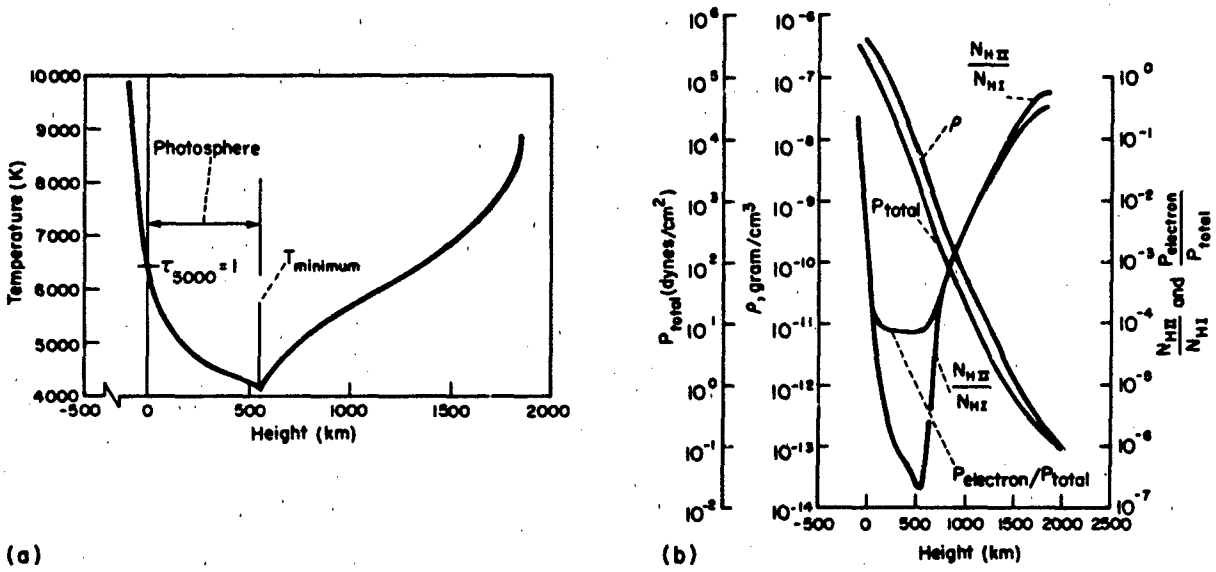


Figure 1-4. Gas properties versus height (HSRA model)[Gingerich, 1971]. (a) Temperature as a function of height. (b) Pressure, density, electron pressure, and hydrogen ionization as functions of height.

the solar gas is opaque. This opacity is primarily due to a small concentration of negative hydrogen ions in the region immediately below the photosphere. These negative ions act as continuous absorbers over a great range of wavelengths, absorbing most of the intense radiation from deeper layers in the sun. The resulting accumulation of thermal energy in this layer sets up a system of convective currents that transports the excess energy through the negative hydrogen barrier. The energy is then reradiated into the relatively transparent gases above the photosphere and, hence, into space. Penetration of these convective elements (*granules*) into the photosphere gives rise to the granular appearance of the solar surface shown in Figure 1-5. Similar opacity barriers which exist deeper in the sun cause other larger scales of convective motions. The most obvious of these is the *supergranulation* shown in Figure 1-6. Figure 1-7 is a schematic diagram of the various scales of motion present in the solar atmosphere.

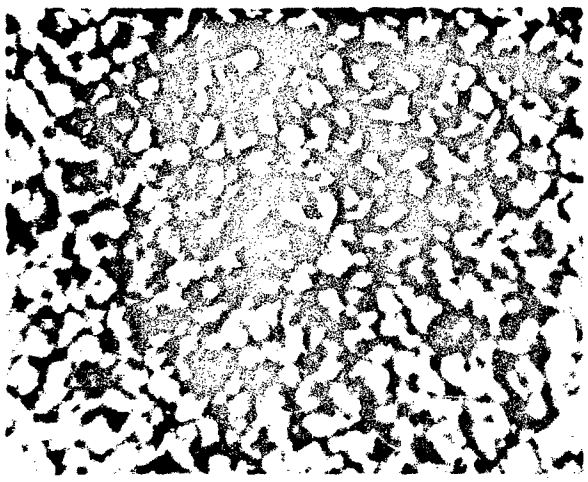


Figure 1-5. Photospheric granulation and sunspot structure photographed with 180 ft focal length tower telescope at the National Solar Observatory at Sacramento Peak.

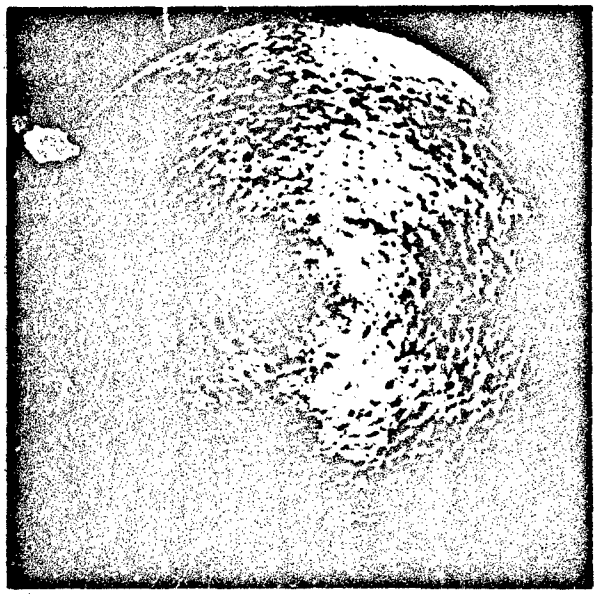


Figure 1-6. Velocity spectroheliogram obtained in Ca 6103 Å, showing supergranulation structure on solar surface. Supergranules have diameters of about 30 Mm and lifetimes of 1-2 days.

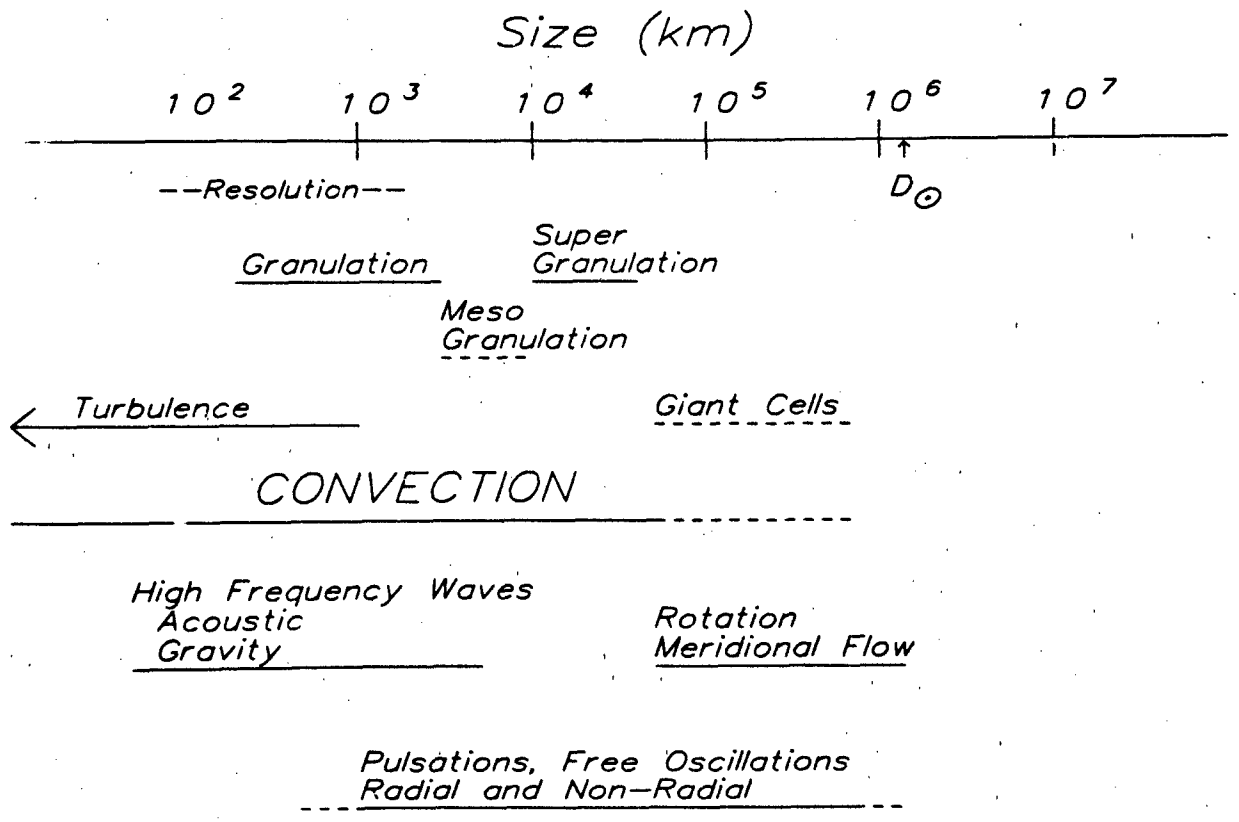


Figure 1-7. This schematic diagram shows the horizontal size of various motions present in the solar atmosphere. Also shown is an approximate range for the resolution limit of ground-based observations. The smaller sizes indicated for resolution are only achieved using special techniques such as speckle interferometry or active optical systems. Typical resolution for conventional telescopes is limited by turbulence in the terrestrial atmosphere to approximate one arcsecond or about 720 km on the sun. The scales shown with the solid lines indicate an approximate range, while the dashed lines indicate sizes either predicted theoretically or that have been only tentatively observed.

Because the opaque gas composing the photosphere absorbs and reradiates approximately as a black body, the photosphere emits an essentially continuous spectrum, or *continuum*. Sharply tuned absorptions by the various atoms in and above the photosphere give rise to thousands of dark lines (Fraunhofer lines) superposed on the continuum. This line absorption spectrum, shown in Figure 1-8, is the source of most information on the abundance of the various elements and the physical state of the solar atmosphere. Although the lines of a number of atoms are unobservable through the terrestrial atmosphere, all the natural chemical elements known on earth probably exist in the sun. Their relative abundances, however, differ radically.

The principal region of line absorption is a layer about 400 km thick just above the visible surface of the photosphere. Spectra with high spatial resolution on the sun as shown in Figure 1-9 show small-scale wiggles due to Doppler shifts induced by mass motions in the line formation region. Analysis of these *wiggly lines* has revealed several physical processes occurring on the sun.

The granules of Figure 1-5 are evidence of convective overshooting into the stable photosphere from deeper layers.

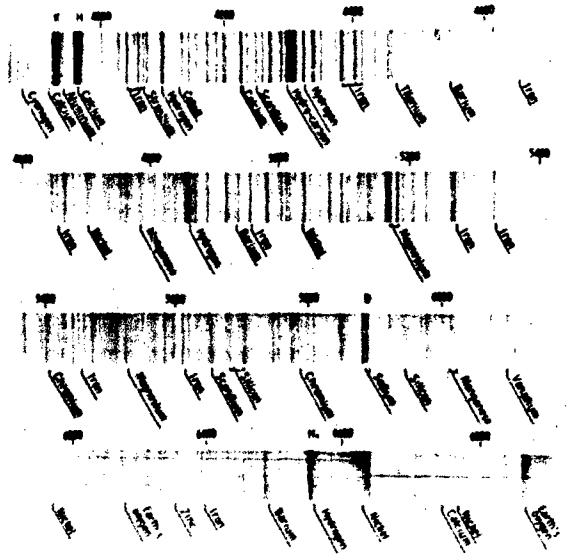


Figure 1-8 The Fraunhofer line absorption spectrum of the solar atmosphere

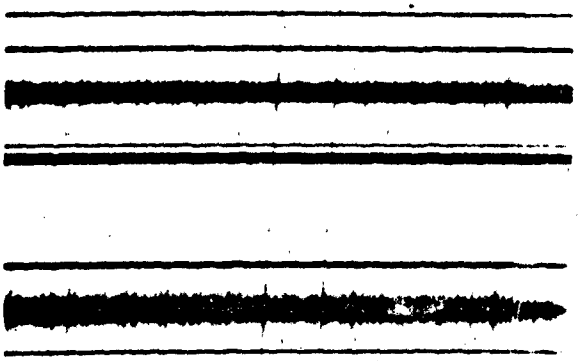


Figure 1-9. The spectral region around the Mg B lines at 5167 Å and 5172 Å is shown. The verticle axis is wavelength and the horizontal axis is distance along the slit. The dark horizontal bars are absorption lines.

An average granule has a diameter of about 1.4 arcsec or 1 Mm. Granules typically live for 8 minutes and reach velocities near 2 km/s. Most granules lose their temperature excess after rising only 50 to 100 km into the photosphere, but their momentum carries them 100 or 200 km higher.

The supergranules shown in Figure 1-6 penetrate higher into the atmosphere but have much smaller vertical motions than granules. They are primarily observed as horizontal flows in the chromospheric layers above the photosphere. Supergranules have diameters of approximately 30 Mm and lifetimes of 1-2 days. Evidence has recently been found for two other scales of convection: mesogranulation [November et al., 1981] with a scale of 5-10 Mm, and giant cells perhaps 100 Mm in diameter.

In addition to revealing various modes of convection, wiggly line spectra show that the sun oscillates in a number of different modes. Oscillatory modes with large coherent wave systems having periods near 300 s are observed in the photosphere and overlying chromosphere. Such data are compared with a theoretical non-radial acoustic spectrum in Figure 1-10. These modes have most of their energy concentrated in the underlying convection zone, and deviations of predicted modes from observed modes can lead to better models of the convection zone. The 300 s oscillations are evanescent in the photosphere (below the temperature minimum shown in Figure 1-4). At the higher temperatures found in the chromosphere, the waves become propagating and higher frequency acoustic waves are observed. The mechanical energy in these waves may be responsible for the temperature increase in the chromosphere and higher layers.

1.1.3.2 The Chromosphere. The chromosphere is a layer, several thousand kilometers thick, of transparent glowing gas above the photosphere. Many of the phenomena occurring in the photosphere also are manifest in the chromosphere. Because the density in the chromosphere continues to decrease with height (Figure 1-4), and is much lower than in the photosphere, magnetic fields and waves can have greater influence on the structure. Thus the chromosphere is even more inhomogeneous than the photosphere. Its continuous spectrum is extremely faint, and like

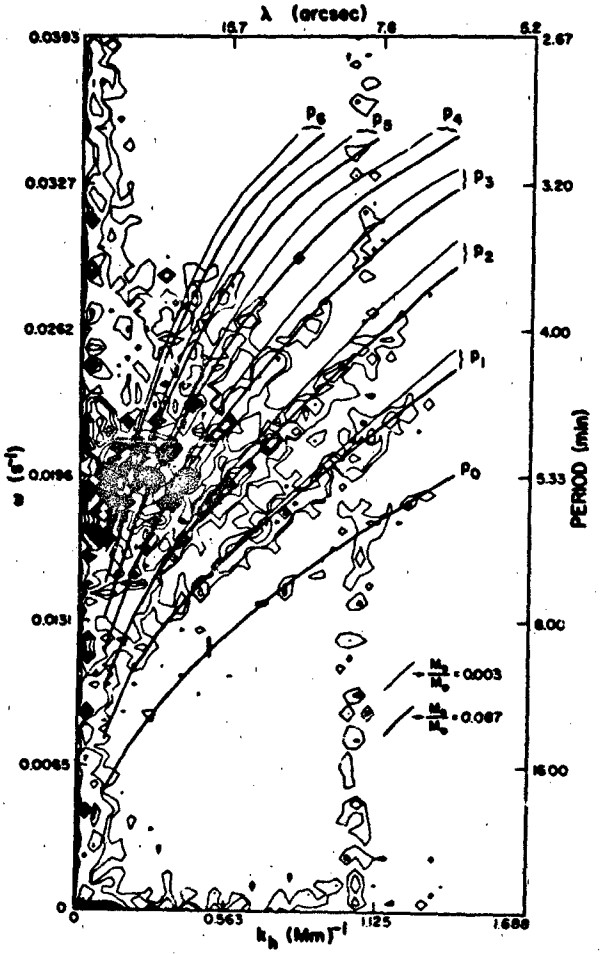


Figure 1-10. Comparison of an observed k-ω diagram with theoretical Eigenfrequencies for two model solar envelopes [Rhodes et al., 1977]. Contour intervals are 0.125, 0.25, 0.75, 1.75, etc. m/s. Observations were made on the magnetically insensitive line Fe I λ 5576.10 using the vacuum tower telescope at the National Solar Observatory at Sacramento Peak. (Reprinted with permission from the American Astronomical Society, © 1977.)

all solar features above the photosphere, it can be observed normally only by isolating the light of one of its strong lines, usually the red Hα line at 6563 Å (Fraunhofer C line) or the line of CA II at 3934 Å (Fraunhofer K).

The rise in chromospheric temperature with height can be understood in terms of a nonradiative energy input into the atmosphere via mechanical energy generated in the convection zone. Acoustical waves propagate outward and can form shocks in the low density chromosphere, creating more energetic collisions between particles. The low density also leads to a relative enhancement in the ability of magnetic fields to carry energy in the form of waves. These mechanisms combine to heat the chromosphere from approximately 4300 K at its base to over 20 000 K in about 2 Mm.

Similar to the granulation patterns observed in the photosphere, a *chromospheric network* can be observed in chromospheric lines. Figure 1-11 shows the appearance of this network in Hα and Ca II. The scale of this network corresponds fairly closely to that of supergranulation, and the bright regions of the chromospheric network correlate well with supergranular boundaries. At sub-photospheric densities the supergranules are capable of pushing magnetic flux tubes to their boundaries while in the chromosphere the fields dominate the gas. If heating occurs preferentially near the field lines, the appearance of the network can be explained. The network is also the location of *spicules* which are brilliant flames of gas that project upward for approximately 10 Mm and live for 2 to 5 minutes. Figure 1-12 shows how these spicules appear at the limb of the sun. Spicules may provide the primary path for moving energy and material into the corona.

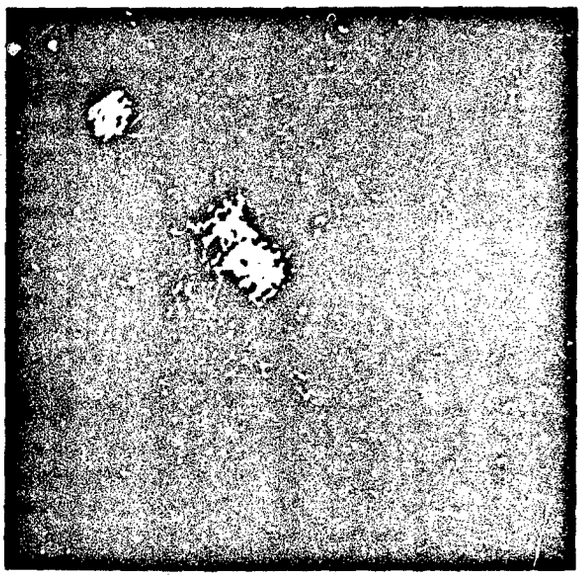


Figure 1-11. Ca II spectroheliograms made on the violet side of Ca II line at 3934 Å. The network of bright emission is clearly evident, and outlines the boundaries of supergranules seen in Figure 1-6.

1.1.3.3 The Transition Region. At the top of the chromosphere, approximately 1.7 Mm above the temperature minimum, the temperature begins a sudden steep rise to coronal values. In the next few megameters it rises from 25 000 K to near 2×10^6 K. This steep temperature rise is called the transition region. Because this transition occurs in such a narrow zone of the solar atmosphere, it is very difficult to measure the physical height boundaries accurately. It is also very hard to observe this region, since its spectral lines are in the EUV and can be studied only from satellite and rocket observations. Thus the amount of data available to date is very limited. We show in Figure 1-13 a profile of temperature vs height from the photosphere to the low corona with some transition zone spectral lines indicated. The slope of the rise and the location of its onset are functions of solar activity, being steeper and lower in the atmosphere in active regions (Section 1.2.3.1) than on the quiet sun. EUV photographs of the transition zone obtained by Skylab are shown in Figure 1-14. For further information on this region, including models, see Jordan [1981], Zirker [1977], Athay [1976], and Bonnet [1981].

1.1.3.4 The Corona. *Physical Properties:* Above the transition region and extending out into interplanetary space is the faintly visible solar corona. Long seen only during the few seconds or minutes of total solar eclipses, the true nature of this outer atmosphere has only in recent years been recognized. Dominated by temperatures of one to two million degrees, magnetically-induced motions, sudden releases of energy, and explosive expension into the surrounding vacuum of space, this entity has strong implications for the energy balance of near-solar space and planetary environments.

The corona is heated from below by mechanical, electrical or magnetic dissipation. The exact nature of the heating has not yet been discovered. Favorite candidates are shock waves produced by the upward surging of granules

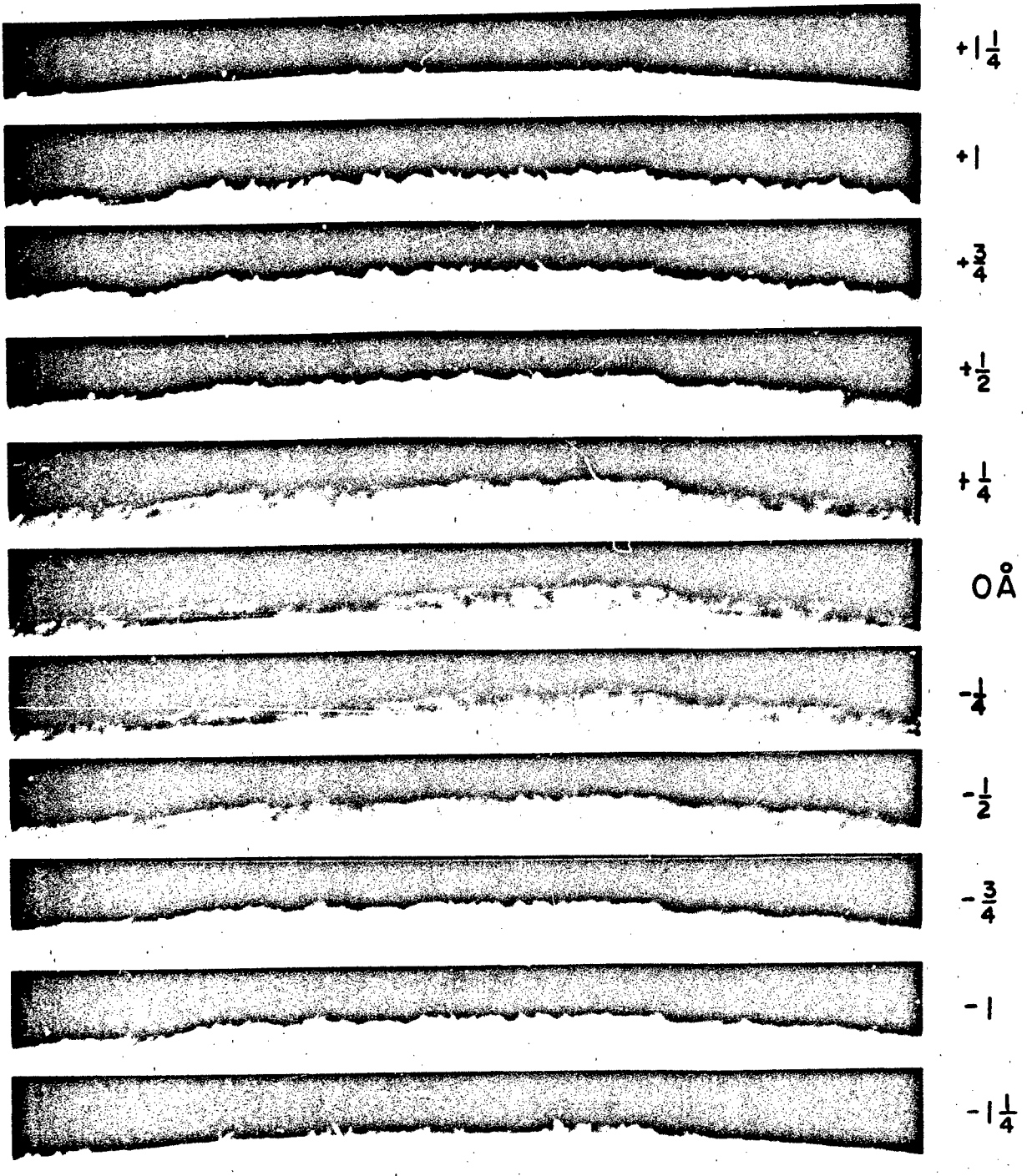
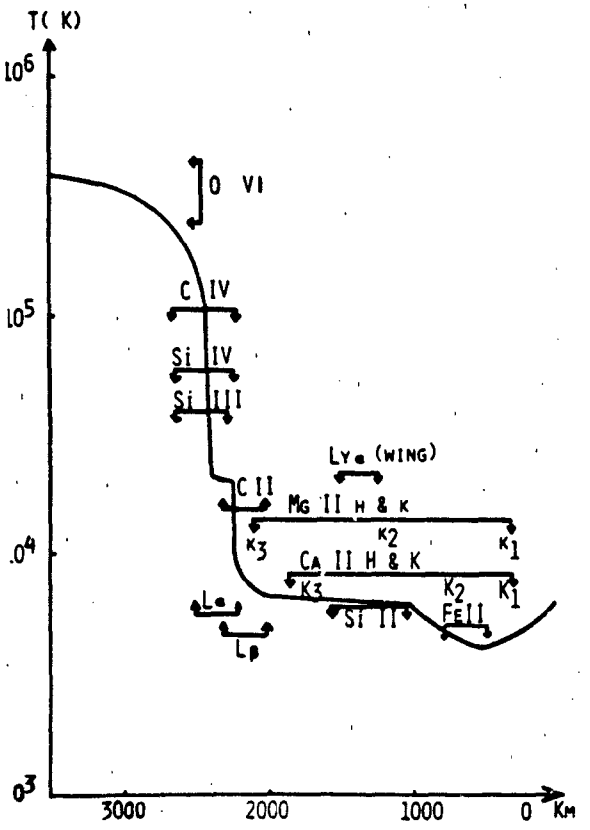


Figure 1-12. Spicules at the limb observed in Hα. These narrow tongue-like flames extend 5-19 Mm above the limb and live for 2–5 minutes.



igure 1-13. An average quiet sun temperature distribution from photosphere (0 km), through the temperature minimum (400 km), chromosphere (500-2000 km), and transition zone (2000-2500 km), and into the low corona (3000 km up), derived from Vernazza et al. [1981].

nto the photosphere or by propagating waves or oscillations nduced by them or by spicules. Other recent speculation as centered on the dissipation of Alfvén waves being propgated upward from the turbulent photosphere along magetic lines of force. The end result of this heating, whatever s source, is an increase in temperature to near 2×10^6 K n the low corona. The temperature decreases slowly above nat level. This extreme heat causes the outer parts of the orona to "boil away" into the near-vacuum of interplanetary pace. Coronal heating is reviewed in Jordan [1981].

Observations of the corona (Figure 1-15) indicate that is highly inhomogeneous. Different regions, such as *treamers* (radially-elongated bright areas—Section 1.2) and *oronal holes* (dark areas—Section 1.2), show small variations in temperature (on the order of 30%) from area to rea, but large variations in density, up to one to two orders f magnitude. These variations become more pronounced /ith height.

Models of the corona are controversial and complex. arly one-dimensional models have now given way to two- nd three-dimensional models representing streamers, coronal holes, magnetic loops, prominence cavities, and time-dependent variations in coronal transients.

Table 1-3 presents a simple model of a coronal hole and the quict corona. Further information should be sought in Zirker [1977], White [1977], Allen [1973], Eddy [1978], and Jordan [1981].

Table 1-3. Model of corona in quiet sun and in coronal hole showing electron density vs height above the solar limb (R_o = 1.0) [Zirker, 1977].

	Quiet Sun ($T = 2.0 \times 10^6$ K)	Coronal Hole ($T = 1.5 \times 10^6$ K)
r/R_o	N_e (cm^{-3})	N_e (cm^{-3})
1.1	1.0×10^8	5.4×10^7
1.2	7.1×10^7	1.6×10^7
1.4	2.3×10^7	2.8×10^6
2.0	2.8×10^6	2.0×10^5
4.0	8.9×10^4	4.0×10^3
10.0	8.0×10^3	

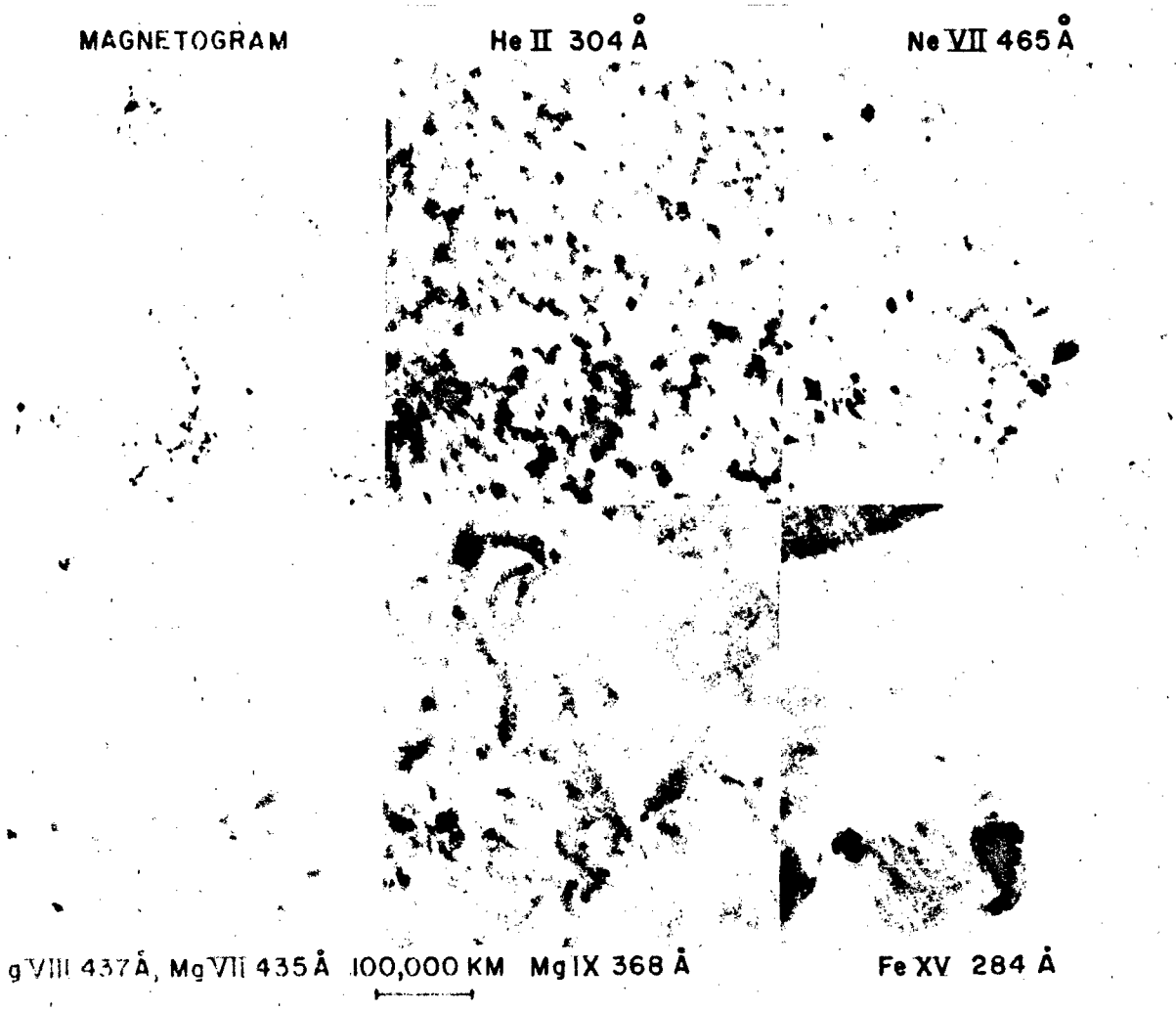
Methods of Observation: Coronal observations are extremely difficult because of the requirement to eliminate the million-times-brighter light from the solar disk. Until the development of the coronagraph in the 1930s, the only observations of the faint coronal light (about as bright as the full moon) were made during total solar eclipses.

In the standard coronagraph, light spilling around an occulting disk is either passed through the narrow slit of a spectrograph to form a spectrum of the faint emission lines or through a narrow-band birefringent filter to form an image of the corona in an emission line, most often the green line of Fe XIV (5303 Å), and at times the red Fe X (6374 Å), or yellow Ca XV (5694 Å) lines. These lines have ionization temperatures of 1.5, 1.0, and 3.4×10^6 K, respectively. Figure 1-16 shows two filtergrams of the corona.

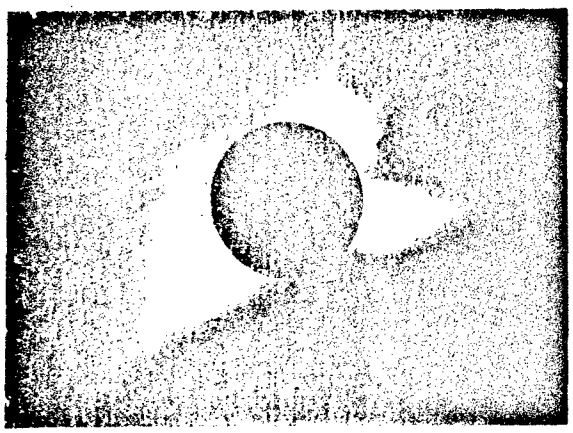
Photoelectric coronal observations from the ground are of two types. Observations of the emission-line corona made with a differential photometer at the National Solar Observatory at Sacramento Peak in New Mexico consist of circular scans around the limb at two or three heights. A sequence of such daily scans can be used to build up a picture of the low corona over a half solar rotation (Figure 1-17). Observations of the intermediate corona using light scattered by coronal electrons (also called the "K" corona) are made daily with the *K-coronameter* on Mauna Loa, Hawaii.

The corona can also be observed with radio telescopes (Chapter 11). Here the corona can be seen directly against the disk (there is no complicating "background light" as in the visible), and different heights can be probed by using different frequencies.

The discovery of coronal holes (Section 1.2.7.1), one of the major advances in solar physics in the past decade,



gure 1-14 Spectroheliograms (EUV) made with a slitless spectrograph of lines in the transition zone obtained with Skylab, 1973–74. The Magnetogram is a ground based observation.



gure 1-15. The Solar Corona. This photograph was taken at a solar eclipse with a filter that allows progressively more light to enter the camera as distance away from the solar limb increases. This greatly enhances the detail to be seen. Venus is visible on the left side (High Altitude Observatory).

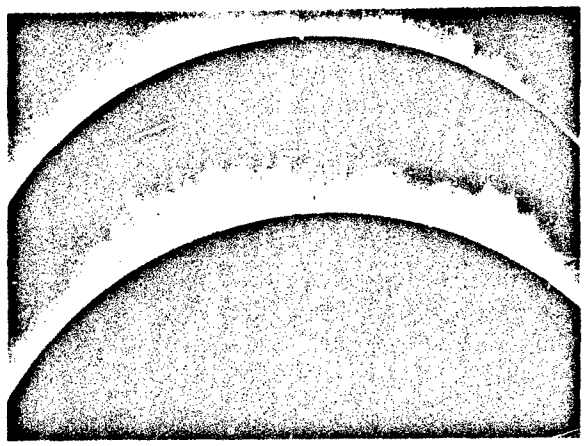
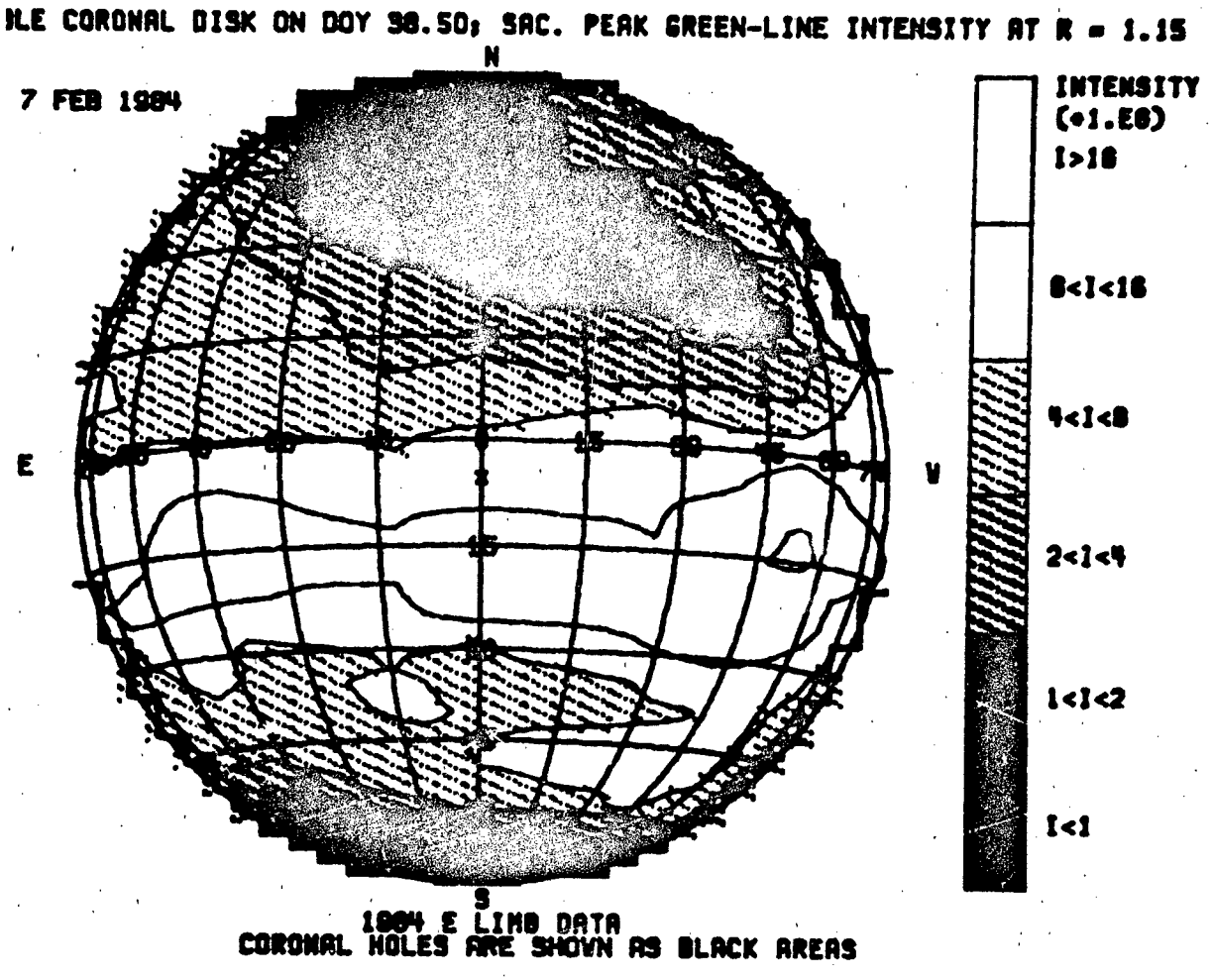


Figure 1-16 Two filtergrams of the solar limb as seen in the coronal lines Fe X, 6374 Å (above) and Fe XIV, 5303 Å (below) obtained with the new "one-shot" coronagraph at the National Solar Observatory at Sacramento Peak.



7. A "picture" of the green-line corona on 7 Feb 1984 obtained by allowing successive east-limb scans to rotate across the disk. Note the (black) polar coronal holes from which high speed solar wind streams flow. (National Solar Observatory at Sacramento Peak).

s made possible by satellite observations. The lack of a ttering and absorbing atmosphere at satellite altitudes de possible the first prolonged high-resolution solar ob- vations in soft x rays. These Skylab observations and ers revolutionized the understanding of coronal pro- ses. Astronomers also use satellite images of the corona ained in the extreme ultraviolet (EUV) with Skylab, and onagraph observations in white light and emi sion lines h from Skylab and the Solar Maximum Mission satellite MM).

Indirect observations of the corona can be made on the s in some spectral lines of helium, most notably that of 30 Å. The population of certain atomic states of chro- spheric helium is partially controlled by conditions in overlying corona, and so coronal holes are marginally ectable in this line. A daily 10830 Å patrol (Figure 1- is operated at the National Solar Observatory facility Kitt Peak, Arizona.

Another source of data concerning the corona and the near-solar interplanetary medium (IPM) is the scintillation of radio stars as observed through the corona/IPM. Semi-empirical relations between the signal fluctuation of a radio star and the turbulence generated at a given point in the IPM by solar wind flow (Chapter 3) have allowed investigators at the University of California, San Diego (UCSD), to produce low-resolution three-dimensional pictures of the solar wind flow within one astronomical unit (the distance from the earth to the sun) of the earth.

1.2 THE ACTIVE SUN

Although the quiet sun is of immense astronomical interest because it is the only star whose spatial features are readily resolvable with earth-bound telescopes, most of the sun's geophysical effects (except for the basic solar irradiance) result from solar activity. In this section we shall discuss phenomena related to the sun's magnetic field, for

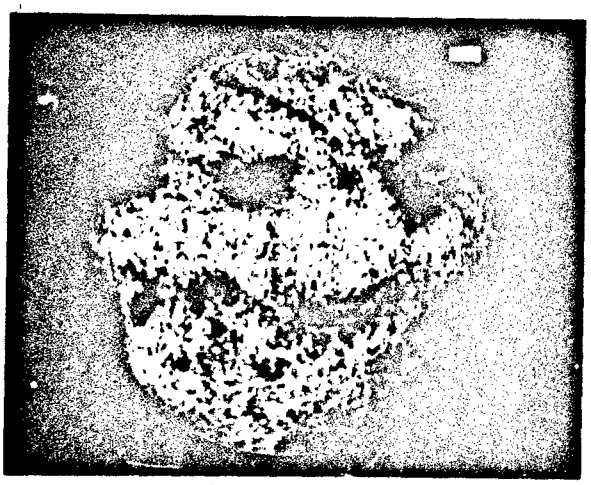


Figure 1-18. A sample 10830 Å scan obtained by the National Solar Observatory at Kitt Peak on 24 Dec 1978. Active Regions and prominences are black and coronal holes are light. Defining a coronal hole usually requires a series of pictures taken over several days as well as information about the magnetic field.

it is probably through the emergence, growth, transport and destruction of magnetic flux that the sun produces the transient behavior which causes large responses in the earth's environment.

1.2.1 The Sunspot Cycle

The most obvious manifestation of solar activity is the 11-year sunspot cycle. The *relative sunspot number R* (or Wolf or Zurich number) remains the single most important index for the general level of solar activity; it is calculated according to the formula

$$R = k(n + 10g),$$

where n is the number of individual spots visible on the solar disk, g is the number of sunspot groups (see below), and k is a station constant, or "personal equation" for the particular observatory.

Records of sunspot counts have been kept since the mid-seventeenth century (Figure 1-19). The cyclic behavior of R with a period averaging 11.4 years is the most obvious feature in the historical record, although an approximately 80-year periodicity may also be present. The anomalous period of low activity, 1645-1700 (*Maunder minimum*), is recognized as a possible reduction in the level of solar activity; its existence has been independently supported by measurements of the carbon-14 content in tree rings formed during that era.

Most sunspots occur in groups that are dominated by a leader spot on the western side of the group and one or more trailer spots to the east (Figure 1-20); the leader spot is often the largest spot within the group. Magnetic fields in large spots are typically 3000 gauss, with leader and trailer spots generally having opposite polarities. During an 11-year cycle all leader spots in the northern hemisphere of the sun have one magnetic polarity, while in the southern hemisphere this polarity is reversed. This pattern persists throughout the rise and decline of the cycle. During the next cycle, however, the entire polarity pattern is reversed. Thus, the fundamental cycle of solar magnetic activity is actually about 23 years in duration. During cycle No. 21 (maximum in 1979-80), leader spots in the northern hemisphere had positive polarity (magnetic vector pointing outward from the solar surface).

The first sunspots of a new cycle typically appear at latitudes of 20° to 25° in both northern and southern hemispheres. As the cycle progresses, additional spots appear at lower latitudes, until they are concentrated near the equator at the end of the cycle (Figure 1-21).

Major variations in coronal structure take place during the solar cycle. Several decades of observations have confirmed that green-line intensities, and to some extent the red-line intensities, follow the sunspot number quite well. In particular, averages of the green-line intensities show large maxima at midlatitudes in both hemispheres that appear near or coincident with sunspot latitudes. These belts move towards the equator with the progression of the solar cycle and finally merge.

The total variation of the green-line intensity during a solar cycle is significant. In addition, the appearance of the two yellow lines is now known to be an excellent indicator of regions which will produce large flares and perhaps even proton events.

1.2.2 The Solar Dipole Magnetic Field

The sun has a weak dipole magnetic field aligned with its axis of rotation. Because of the stronger, complex field patterns associated with spot groups, however, this dipole field is recognizable only at latitudes above 60°, where it measures about 1 gauss on the solar surface. It is believed that the dipole field, although weak in general, becomes concentrated in the midlatitudes due to the dynamo action of the sun's differential rotation. *Magnetic ropes* of strong field may then rise to the surface due to magnetic buoyancy, forming the oppositely-poled leader and trailer portions of an active region where the flux loop intersects the photosphere. This would account for the spot-polarity law discussed above. The loops may eventually rise into the outer corona where they neutralize the original dipole field and ultimately replace it with a new dipole field of opposite sign. The polarity reversal of the dipole field occurs near sunspot maximum although the spot polarities themselves remain consistent with the polarity of the dipole field that existed at the beginning of the cycle.

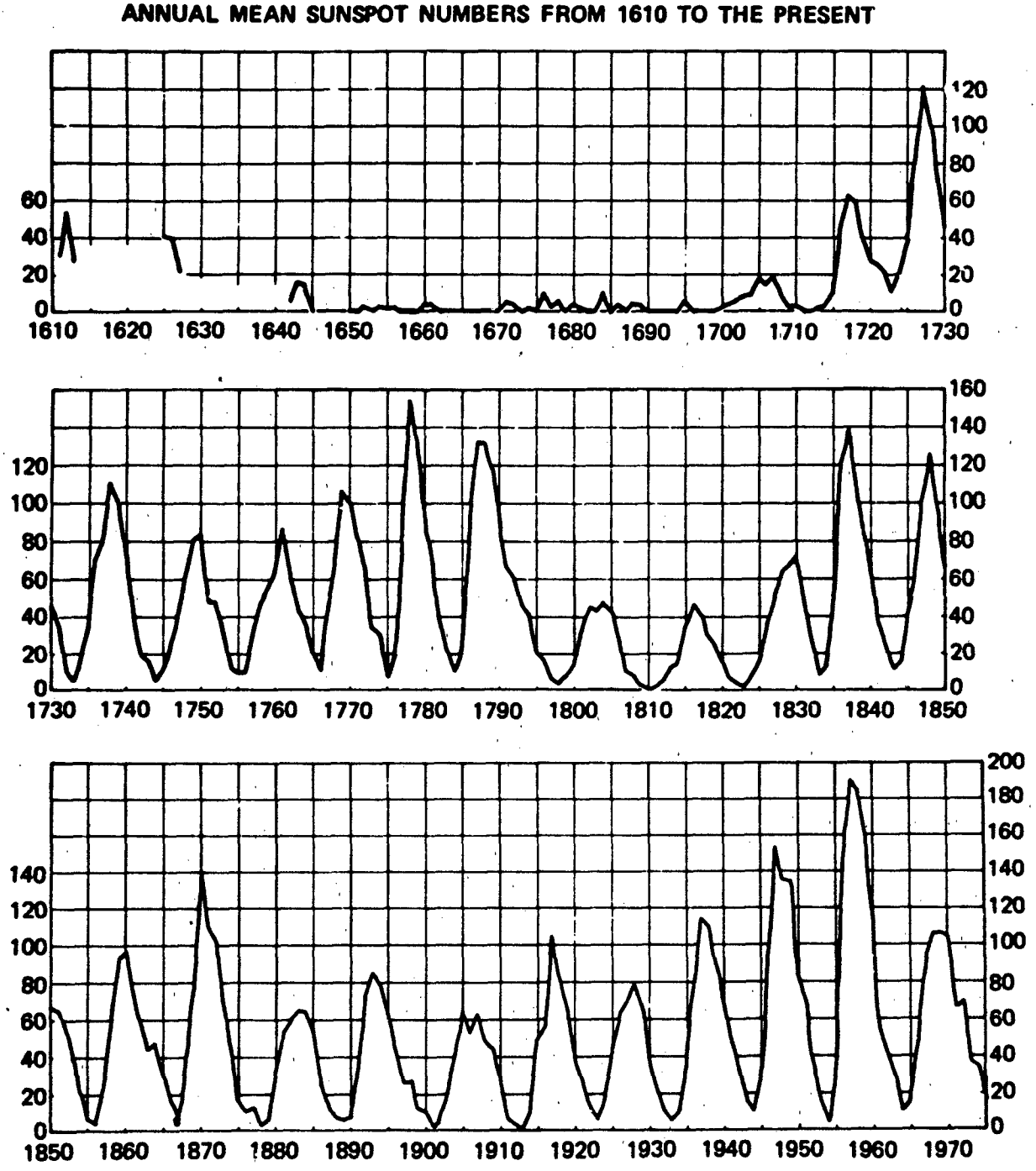


Figure 1-19. Annual mean sunspot numbers from 1610 to 1975 (From Waldmeier [1961] and Eddy [1976] et al.)

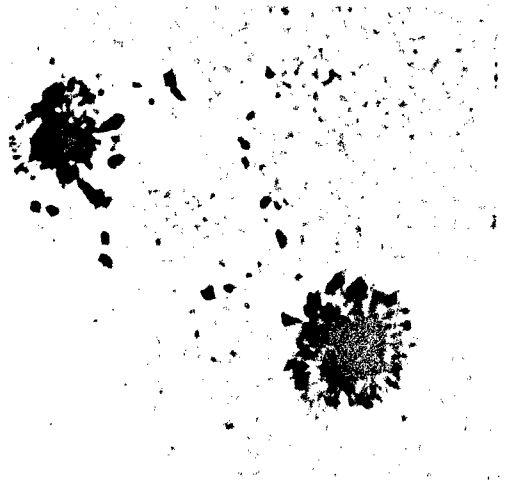


Figure 1-20. Typical bipolar sunspot group showing leader spot (right) and trailer spot. The background pattern of granulation is also visible (National Solar Observatory at Sacramento Peak).

1.2.3 The Emergence and Organization of Magnetic Flux

The solar surface abounds with both large and small-scale magnetic features. The smallest features, known as *filigree* (Figure 1-22) are bright, wiggly structures with lengths of 1 to 5 Mm and narrow widths of 100 to 200 km. They may, in fact, be smaller, but it has been impossible to date to see smaller structures with solar telescopes. Filigree reside in intergranular lanes and are visible with ground-based telescopes only on rare occasions when the atmospheric seeing allows the 1/4 arcsec resolution required to see them. Their locations are presumed to coincide with extremely fine magnetic structures occupying the granule boundaries, perhaps analogous to the somewhat larger flux tubes of the supergranule network (Section 1.1.3.2). Measurements show that these small flux tubes have intense magnetic fields with strengths between 1000 and 2000 gauss.

1.2.3.1 Active Regions. The development of a large-scale sunspot-bearing *active region* usually begins with an increase in the number and intensity of magnetic flux tubes in the photospheric network. In response to this, a brightening in the overlying chromosphere develops. These brightenings, or *plages* are the chromospheric counterparts to the sunspot group that later becomes visible in the photospheric layers. The corona responds also, producing a large area of bright emission visible, for example, in x-ray and radio images (Figure 1-23). Within a day of plage formation a small leader spot emerges which continues to grow in size for several days. The trailer spots usually form within a day of the leader, and many smaller spots often form throughout the entire region (Figure 1-20). Generally a sunspot begins life as a small *pore* which darkens as a con-

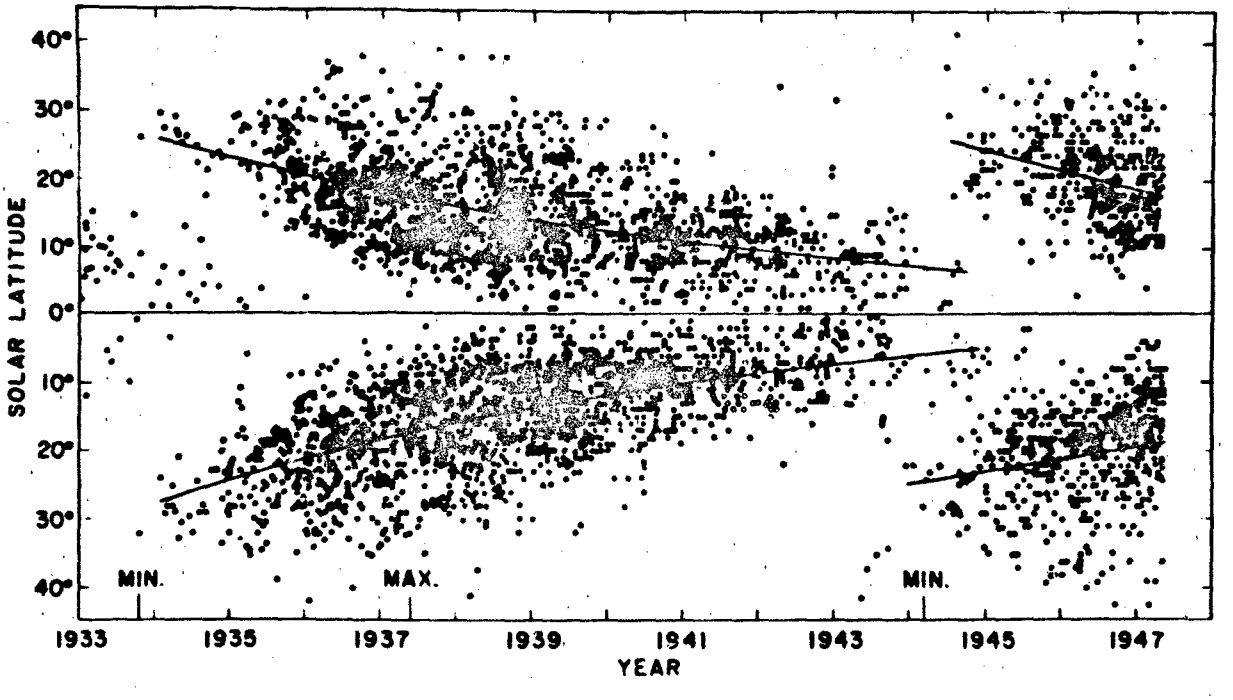


Figure 1-21. Typical "butterfly" diagram showing variation in distribution of sunspots with time. During a sunspot cycle the distribution migrates toward the solar equator (Mount Wilson and Las Campanas Observatories).

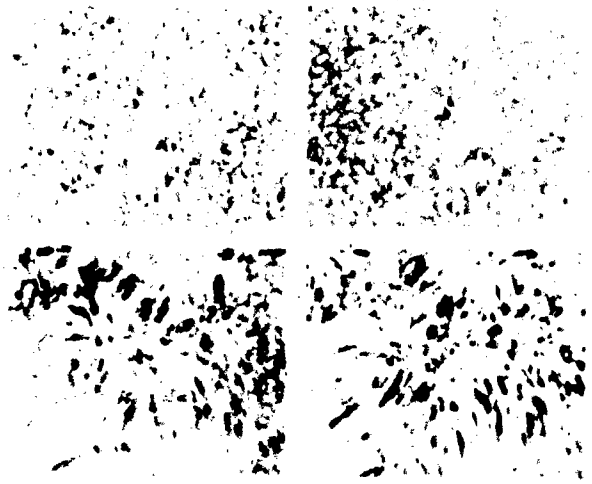


Figure 1-22. The solar filigree are visible as bright, wiggly structures lying in the intergranule zones. This series of filtergrams was taken in different wavelengths (clockwise from upper left): continuum, $H\alpha - 2$ Å, $H\alpha + 2.8$ Å, $H\alpha - 7.8$ Å (National Solar Observatory at Sacramento Peak).

sequence of temperature reduction. The temperature drop is probably caused by the presence of the magnetic field, which may inhibit the supply of convective energy to the photosphere at that location, and/or may enhance the photospheric dissipation of energy at that point. The largest spots attain areas of 3×10^{10} km^2 and temperatures 2000 K lower than the surrounding photosphere.

Sunspot groups usually reach maturity within several days after which they begin a slow decline. The leader spot

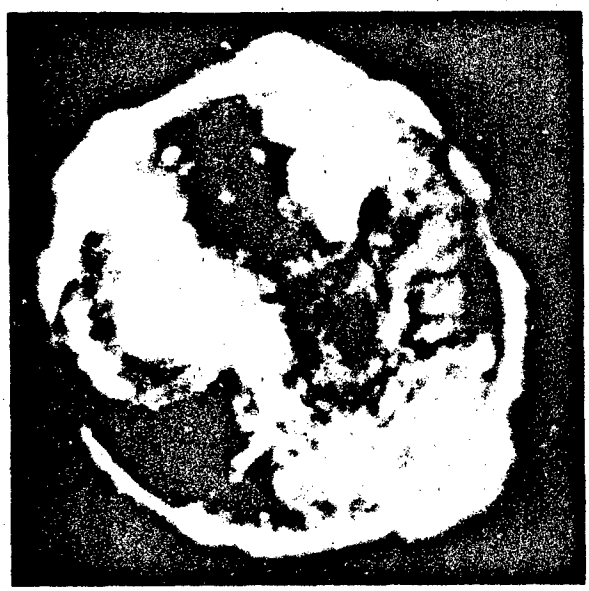


Figure 1-23. Soft x ray photograph of the sun, showing the bright loop structures of active regions (John Davis, American Science and Engineering, Inc.).

Figure 1-24. Complex sunspot group (National Solar Observatory at Sacramento Peak).

is the last to dissolve and occasionally may last for several months. The chromospheric plage persists even after all spots have disappeared.

A small number of active regions violate the simple polarity rule, and instead form complex magnetic structures with reversed or mixed polarities (Figure 1-24). These groups are responsible for nearly all the highly active solar phenomena (Section 1.2.6).

1.2.3.2 Ephemeral Regions. A significant portion of the total magnetic flux on the solar surface resides in *ephemeral active regions*. These are miniature bipolar regions of size ~ 30 Mm with average lifetimes of 12 hours. Although not usually associated with sunspots of their own, they are, nevertheless, legitimate active regions in the sense that they affect the chromospheric structures and develop visible plages. There may be several hundred such regions on the sun at any given moment, although like spots their numbers vary with the 11-year cycle. Their magnetic orientation, however, appears to be far more random than the leader and trailer polarities of sunspot groups.

1.2.4 Prominences

Prominences are large clouds of gas high above the chromosphere, visible as bright flame-like objects with emission spectra against the sky background at the limb, or as long rope-like *dark filaments* with absorption spectra against the solar disk. *Quiescent prominences* (Figure 1-25) are not associated with active regions, but rather form boundaries between very large scale weak magnetic features

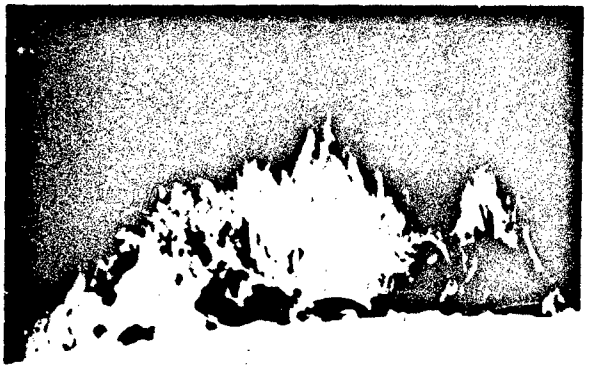


Figure 1-25. Quiescent prominence seen on the solar limb (National Solar Observatory at Sacramento Peak)

Figure 1-26a. Surge prominence associated with a flare seen near the solar limb (National Solar Observatory at Sacramento Peak)

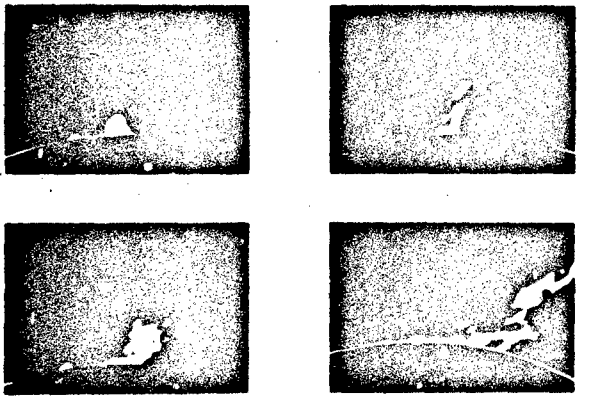


Figure 1-26b. Sequence showing development of a spray prominence associated with a flare (National Solar Observatory at Sacramento Peak)

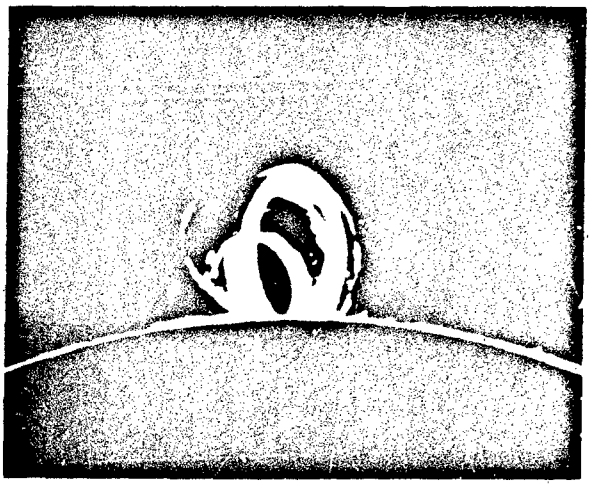


Figure 1-26c. Post flare loop prominence (National Solar Observatory at Sacramento Peak)

in the quiet sun. They are the largest visible structures on the sun, sometimes reaching lengths exceeding a solar radius. Occasionally, quiescent prominences suddenly become active, erupt and blow off to great heights in the corona.

Other types of prominence-related phenomena include *surges*, *sprays* and *loops* (Figure 1-26). These structures are often associated with flares as discussed below. They are, nevertheless, similar to quiescent prominences in the sense that their temperatures (10,000 K to 20,000 K) remain chromospheric even though their structures extend to coronal heights where the ambient temperatures are 100 times higher than their own.

1.2.5 Coronal Magnetic Fields

Careful observations also show a great spatial variation in the nature, if not the magnitude, of the magnetic field that threads its way through the corona. Bright regions tend to correlate with *closed field regions* where models or observations indicate that the magnetic lines of force emerge from and re-enter the photosphere over scales less than a solar radius (Figure 1-23). Conversely, dark regions correlate with *open field regions*, mysterious magnetic monopoles whose field lines may re-enter the sun again only after some convoluted journey through interplanetary space.

1.2.6 Flares

A *flare* is a transient, localized release of energy usually occurring in and above an active region. Flares produce greatly enhanced emissions spanning the entire electromagnetic spectrum and may accelerate particles to energies as high as 50 GeV. The flare eruption affects the corona, chro-

mosphere, and probably even the upper photosphere. The energy released by a flare ranges from 10^{29} to 10^{32} erg; durations vary from 1 to 100 minutes.

Flare *importance* has traditionally been assigned according to the flare's maximum area as observed in the Hα line (Table 1-4). Following the importance class, a letter code (F, N or B) is given to designate the relative brightness of the flare as faint, normal, or bright. In recent years flare importance has been supplemented by the soft x ray classification of the flare (Table 1-5). These x ray classes are only roughly correlated with the optical importances.

Table 1-4. Flare optical importance

Importance	Area (Heliographic Square Degrees)
S (Subflare)	2.06
1	2.06 to 5.15
2	5.15 to 12.4
3	12.4 to 24.7
4	>24.7

Table 1-5. Flare x ray classification

Class	1-8Å x ray flux at 1AU (w/m^2)
C	$10^{-6} \leq E < 10^{-5}$
M	$10^{-5} \leq E < 10^{-4}$
X	$E \geq 10^{-4}$

Historically, flares have been studied in the Hα line, which is formed in the chromosphere. Observations in the EUV and x ray regions, however, have shown that flares may be more correctly characterized as coronal phenomena with the primary energy release probably occurring in the low corona or transition region. It is not surprising, therefore, that flares are best seen in wavelengths originating in the corona. Flare images in soft x rays (1-10 Å) and XUV lines (200-300 Å) obtained by the Skylab Mission show that the basic flare structure is a low-lying coronal loop on the order of 20-50 Mm in length (Figure 1-27), wherein temperatures rise from the normal 1×10^6 K to about 20×10^6 K. Complex flares often reveal the presence of many such loop structures. The loop structures are formed by magnetic fields rooted in the deeper photospheric layers. The chromospheric flare (Figure 1-28), which is probably a secondary response to the coronal energy source, is primarily confined to the footprints of these loops forming intense flare *kernels*, or an arcade of loops may be involved forming a pair of serpentine, parallel *ribbons*. Magnetograph observations reveal opposite magnetic polarities for the two members of a kernel or ribbon pair. Photospheric field strengths for the flare loop footprints are typically 1000-2000 gauss

Figure 1-27. Soft x-ray loop structures in a flare (American Science and Engineering, Inc.).

Figure 1-28. Hα filtergram showing the chromospheric component of a "two-ribbon" flare (National Solar Observatory at Sacramento Peak).

for highly energetic events. Only rarely does the flare disturbance affect layers deep enough in the atmosphere to produce optical continuum (so-called *white-light flares*).

About half of all flares show a sudden or explosive rise to maximum. This *impulsive phase* is most clearly defined by hard x rays (photon energies 10 KeV - 1 MeV) or microwaves (wavelengths 1-30 cm) and may last only a minute in the case of simple flares. Soft x ray and Hα emissions, even in impulsive flares, are usually less abrupt in their onset, and often continue to rise after the impulsive components are in decline. Complex flares show multiple impulsive phases, with each impulse apparently being associated with a particular point in the flare. The impulsive components probably originate in a process which accelerates electrons to energies exceeding 10 KeV, producing hard x rays via electron-proton bremsstrahlung and microwaves via gyrosynchrotron radiation in the ambient magnetic fields (Chapter 11). Type III meter-wave radio bursts (Chapter 11), which may also occur during the impulsive phase, have been attributed to the outward escape of some of these electrons along open field lines. Recently obtained hard x ray images show the early stages of the hard x ray bursts occurring at loop footprints similar to the first optical kernel brightenings. This suggests a common energizing source in the loop, from which electrons are guided downward into the chromosphere along the field lines. It is not clear whether the population of electrons producing the impulsive emissions is best characterized as thermal or non-thermal. Certainly the hard x ray kernels and Type III radio bursts suggest a non-thermal directivity in the particles. If the energy distribution is thermal, then temperatures exceeding 10^8 K, and possibly even 10^9 K, would be required in order to produce the observed hard x ray continuum. The presence of protons with energies exceeding 30 MeV, as evidenced by nuclear gamma-ray lines in the impulsive phases of large flares, lends additional support to acceleration processes which produce non thermal distributions.

A variety of phenomena occur in response to violent heating or magnetic rearrangement associated with flares. *Surges* (Figure 1-26) are spikes or filaments of chromospheric material sent outward (velocity ~ 100 km/s) by a sudden change in the distribution of magnetic field lines in the active region. Surprisingly, many surges are not associated with flares, although conversely a large fraction of flares seem to generate surges. The material pushed upward in a surge usually descends along its original trajectory. The latter characteristic distinguishes surges from *sprays* (Figure 1-26) which are high velocity (exceeding the solar escape velocity, 670 km s) explosive ejections associated with certain flares.

The intense heating in flares usually "evaporates" the chromosphere near the bases of magnetic loops. The heated gas then rises to fill the loop with soft x ray-emitting plasma. Optically, these *post-flare loops* or *loop prominences* are often observed to form in the late stages of large two-ribbon flares (Figures 1-26). The enhanced soft x ray and ultraviolet flare emissions are responsible for a variety of immediate geophysical effects collectively known as *sudden ionospheric disturbances* (SIDs) (Chapter 10). These disturbances are a result of increased ionization at several levels in the ionosphere.

Shock waves produced by explosive heating propagate upward through the corona where they accelerate electrons and protons to relativistic energies. The interaction of shocks, particles, and magnetic fields generates Type II and Type IV meter-wave radio bursts (Chapter 11). The outward-propagating shock (velocity ~ 1000 km/s) is apparently also the origin of solar protons (energies ≥10 MeV) that produce terrestrial *polar cap absorption* (PCA) and solar cosmic ray *ground-level events* (GLE), while the shock wave itself arriving at the earth in about two days is a cause of the sudden commencement phase of geomagnetic storms (Chapter 4).

The origin of flares, particularly the processes of energy storage and release as well as the triggering instability, remains a difficult theoretical problem. Obviously, the local magnetic field is involved. Recent theories point to current-carrying magnetic configurations in which field lines reconnect and simplify. The problem is observationally difficult because the flare plasma parameters cannot be tracked in sufficient spatial and temporal detail. Without the aid of *in situ* measurements we are forced to rely upon observations of secondary effects which detect the flare energy only after disordering processes have already occurred.

1.2.7 Coronal Activity

1.2.7.1. Coronal Holes and Solar Wind. During Skylab, pictures of the sun were taken regularly with a soft x-ray telescope [Timothy et al., 1975], while at the same time 27-day recurrent geomagnetic disturbances were occurring. It was noticed that when areas dark in x rays crossed the central meridian near the sub-earth point, a geomagnetic disturbance occurred a few days later. This relationship persisted, and thus the source of the disturbances, the long-puzzling solar *M-regions*, was finally discovered, solving a riddle that had existed for decades. These dark areas named coronal holes (Figure 1-29) were shown to be sources of high-speed solar wind streams (HSS) that eventually impacted on the magnetosphere and caused the disturbances. For more information on HSS see Chapter 3.

Subsequent work on the nature of coronal holes has shown that they are somewhat cooler (by 500,000 K) than surrounding areas and up to an order of magnitude less dense. See Table 1-3. They also coincide with large areas of open magnetic field. It is thus hypothesized that an open field provides an unrestricted path for the propagation of solar wind into interplanetary space resulting in high solar wind speeds that deplete the matter in the holes and carry away internal energy (thus the lower density and temperature).

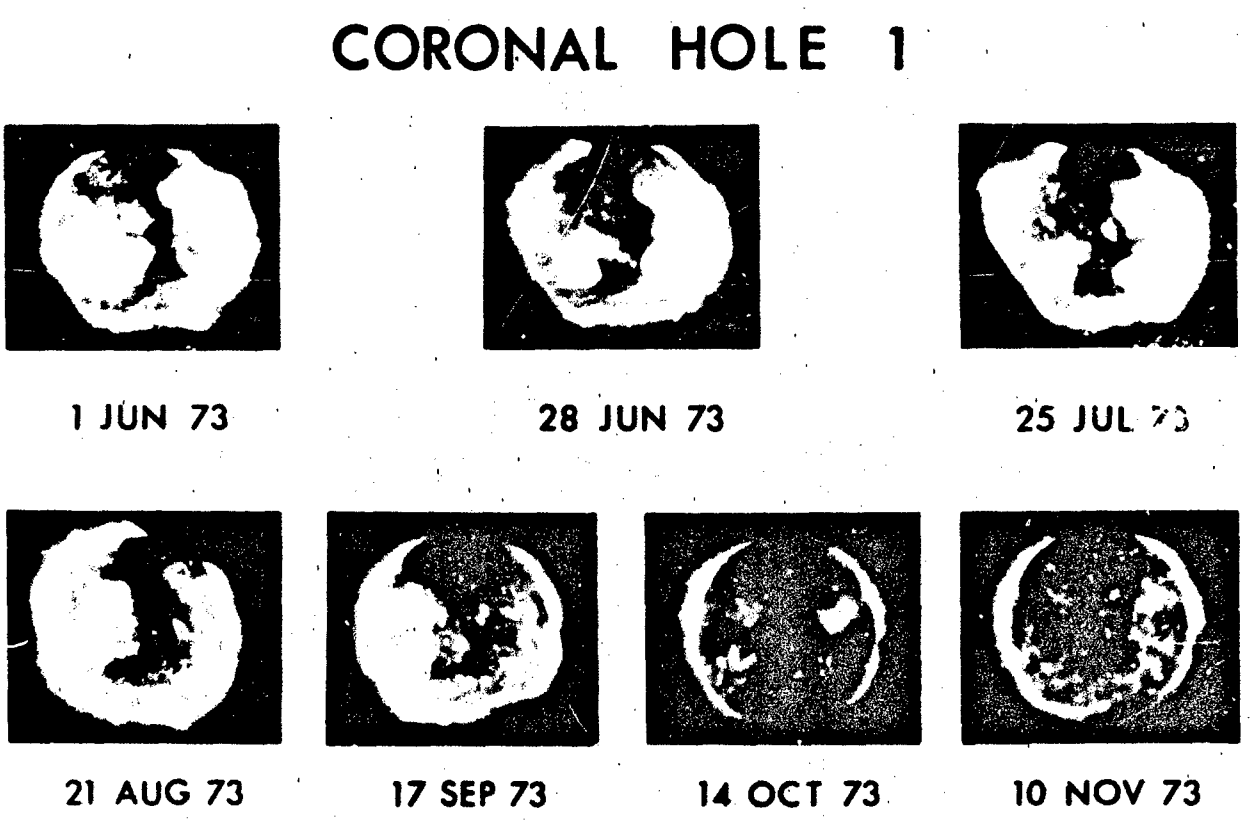


Figure 1-29. The first coronal hole ever observed in extensive detail. This series of pictures from Skylab shows the long-lived nature of the black coronal hole in the center of the disk as seen over several solar rotations (American Science and Engineering, Inc.).

1.2.7.2 Streamers. Earlier speculation on the source of M-regions had centered on *streamers*, which are long, radially-oriented bright regions of the corona. Recent research has shown that streamers can be detected even at the orbit of the earth. The signature of streamers [Borrini et al., 1981] consists of low solar wind velocity (on the order of 350 km/s), an increase in density and a decrease in the He/H ratio. These events appear to be coincident with *sector boundary* passages, which are sudden reversals of the interplanetary magnetic field (IMF) that often separate HSS from one another (Section 1.2.7.3).

1.2.7.3. The Solar Current Sheet. In the upper corona, magnetic fields tend to be mostly radially oriented and to reflect the polarity of the large-scale average field underlying them. Opposing fields in such large quasi-unipolar regions tend to reconnect at low altitudes, resulting in a single polarity remaining at high altitude drawn out from the sun by the solar wind. As the fields are stretched, they remain attached to the rotating sun "beneath" them, and each field line becomes a spiral (Chapter 3, Figure 2).

Although the sun tends to have a dominant polarity in each hemisphere, polarities from one hemisphere often intrude partially into the other. Thus as the sun rotates and this complicated pattern of fields is drawn out into space, a very complex neutral sheet or current sheet separating the two polarities develops (Figure 1-30).

As this pattern sweeps by the earth (or any other point near the ecliptic), it results in several days of one IMF polarity, followed by a sudden reversal and several days of

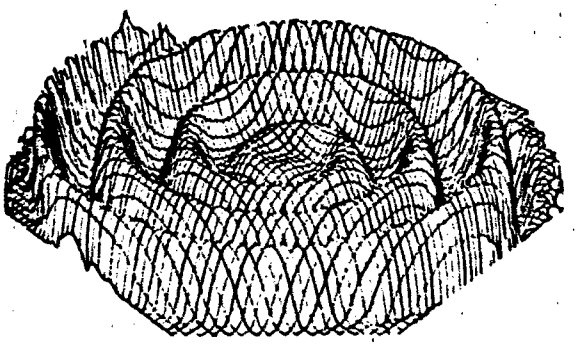


Figure 1-30. Schematic illustation of the warped solar current sheet. This undulating surface divides regions in interplanetary space having magnetic field pointing away from or toward the sun. The region covered by this diagram is approximately 20 astronomical units [Hakamada and Akasofu, 1982].

the opposite polarity. This yields a regular pattern of *magnetic sectors* separated by sector boundaries, which as we saw above are the signatures of coronal streamers sweeping by the earth. Distortions of the IMF also occur as the flare streams overtake slower streams and shocks are generated (Chapter 3). For further information refer to Hakamada and Akasofu [1982].

1.2.7.4 Transients. Skylab observations demonstrated the existence of a phenomenon never before observed outside of radio frequencies in the upper corona: the *coronal transient*. Fast motions had been observed in the lower corona, but the limited height range observable with ground-based coronagraphs gave no indication of their enormous size. Skylab showed huge bubble-shaped disturbances rising through the corona and expanding until their size approached that of the sun itself (Figure 1-31). As they enter the interplanetary medium, the transients may generate a shock wave that sweeps out into space. Their origin has been traced to surges or prominence eruptions.

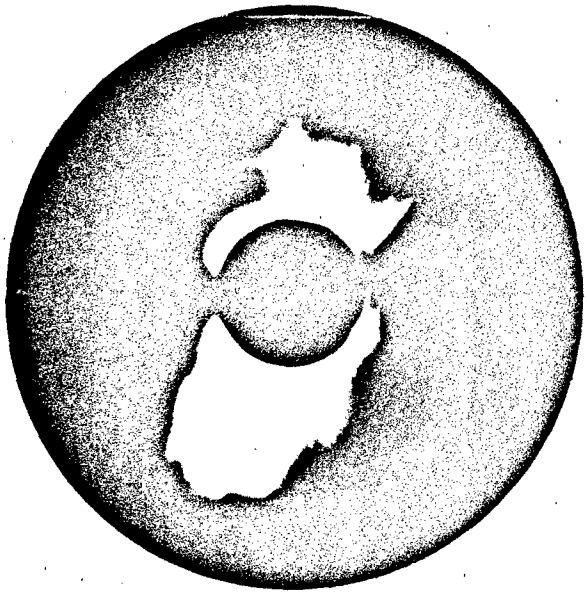


Figure 1-31 White-light photograph of a coronal transient obtained with the High Altitude Observatory white-light coronograph experiment on Skylab. The black circle is not the sun but the occulting disc of the instrument used to block out photospheric light. The size of its image is two solar radii. This particular transient traveled outward from the sun at about 500 km/s.

1.2.7.5 Coronal Induced Disturbances. Direct effects of the corona on communications and electronic systems on and near the earth come mainly through the very short-lived electromagnetic and particle emissions connected with solar flares. However, minor effects may occur due to the variation of radiation from the very highly-ionized elements of the corona. As solar active regions and their associated streamers grow and decay, rotate across the disk, and change in number with the solar cycle, the amount of high-energy radiation falling on the ionosphere varies slowly. Parameters such as the Maximum Usable Frequency (MUF) for long-range communications are partially controlled by the structure of the corona, and knowledge of this structure aids in the prediction of such parameters.

Indirect effects of the corona are manifested through propagation of coronal parameters to the earth by means of the solar wind (Chapter 3). Since solar wind conditions are directly controlled by the corona, the velocity, density, and magnetic field direction and magnitude of the solar wind are all determined at or near the solar surface. It is these parameters that determine whether a geomagnetic disturbance will occur. Geomagnetic disturbances generally indicate the onset of a wide variety of particle and electromagnetic-radiation effects in the geosphere, which together or separately have wide-ranging effects on a number of systems. In addition, Altrock [1980] has found that even if conditions in the solar wind are not conducive to the onset of a geomagnetic disturbance, changes in upper-atmospheric density may occur that affect satellite orbits. The change may be connected with the passage of streamers at the earth. Thus, in order to predict and/or protect against such detrimental effects, it is essential to have a thorough observational and theoretical understanding of processes in the solar corona.

1.2.8 Solar Variability and Climate

While the spectacular and explosive events associated with solar flares have immediate consequences upon the geophysical environment, the radiant energy output of the sun is a fundamental factor controlling the earth's climate. Although we do not yet have a firm, quantitative measure of the sensitivity of the earth's climate and weather to changes in the solar irradiance, climatologists estimate that the earth's mean surface temperature would change by about 1 K in response to a sustained 1% change in the total solar irradiance. The effects of a small change in the mean temperature could be large. An increase of two or three degrees, for example, would suffice to melt the polar icecaps, raise the sea level by several hundred feet and flood large areas of present coastland. On the other hand, a one degree decrease would have a significant adverse impact on the major grain producing regions of Europe and North America, and a decrease of perhaps ten degrees would probably result in total glaciation of the earth.

Fortunately, the total solar irradiance appears to be relatively constant, although this fact has proven rather difficult to verify. Ground-based irradiance measurements require

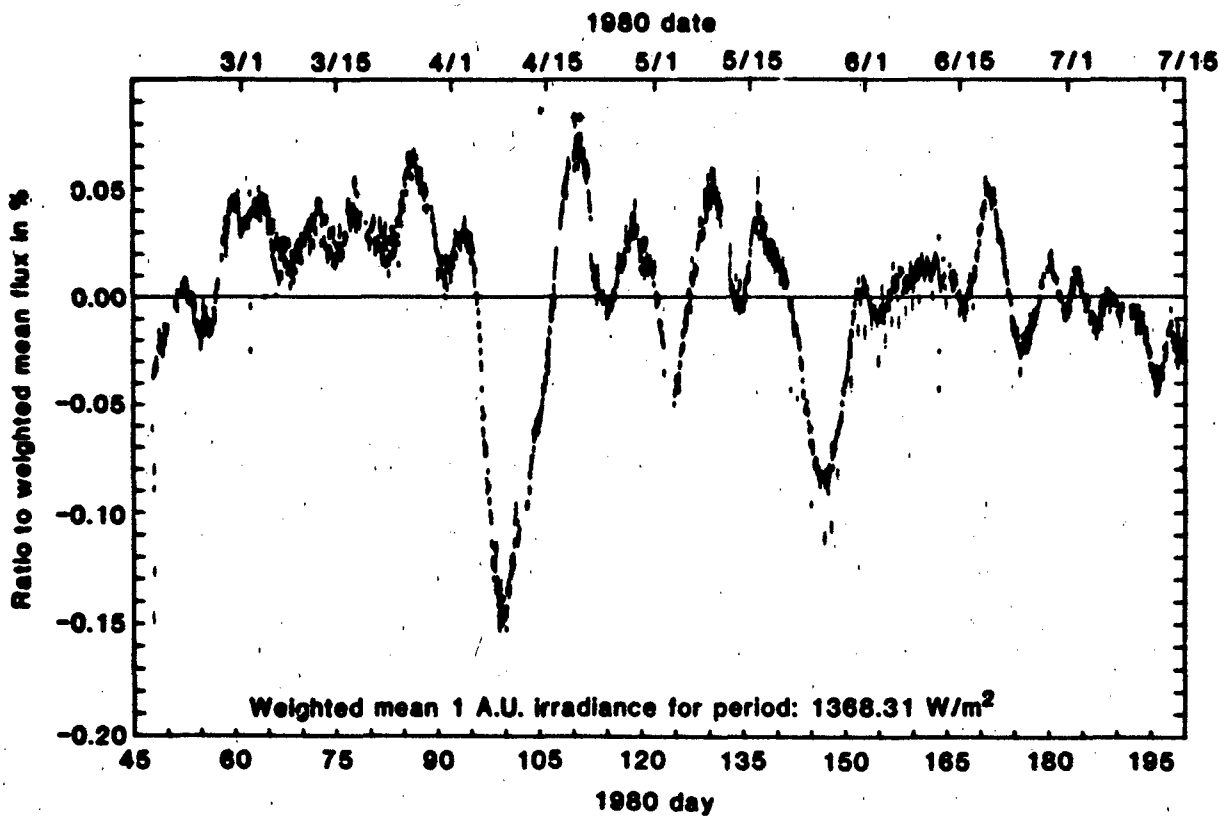


Figure 1-32. Total solar irradiance at 1 AU, shown as a percentage variation about the weighted mean for the first 153 days of SMM.

large corrections for atmospheric absorption, which is both highly selective and variable. Consequently, such measurements cannot achieve an accuracy better than ± 0.2%. Measurements from spacecraft have until only recently been limited to a comparable accuracy due to detector drift problems caused by exposure to the harsh space environment.

The variability of the solar irradiance in the ultraviolet and radio portions of the spectrum is well-established (Chapters 2 and 11) and is clearly associated with various manifestations of solar magnetic activity such as flares. Furthermore, the SMM radiometer measurements have clearly detected short-term fluctuations of 0.1-0.2% in the total irradiance (Figure 1-32) that correlate with the passage of large sunspot groups across the face of the sun. However, longer term trends have been more difficult to establish. Because there is no long term information from space-borne instruments, ground-based measurements still provide the best information concerning such trends. There is some marginal evidence, based on measurements of light reflected from solar system bodies as well as direct radiometry, for a slight increase of about 0.4% in total irradiance between 1968 and 1978. On the other hand, other ground-based radiometric measurements show no evidence for any variation exceeding 0.2% over the 30 year interval from 1923 to 1952. For more information refer to White [1977].

1.2.9 Predictions of Geophysical Disturbances

1.2.9.1 Flare Prediction. Flare-induced geophysical disturbances adversely affect many communications, navigation, and other electronic systems. Therefore daily predictions for the level of flare activity are issued by the joint USAF/NOAA (National Oceanic and Atmospheric Administration) Space Environment Services Center (SESC) located in Boulder, Colorado, working with Air Force Global Weather Central (GWC), Offutt Air Force Base, Nebraska. Present techniques for preparing solar forecasts are based on empirical relationships between flare occurrence and other active region parameters such as sunspot group type, magnetic complexity, and region evolution characteristics.

Most of the data used at SESC is provided by the Solar Electro-Optical Network (SEON) operated by the USAF Air Weather Service. SEON consists both of optical telescopes (SOON - Solar Observing Optical Network) and radio telescopes (RSTN - Radio Solar Telescope Network). The selection of the final SEON site (southeastern Italy) was announced in September 1982. SOON sites are located at five (four sites already operational in 1980) widely-spaced longitudes, providing continuous 24-hour monitoring capability. SOON provides data on solar magnetic configu-

itions, sunspot types and locations, and chromospheric haracteristics. Four SEON sites compose RSTN, which onitors solar background radio emission (10-cm flux) as ell as the type and sizes of flare-associated radio bursts. EON issues immediate reports of flare occurrences, importance, and locations for use at SESC and GWC.

The near-real time acquisition of solar data by SEON nd other observatories, as well as present and planned tellite monitors, is making available daily a large number f parameters for each active region on the solar disk. This ass of data lends itself to computer programs using sohisticated statistical procedures to maximize the predictive capability of the data. One technique, using a variation of multivariate discriminant analysis (MVDA) developed at AFGL, has resulted in an objective flare forecast superior to the presently-used, subjectively-derived forecasts.

1.2.9.2 Coronal Disturbance Prediction. SESC also predicts geomagnetic disturbances due to coronal fluctuations as part of its daily program. Primary data sources come from the National Solar Observatory's Sacramento Peak coronal photometer and Kitt Peak HeI 10830 Å patrol, and the UCSD's radio scintillation program (Section 1.1.3.4).

REFERENCES

Allen, C.W., *Astrophysical Quantities*, Athlone Press, London, 1973.

Altrock, R.C., "Anomalous Satellite Drag and the Green-Line Corona," in *Solar-Terrestrial Predictions Proc.*, edited by R.F. Donnelly, NOAA, Boulder, **4:** E1-E4, 1980.

Athay, R. Grant (ed.), *The Solar Chromosphere and Corona: Quiet Sun, Astrophysics & Space Science Library*, **53**, D. Reidel, Dordrecht, Holland, 1976.

Billings, D.E., *A Guide to the Solar Corona*, Academic Press, New York, 1966.

Bonnet, R.M., "Contribution of OSO-8 to our Knowledge of the Chromosphere and Transition Region," *Space Sci. Rev.*, **29:** 131–200, 1981.

Borrini, G., J.T. Gosling, S.J. Bame, W.C. Feldman, and J.M. Wilcox, "Solar Wind Helium and Hydrogen Structure Near the Heliospheric Current Sheet: A Signal of Coronal Streamers at 1 AU," *J. Geophys. Res.*, **86:** 4565–4573, 1981.

Bray, R.J. and R.E. Loughhead, *Sunspots*, Wiley, New York, 1965.

Bray, R.J. and R.E. Loughhead, *The Solar Chromosphere*, Chapman & Hall, London, 1974.

Christensen-Dalsgaard, J. and D.O. Gough, "Towards a Heliological Inverse Problem," *Nature*, **259:** 89–95, 1976.

Donnelly, R.F. (ed.), *Solar Terrestrial Predictions Proceedings*, Vol. 1-4, U.S. Dept. of Comm., NOAA/ERL, Boulder, 1979–80.

Eddy, J.A. (ed.), *The New Solar Physics*, Westview Press, Boulder, 1978.

Eddy, J.A., P. A. Gilman, and D. E. Trotter, "Solar Rotation During the Maunder Minimum," *Sol. Phys.*, **46:**3–14, 1976.

Gibson, E.G., *The Quiet Sun*, NASA SP-303, NASA, U.S. GPO, Washington, D.C., 1973.

Gingerich, O., R.W. Noyes, and W. Kalkofen, "The Harvard-Smithsonian Reference Atmosphere," *Sol. Phys.*, **18:** 347–365, 1971.

Gough, D.O. (ed.), "Problems in Solar and Stellar Oscillations," Proc. 66th IAU Colloq., Crimean Astrophys. Obs., USSR, 1-5 Sep 1981, *Sol. Phys.*, **82**, 1983.

Hakamada, K. and S.-I. Akasofu, "Simulation of Three-Dimensional Solar Wind Disturbances and Resulting Geomagnetic Storms," *Space Sci. Rev.*, **31:** 3–70, 1982.

Hill, H.A., "Observational Evidence for Global Oscillations of the Sun: A Review," *Lecture Notes in Physics*, 125, Nonradial and Nonlinear Stellar Pulsation Workshop, Tucson, March 12–16, 1977, edited by H. Hill & W. Dziembowski, Springer-Verlag, New York, 1979.

Hill, H. A., R. T. Stebbins, and T. M. Brown, *Proceedings Fifth International Conference Atomic Masses and Fundamental Constants*, Gorden and Breach, New York, 1975.

Howard, R., and B.J. LaBonte, "The Sun is Observed to be a Torsional Oscillator with a Period of 11 Years," *Astrophys. J.*, **239:** L33–L36, 1980.

Jordan, S. (ed.), *The Sun as a Star*, NASA SP-450, 1981.

Kuiper, G.P., *The Sun*, University of Chicago Press, Chicago 1952.

November, L.J., J. Toomre, K.B. Gebbie, and G.W. Simon, "The Detection of Mesogranulation on the Sun," *Astrophys. J.*, **245:** L123–L126, 1981.

Rhodes, E.J., R.K. Ulrich, and G.W. Simon, "Observations of Non-Radial P-Mode Oscillations," *Astrophys. J.*, **218:** 910, 1977.

Sturrock, P. (ed.), *Solar Flares Skylab Workshop*, University of Colorado Press, Boulder, 1980.

Svestka, Z., *Solar Flares*, D. Reidel, Dordrecht, Holland, 1976.

Timothy, A.F., A.S. Krieger, and G.S. Viana, "The Structure and Evolution of Coronal Holes," *Sol. Phys.*, **42:** 135–156, 1975.

Vernazza, J.E., E.H. Avrett, and R. Loeser, "Structure of the Solar Chromosphere III. Models of the EUV Brightness Components of the Quiet Sun," *Astrophys. J. Suppl.*, **45:** 635–725, 1981.

Waldmeier, M. "The Sunspot Activity in the Years 1610–1960," Schulthess, Zurich, 1961.

White O.R. (ed.), *The Solar Output and its Variation*, Colorado Associated University Press, Boulder, 1977.

Wilcox, J.M. and R. Howard, "Differential Rotation of the Photospheric Magnetic Field," *Sol. Phys.*, **13:** 251–260, 1970.

Willson, R.C., S. Gulkis, M. Janssen, H.S. Hudson, and G.A. Chapman, "Observations of Solar Irradiance Variability," *Science*, **211:** 700–702, 1981.

Zirker, J.B. (ed.), *Coronal Holes and High Speed Wind Streams*, Colorado Associated University Press, Boulder, 1977.

Chapter 2

SOLAR ULTRAVIOLET IRRADIANCE

Section 2.1 **L.A. Hall**
Sections 2.2–2.8 **L.J. Heroux and H.E. Hinteregger**

.1 SOLAR UV IRRADIANCE IN THE STRATOSPHERE

.1.1 Solar Spectrum Between 2000 and 3000 Å

Solar radiation incident on the earth's atmosphere is sorbed in the spectral range 2000-3000 Å by stratospheric one in the Hartley band. This absorption is an almost uctureless continuum having a maximum near 2550 Å d falling almost symmetrically on both sides. The abrption of solar radiation in this range provides the energy r stratospheric photochemistry and is the major part of the at input which produces the local temperature maximum the stratopause around 50 km.

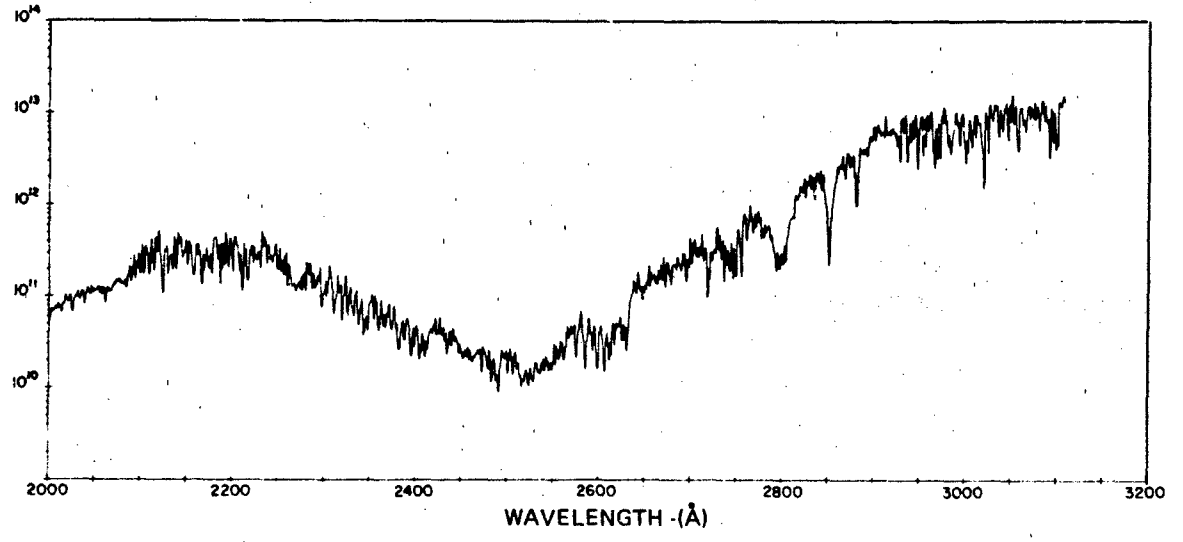
Measurements of the 2000-3000 Å spectrum in the stratphere can be made up to about 40 km by means of high-titude balloons. Only about 3% of the total ozone column nsity remains above 40 km; but at 2550 Å where the sorption coefficient of ozone peaks, only 3% of the indent solar radiation penetrates to 40 km at local noon and ° latitude. This can be seen in Figure 2-1, where the bowlaped depression between 2900 and 2200 Å is caused by one absorption (see Chapter 21).

Table 2-1 gives the solar spectrum in 1-angstrom averages from a measurement made on 19 April 1978 at Holloman AFB, N.M. The data in Table 2-1 were taken at a balloon float altitude of 41.4 km at 1022 h local time. The solar zenith angle was 29.6° and the ozone column density above 41.4 km was about 2.7×10^{17} molecules/cm^2. Further details and a description of the instrumentation and its calibration are given by Hall [1981].

2.2 SOLAR CONSTANT AND SPECTRAL IRRADIANCE

2.2.1 Solar Constant

The solar constant S is the total solar radiative energy at the top of the earth's atmosphere corrected for the mean sun-earth distance. The solar constant or total solar irradiance expressed in units of watts per square meter (W/m^2) is the largest contributor to the earth's energy budget. There have been several measurements of the solar constant as summarized by Frohlich [1977]. The most probable value for S proposed by Frohlich is S = 1373 ± 20 W/m^2.



ure 2-1. The solar irradiance in 1-Å intervals at 39.5 km in the stratosphere. Measurements made on 21 April 1977 with a solar zenith angle of 31° [Hall, 1981].

le 2-1. Solar irradiance at 40 km.

One Angstrom averages centered on the wavelength in units of 10^{11} ph cm^{-2} s^{-1} $Å^{-1}$										
Å	0	1	2	3	4	5	6	7	8	9
00	0.470	0.551	0.519	0.688	0.649	0.623	0.661	0.748	0.674	0.695
10	0.802	0.623	0.727	0.678	0.642	0.821	0.760	0.765	1.03	0.934
20	0.907	0.809	0.897	0.983	1.00	0.648	0.613	0.827	0.861	0.940
30	1.04	1.09	0.909	0.952	0.929	1.13	0.950	1.13	0.905	0.732
40	0.879	1.15	1.21	1.05	1.01	0.968	1.25	1.08	1.13	1.09
50	1.09	1.23	1.09	1.30	1.10	0.943	1.21	1.26	0.981	1.15
60	1.16	1.07	0.754	1.20	1.14	1.12	1.04	1.02	1.12	1.25
70	1.27	1.36	1.35	1.42	1.36	1.54	1.39	1.55	1.29	1.35
80	1.34	1.30	1.32	1.43	1.21	1.35	1.70	1.51	2.22	1.39
90	1.50	1.99	1.86	2.61	1.71	2.06	2.22	2.00	2.72	2.05
100	1.58	2.62	3.75	1.68	1.86	3.56	3.07	3.21	2.41	1.78
10	1.74	3.28	4.27	2.04	3.01	1.95	4.17	4.49	3.22	3.14
20	5.31	3.42	2.79	2.00	1.04	1.78	3.29	3.06	3.18	3.43
30	1.79	2.81	2.39	3.02	2.67	1.84	2.83	4.11	2.89	1.71
40	2.61	5.13	3.78	4.11	4.15	3.33	2.85	2.91	2.98	4.04
50	3.40	2.55	2.81	4.16	3.50	3.51	2.93	1.84	1.80	1.85
60	2.74	3.78	2.94	3.72	2.78	2.99	1.51	1.40	3.07	1.63
70	2.94	2.67	3.51	2.73	2.27	1.76	2.89	2.47	1.54	3.60
80	4.27	4.53	4.40	4.45	2.60	3.76	3.33	1.58	3.92	1.98
90	4.16	2.90	2.27	5.54	4.38	2.85	1.92	3.18	4.66	4.34
200	3.03	1.83	3.80	4.60	4.64	4.01	3.47	3.15	1.68	3.18
10	1.55	1.10	1.74	4.23	2.26	3.34	2.17	1.26	1.60	3.20
20	3.29	3.67	3.85	2.32	3.64	3.13	2.60	2.49	3.40	3.10
30	3.04	2.01	4.06	5.33	4.42	4.65	4.52	4.12	2.75	3.27
40	4.20	2.86	3.25	2.73	2.20	3.48	3.29	4.04	4.00	2.43
50	2.97	2.22	3.37	2.65	2.13	2.82	2.12	2.90	2.12	3.41
60	1.44	2.12	2.62	1.76	1.78	1.45	1.60	1.33	1.57	1.25
70	1.34	1.47	1.40	1.79	1.35	1.92	1.28	1.86	1.90	2.17
80	1.39	2.75	2.87	1.95	2.12	2.83	2.02	1.18	2.30	1.94
90	1.45	1.61	1.75	1.79	1.89	1.77	2.58	1.08	0.858	1.25
300	1.30	1.42	1.85	1.08	1.80	1.98	2.62	2.39	1.87	1.54
10	1.33	0.838	1.03	1.01	0.957	2.18	1.50	1.44	1.68	1.56
20	0.624	1.09	1.40	2.10	2.01	1.18	0.768	0.954	1.42	1.56
30	1.25	1.11	0.914	0.659	1.24	1.13	1.67	0.954	0.577	0.714
40	0.914	1.09	1.19	0.565	0.451	0.626	0.718	0.784	0.489	0.758
50	1.17	0.974	1.24	1.49	1.10	0.653	1.38	1.17	1.17	0.628
60	0.505	0.815	0.785	1.22	0.744	0.601	0.838	0.522	0.719	0.717
70	0.887	0.937	1.11	0.704	0.457	0.626	0.721	0.993	0.846	0.525
80	0.604	0.374	0.281	0.279	0.448	0.631	0.665	0.683	0.430	0.323
90	0.583	0.831	0.583	0.637	0.549	0.279	0.200	0.365	0.542	0.276
400	0.472	0.567	0.479	0.496	0.322	0.193	0.340	0.210	0.430	0.442
10	0.363	0.190	0.346	0.327	0.317	0.428	0.484	0.540	0.516	0.471
20	0.589	0.592	0.702	0.456	0.449	0.479	0.507	0.736	0.529	0.444
30	0.535	0.585	0.436	0.392	0.496	0.242	0.397	0.410	0.518	0.340
40	0.359	0.542	0.507	0.434	0.390	0.448	0.391	0.361	0.271	0.332
50	0.319	0.285	0.234	0.286	0.243	0.299	0.256	0.256	0.325	0.234
60	0.299	0.252	0.206	0.220	0.257	0.234	0.267	0.272	0.291	0.321
70	0.270	0.304	0.294	0.140	0.366	0.307	0.316	0.280	0.259	0.209
80	0.197	0.315	0.239	0.109	0.134	0.211	0.267	0.215	0.122	0.161
90	0.121	0.073	0.165	0.204	0.249	0.340	0.268	0.300	0.271	0.278

Table 2-1. *(Continued)*

One Angstrom averages centered on the wavelength in units of 10^{11} ph cm^{-2} s^{-1} $Å^{-1}$										
Å	0	1	2	3	4	5	6	7	8	9
2500	0.323	0.151	0.241	0.289	0.326	0.292	0.218	0.129	0.245	0.294
10	0.346	0.185	0.235	0.234	0.146	0.159	0.104	0.144	0.138	0.109
20	0.188	0.183	0.168	0.117	0.117	0.172	0.186	0.173	0.129	0.101
30	0.162	0.244	0.229	0.224	0.185	0.200	0.159	0.171	0.229	0.193
40	0.269	0.169	0.224	0.282	0.276	0.259	0.184	0.227	0.254	0.197
50	0.200	0.249	0.410	0.221	0.410	0.270	0.342	0.255	0.349	0.495
60	0.371	0.417	0.297	0.262	0.339	0.394	0.689	0.418	0.564	0.575
70	0.725	0.509	0.593	0.681	0.682	0.413	0.288	0.509	0.594	0.623
80	0.856	0.962	0.781	0.441	0.641	0.370	0.189	0.422	0.470	0.788
90	0.866	0.561	0.558	0.529	0.320	0.685	0.605	0.526	0.258	0.195
2600	0.243	0.510	0.685	0.663	0.515	0.504	0.291	0.191	0.425	0.493
10	0.566	0.393	0.251	0.435	0.300	0.696	0.722	0.554	0.415	0.692
20	0.772	0.696	0.528	0.841	0.796	0.665	0.436	0.666	0.462	0.640
30	0.597	0.346	0.473	0.781	1.23	1.23	1.45	1.60	2.14	2.08
40	1.60	1.66	1.71	2.75	1.99	1.85	1.63	1.84	1.33	1.31
50	1.84	1.63	1.84	2.15	2.05	3.06	1.86	2.49	2.29	2.14
60	2.53	2.43	1.82	2.39	1.63	2.63	2.25	1.90	3.07	2.50
70	1.96	3.74	1.90	1.78	3.40	3.17	3.21	2.18	2.71	1.65
80	2.48	3.16	2.88	2.90	2.96	2.57	3.61	3.30	2.94	3.00
90	2.36	2.90	3.94	3.10	3.76	2.16	1.94	3.44	2.90	2.70
2700	5.25	4.00	2.96	5.02	3.74	5.70	3.86	3.62	4.34	3.19
10	4.45	4.35	3.79	5.92	2.91	3.27	3.98	4.62	3.49	1.21
20	2.44	1.81	3.78	4.08	2.89	3.58	4.13	3.96	3.48	6.37
30	6.98	4.15	5.56	4.42	3.52	4.30	5.25	1.63	4.18	3.32
40	2.47	3.90	3.14	2.20	2.55	4.65	3.55	1.73	4.11	2.38
50	2.03	4.25	6.45	6.19	4.60	4.22	1.80	4.23	8.10	9.87
60	7.34	8.71	4.90	6.28	7.22	11.7	9.70	7.86	7.55	6.31
70	9.44	10.4	7.00	10.3	10.6	9.03	8.83	8.22	5.85	7.13
80	5.79	5.78	7.57	6.18	6.61	7.15	7.38	6.95	5.01	5.86
90	4.79	4.10	3.81	3.31	2.44	3.68	4.43	2.38	2.59	3.59
2800	3.79	3.07	2.56	5.27	2.95	3.97	5.39	5.61	7.01	8.34
10	8.58	8.87	8.80	8.66	12.9	15.7	14.2	13.8	15.0	15.7
20	19.1	17.1	18.0	15.7	22.7	21.1	13.9	21.1	19.5	19.7
30	24.4	24.6	17.8	26.6	23.3	19.1	12.4	24.9	23.9	26.0
40	21.9	19.0	26.7	18.4	15.8	25.4	17.2	14.7	11.9	10.9
50	7.65	5.92	2.60	4.87	7.64	10.8	12.2	15.8	18.7	21.5
60	29.5	24.4	31.8	28.1	30.7	29.3	38.0	27.1	26.5	21.8
70	33.6	40.8	42.1	39.7	24.6	37.9	30.3	33.7	38.4	22.3
80	25.1	12.2	10.2	24.2	29.0	47.5	41.4	41.8	39.2	44.9
90	40.2	38.0	50.4	34.8	47.7	46.7	62.4	60.3	54.6	59.9
2900	73.1	80.5	54.4	69.0	75.2	83.2	67.7	70.0	61.5	64.8
10	65.1	64.6	58.9	101	67.3	63.0	62.3	71.6	55.8	71.9
20	67.4	73.5	72.2	70.9	55.2	65.8	48.9	60.1	59.2	32.3
30	91.7	76.2	92.1	69.4	71.6	98.9	74.4	26.9	62.6	51.6
40	83.4	55.7	62.5	91.3	63.8	66.6	103	70.7	26.1	42.3
50	63.2	87.9	92.0	105	31.6	90.5	87.0	57.2	85.7	102
60	73.1	73.7	97.7	112	87.0	42.9	53.2	26.0	77.7	70.1
70	34.5	69.9	79.8	39.0	71.9	106	81.6	113	123	85.4
80	86.0	67.8	41.8	68.1	34.6	45.3	51.7	63.2	67.3	76.0
90	110	106	87.0	82.9	56.9	40.3	94.0	92.4	110	69.1

Table 2-1. *(Continued)*

One Angstrom averages centered on the wavelength in units of 10^{11} ph cm^{-2} s^{-1} $Å^{-1}$										
Å	0	1	2	3	4	5	6	7	8	9
3000	46.3	25.8	61.7	36.8	57.2	93.5	90.1	54.5	46.8	65.9
10	69.0	110	54.5	101	85.2	82.8	83.4	89.5	54.6	45.2
20	48.5	14.7	62.1	109	76.4	97.4	41.2	108	133	104
30	100	93.5	104	139	120	118	132	78.3	43.3	112
40	117	86.5	66.3	96.3	120	121	138	79.3	50.0	108
50	134	80.8	168	142	79.0	96.0	126	83.0	44.9	44.2
60	90.7	111	117	129	77.0	89.2	101	63.8	83.7	114
70	123	121	96.2	109	122	109	78.6	134	107	107
80	125	115	67.6	103	101	144	150	96.9	77.3	134
90	85.1	81.2	90.6	30.5	73.2	117	113	47.4	96.9	105

Two rocket flight measurements of the solar irradiance were also reported by Willson et al. [1980]. The total solar flux measurement for 29 June 1976 was 1368.1 ± 0.5 W/m^2 and for 16 November 1978 was 1373.4 ± [illegible] ^{2}. These measurements are within the error estim[illegible]hlich.

The solar spectral irradiance S $\Delta\lambda$, defined as the irradiance per wavelength interval $\Delta\lambda$, is plotted in Figure 2-2 for wavelengths longer than 3000 Å. This figure is based on data from Labs and Neckel between 3000 Å and 3 μm, and Thekaekara above 3 μm. These sources of data are from

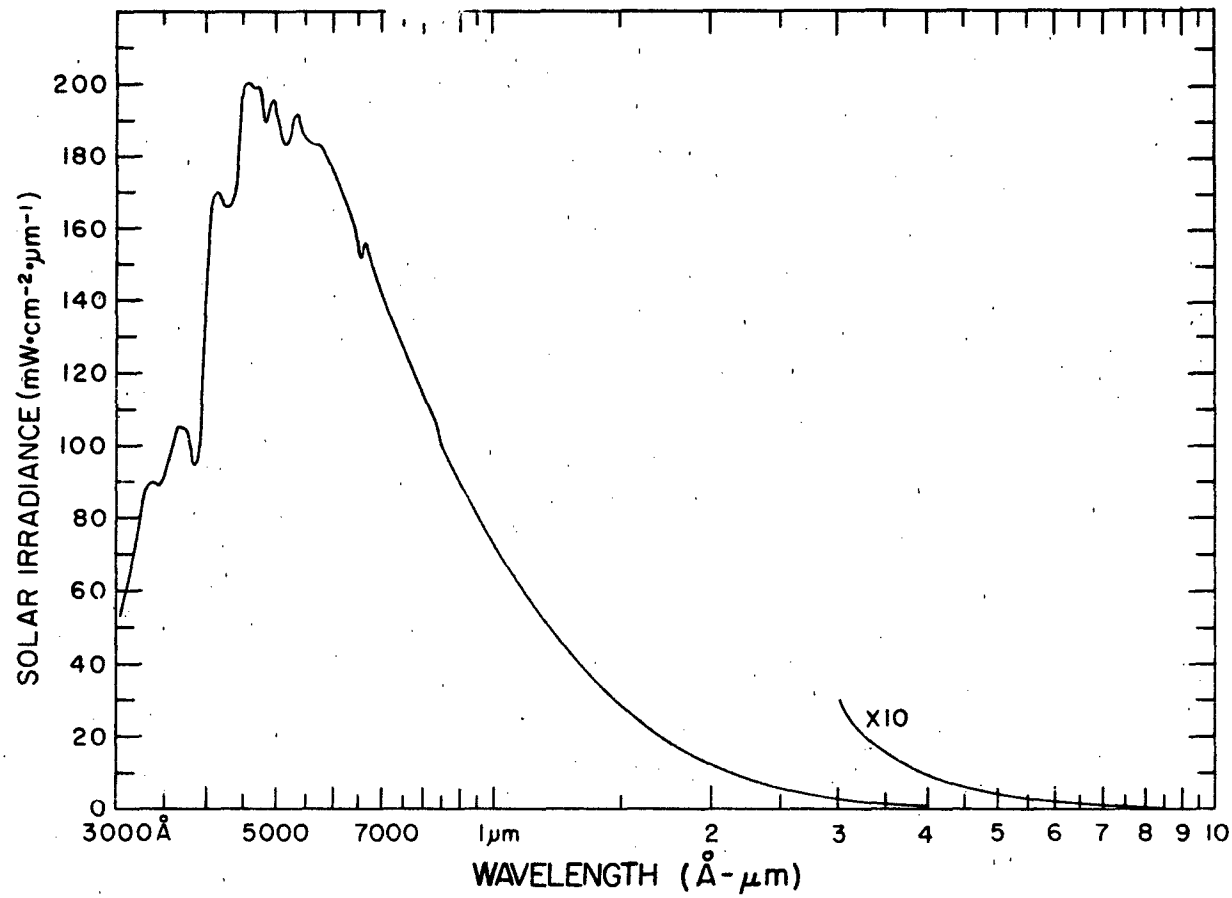


Figure 2-2. The solar spectral irradiance for wavelengths longer than 3000 Å. The data between 3000 Å and 3 μm are from Labs and Neckel and from Thekaekara above 3 μm. These data are from a summary given by Pierce and Allen [1977].

a summary given by Pierce and Allen [1977]. The spectral region above 3000 Å accounts for approximately 99% of the solar constant. Nearly all of this energy reaches the surface of the earth since atmospheric absorption of the radiation is negligible. The irradiance in this wavelength region does not appear to vary with solar activity. Because this region is the principal source of solar energy, the variation in the total solar irradiance or solar constant is also small and is now believed to be less than ± 1%.

2.2.2 Solar Irradiance Between 1 and 3000 Å For Solar Minimum

The solar spectral irradiance for wavelengths below 3180 Å at the top of the earth's atmosphere is tabulated in Table 2-2 for 10 Å wavelength intervals. The values of irradiance for the wavelength range 1 to 2300 Å are for levels of solar activity near the minimum value of activity for the 11-year solar cycle. The irradiance values between 1 and 10 Å are from Kreplin [1977], between 10 and 50 Å from Manson [1977], and between 50 and 1850 Å from the solar minimum reference spectrum of Heroux and Hinteregger [1978]. The irradiance between 1850 and 2090 Å from Semain and Simon and between 2090 and 2300 Å from Simon is from a summary of Delaboudiniere et al. [1978]. The irradiance between 2300 and 3150 Å was measured by Mount and Rottman [1981] in a rocket flight flown near solar maximum of the present solar cycle 21. However, because the spectral irradiance in this longer UV wavelength region does not vary significantly with solar activity, these irradiance values should also represent solar minimum conditions. In Table 2-2, the solar spectral irradiance is given in units of photons $cm^{-2}s^{-1}$ and in units of mW/m^2, which is also equivalent to ergs $cm^{-2}s^{-1}$. The fourth column in Table 2-2 gives the percentage of the total solar irradiance, s = 1373 W/m^2,

Table 2-2. Solar irradiance and cumulative irradiance for quiet sun at the top of the earth's atmosphere. S = 1373 W/m^2, 1 mWm^{-2} = 1 erg cm^{-2} s^{-1}

Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^9 photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_0^\lambda S_{\Delta\lambda}/S$ %
1–10	0.0005	0.002	0.15×10^{-6}
10–20	0.0014	0.002	0.29
20–30	0.0076	0.006	0.73
30–40	0.088	0.050	4.4
40–50	0.095	0.042	7.4
1–50	0.193	0.102	
50–60	0.055	0.020	8.9×10^{-6}
60–70	0.069	0.021	1.0×10^{-5}
70–80	0.073	0.019	1.2
80–90	0.106	0.025	1.4
90–100	0.094	0.020	1.5
50–100	0.397	0.105	
100–110	0.048	0.009	1.6×10^{-5}
110–120	0.017	0.003	1.6
120–130	0.014	0.002	1.6
130–140	0.009	0.001	1.6
140–150	0.062	0.008	1.7
100–150	0.150	0.023	
150–160	0.082	0.011	1.8×10^{-5}
160–170	0.115	0.14	1.9
170–180	0.782	0.089	2.5
180–190	0.787	0.085	3.1
190–200	0.602	0.062	3.6
150–200	2.37	0.261	
200–210	0.292	0.029	3.8×10^{-5}
210–220	0.266	0.024	4.0

Table 2-2. *(Continued)*

Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^9 photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_0^\lambda S_{\Delta\lambda}/S$ %
220–230	0.268	0.024	4.1
230–240	0.206	0.017	4.3
240–250	0.531	0.043	4.6
200–250	1.56	0.137	
250–260	1.233	0.095	5.3×10^{-5}
260–270	0.204	0.015	5.4
270–280	0.426	0.031	5.6
280–290	0.262	0.018	5.7
290–300	0.225	0.015	5.8
250–300	2.35	0.174	
300–310	7.700	0.504	9.5×10^{-5}
310–320	0.375	0.023	9.7
320–330	0.020	0.001	9.7
330–340	0.140	0.008	9.8
340–350	0.430	0.025	
300–350	8.67	0.561	
350–360	0.110	0.006	1.0×10^{-4}
360–370	0.840	0.045	1.0
370–380	0.000	0.000	1.0
380–390	0.000	0.000	1.0
390–400	0.014	0.001	1.0
350–400	0.964	0.052	
400–410	0.161	0.008	1.0×10^{-4}
410–420	0.027	0.001	1.0
420–430	0.002	0.000	1.0
430–440	0.187	0.009	1.0
440–450	0.005	0.000	1.0
400–450	0.382	0.018	
450–460	0.009	0.000	1.0×10^{-4}
460–470	0.305	0.013	1.1
470–480	0.026	0.001	1.1
480–490	0 053	0.002	1.1
490–500	0.173	0.007	1.1
450–500	0.566	0.023	
500–510	0.163	0.006	1.1×10^{-4}
510–520	0.025	0.001	1.1
520–530	0.111	0.004	1.1
530–540	0.120	0.004	1.1
540–550	0.021	0.001	1.1
500–550	0.440	0.016	
550–560	0.786	0.028	1.1×10^{-4}
560–570	0.093	0.003	1.1
570–580	0.031	0.001	1.1

Table 2-2. (*Continued*)

Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^9 photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_0^{\lambda} S_{\Delta\lambda}/S$ %
580–590	1.270	0.043	1.1
590–600	0.155	0.005	1.1
550–600	2.34	0.080	
600–610	0.530	0.017	1.1×10^{-4}
610–620	0.017	0.001	1.1
620–630	1.832	0.058	1.2
630–640	0.021	0.001	1.2
640–650	0.062	0.002	1.2
600–650	2.46	0.079	
650–660	0.026	0.001	1.2×10^{-4}
660–670	0.014	0.000	1.2
670–680	0.014	0.000	1.2
680–690	0.131	0.004	1.2
690–700	0.044	0 001	1.2
650–700	0.229	0.006	
700–710	0.369	0.010	1.2×10^{-4}
710–720	0.072	0.002	1.2
720–730	0.015	0.000	1.2
730–740	0.019	0.001	1.2
740–750	0.024	0.001	1.2
700–750	0.499	0.014	
750–760	0.117	0.003	1.2×10^{-4}
760–770	0.340	0.009	1.2
770–780	0.323	0.008	1.2
780–790	0.587	0.015	1.2
790–800	0.516	0.013	1.2
750–800	1.88	0.048	
800–810	0.111	0.003	1.2×10^{-4}
810–820	0.143	0.003	1.2
820–830	0.184	0.004	1.2
830–840	0.857	0.020	1.3
840–850	0.305	0.007	
800–850	1.60	0.037	
850–860	0.392	0.009	1.3×10^{-4}
860–870	0.504	0.012	1.3
870–880	0.643	0.015	1.3
880–890	0.833	0.019	1.3
890–900	1.071	0.024	1.3
850–900	3.44	0.079	
900–910	1.487	0.033	1.3×10^{-4}
910–920	0.502	0.011	1.4
920–930	0.274	0.006	1.4
930–940	0.447	0.009	1.4
940–950	0.408	0.009	1.4

Table 2-2.' (*Continued*)

Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^9 photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_0^\lambda S_{\Delta\lambda}/S$ %
900–950	3.12	0.068	
950–960	0.048	0.001	1.4×10^{-4}
960–970	0.058	0.001	1.4
970–980	5.069	0.103	1.5
980–990	0.253	0.005	1.5
990–1000	0.439	0.009	1.5
950–1000	5.87	0.119	
1000–1010	0.119	0.002	1.5×10^{-4}
1010–1020	0.222	0.004	1.5
1020–1030	3.671	0.071	1.5
1030–1040	3.794	0.073	1.6
1040–1050	0.244	0.005	1.6
1000–1050	8.05	0.155	
1050–1060	0.292	0.006	1.6×10^{-4}
1060–1070	0.405	0.008	1.6
1070–1080	0.528	0.010	1.6
1080–1090	1.021	0.019	1.6
1090–1100	0.599	0.011	1.6
1050–1100	2.85	0.054	
1100–1110	0.083	0.001	1.6×10^{-4}
1110–1120	0.022	0.000	1.6
1120–1130	0.679	0.012	1.6
1130–1140	0.045	0.001	1.6
1140–1150	0.077	0.001	1.6
1100–1150	0.906	0.015	
1150–1160	0.129	0.002	1.6×10^{-4}
1160–1170	0.219	0.004	1.6
1170–1180	2.872	0.049	1.7
1180–1190	0.493	0.008	1.7
1190–1200	0.675	0.011	1.7
1150–1200	4.39	0.074	
1200–1210	4.855	0.080	1.7×10^{-4}
1210–1220	251.774	4.114	4.7
1220–1230	0.636	0.010	4.7
1230–1240	1.480	0.024	4.8
1240–1250	0.640	0.010	4.8
1200–1250	259.39	4.24	
1250–1260	1.080	0.017	4.8×10^{-4}
1260–1270	0.930	0.015	4.8
1270–1280	0.750	0.012	4.8
1280–1290	0.500	0.008	4.8
1290–1300	0.860	0.013	4.8

Table 2-2. (*Continued*)

Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^9 photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_0^{\lambda} S_{\Delta\lambda}/S$ %
1250–1300	4.12	0.065	
1300–1310	4.585	0.070	4.9×10^{-4}
1310–1320	0.780	0.012	4.9
1320–1330	0.780	0.012	4.9
1330–1340	5.300	0.079	4.9
1340–1350	0.920	0.014	4.9
1300–1350	12.37	0.187	
1350–1360	1.47	0.022	5.0×10^{-4}
1360–1370	1.05	0.015	5.0
1370–1380	1.13	0.016	5.0
1380–1390	1.04	0.015	5.0
1390–1400	2.70	0.038	5.0
1350–1400	7.39	0.106	
1400–1410	2.70	0.038	5.1×10^{-4}
1410–1420	1.59	0.022	5.1
1420–1430	1.90	0.026	5.1
1430–1440	2.10	0.029	5.1
1440–1450	2.10	0.029	5.1
1400–1450	10.39	0.144	
1450–1460	2.30	0.031	5.2×10^{-4}
1460–1470	3.00	0.041	5.2
1470–1480	3.80	0.051	5.2
1480–1490	3.70	0.049	5.3
1490–1500	3.40	0.045	5.3
1450–1500	16.20	0.217	
1500–1510	3.90	0.051	5.3×10^{-4}
1510–1520	4.50	0.059	5.4
1520–1530	5.50	0.072	5.4
1530–1540	5.90	0.076	5.5
1540–1550	9.50	0.122	5.6
1500–1550	29.30	0.380	
1550–1560	8.80	0.113	5.6×10^{-4}
1560–1570	9.00	0.114	5.7
1570–1580	7.70	0.097	5.8
1580–1590	7.10	0.089	5.9
1590–1600	7.20	0.090	5.9
1550–1600	39.80	0.503	
1600–1610	7.90	0.098	6.0×10^{-4}
1610–1620	9.40	0.116	6.1
1620–1630	11.10	0.136	6.2
1630–1640	12.00	0.146	6.3
1640–1650	15.30	0.185	6.4

Table 2-2. (*Continued*)

Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^9 photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_{0}^{\lambda} S_{\Delta\lambda}/S$ %
1600–1650	55.70	0.681	
1650–1660	25.00	0.300	6.6×10^{-4}
1660–1670	18.40	0.220	6.8
1670–1680	23.00	0.273	7.0
1680–1690	26.99	0.318	7.2
1690–1700	37.02	0.434	7.6
1650–1700	130.41	1.55	
1700–1710	43.00	0.501	7.9×10^{-4}
1710–1720	43.99	0.510	8.3
1720–1730	45.98	0.530	8.7
1730–1740	41.99	0.481	9.0
1740–1750	50.02	0.569	9.4
1700–1750	224.98	2.59	
1750–1760	56.99	0.645	9.9×10^{-4}
1760–1770	61.97	0.698	1.0×10^{-3}
1770–1780	74.02	0.828	1.1
1780–1790	80.99	0.901	1.2
1790–1800	82.96	0.908	1.2
1750–1800	356.93	3.99	
1800–1810	103.00	1.13	1.3×10^{-3}
1810–1820	126.00	1.38	1.4
1820–1830	132.02	1.44	1.5
1830–1840	131.96	1.43	1.7
1840–1850	110.95	1.20	1.7
1800–1850	603.93	6.57	
1850–1860	190.51	2.04	1.9×10^{-3}
1860–1870	237.55	2.53	2.1
1870–1880	265.25	2.81	2.3
1880–1890	279.95	2.95	2.5
1890–1900	293.84	3.08	2.7
1850–1900	1267.10	13.41	
1900–1910	293.48	3.06	2.9×10^{-3}
1910–1920	332.62	3.45	3.2
1920–1930	347.92	3.59	3.4
1930–1940	254.26	2.61	3.6
1940–1950	446.52	4.56	4.0
1900–1950	1674.80	17.27	
1950–1960	427.16	4.34	4.3×10^{-3}
1960–1970	485.73	4.91	4.6
1970–1980	487.21	4.90	5.0
1980–1990	492.68	4.93	5.3
1990–2000	552.41	5.50	5.7
1950–2000	2445.19	24.58	

Table 2-2. (*Continued*)

Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^9 photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_0^{\lambda} S_{\Delta\lambda}/S$ %
2000–2010	624.83	6.19	6.2×10^{-3}
2010–2020	629.98	6.21	6.8
2020–2030	642.27	6.30	7.1
2030–2040	763.27	7.45	7.6
2040–2050	895.71	8.70	8.3
2000–2050	3556.06	34.85	
2050–2060	919.75	8.89	8.9×10^{-3}
2060–2070	964.77	9.28	9.6
2070–2080	1128.23	10.80	1.0×10^{-2}
2080–2090	1270.13	12.10	1.1
2090–2100	2562.98	24.30	1.3
2050–2100	6845.86	65.37	
2100–2110	$2.72 \times 10^{+3}$	25.6	1.5×10^{-2}
2110–2120	3.52	33.1	1.7
2120–2130	2.92	27.3	1.9
2130–2140	3.81	35.4	2.2
2140–2150	4.81	44.5	2.5
2100–2150	$17.78 \times 10^{+3}$	165.90	
2150–2160	$3.74 \times 10^{+3}$	34.5	2.8×10^{-2}
2160–2170	3.48	31.9	3.0
2170–2180	3.72	34.0	3.2
2180–2190	5.07	46.1	3.6
2190–2200	5.60	50.7	3.9
2150–2200	21.61×10^{3}	197.2	
2200–2210	5.42×10^{3}	48.8	4.3×10^{-2}
2210–2220	4.19	37.6	4.6
2220–2230	5.80	51.8	5.0
2230–2240	7.96	70.7	5.5
2240–2250	7.17	63.4	5.9
2200–2250	30.54×10^{3}	272.3	
2250–2260	6.95×10^{3}	61.2	6.4×10^{-2}
2260–2270	5.34	46.8	6.7
2270–2280	5.50	48.0	7.1
2280–2290	8.88	77.2	7.6
2290–2300	7.18	62.1	8.1
2250–2300	33.85×10^{3}	295.3	
2300–2310	6.88×10^{3}	59.3	8.5×10^{-2}
2310–2320	5.94	51.0	8.9
2320–2330	6.54	55.9	9.3
2330–2340	5.42	46.1	9.6
2340–2350	4.87	41.3	9.9
2300–2350	29.65×10^{3}	253.6	

Table 2-2 (*Continued*)

Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^{n} photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_{0}^{\lambda} S_{\Delta\lambda}/S$ %
2350–2360	6.78×10^{1}	57.2	1.0×10^{-1}
2360–2370	6.18	51.9	1.1
2370–2380	6.15	51.4	1.1
2380–2390	5.36	44.6	1.1
2390–2400	5.85	48.5	1.2
2350–2400	30.32×10^{1}	253.6	
2400–2410	5.10×10^{1}	42.1	1.2×10^{-1}
2410–2420	7.00	57.6	1.3
2420–2430	9.39	76.9	1.3
2430–2440	8.49	69.3	1.4
2440–2450	7.73	62.8	1.4
2400–2450	37.71×10^{1}	308.7	
2450–2460	6.46×10^{1}	52.3	1.4×10^{-1}
2460–2470	6.79	54.7	1.5
2470–2480	7.50	60.2	1.5
2480–2490	5.43	43.4	1.6
2490–2500	8.18	65.1	1.6
2450–2500	34.36×10^{1}	275.7	
2500–2510	7.66×10^{1}	60.7	1.6×10^{-1}
2510–2520	5.69	44.9	1.7
2520–2530	5.55	43.7	1.7
2530–2540	7.16	56.1	1.8
2540–2550	8.17	63.8	1.8
2500–2550	34.23×10^{1}	269.2	
2550–2560	11.2×10^{1}	87.1	1.9×10^{-1}
2560–2570	14.6	113	1.9
2570–2580	16.9	130	2.0
2580–2590	16.1	124	2.1
2590–2600	12.0	91.9	2.2
2550–2600	70.8×10^{1}	546	
2600–2610	11.9×10^{1}	90.7	2.3×10^{-1}
2610–2620	12.9	98.0	2.3
2620–2630	13.5	102.	2.4
2630–2640	25.1	189	2.5
2640–2650	32.4	243	2.7
2600–2650	95.8×10^{1}	723	
2650–2660	34.2×10^{1}	256	2.9×10^{-1}
2660–2670	31.7	236	3.1
2670–2680	33.0	245	3.3
2680–2690	31.8	235	3.4
2690–2700	32.4	238	3.6
2650–2700	163×10^{1}	1210	

Table 2-2. (*Continued*)

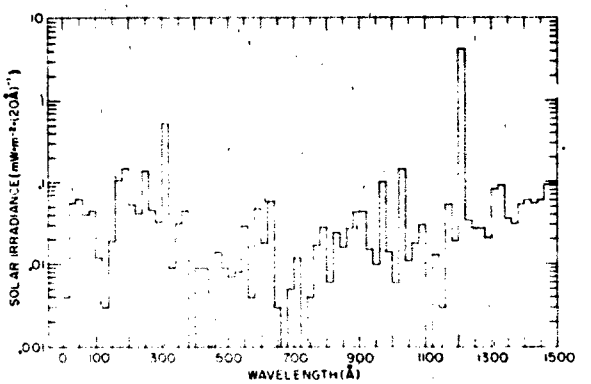
Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^{9} photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_{0}^{\lambda} S_{\Delta\lambda}/S$ %
2700–2710	35.7×10^{3}	262	3.8×10^{-1}
2710–2710	29.3	214	4.0
2720–2730	26.0	190	4.1
2730–2740	28.6	208	4.2
2740–2750	18.2	132	4.3
2700–2750	138×10^{3}	1006	
2750–2760	21.2×10^{3}	153	4.4×10^{-1}
2760–2770	32.1	231	4.6
2770–2780	34.6	248	4.8
2780–2790	24.9	178	4.9
2790–2800	13.3	95	5.0
2750–2800	126×10^{3}	905	
2800–2810	13.1×10^{3}	92.8	5.1×10^{-1}
2810–2820	28.9	204	5.2
2820–2830	40.1	282	5.4
2830–2840	44.0	308	5.6
2840–2850	33.9	237	5.8
2800–2850	160×10^{3}	1124	
2850–2860	22.6×10^{3}	157	5.9×10^{-1}
2860–2870	46.9	325	6.2
2870–2880	45.7	316	6.4
2880–2890	44.4	306	6.6
2890–2900	63.9	438	7.0
2850–2900	224×10^{3}	1542	
2900–2910	81.6×10^{3}	558	7.3×10^{-1}
2910–2910	77.1	525	7.7
2920–2930	68.3	464	8.1
2930–2940	73.9	500	8.4
2940–2950	69.2	467	8.8
2900–2950	370×10^{3}	2514	
2950–2960	68.1×10^{3}	458	9.1×10^{-1}
2960–2970	73.6	493	9.5
2970–2980	61.1	408	9.8
1980–2990	64.0	426	1.0×1
2990–3000	64.4	427	1.0
2950–3000	331×10^{3}	2212	
3000–3010	50.8×10^{3}	336	1.1
3010–3020	59.3	391	1.1
3020–3030	63.5	417	1.1
3030–3040	77.7	509	1.2
3040–3050	74.9	489	1.2
3000–3050	326×10^{3}	2142	

Table 2-2. (*Continued*)

Wavelength Interval Å	$\Phi_{\Delta\lambda}$ (10^9 photons cm^{-2} s^{-1})	$S_{\Delta\lambda}$ mWm^{-2}	$\sum_0^\lambda S_{\Delta\lambda}/S$ %
3050–3060	73.5×10^3	478	1.2
3060–3070	71.9	466	1.3
3070–3080	79.0	510	1.3
3080–3090	76.4	492	1.3
3090–3100	58.5	375	1.4
3050–3100	359×10^3	2321	
3100–3110	78.5×10^3	502	1.4
3110–3120	87.9	561	1.4
3120–3130	78.9	502	1.5
3130–3140	85.3	540	1.5
3140–3150	74.6	471	1.6
3100–3150	405.2×10^3	2576	
3150–3160	66.5×10^3	419	1.6
3160–3170	77.8	488	1.6
3170–3180	105	657	1.7

contributed by the wavelength region shorter than the listed wavelength. For example, the solar irradiance for the entire wavelength region below 3000 Å is 1% of the solar constant.

The solar irradiance based on Table 2-2 is plotted for 20 Å wavelength intervals in Figure 2-3. The pronounced structure for the wavelength Region 1 to 1500 Å arises from emission lines of elements abundant in the sun. The high levels of irradiance in the intervals 300-320 Å and 1200-1220 Å originate primarily from the solar emission lines of He II at 303.78 Å and hydrogen Lyman-α (H Ly-α) at 1215.67 Å, respectively. The solar spectrum above about 1500 Å is essentially a continuum with absorption and emission lines superimposed on the continuum.



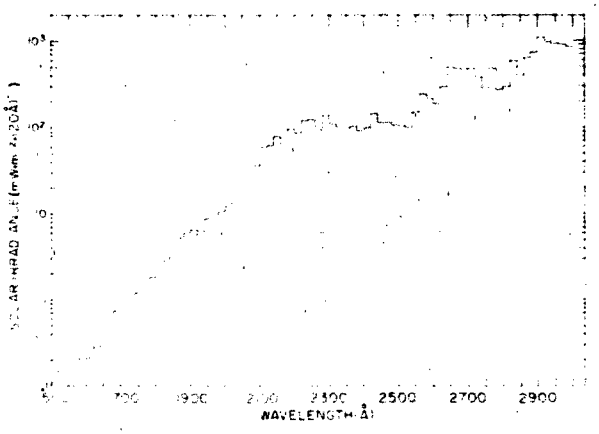


Figure 2-3. The solar spectral irradiance between 1 and 3000 Å for a quiet sun at the top of the earth's atmosphere. The plotted values are averages over 20 Å intervals of wavelength.

2.2.3 Solar Irradiance Below 1200 Å

The temperature of the visible surface of the sun is about 6000 K. If the temperature is measured as a function of distance outward from the surface of the sun, the temperature decreases to about 4500 K at the top of the photosphere (near 500 km) and increases gradually in the chromosphere which extends to about 2000 km. As can be seen in Figure 2-4, the temperature increases rapidly from about 10^4 to 5×10^5 K in a very narrow transition region near 2000 km, and then continues to increase gradually in the corona where the temperature can exceed 2×10^6 K. The solar UV spectrum between 1 and 3000 Å is produced throughout this spatial region which encompasses a temperature range from the photospheric minimum temperature of 4500 K to the coronal temperature of 2×10^6 K.

The solar spectrum between 3000 and 1300 Å is pro-

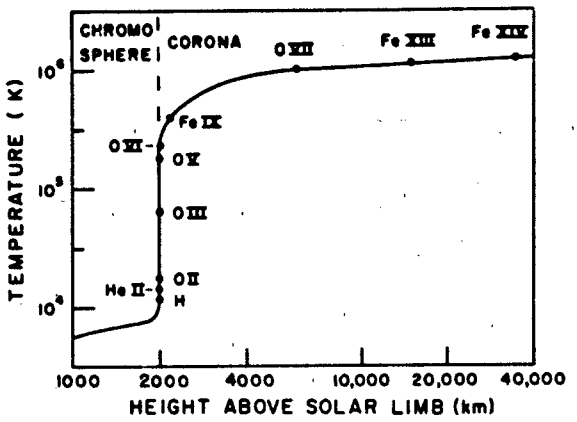


Figure 2-4. The spatial distribution of temperature in the solar chromosphere and corona. The temperature region of production of several multiply-ionized species is also given.

duced predominantly in the upper photosphere and lower chromosphere. The dominant feature in this wavelength region is a continuum that becomes apparent near 1300 Å and increases in intensity toward longer wavelengths as is evident in Figure 2-3. There are numerous emission and absorption lines superimposed on the continuum. At wavelengths shorter than about 1300 Å, the solar spectrum consists predominantly of sharp emission lines that originate from transitions in multiply ionized atomic species that are produced over the wide range of temperatures established in the chromosphere and corona. Figure 2-4 indicates the temperature regions in which several of the ion species observed in the solar spectrum are produced. An example of the solar UV spectrum between 1300 and 270 Å, measured on the NASA OSO 3 satellite by Hall and Hinteregger [1970], is given in Figure 2-5. The scattered light background beneath the spectrum is estimated from a pre-flight calibration to decrease gradually from 150 counts near 120 Å to 50 counts

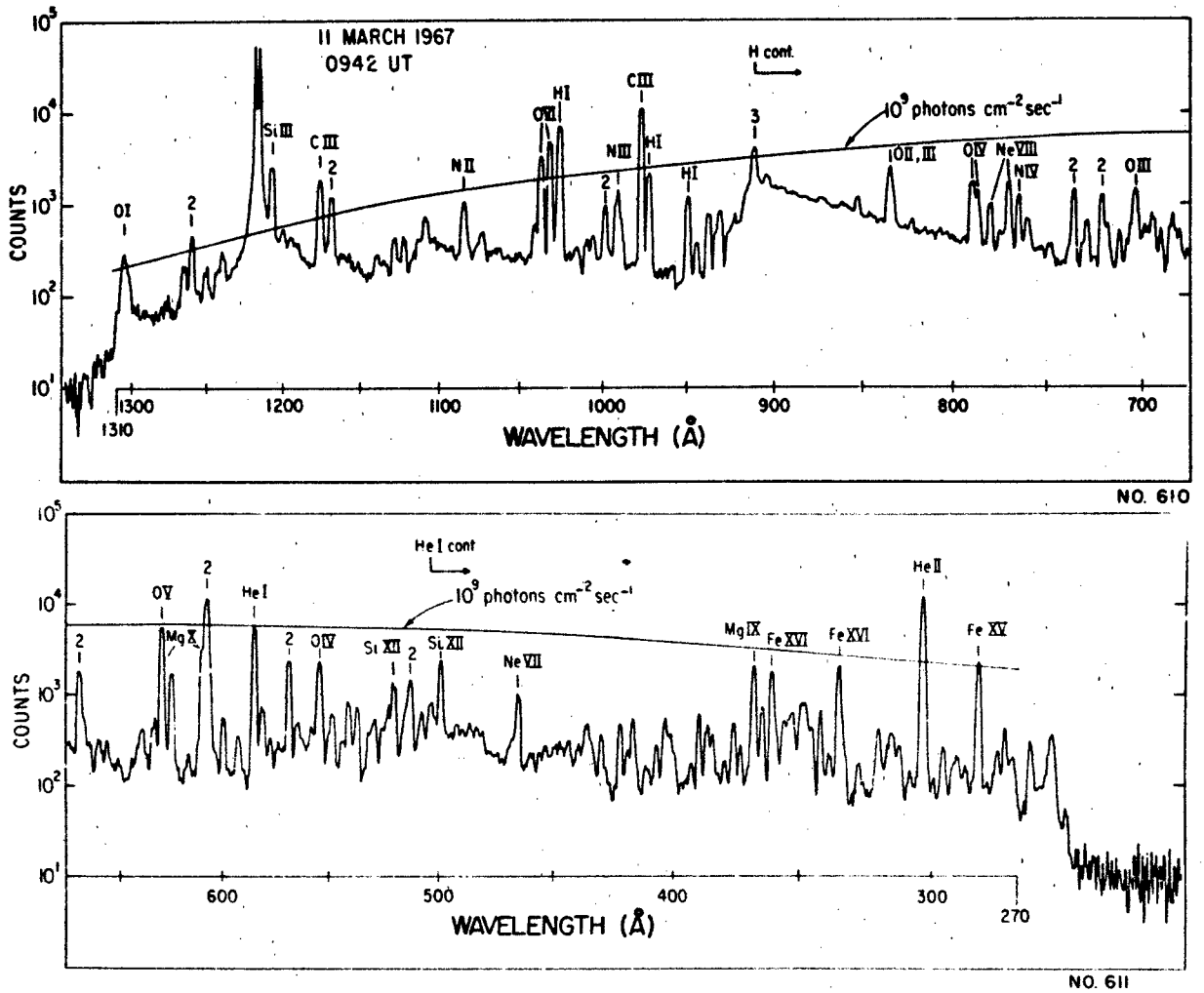


Figure 2-5. The solar UV spectrum between 1300 and 270 Å observed from the satellite OSO 3. The folding of the data for H Ly α line at 1216 Å is introduced by the pulse counting circuitry [Hall and Hinteregger, 1970].

ar 270 Å. The reversal of the center of the intense hydrogen Lyman- α line at 1216 Å is instrumental, caused by erflow of the photon counter. The hydrogen continuum ıich extends from about 911 Å toward shorter wavelengths the only significant continuum in the spectrum for wavelgths shorter than 1300 Å. A detailed discussion of the ectrum shown in Figure 2-5 is given by Hall and Hinterger [1970].

The solar UV spectrum between 300 and 100 Å meared in a rocket flight on 4 April 1969 is given in Figure 6. The figure is from a paper by Malinovsky and Heroux 973] where the top of the atmosphere intensities of the ıes are also tabulated. For the wavelength interval 300-:0 Å, one can see that this small interval of wavelength ıntains multiplets of ions produced over a wide temperature nge extending from 2×10^4 to 2×10^6 K, corresponding to the temperature region of formation of He II and Fe XVI, respectively. The reversal of the intense emission line of Fe XV at 284 Å is again instrumental, caused by overflow of the photon counter.

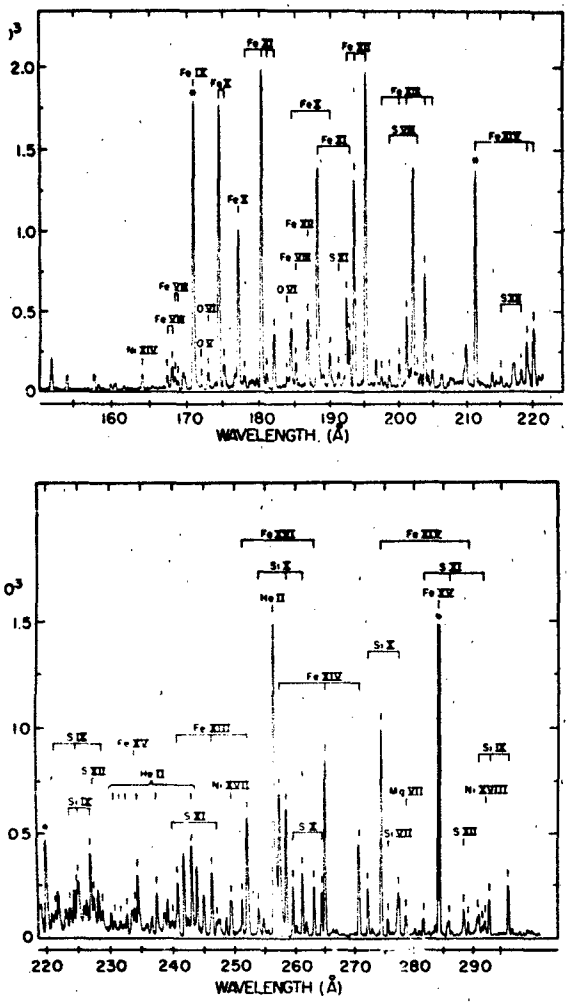


Figure 2-6. The solar UV spectrum between 100 and 300 Å obtained from a rocket flight on 4 April 1969. The folding of the data for the Fe XV line at 284 Å is introduced by the pulse counting circuitry [Malinovsky and Heroux, 1973].

2.3 ABSORPTION OF SOLAR UV

2.3.1 Absorption Below 3200 Å

Solar radiation at wavelengths shorter than about 3200 Å is totally absorbed in the earth's upper atmosphere. Although this ultraviolet region accounts for less than 2% of the total solar irradiance, it is the principal source of energy in the upper atmosphere and controls the neutral and ion composition, temperature, and photochemistry in the stratosphere, mesosphere, and thermosphere. The absorption of solar UV in the earth's atmosphere for an overhead sun is illustrated in Figure 2-7, which gives the altitude at which the rate of absorption at a particular wavelength is a maximum. At these altitudes, solar radiation is reduced by a factor of $e^{-1} = 0.37$ of its value above the earth's atmosphere. The primary atmospheric constituents that absorb the radiation in the different wavelength regions are also given in the figure.

2.3.2 UV Atmospheric/Ionospheric Processes

The absorption of solar UV radiation at wavelengths shorter than 1027 Å ionizes the primary atmospheric constituents O_2, N_2, and O. This radiation is absorbed at altitudes above about 100 km. The rate at which the ions O_2^+, N_2^+ and O^+ are produced as a function of altitude is shown in Figure 2-8. These rates were calculated by using rocket measurements of solar UV intensities for 23 August 1972. The corresponding total production rates of the three species of ions produced by radiation within several wavelength intervals of the solar UV spectrum are given in Figure 2-9. A detailed account of the calculations used to obtain these figures is given by Heroux et al. [1974].

2.4 SOLAR UV IRRADIANCE VARIABILITY

2.4.1 Sunspot Number and 10.7cm Solar Radio Flux

The sunspot number and the 10.7 cm (2800 MHz) solar radio flux are used extensively as indices of solar activity. Eddy [1977] gives a review of sunspot observations that have been made for several hundred years. These records clearly indicate that the sunspots are indicators of solar activity. An additional diagram of the latitudinal distribution of sunspots during the period 1874-1976 is given by Yallop and Hohenkerk [1980]. The sunspot numbers for the present solar cycle 21 are plotted in Figure 2-10. The values are the monthly mean sunspot numbers obtained from the Solar-Geophysical Data Reports published monthly by the Na-

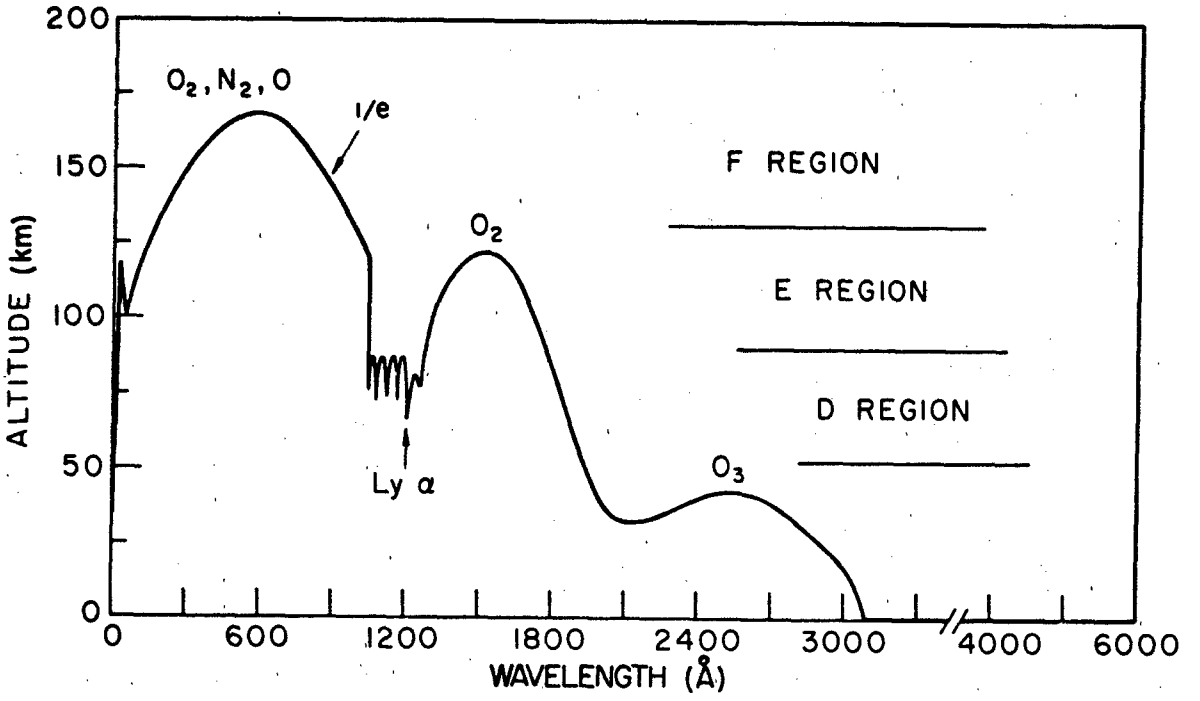


Figure 2-7. The altitude at which the rate of absorption of solar UV radiation is at maximum. The principal atmospheric constituents that absorb the radiation in the different wavelength bands are indicated.

tional Oceanic and Atmospheric Administration (NOAA). These reports also give the daily values of the sunspot numbers. The smoothed mean values of the sunspot numbers for cycles 8-20 are also included in the figure. Each cycle's beginning minimum has been shifted to coincide with the minimum of cycle 21. For comparison, the monthly mean values of the 10.7 cm radio flux for cycle 21 obtained from the NOAA reports are also given in Figure 2-10. Both indices show a similar increase with increasing activity for cycle 21, which began in July 1976. The daily variations

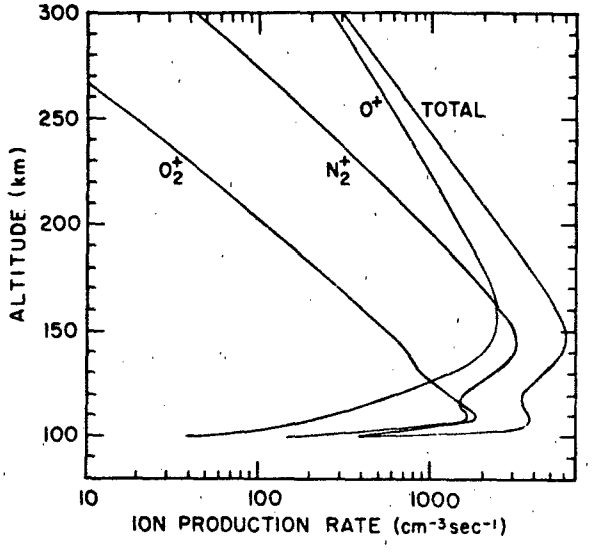


Figure 2-8. Production rates of O^+, O_2^+, and N_2^+ and their totals as a function of altitude. The rates were calculated by using rocket measurements of solar intensities between 1027 and 52 Å of 23 August 1972 [Heroux et al., 1974].

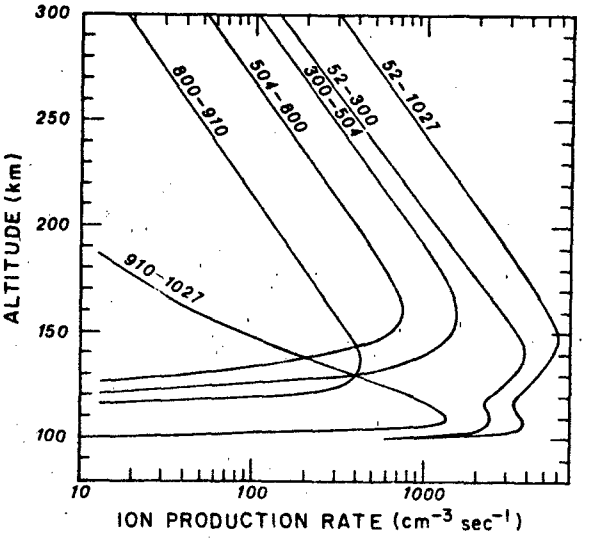


Figure 2-9. Total production rates of O^+, O_2^+, and N_2^+ produced by several wavelength regions of the solar UV spectrum. The rates were calculated by using the solar intensities of 23 August 1972 [Heroux et al., 1973].

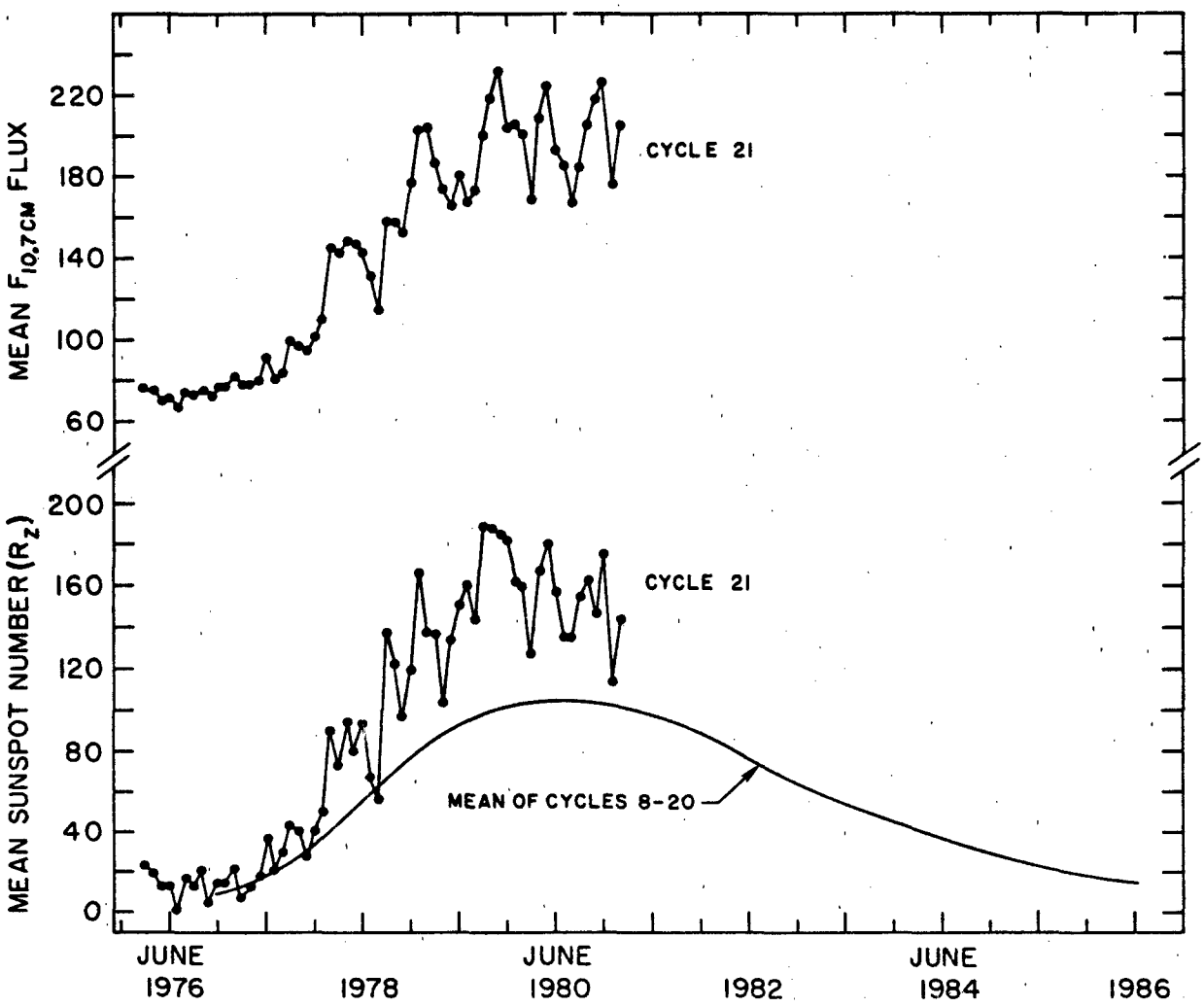


Figure 2-10 Monthly mean sunspot numbers and 10.7 cm radio flux for solar cycle 21, which began in July 1976. The smoothed mean values of the sunspot numbers for cycles 8-20 were obtained by shifting each cycle's minimum value to coincide with the minimum of cycle 21.

in both the sunspot number and the 10.7 cm flux can be significant. For November 1979, the mean value of the sunspot number is 185, while the daily values during the month varied from a minimum value of 98 to a maximum value of 302. For the same month, the mean value of the 10.7 cm flux was 232, while the minimum and maximum values were 154 and 374, respectively.

Although the 10.7 cm flux is a reliable indicator of the general level of solar activity, the correlation between the solar UV flux in narrow wavelength regions and the daily 10.7 cm flux is highly unreliable. Solar UV emissions originate from a wide range of temperatures, densities, and heights in the solar atmosphere extending from the lower chromosphere, the chromosphere-corona transition region, and the corona. The 10.7 cm flux, however, originates predominantly from the upper chromosphere and lower corona. Therefore, it is not surprising that the UV flux emitted from a wide range of regions in the solar atmosphere does not correlate well with the daily 10.7 cm flux emitted from a limited region in the solar atmosphere.

2.4.2 Solar Cycle and UV Variability Below 2000 Å

The variation of the spectral irradiance from the minimum through the maximum level of activity of a solar cycle can be significant for wavelengths shorter than about 1800 Å. This variation is illustrated in Figure 2-11, where the ratio of the irradiance near solar maximum during January 1979 to the irradiance near solar minimum during July 1976 for solar cycle 21 is plotted for 25 Å intervals of wavelength. The data were obtained by Hinteregger [1981] from the

AFGL spectrometer on the NASA Atmosphere Explorer (AE-E) satellite. Figure 2-11 shows that the variation decreases with increasing wavelength toward 1800 Å. At wavelengths longer than about 2300 Å the ratio approaches unity, and therefore solar activity has a negligible effect on the spectral irradiance at these longer wavelengths. The large ratios that are apparent at wavelengths below 500 Å originate from highly ionized atomic species produced in the solar corona. Individual lines that fall within the 25 Å intervals can vary by a significantly higher ratio than that plotted in Figure 2-11 since the plotted ratio represents the irradiance averaged over several lines that fall within the wavelength interval. For example, the ratio of the coronal emission line of Fe XVI at 335.41 Å increases by a factor greater than 100 from solar minimum to maximum, although the ratio of the averaged irradiance in the interval 325-350 Å is a factor of 12.

Figure 2-12 illustrates the relative variation of the solar spectral irradiance for several 50 Å intervals in the 1400-1750 Å range, along with the hydrogen Lyman-α line, at 1216 Å. The variations for the Fe XV 284 Å line emission and for two intervals of the 170-205 Å range are shown in Figure 2-13. These data were obtained on the AE-E satellite by Hinteregger [1980]. Each point represents the average of all measurements for the month. For comparison, the monthly averages of the sunspot number R_z and the 10.7 cm radio emission are included in the figures. These data illustrate that the variation of the irradiance with solar ac-

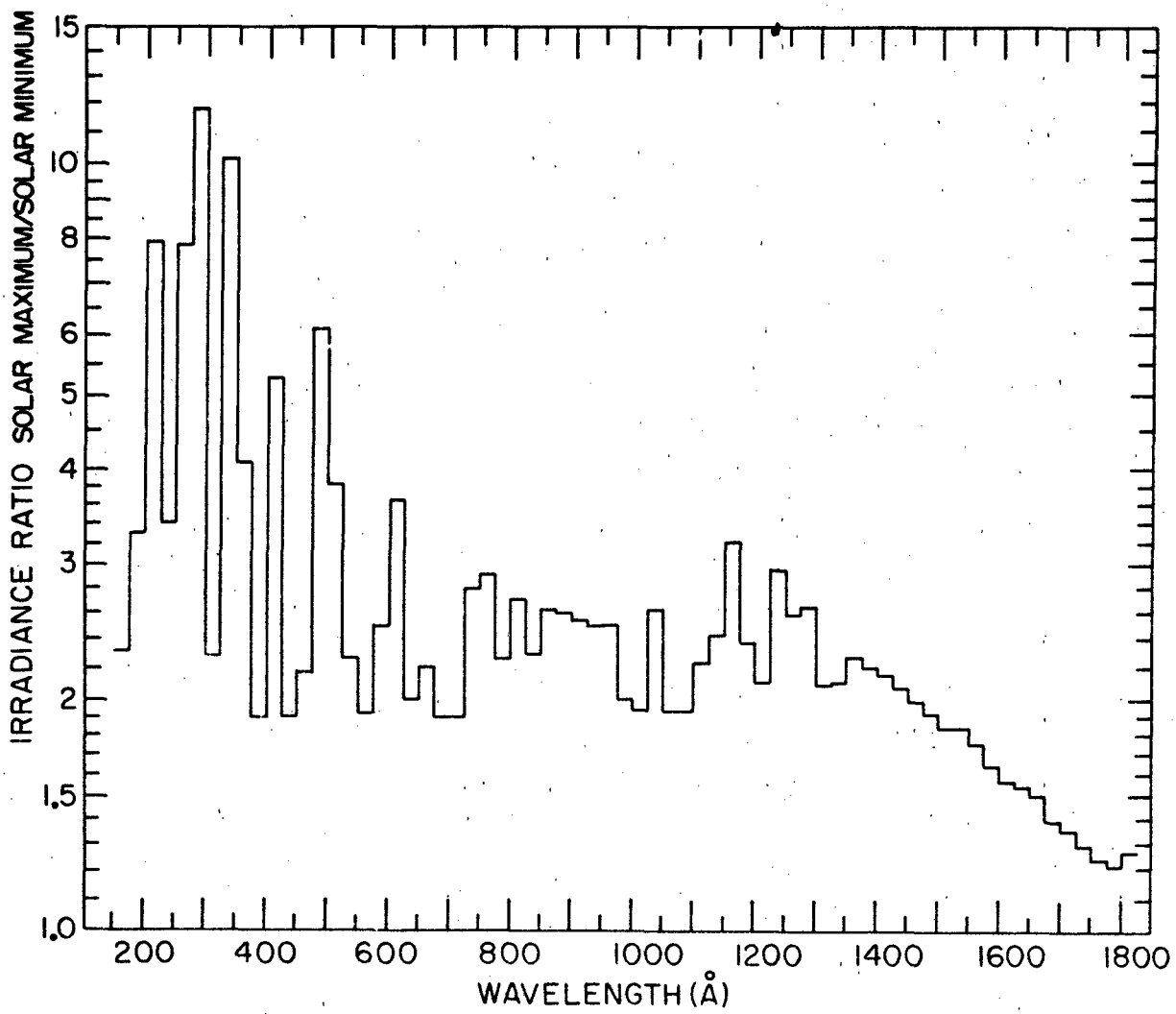


Figure 2-11. The ratio of the solar spectral irradiance near solar maximum during January 1979 to the irradiance near solar minimum during July 1976 for solar cycle 21. The ratios are plotted for 25 Å intervals of wavelength [Hinteregger, 1981].

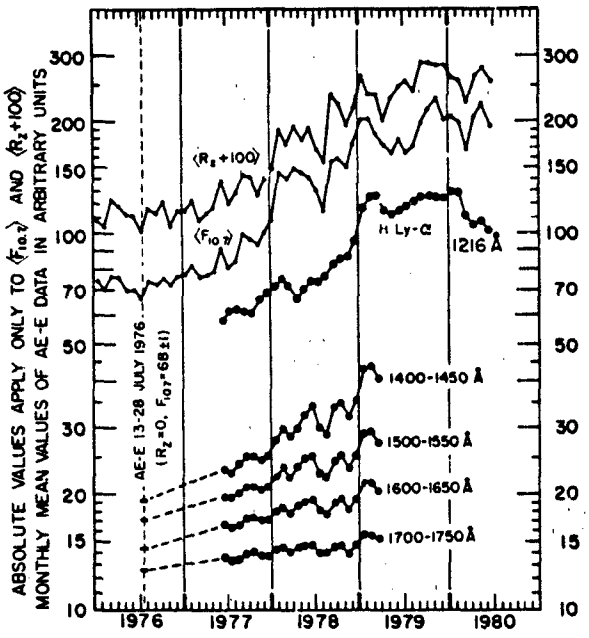


Figure 2-12. Irradiance variations at wavelengths from 1216 to 1750 Å from AE-E satellite measurements during solar cycle 21 [Hinteregger, 1980].

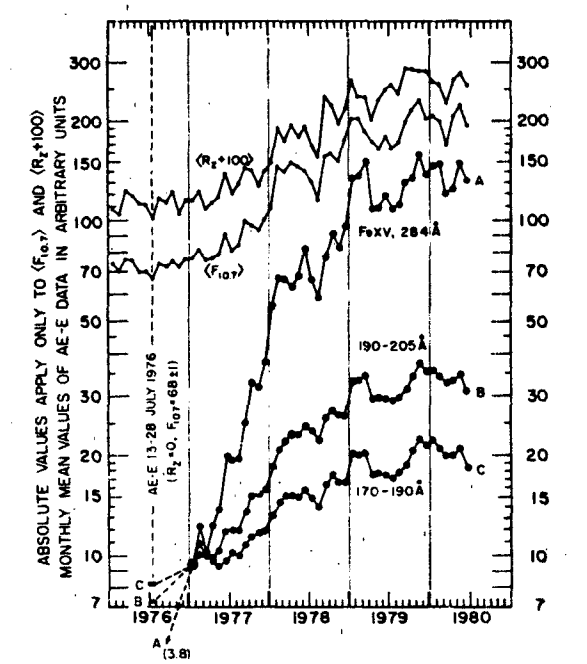


Figure 2-13. Irradiance variations of some solar coronal emissions from AE-E satellite measurements during solar cycle 21 [Hinteregger, 1980].

tivity between 1977 and 1980 depends upon the region of production of the radiation in the solar atmosphere. The variation of the H Lyman-α line that is produced in the lower chromosphere is a factor of approximately 2.3, while the variation of the Fe XV line produced in the corona is approximately 8.0 The variation of the photospheric radiation in the wavelength interval 1700-1750 Å is only 1.16. This variation approaches 1.0 for wavelengths greater than about 2200 Å.

In addition to the variability of the solar UV irradiance over an 11-year cycle, there is also a 27-day variation of solar UV associated with the period of solar rotation. The magnitude of the variation at different wavelengths depends upon the region of the solar atmosphere emitting the radiation. The variations of the emission lines from highly ionized species produced in the corona are significantly greater than the variations of lines produced in the lower chromosphere. This is illustrated in Figure 2-14 where measurements are plotted of the 27-day variation of several solar emission lines for May and June 1967 from an experiment on the OSO 3 satellite [Hall and Hinteregger [1970]. The variation of the coronal Fe XVI lines at 335 Å can be seen to be significantly greater than the variations of the other chromospheric lines shown in the figure. The 27-day variation of the unresolved He II and Si XI lines near 304 Å has also been measured over a period of almost two years (1967 to 1969) from an instrument flown on the OSO 4 satellite by Timothy and Timothy [1970].

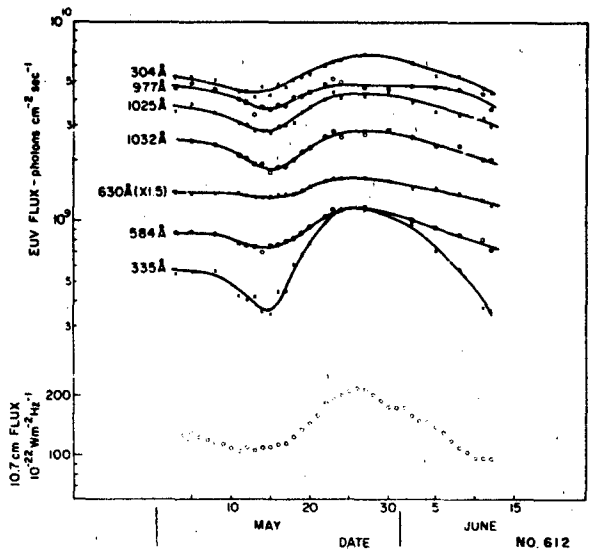


Figure 2-14. The 27-day variations in several UV emission lines observed from the OSO 3 satellite in May and June 1967. The 10.7 cm radio flux for the same period is also given for comparison [Hall and Hinteregger, 1970].

REFERENCES

Delaboudinier, J.P., R.F. Donnelly, H.E. Hinteregger, G. Schmidtke, and P.C. Simon. "Intercomparison/ Compilation of Relevant Solar Flux Data Related to Aeronomy" International Council of Scientific Unions, Committee on Space Research, *COSPAR Technique Manual Series*, **7**:48–51, Paris, 1978.

Eddy, J.A., "Historical Evidence for Existence of the Solar Cycle," in *The Solar Output and Its Variation*, edited by O.R. White, Colorado Associated University Press, Boulder, 1977.

Frohlich, C., "Contemporary Measures of the Solar Constant," in *The Solar Output and its Variation* edited by O.R. White, Colorado Associated University Press, Boulder, 1977.

Hall, L.A., "Solar Ultraviolet Irradiance at 40 Kilometers in the Stratosphere," *J. Geophys. Res.*, **86:** 555, 1981.

Hall, L.A. and H.E. Hinteregger, "Solar Radiation in the Extreme Ultraviolet and Its Variation with Solar Rotation," *J. Geophys. Res.*, **75:** 6959, 1970.

Heroux, L. and H.E. Hinteregger, "Aeronomical Reference Spectrum for Solar UV Below 2000 Å," *J. Geophys. Res.*, **83:** 5305, 1978.

Heroux, L., M. Cohen, and J.E. Higgins, "Electron Densities Between 110 and 300 km Derived from Solar EUV Fluxes of August 23, 1972," *J. Geophys. Res.*, **79:** 5237, 1974.

Hinteregger, H.E., "Representations of Solar EUV Fluxes for Aeronomical Applications," *Adv. Space Res.*, **1:** 39, COSPAR, 1981.

Hinteregger, H.E., "AE-E Experiences of Irradiance Monitoring for 1250-1850 Å," in *Proceedings of the Workshop on Solar UV Irradiance Monitors*, NOAA Environmental Research Laboratories (ERL), Boulder, 1980.

Kreplin, R.W., K.P Dere, D.M. Horan, and J.F. Meekins, "The Solar Spectrum Below 10 Å," in *The Solar Output and Its Variation*, edited by O.R. White, Colorado Associated University Press, Boulder, 1977.

Malinovsky, M. and L. Heroux, "An Analysis of the Solar Extreme-Ultraviolet Spectrum Between 50 and 300 Å," *Astrophys. J.*, **181**: 1009, 1973.

Manson, J.E., "The Solar Spectrum Between 10 and 300 Å," in *The Solar Output and Its Variation*, edited by O.R. White, Colorado Associated University Press, Boulder, 1977.

Mount, G.H. and G.J. Rottman, "The Solar Spectral Irradiance 1200-3184 Å Near Solar Maximum: July 15, 1980," *J. Geophys. Res.*, **86:** 9193–9198, 1981.

Pierce, A.K. and R.G. Allen, "The Solar Spectrum Between 0.3 and 10 μm," in *The Solar Output and Its Variation*, edited by O.R. White, Colorado Associated University Press, Boulder, 1977.

Timothy, A.F. and J.G. Timothy, "Long-Term Intensity Variations in the Solar Helium II Lyman Alpha Line," *J. Geophys. Res.*, **75:** 6950, 1970.

Willson, R.C., C.H. Duncan, and J. Geist, "Direct Measurement of Solar Luminosity Variation," *Science*, (London) **207:**177–179, 1980.

Yallop, B.D. and C.Y. Hohenkerk, "Distribution of Sunspots 1874–1976," *Sol. Phys.*, **68:** 303, 1980.

Chapter 3

SOLAR WIND

J. Feynman

The solar wind or interplanetary medium is a fully ionized, electrically neutral plasma that carries a magnetic field and streams outward from the inner solar corona at all times. The wind is highly variable in both time and space and fills the interstellar region in the vicinity of the sun. This chapter describes the observed properties of the solar wind in the vicinity of the earth, including the undisturbed wind and the magnetosheath and bow shock formed as the solar wind flows around the earth's magnetic field.

3.1 BASIC THEORY

3.1.1 Simple Spherically Symmetric Treatment

The solar wind is the outer part of the sun's corona, streaming past the earth. Chapman [1957] made an early estimate of the particle density expected in space at the orbit of the earth. He considered the expansion of a corona with a temperature at the base on the order of 10^6 K. He calculated the coronal density profile assuming that the corona was in hydrostatic equilibrium and found that the electron density at the earth's orbit would be 10^2 or 10^3 particles/cm^3. He also found that the pressure at large heliocentric distances would be of the order of 10^{-5} dyn/cm^2, whereas estimates of the interstellar pressure were in the range of 10^{-12} or 10^{-13} dyn/cm^2. Thus, the hydrostatic corona could not be matched with its boundary conditions in the distant heliosphere. This treatment, although describing the sun in a very simplified manner, led Parker to postulate that the corona was not in hydrostatic equilibrium and that, in fact, the coronal plasma was flowing out from the sun continuously. Parker [1963] treated the problem as that of a steady state spherically symmetric fluid expansion. The equations that govern the behavior of such a system are the equations of mass, momentum, and energy conservation. The mass conservation is expressed by

$$\frac{1}{r^2}\frac{d}{dr}(r^2 \rho u) = 0, \tag{3.1}$$

where r is the radial distance, ρ the mass density, and u the radial velocity. The momentum equation for a fluid is written as

$$\rho u \frac{du}{dr} = \frac{-dP}{dr} - \rho \frac{GM_\odot}{r^2}, \tag{3.2}$$

where P is the scalar pressure, G the gravitational constant, and $M_\odot$ the mass of the sun. Energy conservation is given by

$$\frac{1}{r^2}\frac{d}{dr}\left[r^2 \rho u \left(\frac{1}{2}u^2 + \frac{3}{2}\frac{P}{\rho}\right)\right] = \frac{-1}{r^2}\frac{d}{dr}(r^2 P u) - \rho u \frac{GM_\odot}{r^2} + S(r) \tag{3.3}$$

where S(r) is an energy source or sink term. The corona is assumed to be a neutral electron proton gas so that

$$\rho = n(m_p + m_e) = nm \tag{3.4}$$

where m_p and m are the proton mass and m_e the mass of the electron. For equal electron and proton temperature the pressure is given by

$$P = 2nkT. \tag{3.5}$$

Parker further simplified the problem by assuming that the pressure and temperature were related by the polytropic law

$$P = P_\odot \left(\frac{\rho}{\rho_\odot}\right)^\alpha \tag{3.6}$$

where α is the polytropic index. An isothermal gas corresponds to $\alpha = 1$ and an adiabatic process to $\alpha = 5/3$. The assumption of the polytropic law effectively assumes a solution to the energy conservation equation with an implied source term. Parker examined the solutions of the equations for a range of selections for α. See Parker [1963] or Hundhausen [1972] for further discussion and details.

CHAPTER 3

The simplest case, but one that illustrates important points concerning the behavior of the solutions, is that of the isothermal corona. Near the sun the isothermal assumption is probably valid. Using the expressions for the density and the polytrope law with $\alpha = 1$, the momentum conservation equation becomes

$$nmu\frac{du}{dr} = -2kT\frac{dn}{dr} - \frac{nmGM_o}{r^2}. \qquad (3.7)$$

Equation (3.1) for mass conservation is integrated to give

$$4\pi nur^2 = \text{constant}, \qquad (3.8)$$

that is, the flux through a sun centered sphere is a constant. Equation (3.8) can be used to eliminate the number density from Equation (3.7)

$$\frac{1}{u}\frac{du}{dr}\left(u^2 - 2\frac{kT}{m}\right) = \frac{4kT}{mr} - \frac{GM_o}{r^2}. \qquad (3.9)$$

Assume that the temperature of the corona is such that

$$T < \frac{GM}{r_o}\frac{m}{4k},$$

where r_o is the base of the corona. The right hand side of Equation (3.9) is negative in the range of $r_o < r < r_c$ where

$$r_c = \frac{GM_o m}{4kT}, \qquad (3.10)$$

and r_c is called the "critical radius." For $r > r_c$, the right hand side of Equation (3.9) will be positive. For $r = r_c$, the right hand side must be zero. In that case, to make the left hand side zero either

$$u^2(r_c) = \frac{2kT}{m} \qquad (3.11a)$$

or

$$\left.\frac{1}{u}\frac{du}{dr}\right|_{r = r_c} = 0. \qquad (3.11b)$$

These two conditions expressed in Equation (3.11a and b) define four classes of solutions to the set of simultaneous equations governing the model corona. If Equation (3.11a) is satisfied then du/dr has the same sign for all r, that is, u either increases or decreases monotonically with distance from the sun. If Equation (3.11b) is satisifed, then u is either maximum or minimum near r_c. These four classes of solutions for the velocity are shown in Figure 3-1. The physical solutions for the solar wind are members of the class of solutions that satisfy the physically relevant boundary conditions. Class 1 solution velocities approach zero near the sun and at great distances. Class 2 is the unique solution that has a low velocity near the sun, passes through the critical point, and has high velocity at great distances. Both class 3 and 4 solutions have high velocity near the sun. Class 3 is the unique solution passing through the critical point and having low velocity at great distances. whereas class 4 solutions have high distant velocities. Since, for the physical solutions, the velocity near the sun is small, class 3 and 4 solutions are immediately eliminated. Class 1 solutions would give finite pressures at large distances and would result in the same problem of matching to the interstellar medium as did the hydrostatic corona. For class 2, however,it can be shown that the expansion speed continues to increase slowly as $r \to \infty$, so that the density and hence the pressure approach zero at infinity. Thus the class 2 solution satisfies the boundary conditions and represents the physically existing solar wind. Since according to Equation (3.11a) the velocity at r_c is equal to the velocity of sound, this solution is supersonic for $r > r_c$.

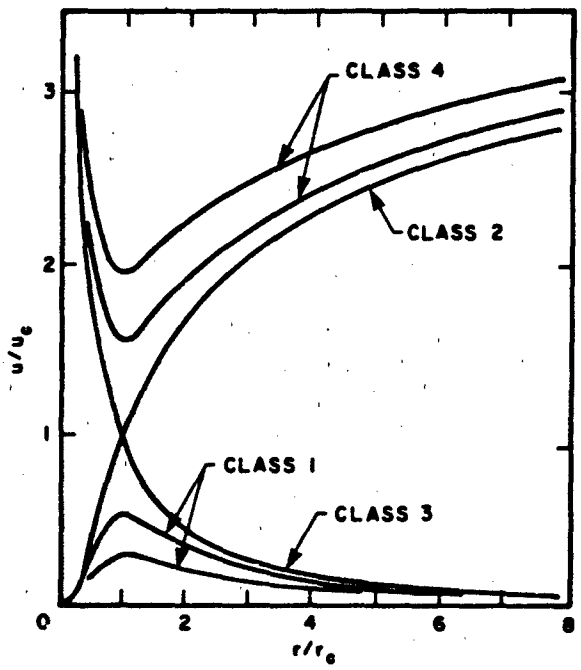


Figure 3-1. Topology of solutions to the solar wind equations for the spherically symmetric isothermal case. The only solution that can satisfy the boundary conditions of low velocity at the sun and low pressure at large distances is the class 2 solution [Hundhausen, 1972].

Although the dicussion above is highly simplified [fol-

lowing Hundhausen, 1972] the essential characteristics are maintained in all more sophisticated treatments. That is, a family of solutions in every treatment exists but the imposition of the boundary conditions selects a finite number of solutions that have low velocities at the sun, pass through critical points where the flow becomes supersonic, and result in a high finite velocity and zero pressure at infinity.

The configuration of the solar wind magnetic field was also treated by Parker [1963]. Since the conductivity of the solar wind is extremely high, there is essentially no diffusion of the plasma transverse to the magnetic field and the field is said to be "frozen in" to the flow. The plasma carries the field with it into space. If the sun were not rotating, the field would extend straight outward in all directions but the rotation causes the field to appear wound up as shown in Figure 3-2. In a spherical system of coordinates (r, ϕ, θ)

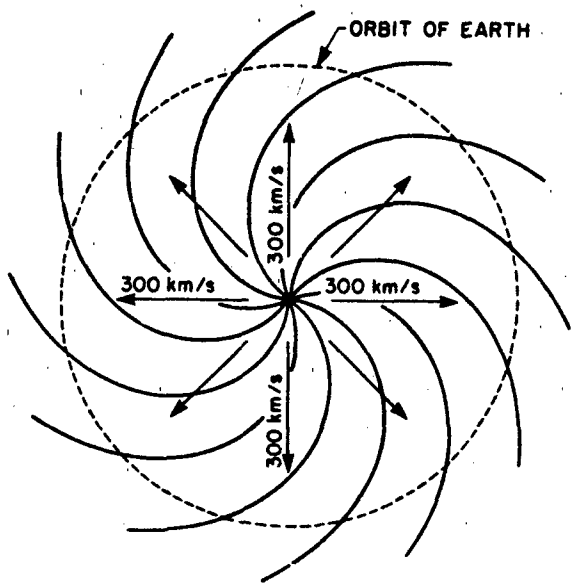


Figure 3-2. The configuration of the interplanetary magnetic field in the equatorial plane for a steady solar wind. The field winds up because of the rotation of the sun. Fields in low speed wind will be more wound up than those in high speed wind [Parker, 1963]. (Reprinted with permission from J. Wiley and Sons © 1963.)

rotating with the sun, the velocity of a fluid element carrying the field is given by

$$U_r = u$$
$$U_\phi = -\omega r \sin\theta \qquad (3.12)$$
$$U_\theta = 0$$

where u is the speed of the wind and ω is the angular velocity of solar rotation ($\omega = 2.7 \times 10^{-6}$ rad/s).

The velocity of the wind changes only slowly with radius beyond a few times the critical radius and will be approximated by a constant. The path followed by the fluid element is determined by

$$\frac{1}{r}\frac{dr}{d\phi} = \frac{U_r}{U_\phi} = \frac{u}{-\omega r \sin\phi}. \qquad (3.13)$$

For constant u beyond r_o this is integrated to give

$$r - r_o = \frac{-u}{\omega \sin\theta}(\phi - \phi_o) \qquad (3.14)$$

Then for a spherically symmetric geometry $\nabla \cdot \mathbf{B} = 0$ gives

$$B_r(r, \phi, \theta) = B(r_o, \phi_o, \theta)\left(\frac{r_o}{r}\right)^2$$
$$B_\phi(r, \phi, \theta) = -B(r_o, \phi_o, \theta)\frac{\omega r_o^2}{ur}\sin\theta \qquad (3.15)$$
$$B_\theta = 0.$$

The configuration near the solar equatorial plane is shown in Figure 3-2.

Transforming to a stationary coordinate system leaves the magnetic field configuration unchanged but in the stationary system there is an electric field

$$\mathbf{E} = -\mathbf{u} \times \mathbf{B} \qquad (3.16)$$

so that the direction of the plasma flow is radial rather than along the field lines.

3.1.2 Further Considerations

This simple spherically symmetric solar wind theory was remarkably successful in predicting the existence of the wind and estimating the velocity and density. However, it soon became apparent that the predicted velocity was lower than that observed and the predicted density was too large. Refinements of solar wind theory have taken many approaches, none of them as yet completely successful. The earliest attempt at improvement was to consider the electrons and protons as two fluids each expanding separately from the sun. The particles within each fluid were interacting with each other through large scale electric fields. This two-fluid model produced results that were again too slow and too dense, and the predicted total energy flux at 1 AU was twice that observed. Later, observations showed that the solar wind came from restricted areas of the sun so that the expansion of the wind was faster than the r^2 appropriate to the

spherical case discussed by Parker [1963]. Studies of the effect of this geometrical difference, however, showed that although the acceleration of the solar wind at small distances from the sun was increased markedly, the speed at 1 AU was not raised appreciably. A possible cause for the speed of the wind being higher than expected from simple theory is the effect of magnetohydrodynamic waves on the wind in the supersonic region. Theory indicates these waves will exert a force on the solar wind tending to accelerate it. At 1 AU, outward propagating Alfvén waves (Section 3.2.5) are the predominant mode observed. Fast mode MHD waves (Section 3.2.5) propagating outward near the sun would also accelerate the wind. Since this latter mode is strongly damped, it is quite reasonable that if there was a wave flux near the sun it would be lost before the wind propagated to 1 AU. This consideration is important because fast mode MHD waves are rarely seen at 1 AU. For both these wave acceleration mechanisms, an important problem is that the existence of the required waves in the required region has not been demonstrated. An alternative idea for wind acceleration is that the electron thermal conductivity has been overestimated by using the classical values, and at the same time, the thermal coupling between the electrons and protons has been underestimated. As with several other proposals listed here, this seems to be an attractive possibility but the observations and analysis have not been carried far enough to make a definitive test.

3.2 PARTICLE AND FIELD OBSERVATIONS

3.2.1 Coordinates

There are two coordinate systems commonly used to describe observations of the interplanetary medium [Russell, 1971]: geocentric solar ecliptic (GSE) and the geocentric solar magnetospheric (GSM). In the GSE system, X points from the earth towards the sun, Y is in the ecliptic plane and negative in the direction of planetary motion, and Z is parallel to the ecliptic pole. In GSM, X is along the earth sun line and positive in the direction of the sun, Z is positive toward the north and is perpendicular to X and in the plane which contains X and the earth's magnetic dipole axis, and Y completes the right handed system of coordinates.

GSM coordinates are usually used to describe interactions between the solar wind and the earth's magnetic field whereas GSE coordinates are used to describe the undisturbed solar wind. A second system sometimes used for the undisturbed solar wind is the geocentric solar equatorial system in which the X axis remains in the ecliptic plane pointing from the earth to the sun, but Y is parallel to the sun's equatorial plane and Z completes the system. The sun's pole of rotation is tipped at 7° to the plane of the ecliptic. The transformations among the systems are given in Russell [1971].

3.2.2 Magnetic Fields

Observations of the magnetic field in the vicinity of the earth show that the field is, on the average, in the spiral direction predicted by the steady state model but that there are often large variations. The variations are ascribed to several sources: waves, discontinuities, and certain large scale disturbances. These topics will be described in Section 3.3. Here the basic large scale structure and average properties will be described.

The magnetic field in the solar wind is, as expected, drawn out in a spiral such that the angle the field makes with the earth-sun line is a function of the velocity, low solar wind speeds resulting in a tight spiral and faster speeds in a more radially aligned field. The angle increases with distance from the sun. For the observed velocities of the solar wind, the magnetic field has a spiral angle of about 45° at the earth. The magnitude of the field is on the order of a few nanotesla (earlier literature used the term gamma, $1\gamma = 1$ nanotesla). An example of observed magnetic field distributions is given in Figure 3-3a, b, and c [Hirshberg, 1969]. The direction of the radial component (often referred to as the x component) is, of course, either towards or away from the sun. Wilcox and Ness [1965] found that the field typically maintains one of these orientations for many days and then rapidly changes to the other. The region of space in which an orientation is maintained is called a "sector". The sector orientation is usually determined on the basis of one-hour or three-hour averages of the radial field component, but the maintenance of the sign of B_x is typically seen on a smaller time scale. Transitions from one sector to another take place over a period of minutes or hours. At transitions the direction of the interplanetary field usually changes sign rapidly while maintaining more or less the same magnitude. Sometimes the sector boundary on the scale of minutes or tens of minutes displays only a single change of sign. Sometimes the sign changes back and forth many times and over many hours before the field settles down to the new orientation. When this happens the position of the boundary may be hard to define. An example of the observed structure is shown in Figure 3-4. Typically, two or four sectors occur during a solar rotation. The sector structure can also be inferred from high latitude surface magnetic variations [Svalgaard, 1973]. The accuracy of the procedure varies with season and station latitude. The sector structure is stable over many months during the declining phase of the solar cycle, but changes from month to month during the increasing phase. The sector structures and/or the high speed streams (Section 3.3.2) apparently order a large amount of solar wind phenomenology. Studies of the typical behavior of the interplanetary medium within sectors

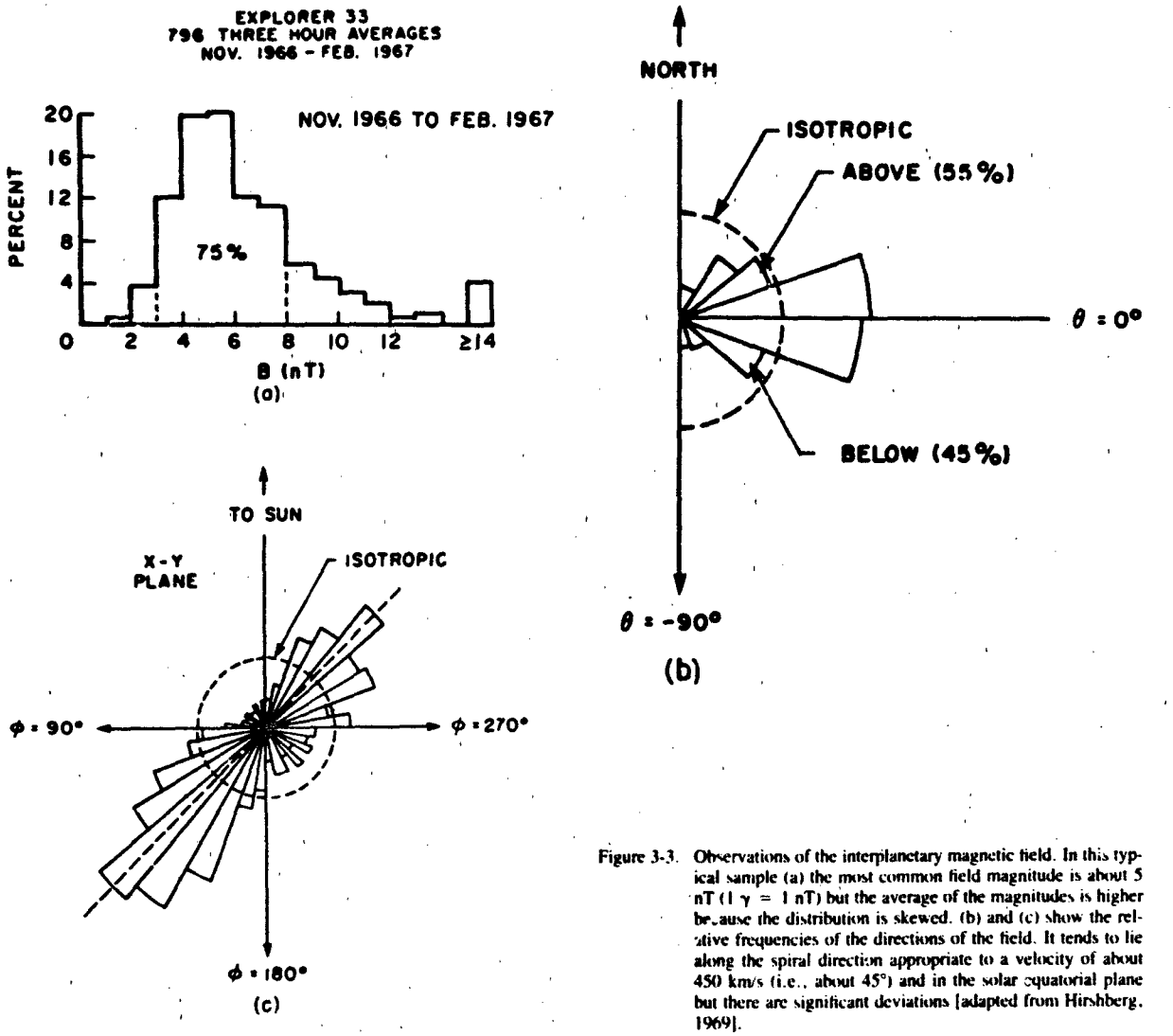


Figure 3-3. Observations of the interplanetary magnetic field. In this typical sample (a) the most common field magnitude is about 5 nT (1 γ = 1 nT) but the average of the magnitudes is higher because the distribution is skewed. (b) and (c) show the relative frequencies of the directions of the field. It tends to lie along the spiral direction appropriate to a velocity of about 450 km/s (i.e., about 45°) and in the solar equatorial plane but there are significant deviations [adapted from Hirshberg, 1969].

are relatively simple to carry out using sector boundaries as convenient time markers. The results of some of these studies are described in Section 3.3 on large scale disturbances.

Observations from spacecraft that traveled to solar latitudes of over 10° from the solar equatorial plane showed that the sector structure was dependent on latitude. The structure tends to disappear at higher latitudes, and the interplanetary magnetic field then has the sign of the solar magnetic field of the appropriate solar pole. The picture that has developed to explain these observations is shown in Figure 3-5. Plasma leaves the sun predominantly at high latitudes and flows out and towards the equator where a current sheet is formed corresponding to the change in magnetic field polarity. The current sheet is tipped relative to the sun's plane of rotation and warped so that as the sun rotates the earth passes through the current sheet, experiencing periods of alternating polarities, that is, sectors.

In the vicinity of the earth a typical magnetic field intensity $|B|$ is about 6 nT, but hourly average values vary from about 1 nT to a high of 37 nT. The highest recorded single value was 170 nT observed on 4 August 1972 [Burlaga and King, 1979]. Figure 3-6 shows the distribution of over 70 000 hourly values collected from satellites at 1 AU between 1963 and 1977. The distribution of the magnitude $|B|$ is skewed towards high values. However, the distribution of the log $|B|$ is distributed normally with a mean of 0.75 ($|B| = 5.6$ nT) and an rms deviation of 0.18. For about 3% of these hours log $|B|$ exceeded the mean by at least 20, that is $|B| \geq 13$ nT. When the field is high it typically remains so for many hours, forming distinct events. There were about 100 such events observed between 1963 and 1974 of which 75% were associated with shocks or interfaces between interacting solar wind streams (Section 3.3.2). Periods of high-intensity field may persist for as long as 24

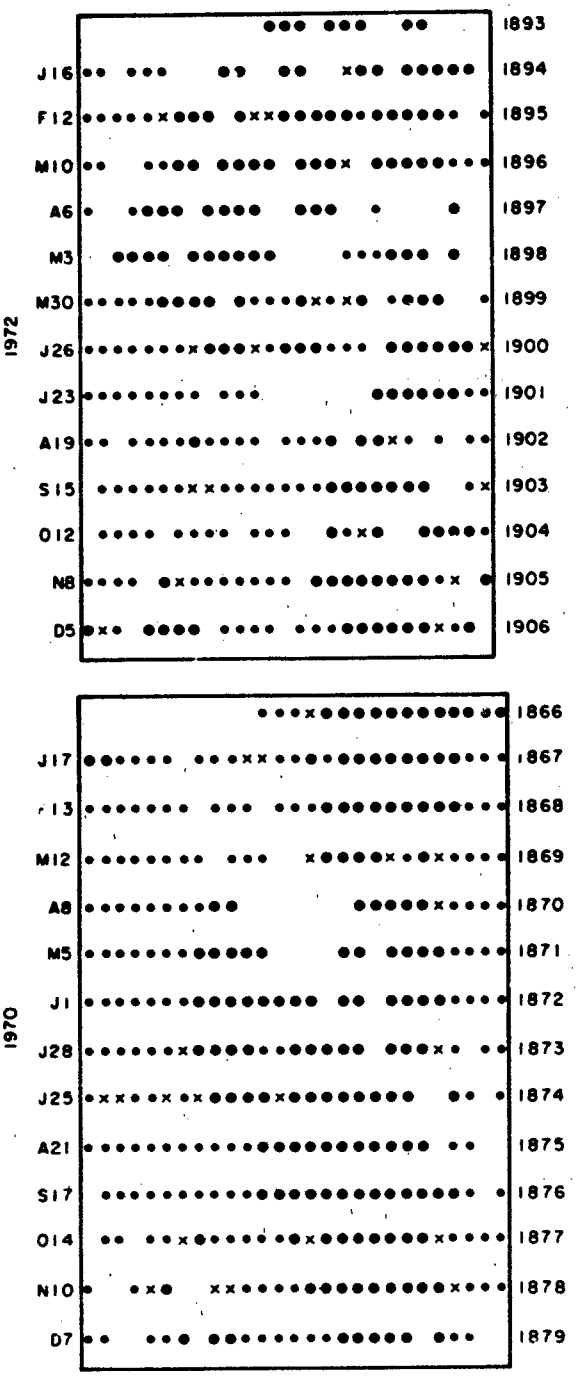


Figure 3-4 Two years of sector structure polarity. A small dot indicates a day with polarity directed away from the sun; a big dot, towards the sun; a cross, an ambiguous day; and a blank, a data gap. Each line represents one solar rotation or 27 days. The date of the first day of the rotation is given on the left-hand scale and the number of the rotation on the right-hand scale. Note the strong tendency for sector structure to repeat on subsequent rotations [adapted from Wilcox et al., 1975].

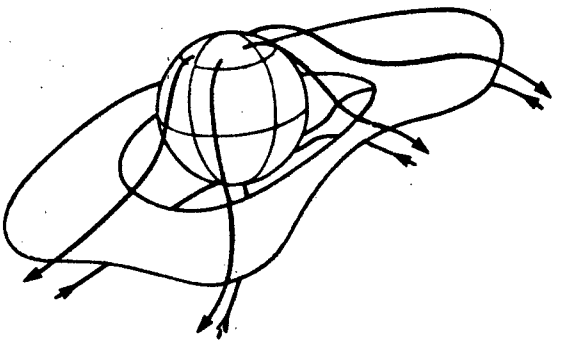


Figure 3-5. Sketch of the current sheet responsible for the sector structure. The sheet is warped and tilted relative to the plane of solar rotation [Smith, 1979].

hours. The persistence of a given level of |B| can be found from the autocorrelation function which falls to 0.5 at about a 15-hour lag [Burlaga and King, 1979].

An important magnetic field parameter is the component of the field perpendicular to the X-Y plane. This component, B_z, plays a central role in the coupling of the wind and the earth's magnetic field such that when B_z is southward and

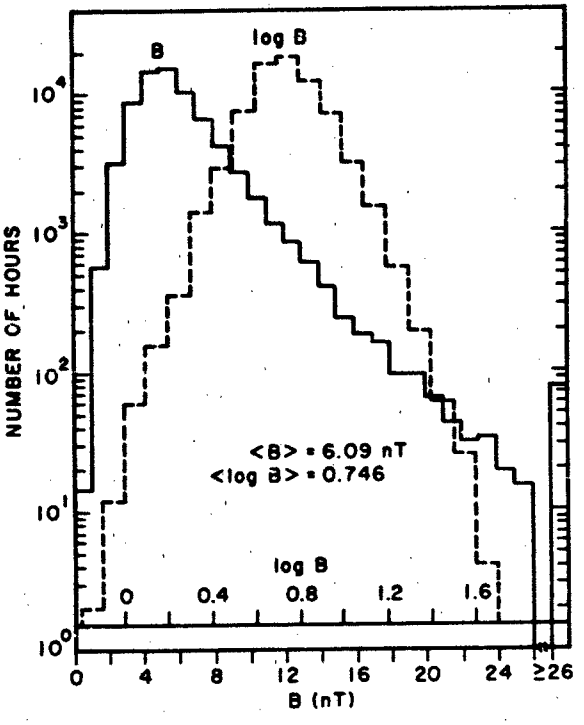
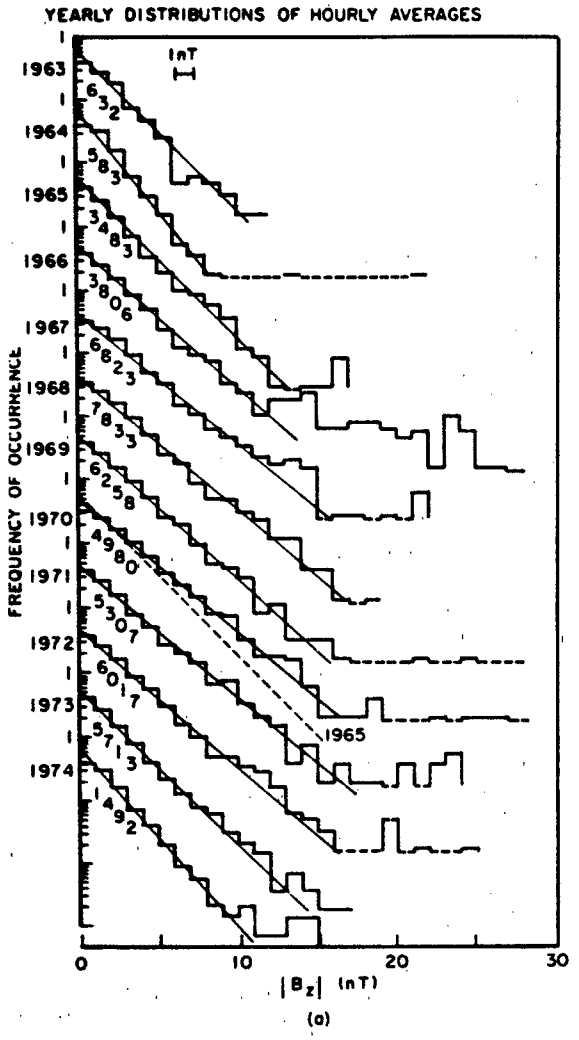


Figure 3-6. Distribution of 71431 hourly averaged values of the interplanetary magnetic field magnitude |B|, and of log |B|. These data were collected between 1963 and 1977. The average value of B is 6.09 nT but the average of log B corresponds to a value of 5.6 nT. The most common value of B is about 5 nT [Burlaga and King, 1979].



large, geomagnetic disturbances occur. Figure 3-7a shows distributions of GSE B_z observed for each year from 1963 to 1974 [Siscoe et al., 1978]. The figure gives the number of times $|B_z|$ fell within each 1 nT bin. The number of values used in each histogram is also given. The distributions are generally exponential, but with a high intensity tail. The exponent changes in a systematic way with the solar cycle. These data are collected in Figure 3-7b which gives the probability that an hour chosen at random from the set of over 50 000 hours of observation will fall in a particular 1 nT bin. Of course, a figure of this type made using different years of observation would differ somewhat from that shown. The finding that the most probable value of B_z is less than 1 nT is in general agreement with the notion that a steady spherically expanding solar wind would have no B_z component when measured in the sun's equatorial plane. The B_z component is primarily due to disturbances such as waves or interacting parcels of the interplanetary medium.

3.2.3 Protons and Electrons

The solar wind plasma consists of protons, doubly charged helium ions, a small number of other positively charged particles, and enough electrons so that the plasma is electrically neutral. The Debye length is about 10 m. Since commonly 95% of the positively charged particles are protons, it is the protons that are usually referred to when statements are made about the solar wind bulk properties such as velocity, density, and temperature. In this section the large scale general properties of the wind will be described for protons and electrons. Alpha particles and heavy ions will be dealt with in Section 3.2.4.

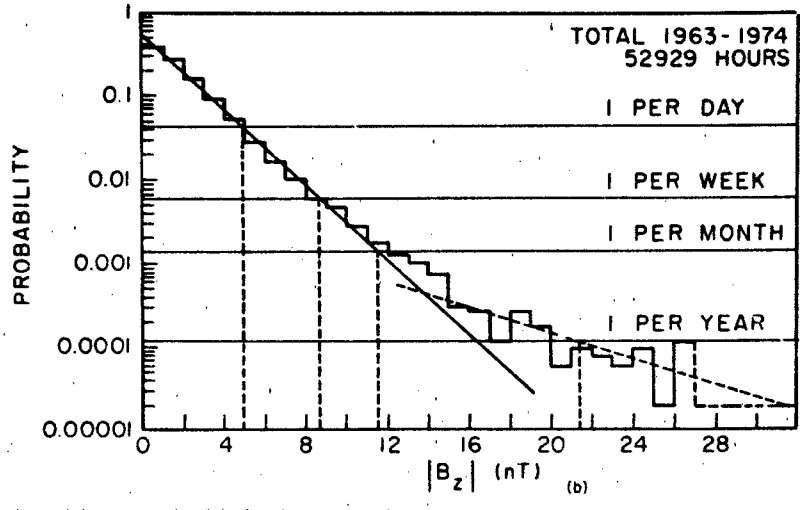
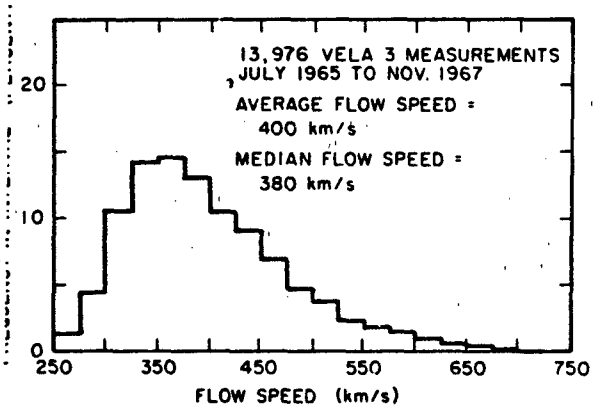


Figure 3-7. (a) The distributions of the magnitude of the hourly average value of the component of the interplanetary magnetic field perpendicular to the ecliptic plane, $|B_z|$. Data for each year from 1963 to 1974 are plotted. The data coverage varied from year to year. Note the changes in slope and the high intensity tail. (b) Histogram of the occurrence frequency of $|B_z|$ based on the 52 929 values used in the above distributions. The ordinate gives the probability that an entry selected at random will fall in the particular 1 γ bin. Levels for selected average recurrence periods are marked. For example, the average time between observations of 13γ < B_z < 12γ is about one month [Siscoe et al., 1978].

HAPTER 3

Solar wind properties do not vary randomly on a time cale of hours. Instead the wind is organized mainly into eriods of slow plasma and fast plasma, with the slow and ist regions interacting. The structure of these streams will e discussed in Section 3.3.2. Here the interest will focus n typical values of physical parameters.

Early plasma probes detected protons and other posively charged particles but did not permit detailed descripons of particle distributions to be made. However, the first iree moments of the distribution could be determined, that , the proton velocity, density, and (less exactly) the temerature. A typical distribution of speeds is shown in Figure -8. The direction of flow is almost always from the sun



igure 3-8. A frequency distribution of solar wind speed [Hundhausen, 1972]. Both the average speed and the shape of the distribution differ from year to year.

within a few degrees). The major causes of non-radial flow re aberration and interaction of fast and slow solar wind tructures. The amount of deviation from radial flow due to he angular velocity of the source of the solar wind at the un is still a matter of debate. Although there is considerable ariation in the velocity distributions for different time peiods, they all have nearly the same range of values. In igure 3-8 there are few measurements with velocities lower han 275 km/s and in data sets from several other years there vere none. In the sample shown, values higher than 650 m/s were rare but in other years the upper frequency bin vith non-negligible sample number was 775 km/s. Velocties of over 1000 km/s occasionally have been reported fter extremely energetic flares occurred on the sun. Auocorrelation among speed measurements is high for lags of ess than 40 h, reflecting the stream structure (Section 3.3.2). Table 3-1 shows average values and ranges for 3-hour avrage speeds for about 3.5 years of data [Feldman et al., 1977]. The average velocity given there is 468 km/s, somewhat higher than that shown in Figure 3-8.

The density of the proton component is also ordered by he fast and slow velocity streams with an average of 8.7 particles/cm^3. Most of the extreme density values are due to rarefaction and compression arising through stream interactions or solar flare disturbances discussed in Section 3.3.3. The proton temperature T_p, which increases with the velocity, varies by a factor of twenty and has an average of 1.2×10^5 in this data set.

The proton distribution function varies widely and is usually not well represented by a single Maxwellian. Figures 3-9a and b show examples of two-dimensional proton velocity distributions. Figure 3-9a shows a simple distribution with a single peak, corresponding to a single velocity and density. Figure 3-9b shows a distribution with a strong second peak corresponding to a plasma having two beams with different velocities and densities. The most usual distribution is between these two and would show a single major peak with evidence for a high velocity tail. Note that in both main peaks the spread in contours is greater in one direction than at right angles, that is, the temperatures parallel and perpendicular to some direction (the direction of the magnetic field) are not equal and are often given separately. The temperature T_p is $1/3\ (T_\parallel + 2T_\perp)$ where $\parallel$ and $\perp$ denote components parallel and perpendicular to the magnetic field. The average value of $T_\parallel/T_\perp$ is 1.5 with a range of 0.7 to 2.8. Distributions that have two velocity peaks can be described as two streams of protons moving relative to one another. The higher density streams are usually the lower velocity ones.

Solar wind electrons are more difficult to observe and their properties have been described less fully than those of the protons. In contrast to protons, electrons are subsonic at earth. Observed distributions of electrons are shown in Figure 3-10 [Feldman et al., 1975]. Note that because of the subsonic velocity, the spreads in particle speeds are much larger than the bulk speed. Very low velocity electrons, which would allow an actual measurement of the bulk speed to be made, cannot be measured at all because of observational difficulties. The figure shows four cuts through the distribution. Figure 3-10a shows particle fluxes in the direction parallel and anti-parallel to the electron heat flux. Since the electron heat flux is along the magnetic field direction these two cuts are also in directions parallel and antiparallel to the magnetic field. Note that fluxes for each speed are not the same in two directions, indicating that a substantial heat flux is carried by the electrons. In fact, it is the electrons that regulate the heat flux in the wind as a whole. The curves in the upper panel represent fits to the points, the fits to the lower speeds being Maxwellians. Note that although the two Maxwellians are excellent fits at low speeds, they fail decisively at high speeds. In much of the data reported in the literature, the higher speed particles are fit with a second set of two Maxwellians or a somewhat more complex function as shown here. In Figure 3-10b the data shown are from a plane perpendicular to the heat flux. Two data sets are shown: one for positive and one for negative directions along a line in the plane. The fluxes are the

able 3-1. Average solar wind parameters

Parameter	Mean	σ	5–95% Range Limit
speed (km/s)	468	116	320 to 710
density (proton/cm^3)	8.7	6.6	3.2 to 20
M_A	10.7	4.8	4.4 to 20
M_S	7.7	1.5	5.6 to 10
$\|B\|$ (nT)	6.6	2.9	2.2 to 9.9
T_p (K)	1.2×10^5	9×10^4	1×10^4 to 3×10^5
T_e (K)	1.4×10^5	4×10^4	9×10^4 to 2×10^5
T_e/T_p	1.9	1.6	0.37 to 5.0
n_α/n_p	0.047	0.019	0.017 to 0.078
T_α (K)	5.8×10^5	5×10^5	6×10^4 to 15.5×10^5
T_α/T_p	4.9	1.8	2.3 to 7.5
β_p	0.7	0.9	0.11 to 1.7
β_e	1.3	1.9	0.3 to 3.1
$\beta_p + \beta_e$	2.1	2.4	0.5 to 4.6
C_A(km/s)	50	24	30 to 100
C_S(km/s)	63	15	41 to 91

nformation from Feldman et al. [1977])

ame in these two directions showing that the electron distribution function is symmetric about the magnetic field irection. Again the data are fit with Maxwellians at low peeds and more complex functions are fit to the residual luxes at high speeds. The low speed portions of electron listribution functions are collision dominated, whereas the ery high speed portions are collisionless. Observed distribution functions are consistent with a theoretical model in vhich the entire distribution at all velocities arises from the ffects of Coulomb collisions [Scudder and Olbert, 1979].

Although there is not yet general agreement concerning he best method for describing and interpreting electron bservations, the data presentation that is used here is that of Feldman et al. [1977], in which the measured flux distribution is broken up into low and high speed parts as shown n Figure 3-10. The electron and proton densities and bulk velocities are set equal, consistent with the observations. The electron distribution function is described as a superposition of two distributions, one a cold component f_c and he other a hot component f_h, so that $f = f_c + f_h$. The cold component electrons below 50 eV that can be well fit with a bimaxwellian function are also known as the core electrons. The hot component or halo electron component is defined as the remainder of the measured distribution function after the cold component is subtracted. In this scheme from 3 to 10% of the electrons belong to the hot component. The energy at which f_h begins to rise above f_c is about 60 eV. The hot and cold components generally move relative to one another along the magnetic field direction so that no net current flows in the frame of reference moving with the protons. The total bulk velocity is observed to be within about 50 km/s of the proton velocity. The total heat flux in the solar wind is primarily carried by the hot electron component. If both halo and core components are fit with bimaxwellian distributions then the average core electron temperature is 1.2×10^5 K, whereas the average halo electron temperature is 6.9×10^5 K. The value of 1.4×10^5 K given in Table 3-1 is the temperature derived from the fit of the distribution function as a whole. As can be seen from Table 3-1, the electron temperature does not vary as widely

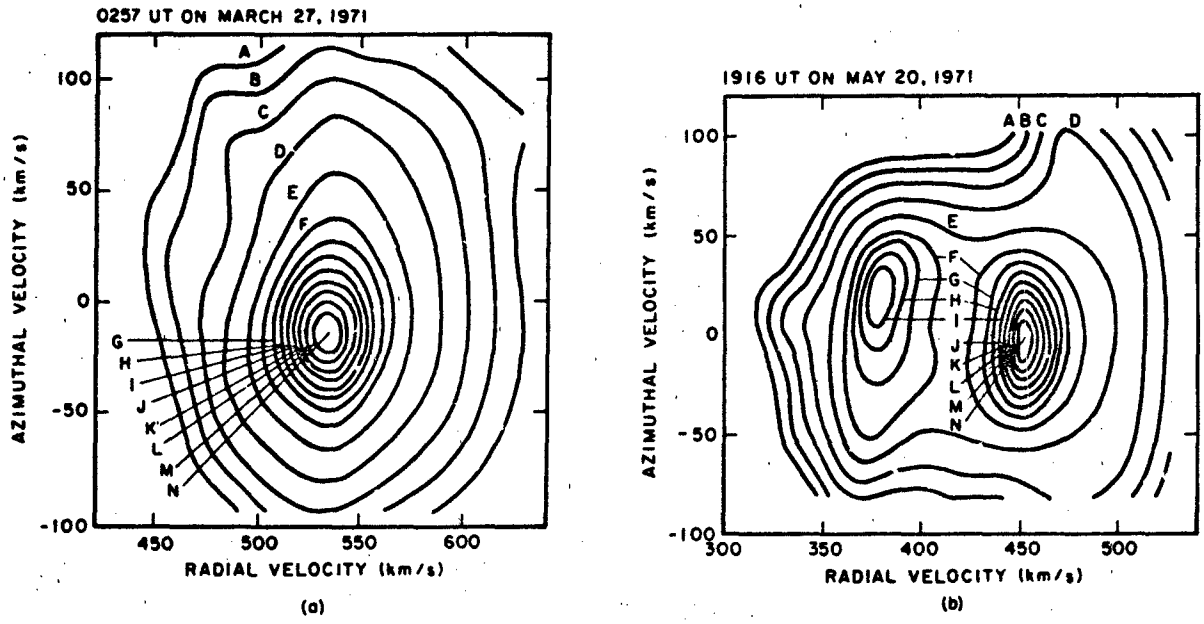


Figure 3-9. Proton velocity distributions. Two examples (a,b) of two-dimensional proton velocity distributions contoured for equal number density. Contours A, B, C and D correspond to 0.001, 0.0032, 0.01 and 0.032 and contours E through N to 0.1 to 1.0 in steps of 0.1 of the peak contribution to the proton number density [Feldman et al., 1974].

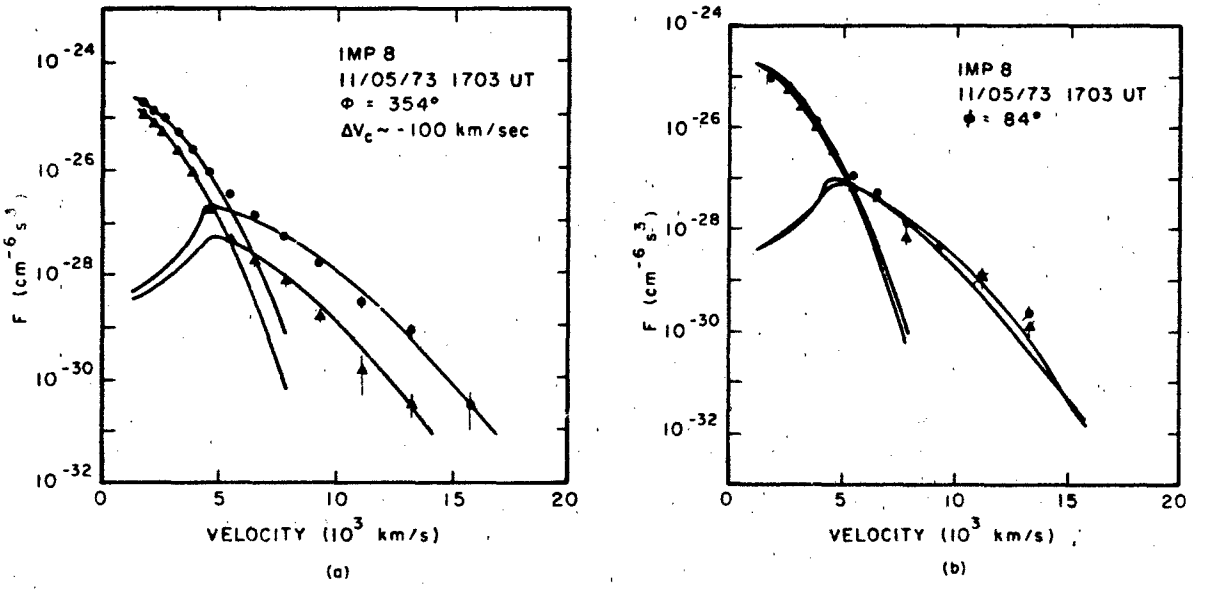


Figure 3-10. Electron flux as a function of speed [Feldman et al., 1975]. (a) shows the flux parallel and anti-parallel to the direction of the electron heat flux. (b) shows the fluxes in a perpendicular plane.

as the proton temperature. The electron thermal pressure is usually, but not always, greater than the magnetic field pressure ($\beta = 8\pi nkT/B^2$, $\beta_e > 1$). Thus, although the proton thermal pressure is usually less than the magnetic field pressure ($\beta_p < 1$), the average β calculated using the total thermal pressure is about 2 (Table 3-1).

3.2.4 Other Ions

Since helium is the second most abundant solar constituent, it was expected to be present in the solar wind in quantities that could be detected by the early solar wind probes. This expectation was fulfilled with the flight of Mariner II in 1962 [Neugebauer and Snyder, 1966]. However, the measured values of the abundance, velocity, and temperature of the helium component of the wind were so different from what would be expected on the basis of simple theory that the behavior of helium is still not well understood. For a recent review of helium observations see Neugebauer [1981]. Observationally, the helium properties are ordered by large scale structures such as sectors, velocity streams, and flare disturbances. These properties will be discussed further in Section 3.3. The simplest helium parameters to measure are relative abundance, velocity, and temperature. The major unexpected results are that in the vicinity of the earth, relative abundances of helium by number, n_α/n_p, vary from less than 10^{-3} to more than 0.3, helium velocities tend to be a few km/s larger than hydrogen velocities, and the temperature of helium is usually about 3 or 4 times that of hydrogen.

Helium in the solar wind is almost always completely doubly ionized, as would be expected because of the high coronal electron temperatures. However, a few events have been reported in which He^+/He^{++} reached levels as high as 0.1 [Schwenn et al., 1980] and 0.3 [Gosling et al., 1980]. These events have been interpreted as the inclusion into the wind of material from relatively cool chromospheric prominences.

The distribution of relative abundance values of He^{++} by number measured by the Vela 3 satellites is shown in Figure 3-11 [Robbins et al., 1970]. For this data set the values ranged from 8.1×10^{-4} to 4.2×10^{-1}. The yearly average abundance between 1962 and 1976 varied from 3% to something over 5%. The helium abundances are not distributed randomly in the plasma but instead show systematic behavior that has not yet been satisfactorily explained theoretically. Most values of $n\alpha/np \geq 0.15$ in Figure 3-11 were observed in solar wind reliably associated with major solar flares [Hirshberg et al., 1972]. Although most helium abundance enhancements are associated with flares, regions of solar wind containing between 9% and 15% helium plasma are sometimes seen in low speed wind with no association with major flares or interplanetary shocks (Section 3.3.3). In general, the data suggest that unusually high solar wind helium abundances may be associated with impulsive injections into the wind of hot or cool coronal plasma as a result of solar activity but there does not appear to be a one to one correlation. Helium abundance also shows a trend toward larger values in higher velocity winds but this trend is time dependent and may change with the solar cycle. The average abundances at velocities of 300 km/s vary between 3% and 4%, and at 500 km/s average abundances between 5% and 9% have been reported [Neugebauer, 1981].

The helium and hydrogen components of the solar wind have almost the same speed at 1 AU. However, differences of up to a few tens of kilometers are common, and the helium is usually moving faster than the hydrogen. The average reported value of the ratio of the helium to hydrogen speeds $\langle v_\alpha/v_p \rangle$ ranges from 1.001 to 1.035 and is always positive. The direction of helium motion relative to hydrogen is generally along the direction of the magnetic field. At 1 AU the velocity difference $v_\alpha - v_p$ is in general proportional to, but less than, the Alfvén wave speed. Also,

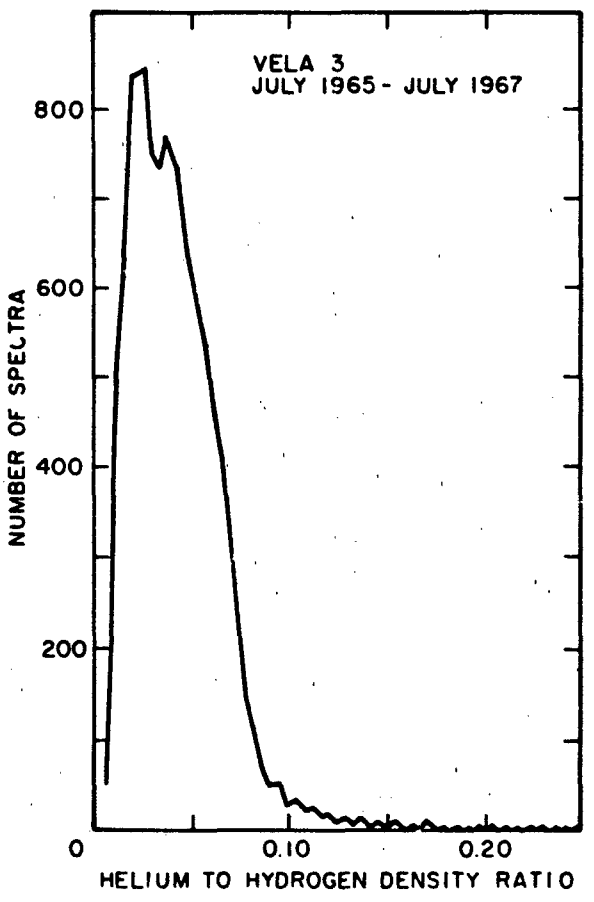


Figure 3-11. The distribution of solar wind helium to hydrogen number density ratios [Robbins et al., 1970]. The helium enhancement events reflected in the high abundance portion are associated with major solar flares.

$v_\alpha - v_p$ is strongly dependent on the ratio of the solar wind expansion time to the Coulomb collision slowing down time. The expansion time scale is defined as the time required for a solar wind ion to move through one density scale height. It has been suggested that the helium is accelerated to higher velocities than hydrogen by waves in the interplanetary medium but as yet there is no direct observational confirmation.

The temperature of helium is almost always higher than that of hydrogen, and it appears that there is a tendency for the thermal velocities of the two components to be equal rather than for the temperatures to be equal. The most probable value of T_α/T_p is between 2 and 4 but values as high as 10 are occasionally observed. Average values of T_α/T_p calculated from different data sets have varied from about 3 to 5. It is not known if this difference is due to actual changes in the solar wind or to inaccuracies in the measurements. There is a positive correlation between T_α/T_p and $v_\alpha - v_p$. The behavior of T_α is not understood theoretically but it has been suggested that the waves that can preferentially accelerate helium may also heat the ions by an amount roughly proportional to their masses.

In addition to ^4He ions, solar wind ^3He ions have been detected. The $^4He^{++}/^3He^{++}$ ratio has been measured both from foils left on the lunar surface and directly using an ion composition experiment on ISEE 3. The results of the two techniques are in agreement. The ISEE experiment found on average $^4He^{++}/^3He^{++}$ of $2{\cdot}1 \pm 0.2 \times 10^3$ using more than 4000 observations; however, the distribution was very broad. Low ratios, which lasted a few hours, had a tendency to characterize periods of low $^4He^{++}$ flux [Ogilvie et al., 1980].

Some ions heavier than helium have been observed in the interplanetary medium. The observations are difficult to make because of the low ionic abundances. However, neon has been collected on foils on the moon, and plasma probes have been able to resolve oxygen, iron, and silicon at times when the plasma temperatures are low. The relative abundances are within a factor of ten of coronal abundances, and the speeds tend to be the same as the speeds of the helium observed at the same time. There does not appear to be any further acceleration of heavy ions. The temperatures however, apparently increase further with higher ion mass and measurements of oxygen temperatures suggest that the thermal speeds of all constituents tend to be equal in the solar wind [Ogilvie et al., 1980].

The ionization state of heavy ions provides important information on the electron temperature at their source in the corona. The ionization state of the atom in the collision dominated corona is determined by a balance between ionization by electron collisions and recombination due to radiative and dielectronic processes. As the solar wind moves away from the sun, collisions become more rare, and beyond some heliocentric distance the ionization state of the particle remains unchanged. This process, called "freezing in," permits the particles to be used as tracers of their sources. (Note that the term "freezing in" has already been used in another context in Section 3.1.1. The two uses are so different they almost never cause confusion.) The heliocentric distance at which freezing in occurs depends somewhat on the type of particle involved because of the details of the interactions. In slow, low temperature regions of the wind the abundances of O^{6+} to O^{7+} ions in the available data set indicate a freezing in temperature of 2.1×10^6 K and the abundances of Fe^{7+} through Fe^{13+} yield 1.5×10^6 K. Other plasma, associated with solar wind accelerated by flares showed oxygen and iron freezing-in temperatures of 3×10^6 K [Bame et al., 1979], and iron charge states have indicated temperatures as high as 1.6×10^7 K.

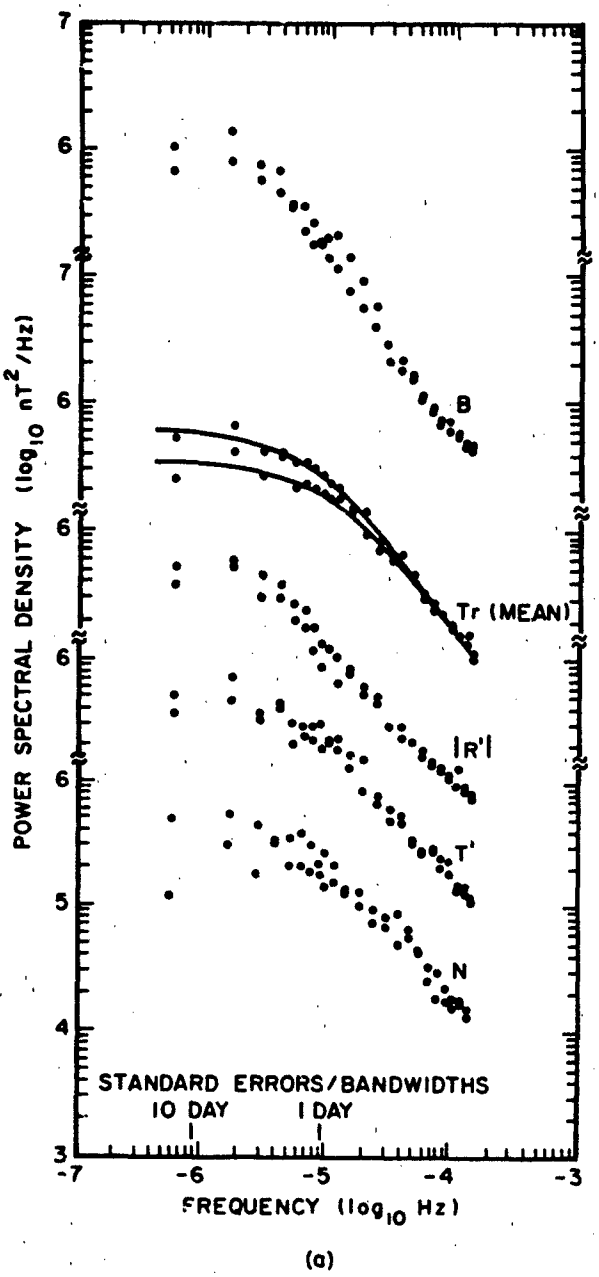
3.2.5 Waves and Turbulence

Observations of the solar wind show that there are variations in almost all quantities on all time scales. For example, as discussed in Section 3.2.2 the interplanetary magnetic field of a uniform undisturbed wind should not have any Z component; however, hourly average Z components are about 1/3 the intensity of the field magnitude. This Z component is due to the presence of waves and discontinuities in the plasma.

Considering the solar wind as a turbulent medium, power spectra have been calculated and compared with the theoretical predictions based on magnetohydrodynamic turbulence theory. Spectra of the magnetic field, particle density, and particle velocity variations have been given in the literature.

Magnetic field data lend themselves to these studies because of the good time resolution and accuracy of the measurements. Examples of spectra in the 10^{-7} to 10^{-1} Hz range are given in Figures 3-12a and 3-12b [Hedgecock, 1975]. If the field could be described as that ideally present in a spherically expanding constant velocity solar wind without fluctuations, there would be no components in the two transverse directions. The top curve in Figure 3-12a gives the power in the variations of magnitude. The power levels in the magnitude are much lower than the levels in the components. The spectra of the two transverse components (T and N) are very similar and have been averaged to produce the transverse spectrum Tr. A composite spectrum of the transverse power is shown in Figure 3-12b. Note the different spectral indices used to fit the data in the different frequency ranges. The magnetic field power spectrum for periods greater than a day tends to flatten. The power spectrum slope varies from day to day, or for lower frequencies from sample to sample. In the 10^{-5} to 10^{-3} Hz range both the slopes and the absolute value of the power variations are consistent with a factor of 10 variation. [Hedgecock, 1975]. When the spectrum was fitted with a line given by F^{-N}, the average spectral exponent N for 17 periods in 1972 was 1.27.

Examples of power spectra of solar wind fluctuations in positive ion flux are shown in Figure 3-13a [Neugebauer,



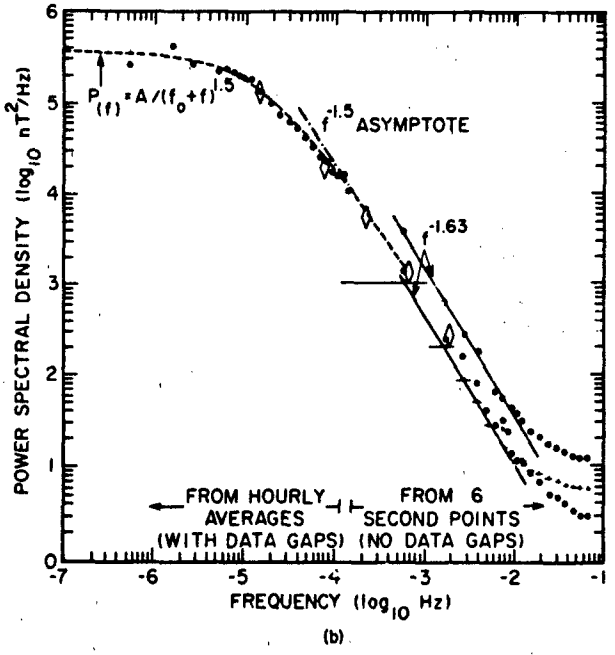


Figure 3-12. (a) Power spectral density (PSD) of the interplanetary magnetic field. The R direction is along the average spiral direction of the field, N is in the solar meridian plane and T completes the system. B is the magnitude of the field. The components are as marked. T_R is the average of the transverse (N, T) spectra [adapted from Hedgecock, 1975]. (b) A composite mean transverse power spectrum for 1972 [Hedgecock, 1975].

1975]. Curves A and D are for the noisiest and quietest 100 spectra of the 1728 spectra in the data set used in this study. In the 10 second region these two power densities differ by more than a factor of 100. Curves B and C are two different types of averages of the entire data set. In curve B the values of the power density are simply averaged. In curve C the logarithms of the power spectral density are averaged to avoid giving too much weight to the shape of the spectra from the most turbulent intervals. The error bars represent 90% uncertainty levels. The enhancement at high frequencies is probably due to fluctuations at the proton thermal gyroradius. A long period velocity spectrum is shown in Figure 3-13b from Coleman [1968].

Frequently, fluctuations in the solar wind can be recognized as distinct types of waves and/or discontinuities. Discontinuities will be discussed in Section 3.3. Although a kinetic theory treatment is required for a proper description of the behavior of waves, the magnetohydrodynamic theory is useful for classifying the waves [Barnes, 1979]. There are three types of MHD waves, the transverse Alfvén (or intermediate) wave, and fast and slow magnetoacoustic waves. The transverse Alfvén wave is non-compressive with constant magnitude of the magnetic field, but not constant direction. The fluctuating magnetic field and velocity are related by

$$\mathbf{b} = \pm (4\pi\rho)^{1/2}\mathbf{v} \qquad (3.17)$$

where b is the magnetic perturbation, v is the velocity perturbation, and ρ is the constant density. The velocity of an Alfvén wave is given by $C_A = B/(4\pi\rho)^{1/2}$. Typical values of C_A are given in Table 3-2.

The two magnetoacoustic modes are compressive and their magnetic fluctuations are nearly linearly polarized. The velocities of the fast and slow waves are given by

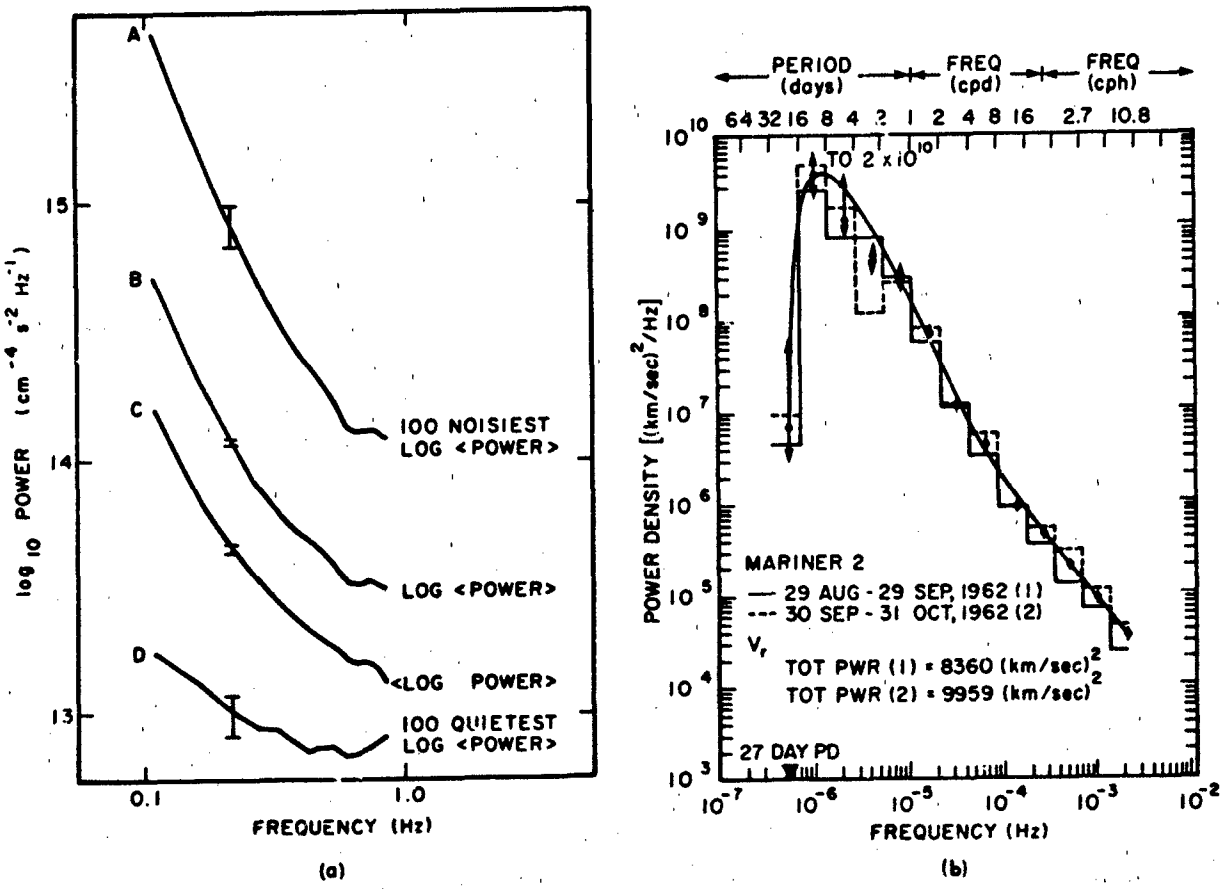


Figure 3-13. (a) Four different types of averages of the power spectral density of variations in the total charge flux of positive ions [Neugebauer, 1975]. (b) Power spectra of the solar wind velocity in 1962 [Coleman, 1968]. (Reprinted with permission from the American Astronomical Society © 1968.)

$$v_m = \frac{1}{2}\left[(C_A^2 + C_S^2) \pm \left(C_A^4 + C_S^4 - 2C_A^2C_S^2\cos 2\theta\right)^{1/2}\right] \quad (3.18)$$

where C_S is the speed of sound, θ is the direction of the wave normal relative to the field, and the plus is taken for the fast mode and minus for the slow mode. Typical values of C_S are shown in Table 3-1 where $C_S = [k(T_e + 3T_p)/m]^{1/2}$. This expression assumes the electrons respond isothermally and the protons adiabatically to one dimensional sound waves as expected from the electrons role in heat conduction. In the limit of small amplitude waves, fluctuations can be thought of as a superposition of these three types of MHD waves. However, the waves in space are not small amplitude so that their behavior is more complex than described by small amplitude MHD theory. The magnetoacoustic modes will be damped and so will not be expected to be observed far from their place of origin. The Alfvén wave, however, is not damped.

By far the most commonly observed wave in the solar wind is the Alfvén wave. An example is shown in Figure 3-14. In this figure R is the radial direction and N and T are transverse directions. The velocity and field component variations are approximately related according to the expression above. The magnetic field magnitude and the plasma density show very little variation. Interplanetary fluctuations are dominated by large amplitude aperiodic Alfvén waves more than 50% of the time [Belcher and Davis, 1971]. The clearest examples occur in high velocity streams (Section 3.3.2) and their leading edges. These waves are propagating outward in the rest frame which implies that they are produced in or near the sun below the critical point of the flow r_c. They have not yet been actually observed in the vicinity of the sun and this is a very difficult observation to make without *in situ* detectors. In the vicinity of the earth, changes in the magnitude of B are small with $\Delta|B|/|B| \sim 0.06$ whereas the variation of the components is relatively large with $\Delta B_i/|B| \sim 0.4$ where i indicates a component [Burlaga and Turner, 1976]. The direction of minimum variance of the field is

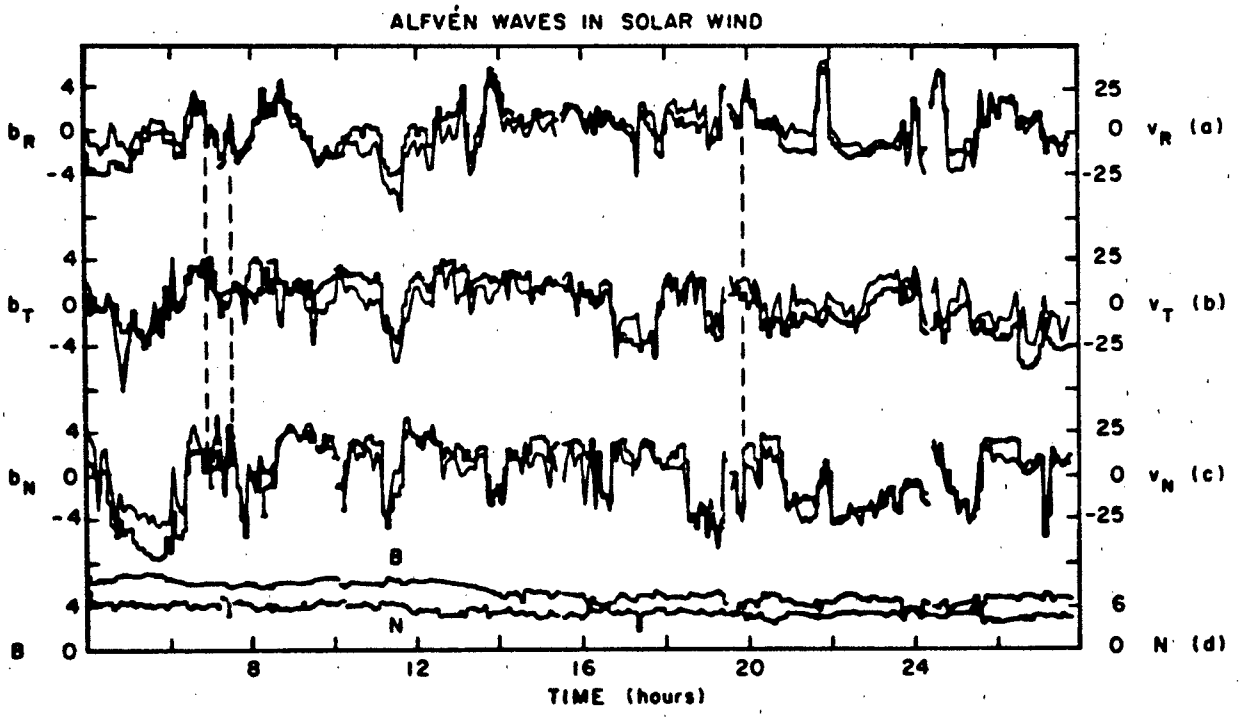


Figure 3-14. Alfvén waves in the solar wind. Curves a, b & c show that the strong variations of the magnetic field components occur simultaneously with like variations of the velocity. Curve d shows that the field magnitude and density were both very steady [Belcher and Davis, 1971]

along the main field. The observed frequencies of the waves are from the order of minutes to hours. In Figure 3-12a the low power levels in the spectrum of |B| compared to the spectrum of the variations of components is due to the dominance of Alfvén waves.

Observations also show fluctuations in the solar wind that are non-Alfvénic. For example, the leading edges of high velocity streams contain large amplitude Alfvén waves, a non-Alfvénic component, and also perhaps an Alfvénic component propagating toward the sun [Belcher and Davis, 1971]. These latter two types of waves are probably of local origin.

There are very few unambiguous identifications of interplanetary magnetoacoustic waves. One case reported by Burlaga [1968] is probably a fast magnetoacoustic wave propagating across the magnetic field [Barnes, 1979].

3.3. LARGE SCALE STRUCTURES AND DISCONTINUITIES

3.3.1 Discontinuities

Before discussing the types of disturbances that appear at 1 AU in the solar wind, it is helpful to describe the classifications of discontinuities more completely. For this purpose the hydromagnetic approximation is used, that is, the plasma flow will be approximated by the flow of a dissipationless perfect gas. The magnetic field will be assumed to follow the flow. The approximation will be sufficiently accurate to organize the observations of interplanetary disturbances and discontinuities discussed here. The quantities that must be conserved across a hydromagnetic discontinuity are mass, momentum, energy, the component of the magnetic field perpendicular to the discontinuity, and the component of the electric field tangential to the discontinuity. The jump conditions have been given by Spreiter and Alksne [1969].

For a tangential discontinuity:

$$v_n = H_n = 0 \qquad (3.19)$$
$$[v_t] \neq 0,\ [H_t] \neq 0,\ [\rho] \neq 0,\ [P + H^2/8\pi] = 0,$$

for a contact discontinuity:

$$v_n = 0,\ H_n \neq 0,\ [v] = [B] = [P] = 0,\ [\rho] \neq 0, \qquad (3.20)$$

for a rotational discontinuity:

$$v_n = \pm H_n/(4\pi\rho)^{1/2}, \qquad [v_t] = [H_t]/(4\pi\rho)^{1/2} \qquad (3.21)$$
$$[\rho] = [P] = [v_n] = [v^2] = [H^2] = [H_n] = 0,$$

and fast and slow shock waves satisfy

$$v_n \neq 0,\ [\rho] > 0,\ [P] > 0,\ [H_n] = 0$$

$$(\rho v_n)_{\text{fast}} \geq (\rho v_n)_{\text{rot}} \geq (\rho v_n)_{\text{slow}} \qquad (3.22)$$

H_t and H^2 $\begin{pmatrix}\text{increase}\\ \text{decrease}\end{pmatrix}$ through $\begin{pmatrix}\text{fast}\\ \text{slow}\end{pmatrix}$ shock waves.

Here v, ρ, P, and H are the velocity, density, pressure, and magnetic field intensity respectively in a frame of reference at rest with the discontinuity. Subscripts t and n refer to components tangent to and normal to the discontinuity surface and [] denotes the difference in the enclosed quantity on the two sides of the discontinuity. The rotational discontinuity is the steepening of an Alfvén wave. Fast and slow shocks are the steepening of fast and slow magnetoaccoustic waves. Contact and tangential discontinuities do not propagate in a frame of reference at rest with the undisturbed plasma. The three other discontinuities do propagate in that frame. Contact discontinuities are not expected to be observed in the solar wind because they are difficult to form and would decay away quickly. Tangential and rotational discontinuities and fast and slow shocks have been observed. Slow shocks are very rare, and if not specifically stated otherwise, the word "shock" will be used here to refer to a fast shock.

At earth, three types of discontinuities are common, fast shocks and rotational and tangential discontinuities. Shocks are formed in the solar wind by the steeping of the interaction region between fast and slow solar wind parcels. They are commonly associated with impulsive high speed solar wind from flare events. The interaction region between long lived streams of slow and fast solar wind rarely steepens into shocks by the time the solar wind regions arrive at earth. Rotational and tangential discontinuities are commonly observed at 1 AU and are apparently formed close to the sun but the mechanism of formation is not known [Tsurutani and Smith, 1979]. Since the magnetic fields are measured in space with a smaller sampling interval than particles, most studies of discontinuities have used observational criteria based on changes in the magnetic fields. The frequency of occurrence of discontinuities will depend on the selection criterion but with carefully designed criteria it is found that tangential and rotational discontinuities occur at a rate of one or two per hour. The mean time interval between arrivals of discontinuities can typically be described by a Poisson distribution over the interval of from perhaps 10 minutes to an hour, but the rate of occurrence varies strongly on a time scale of several days. Even on a time scale of several solar rotations, the average rate varies by a factor of more than 2 [Tsurutani and Smith, 1979]. Two important parameters that characterize the discontinuities are the angle through which the magnetic field rotates and the percentage change in the field magnitude. The distributions of these quantities are shown in Figure 3-15, adapted from Tsurutani and Smith [1979]. Small angle and magnitude changes occur most frequently. Discontinuities are usually quite abrupt. A study of Mariner 4 magnetic field data found that 50% of the discontinuities analyzed had thicknesses less than 3500 km and 10% less than 500 km [Siscoe et al., 1968].

A class of discontinuities called directional discontinuities (DD) has been defined as events that are either rotational or tangential discontinuities and show abrupt changes in magnetic field of at least 30° in consecutive 30-second magnetic field averages. It is not possible to unambiguously distinguish between tangential discontinuities and rotational discontinuities among DDs, but is was found that in low velocity solar wind tangential discontinuities are most common whereas DDs consistent with outward propagating rotational discontinuities become the predominant mode in high velocity winds [Solodyna et al., 1977].

Shock discontinuities are also observed at earth and are usually associated with solar events such as flares or perhaps errupting prominences. The fast solar wind from such an event accelerates the ambient wind in front of it. If the difference in velocity between the fast wind and the ambient wind is larger than the speed of a fast magnetohydrodynamic wave, a shock will form in the ambient wind at a position well in front of the boundary between the ambient wind and the flare plasma. This shock is known as the forward shock. As the fast plasma accelerates the ambient plasma, conservation of energy and momentum require the fast plasma itself to slow down. If the slowing down is severe enough another shock called the reverse shock will form within the fast plasma on the sun side of the boundary. When a detector passes through a forward shock, there will be a sudden increase in magnetic field, density, and temperature and a decrease in velocity. Passage through a reverse shock will appear as a decrease in magnetic field, density, and temperature and an increase in velocity. Both types of shocks have been observed as shown in Figure 3-16 [Burlaga and King, 1979]. Forward and reverse shocks are found in flare associated events and reverse shocks occasionally are formed during the interaction of long lived streams of slow and fast solar wind. However, shocks are not as common as DDs and occur once every few days at sunspot maximum and less frequently at low sunspot number.

The jump conditions at shocks (Rankine Hugoniot conditions) allow the strength of the shock and the direction of the normal to the shock to be determined if the plasma and field properties are known accurately enough. For strong shocks the direction of the normal to the shock can be found from

$$\mathbf{n} = \pm (\mathbf{B}_1 \times \mathbf{B}_2) \times (\mathbf{B}_2 - \mathbf{B}_1), \qquad (3.23)$$

where $\mathbf{B}_1$ and $\mathbf{B}_2$ are the observed fields before and after shock passage and the sign depends on whether the shock is forward or reverse. A wide range of shock normals has

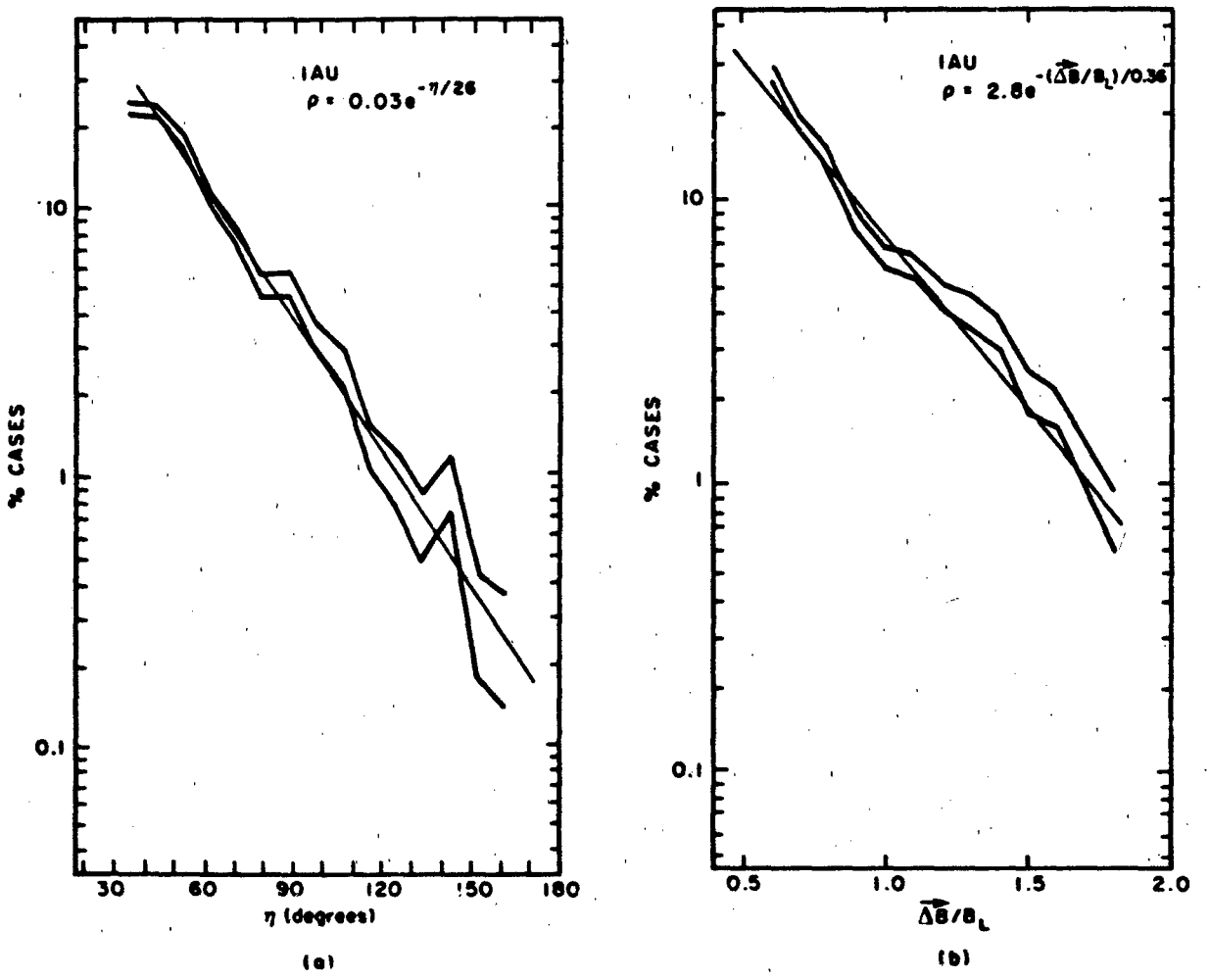


Figure 3-15 Normalized distributions of (a) the field rotation angle η across discontinuities and of (b) the relative field change $|\Delta B|/B_L$, where B_L is the larger of the field magnitudes on either side of the discontinuity. The two plots in each panel are the normalized distribution plus or minus one standard deviation [Tsurutani and Smith, 1979].

been observed, and although most normals lie within 45° or so of the earth-sun line, almost any orientation can occur. In addition, the flares that cause the strongest shocks tend to occur within 45° of the central meridian of the sun but sometimes flares occurring near the solar limbs cause shocks to appear at earth's orbit. The Mach numbers of most of the few shocks for which determinations have been made tend to be rather low, about 2 or 3, but stronger shocks do occur [Hundhausen, 1972].

3.3.2 Solar Wind Streams

The earliest continuous observations of the solar wind were those from Mariner 2 [Neugebauer and Snyder, 1966] which found that the solar wind speed and density were variable, but both quantities also exhibited a good deal of order (Figure 3-17). During a 27-day rotation period of the sun, several intervals of higher than average speed occured. As seen in the figure, the high speed intervals tended to recur on the next rotation. The proton density typically maximized before the velocity. These high-speed corotating (that is, rotating with the sun) streams are one of the outstanding organized features of the solar wind. A particular stream may persist for many months or longer than a year. They are responsible for recurrent geomagnetic storms that characterize the declining phase of the sunspot cycle. The streams are imbedded within the sector structure, that is, magnetic sector boundaries typically precede the velocity rise by hours or days. The solar sources of the streams are the low temperature, open magnetic field regions known as coronal holes [Neupert and Pizzo, 1974].

The properties of the stream structure have been exten-

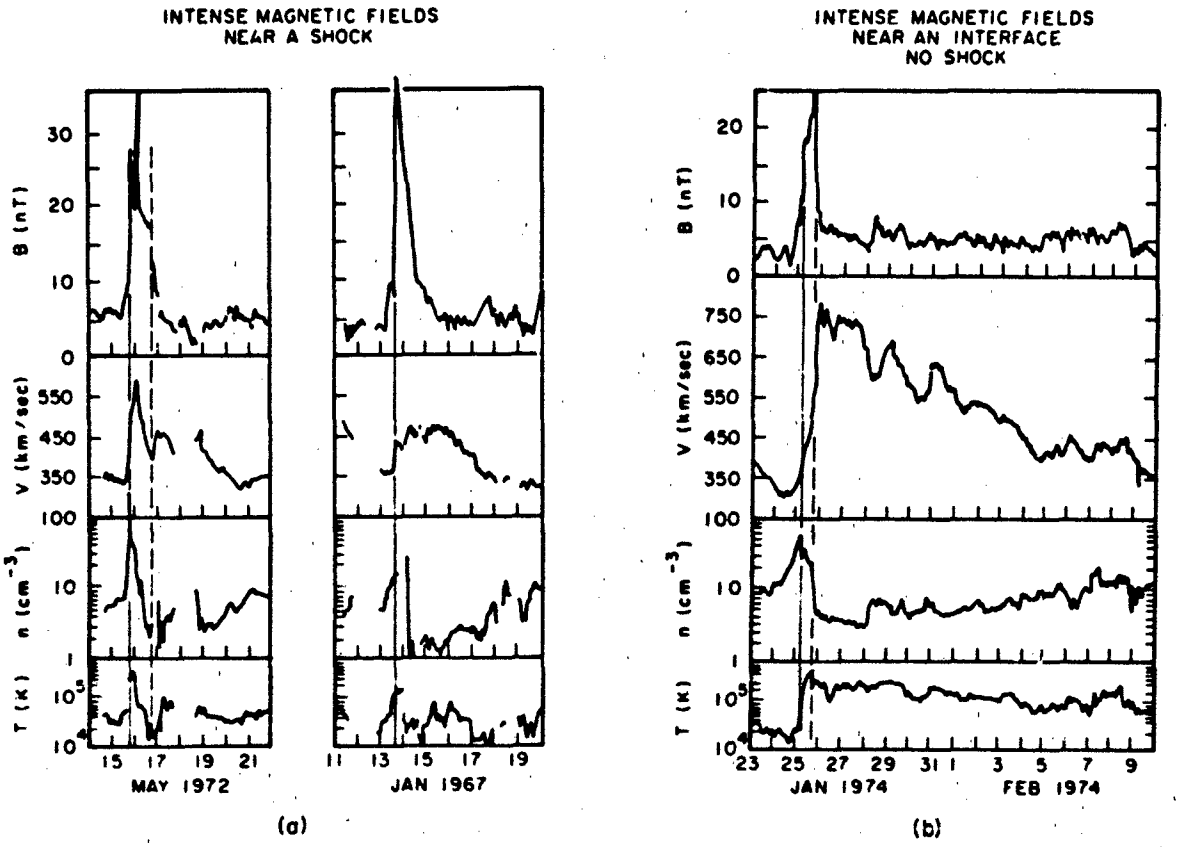


Figure 3-16 Shocks and stream interfaces. (a) gives hourly average data for two examples of shocks detected in the solar wind. Note the rise in velocity, temperature, density and magnetic field at the solid vertical line. The dashed line indicates the end of the shocked region. (b) shows a stream interface. The interface region is characterized by an abrupt drop in density and rise in temperature. It is the region between the density and temperature maxima and is bounded by the vertical lines. Although both the shocks and the interface region are characterized by very high magnetic fields in these examples, that feature does not always occur (see Figure 3-23 for magnetic field behavior across a shock). [Burlaga and King, 1979]

sively analyzed. The structure observed at 1 AU is partly due to the structure of the wind as it leaves the sun, and partly to the evolution of that structure due to the interaction of low and high speed streams emmitted from a rotating sun. An intuitive idea of the interaction is easily developed by the following picture. Assume low speed wind leaves the sun's neighborhood to the west of the source of high speed wind, that is, at a position further along in the direction of rotation. This means that as the sun rotates the high speed region rotates to the angular position previously occupied by the low speed source region. Then, viewed along a radius vector to the sun, the low speed parcel of wind will have a high speed parcel of wind overtaking and compressing it. As the high speed material increases the velocity of the low speed material, the low speed material will become compressed. At the same time, due to conservation of momentum, the high speed material will be slowed and its density will increase. Thus a ridge of high density material will form on both sides of the boundary between the streams.

The boundary between low and high speed streams has a distinctive character and is called a stream interface region. It is characterized by an abrupt drop in density, a sharp rise in temperature, and a rise in velocity that may be small. An example of a stream interface is given in Figure 3-16 [Burlaga and King, 1979]. This figure also shows an example of an interface associated with a very intense magnetic field. Magnetic field change is not a defining characteristic of interfaces and often is not present.

A study of the characteristics of selected solar wind streams was made by Gosling et al. [1978] who superposed data using 23 very abrupt interfaces as the zero time point for the data. The data were extremely well ordered by this procedure, as is seen in Figure 3-18. The figure gives the one hour average data for 100 hours both before and after the interfaces. The error bars are the estimated error of the

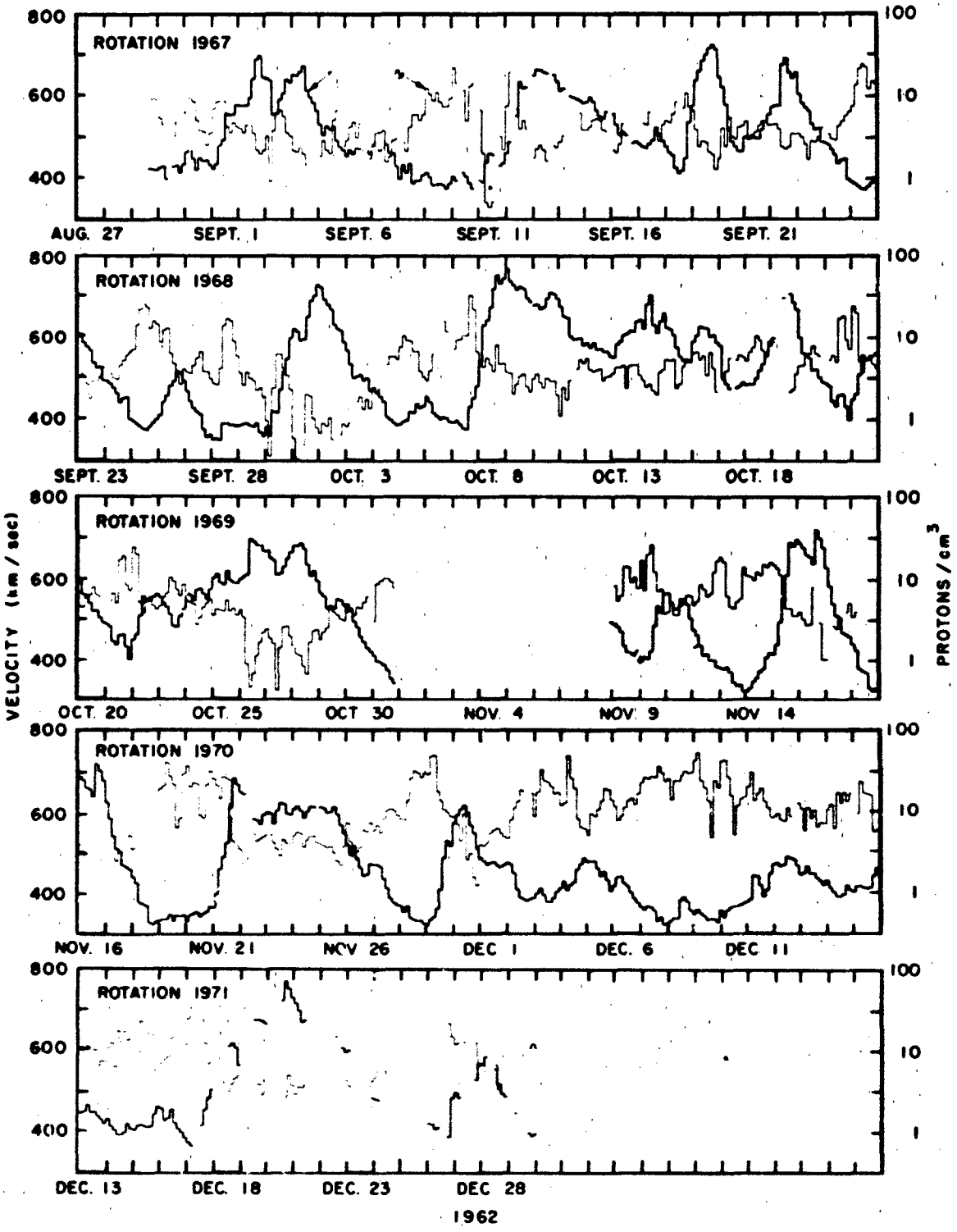
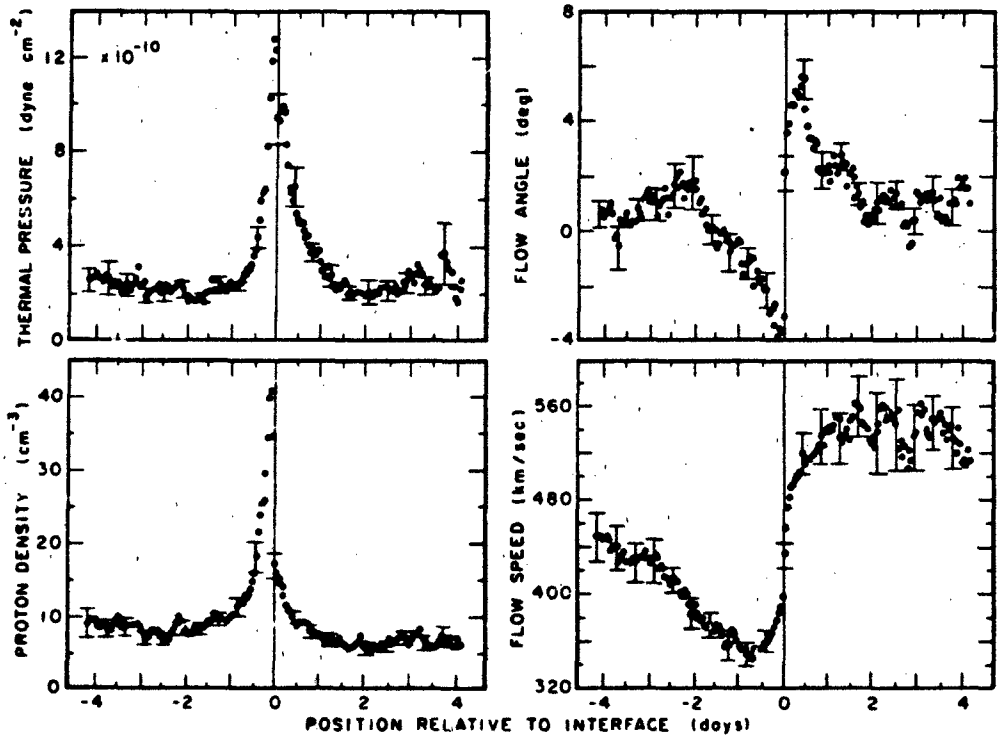
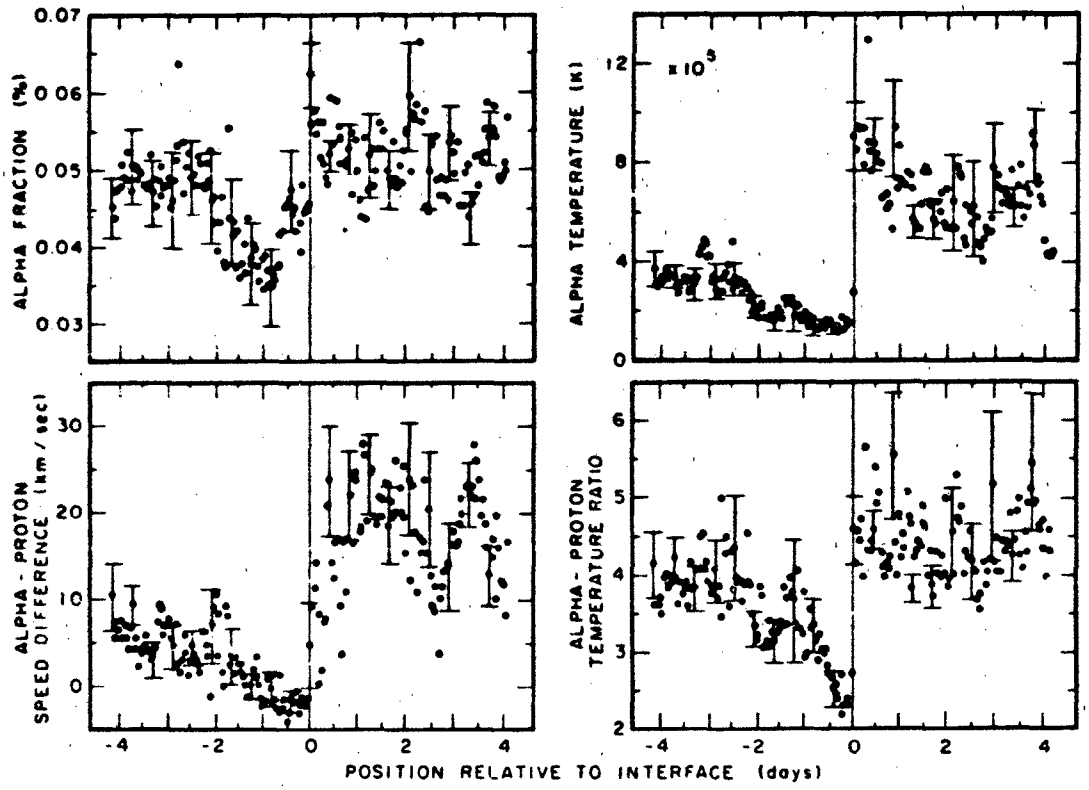


Figure 3-13 Three-hour averages of the solar wind proton density (light line) and flow speed (heavy line) from Mariner 2 on its flight toward Venus in 1962. Each panel presents 27 days of data. Note the velocity stream structure and the tendency for recurrence of streams on each solar rotation [Neugebauer and Snyder, 1966]


THERMAL PRESSURE (dyne cm^{-2})
x10^{-10}
FLOW ANGLE (deg)
PROTON DENSITY (cm^{-3})
FLOW SPEED (km/sec)
POSITION RELATIVE TO INTERFACE (days)

(a)


ALPHA FRACTION (%)
ALPHA TEMPERATURE (K)
x10^5
ALPHA - PROTON SPEED DIFFERENCE (km / sec)
ALPHA - PROTON TEMPERATURE RATIO
POSITION RELATIVE TO INTERFACE (days)

(b)

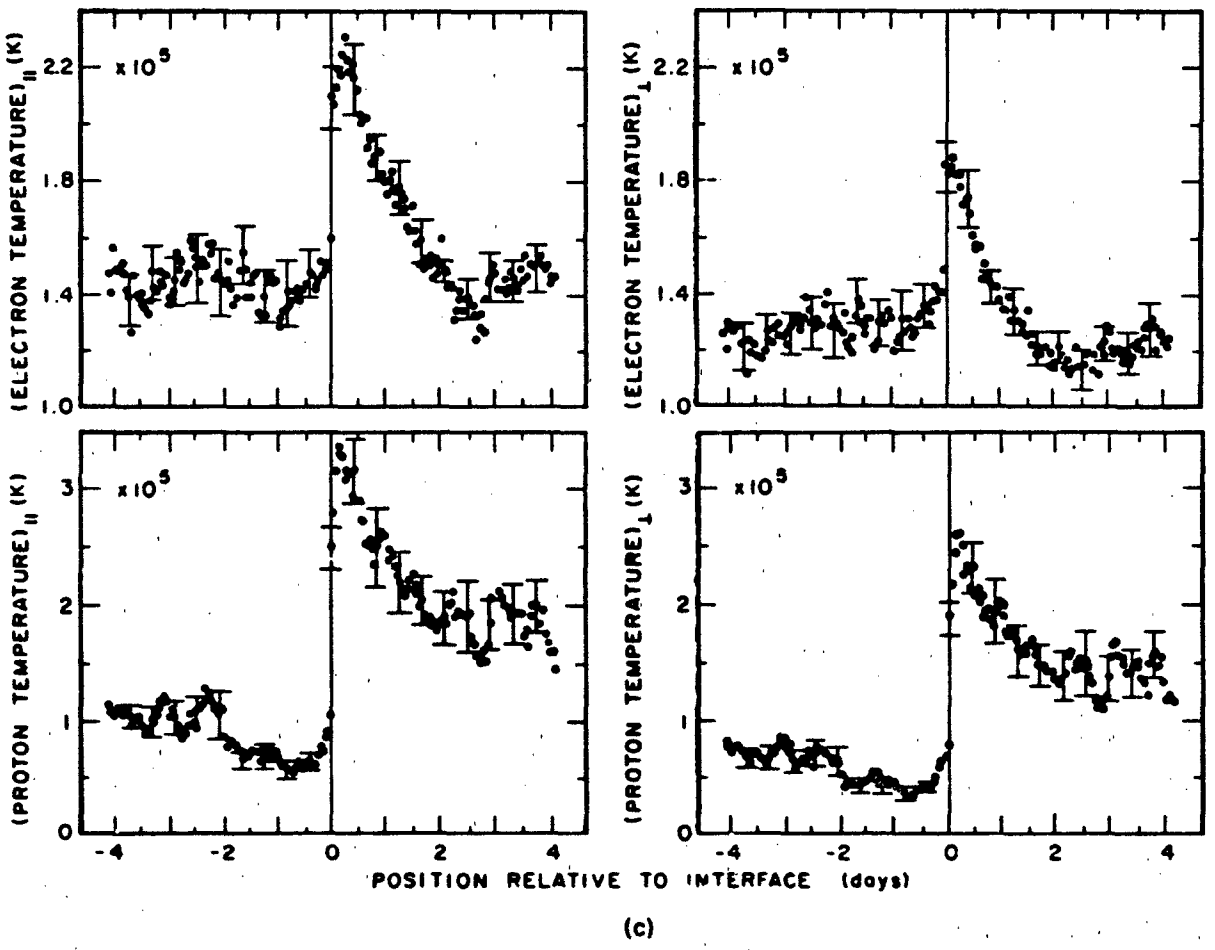


Figure 3-18. Superposed epoch analysis of one-hour average plasma properties in 23 streams. The zero time is defined as being the time of an abrupt interface. No streams with more gradual interfaces (Figure 3-16) were included in the analysis [Gosling et al., 1978].

mean. Figure 3-18a gives the proton thermal pressure, the proton density, the flow angle in the ecliptic plane and the flow speed. The interface occurs when the speed is still quite low, about 400 km/s. Three hours after interface passage, the average speed is about 480 km/s, but the maximum average speed of 560 km/s is not reached until 1.5 days later. The plasma near the interface is accelerated by the stream-stream interaction. Since the interaction region has the shape of an Archimedes spiral (for the same reasons discussed in Section 3.2.2, for the magnetic field direction) the low speed plasma is deflected toward the west whereas the high speed plasma is deflected towards the east. A shear flow occurs at the interface. Magnetic sector boundaries occur in the low speed wind preceding the interfaces. They typically appear from 1.5 days to 1.5 hours before the interfaces. The total pressure (that is, particle plus magnetic field) peaks approximately at the interface although the proton density and proton thermal pressure are already declining at that position. On the average there is discontinuity in density of about 7 cm^{-3} at the interface. This discontinuity takes place on a time scale of minutes.

Figure 3-18b gives the temperatures of the protons and electrons parallel and perpendicular to the magnetic field. All four quantities change discontinuously at the interface and both electron temperatures return to near their low speed values three hours after the interface passes, but both proton temperatures remain elevated. Note that the temperature scales are not the same for the protons and electrons and that the electron temperature increase is relatively small.

The properties of the helium in the solar wind also show organization when analyzed according to stream structure. Figure 3-18c shows that the abundance of helium minimizes about a day before the interface and that both the helium temperature and the ratio of alpha particle temperature to proton temperature are discontinuous. In most of the stream structure the helium is, on average, flowing faster than the hydrogen, but during the day before the interface is crossed the relative velocity changes sign, and in this region, the

proton speed exceeds the alpha particle speed by a few cm/s.

3.3.3 Solar Flare Disturbances

The second major interplanetary structure that orders the solar wind is the solar flare disturbance. Major flares eject high speed plasma into the wind, which results in a disturbance propagating to the earth and beyond. The disturbance is shown schematically in Figure 3-19 [Bame et al., 1979] The rapidly moving flare plasma compresses the ambient wind before it and a shock is formed. Studies indicate that for major disturbances the shock expands on a broad front so that its shape when the shock arrives at 1 AU can be approximated by a half circle of radius 0.6 AU and center 0.4 AU [Hirshberg, 1968; Rickett, 1975]. At the shock the density, proton temperature, magnetic field, and velocity increase abruptly as the shocked gas is entered. The density and velocity may or may not continue to increase in the post shock region depending on the details of the delivery of energy from the flare site to the solar wind and also on the structures through which the shock propagated on its way from the sun to the earth. For example, if a shock is observed in the low speed wind preceeding or following a stream, the density and velocity change at the shock will be relatively large. However, if observed near the density peak or in the high speed wind, a disturbance with the same initial conditions at the sun will show a relatively small change in velocity, although the change in density may still be considerable [Hirshberg et al., 1974]. A discontinuity between the shocked ambient plasma and the driven gas is shown in Figure 3-19, but its exact position is often very difficult to identify observationally. A reverse shock may form behind the discontinuity.

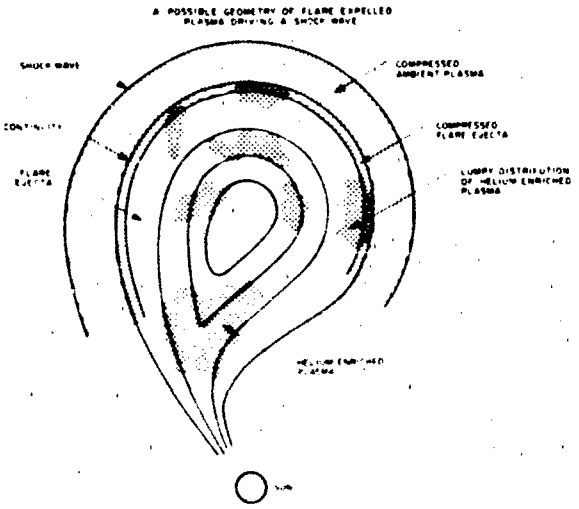


Figure 3-19 A schematic diagram of a possible geometry for the solar wind disturbance caused by a major solar flare [Bame et al., 1979]

The driver gas itself, which typically arrives at earth about 5–10 hours after the shock, is often characterized by enhancement in the helium abundance that commonly exceeds 15% helium by number and may reach more than 20%–25% [Hirshberg et al., 1972]. On some occasions the helium enhancement is not seen or it may be seen later in the event. In Figure 3-19 this is interpreted as indicating a patchy distribution of helium enhanced plasma. The number of particles released into interplanetary space by major flare events is estimated to be on the order of 10^{39}–10^{40} particles or 10^{15}–10^{16} g. The driver gas is often also characterized by low proton [Gosling et al., 1973] and/or electron temperatures and bidirectional streaming of electrons. These observations suggest that the magnetic field can form closed loops as shown in Figure 3-19. Such loops have apparently been observed [Burlaga et al., 1981]. Examinations of the ionization states of heavy ions in plasma associated with flares indicate that at least part of the flare ejecta originates in coronal regions with temperatures of 3–4 $\times$ 10^6 K. The driver gas has been slowed down on its way from the sun to the earth by interaction with the ambient solar wind. Thus the observed velocity of the driven gas will be less than the mean transit velocity of that gas from the sun to the earth. Typical values of the ratio of observed velocities to mean transit velocities for a series of major flare-shock events is about 0.8 [Hirshberg et al., 1972]. The interplanetary magnetic field during flare disturbances often reaches very high values (30–40 nT) in the direction perpendicular to the ecliptic plane. Since the velocity of the driver and disturbed gas is high, and if these strong fields are southward, then the solar wind will interact strongly with the earth's magnetic field and will drive impressive geomagnetic storms and auroras. The storms reflect the structure of the flare disturbance. The classical storm begins with a sudden commencement caused by the abrupt increase in dynamic pressure on the magnetopause that occurs when the shock is crossed. The initial phase of the storm, during which the earth's field is enhanced but not disturbed is due to the shocked region flow when the magnetic field in that region is northward. When and if the field turns southward the main phase of the storm develops.

3.4 LONG TERM VARIATIONS

3.4.1 Solar Cycle

It was expected that the interplanetary medium varied with the solar cycle even before *in situ* observations of the solar wind began. Fluxes of particles from the sun were believed to be the cause of geomagnetic activity and aurora and it was well established that these phenomena exhibited

lar cycle variation. The variation of the solar output ıg the eleven year solar cycle has many manifestations; :xample, sunspot numbers, major solar flares, and the · wind velocity. Distinctions should be made between olar cycle as expressed in each of the various quantities. ough all solar cycle variations will have the same period may have different phases, relative amplitudes, and es. It is common practice to use the variations of sunspot bers as the standard for the solar cycle and to compare :hanges in all other quantities to the sunspots. Here the "solar wind cycle" will be used to denote the changes ıe various interplanetary parameters.

Before direct observations were begun, studies of geo-netic activity had shown that the variations in the in-lanetary medium were such that minimum solar cycle litions at earth's orbit occurred (on the average) about 1.5 years after sunspot minimum. Also, geomagnetic ms characterized by sudden beginnings were most com-ı at sunspot maximum, but storms tending to recur on :essive solar rotations were more common about three rs after maximum. The sudden commencements are due olar wind shocks which are most common at sunspot timum. Recurrent storms are caused by long-lived rel-ely steady high speed streams in the solar wind such as se observed in 1962 and 1973–74, that is, during the se of decreasing sunspot numbers.

The solar wind has now been observed at least inter-tently for almost 20 years and solar cycle variations of eral parameters have been established. Figure 3-20 ynman, 1983] shows the variations of some of these ntities. Figure 3-20a gives the annual mean sunspot num-from 1962 to 1978. Figure 3-20b shows the annual mean ocity of the solar wind. The first few values are joined a dashed line to indicate that the data sample was too all to give an accurate value.

There is also a solar cycle variation in the magnitude of interplanetary field that is most clearly seen in the yearly an value of the logarithm of |B| as shown in Figure 0c. The data exhibit relatively low mean values, cor-ponding to |B| about 5 nT in 1964, 1965, 1975, 1976 1977. The mean log |B| is remarkably constant through-t the cycle between 1966 and 1974 corresponding to |B| ut 5.8 nT. The large rise to |B| about 6.9 nT appearing the beginning of the new sunspot cycle shows that the ar wind magnetic field can differ from cycle to cycle.

The yearly averages of the hourly averaged magnitude B_z, ($|B_z|$), are also compared to the sunspot number in gure 3-20d. Here B_z is defined in GSE coordinates and hourly average is taken by including the sign of B_z, that if B_z were of constant magnitude and negative for half the hour and positive for the rest, $|B_z|$ would be zero. ere are two maxima in ($|B_z|$), one before and one after nspot maximum. It is not known if this double maximum a common form of the ($|B_z|$) solar cycle variation.

The variation in the yearly average solar wind speed (v) om 1966 to 1978 is shown in Figure 3-20b. Data before

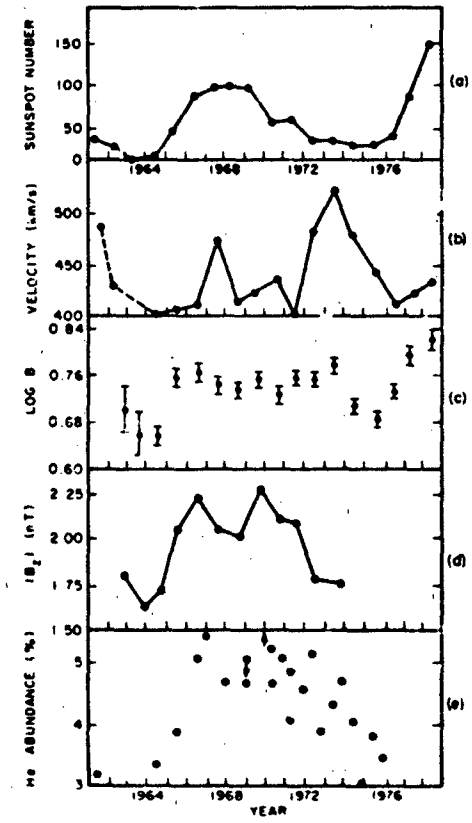


Figure 3-20. Long term variations of selected quantities. The top panel (a) shows the annual mean sunspot number. The second panel (b) gives the annual mean velocity. The third panel (c) shows the yearly averaged values of logarithms of hourly interplanetary magnetic field magnitudes. The observed distribution of hourly IMF is log normal. The fourth panel (d) shows the yearly average of hourly values of the magnitude of the z component of the IMF. The final (e) panel gives the percentage abundance of helium by number. The arrows indicate that only limits on the values could be determined [Feynman, 1983]

1966 were sparse but what information there is indicates that the mean velocity observed in 1962, during the declining leg of solar cycle 19, was high with a velocity of about 490 km/s whereas in 1964 and 1965 it was near 400 km/s. The distinction between the shape of the sunspot cycle variation and the solar wind cycle variation is clearly shown by the velocity. Here the highest yearly average velocities are observed during years dominated by recurring high speed solar wind streams. That this is nonetheless a solar cycle variation is shown by the fact that the declining phase of each sunspot number cycle exhibits such a series of recurrent geomagnetic storms.

The averaged helium abundance in Figure 3-20e shows a solar cycle variation much more like that of the sunspot number. The abundance near sunspot minimum in 1962–65 and 1976 is just above 3%, whereas 5% was the abundance typical of sunspot maximum conditions.

3.4.2 Secular Variations

Little information exists on the variation of the solar wind for time scales longer than a solar cycle. In Figure 3-20c it appears that the interplanetary magnetic field intensity in 1979 near the maximum of the current solar cycle was larger than at any time in the past, but Figure 3-20b shows that the velocity was quite close to values typical of the preceding maximum. It has also been inferred from geomagnetic records that the solar wind was in some sense weaker during the early years of this century than it has been recently [Feynman and Crooker, 1978], and from auroral records that it was also weak in the early years of the 19th century [Feynman and Silverman, 1980]. The lack of aurora during the Maunder minimum in the 17th century again indicates a weak wind [Suess, 1979]. A revaluation of the auroral record from 450 AD to 1450 AD convincingly shows a periodicity in occurrence frequency with an average period of about 87 years [Siscoe, 1980] which almost certainly reflects a solar wind variation.

3.5 THE CLOSE NEIGHBORHOOD OF THE EARTH

3.5.1 Magnetosheath

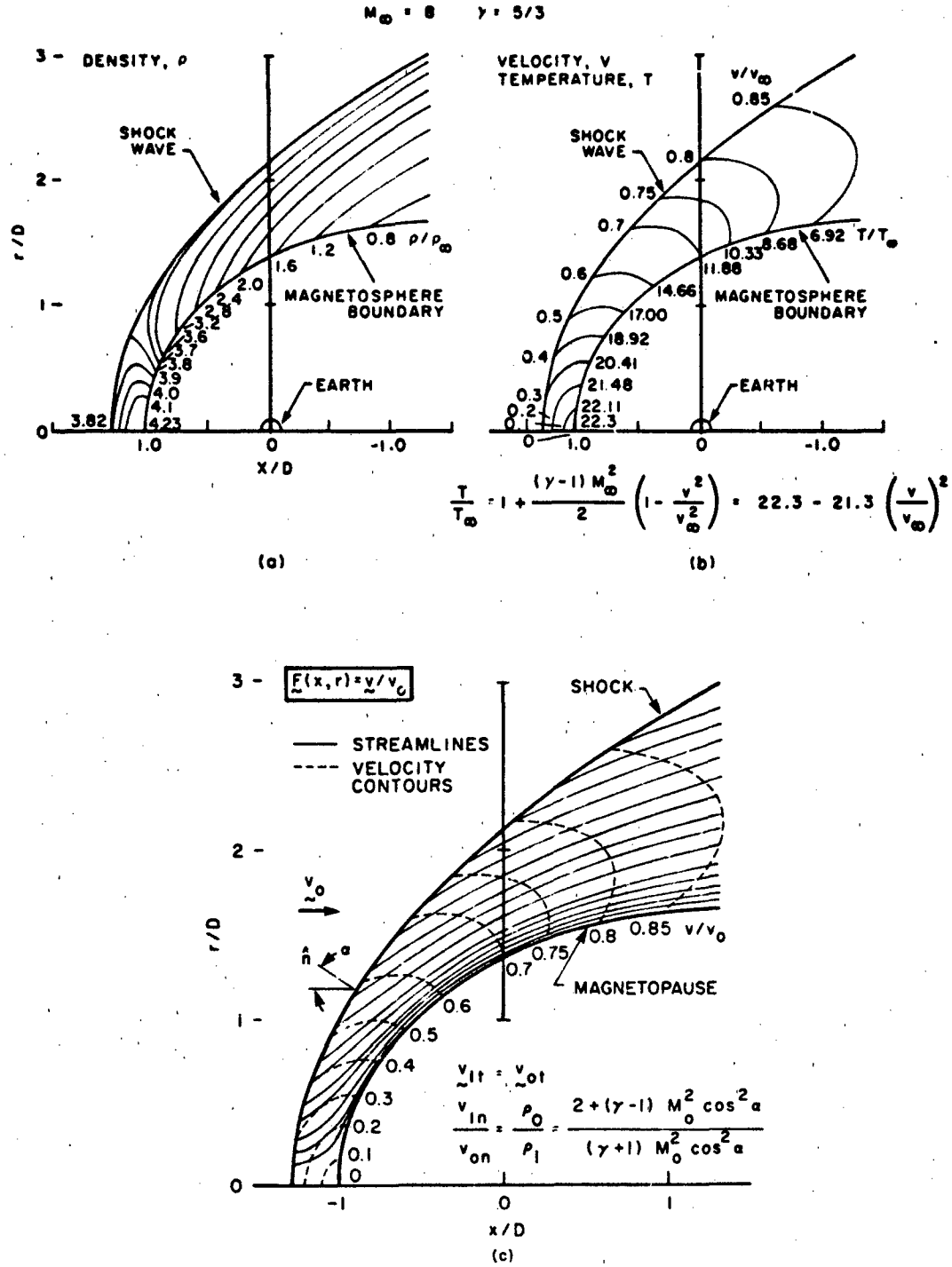
As the magnetized solar wind flows past the earth, the plasma interacts with the earth's magnetic field and confines the field to a cavity, the magnetosphere. Since the proton flow of the solar wind is almost always highly supersonic and super-Alfvénic, a shock, called the bow shock, forms sunward of the magnetospheric cavity. The configuration is shown schematically in Figure 5 of Chapter 8. The solar wind flows across the bow shock in front of the earth where it is slowed to subsonic velocities. It is deflected by the obstacle presented by the earth's field at the boundary designated as the magnetopause in Figure 8-5. The wind then flows generally antisunward down the flanks of the magnetosphere carrying the interplanetary field with it. A typical distance between the earth's center and the subsolar point on the magnetopause (the boundary between the solar wind and the magnetosphere) is 10 earth radii (R_I). The shock is about 3 R_I sunward of the magnetopause at the subsolar point. The region between the shock surface and the surface of the magnetosphere is called the magnetosheath.

The problem of calculating the flow of the solar wind plasma past the earth using the full plasma equations is prohibitively difficult. Even in the magnetohydrodynamic approximation, the problem has not been treated in general. The high conductivity of the solar wind and the large kinetic energy density of the particles permits the approximation to be made that the magnetic field configuration can be treated as though the field is frozen into the flow. Although the solar wind proton mean free path for Coulomb collision is typically 1 AU at earth's orbit, the plasma acts as a fluid on a much smaller spatial scale, because plasma instabilities and wave particle interactions produce a small equivalent mean free path. For many applications it is good approximation to treat the solar wind as a magnetohydrodynamic fluid interacting with the earth's field. For a full discussion of the interaction of the wind and the magnetosphere the process of magnetic reconnection or merging is important; however, its major effect is to feed energy into the magnetosphere rather than to affect the flow around the magnetosphere and so it is beyond the scope of this chapter. See Chapter 8. A further simplification is usually made in the calculations by assuming the Alfvén Mach number of the solar wind flow is large. This decouples the fluid equations from the magnetic equations so that the properties of the flow can be determined by solving the gas dynamic equations, and subsequently the magnetic field can be found from the flow.

The shape of the forward part of the magnetosphere is calculated by neglecting the magnetic field pressure in the solar wind outside the magnetosphere boundary and the particle pressure inside the boundary. The remaining pressures are then set equal on the two sides, that is, the dynamic pressure of the solar wind flow outside is balanced by the magnetic field pressure inside.

The flow around the magnetosphere can then be calculated. To simplify the calculations further the magnetosphere is often approximated as an axisymmetric body formed by the rotation of the equatorial trace of the magnetosphere boundary around the earth sun line. The resultant flow is quite sensitive to the ratio of specific heats γ. Observations favor a γ of 5/3 [Auer, 1974] and this value was used in the model calculations discussed here.

The model flow in the high Mach number approximation is illustrated by Figure 3-21 [Spreiter and Alksne, 1969]. In Figure 3-21a, the lines of constant ratio of magnetosheath density to solar wind density are shown. The sheath density is high in the forward part of the magnetosphere reaching a maximum in the subsolar region. As the plasma flows back toward the magnetospheric flanks, it becomes less dense and values less than one begin to appear near the magnetopause at about 7 R_I behind the earth. Figure 3-21b shows both lines of equal velocity ratio and temperature ratio. Since the fluid approximation is used, the temperatures of the electrons and protons are assumed to be the same, and for comparison with observation they are assumed equal to the measured proton temperatures. The numbers labeling the lines and written on the magnetopause side of the magnetosheath are the temperature ratios, and the numbers on the bow shock side are the velocity ratios. The stagnation point is labeled at zero velocity, which of course is unrealistic as far as expectations of observations are concerned. However, slowing and heating of the plasma in the subsolar stagnation region is observed. As the plasma moves back along the flanks it begins to cool and increase in speed. Figure 3-21c shows the stream lines of the flow.



$$\frac{T}{T_\infty} = 1 + \frac{(\gamma-1)M_\infty^2}{2}\left(1 - \frac{v^2}{v_\infty^2}\right) = 22.3 - 21.3\left(\frac{v}{v_\infty}\right)^2$$

$$\frac{v_{1t} = v_{0t}}{\frac{v_{1n}}{v_{0n}} = \frac{\rho_0}{\rho_1} = \frac{2+(\gamma-1)M_0^2\cos^2\alpha}{(\gamma+1)M_0^2\cos^2\alpha}}$$

Figure 3-21 Flow in the magnetosheath in the gas dynamic approximation for the high Mach number case (M = 8). (a) gives the density and (b) the temperature and velocity contours. The subscript ∞ denotes quantities in the undisturbed wind at "infinity". The non-subscripted quantities are the parameters within the magnetosheath. Lines of constant ratio of magnetosheath in undisturbed quantities are given. The shape of the contours for velocity and temperature are the same but the ratios differ as marked. The jump relation for the temperature is given in (b). In (c) the velocity contours are given again, along with the stream lines. The formula at the bottom of (c) gives the jump relations across a shock of arbitrary angle with the incident flow. Subscript 1 indicates values just behind the shock. The abscissas are normalized so the distance from the center of the earth to the subsolar point on the magnetopause is taken as one [Spreiter and Alksne, 1969; Spreiter et al., 1966].

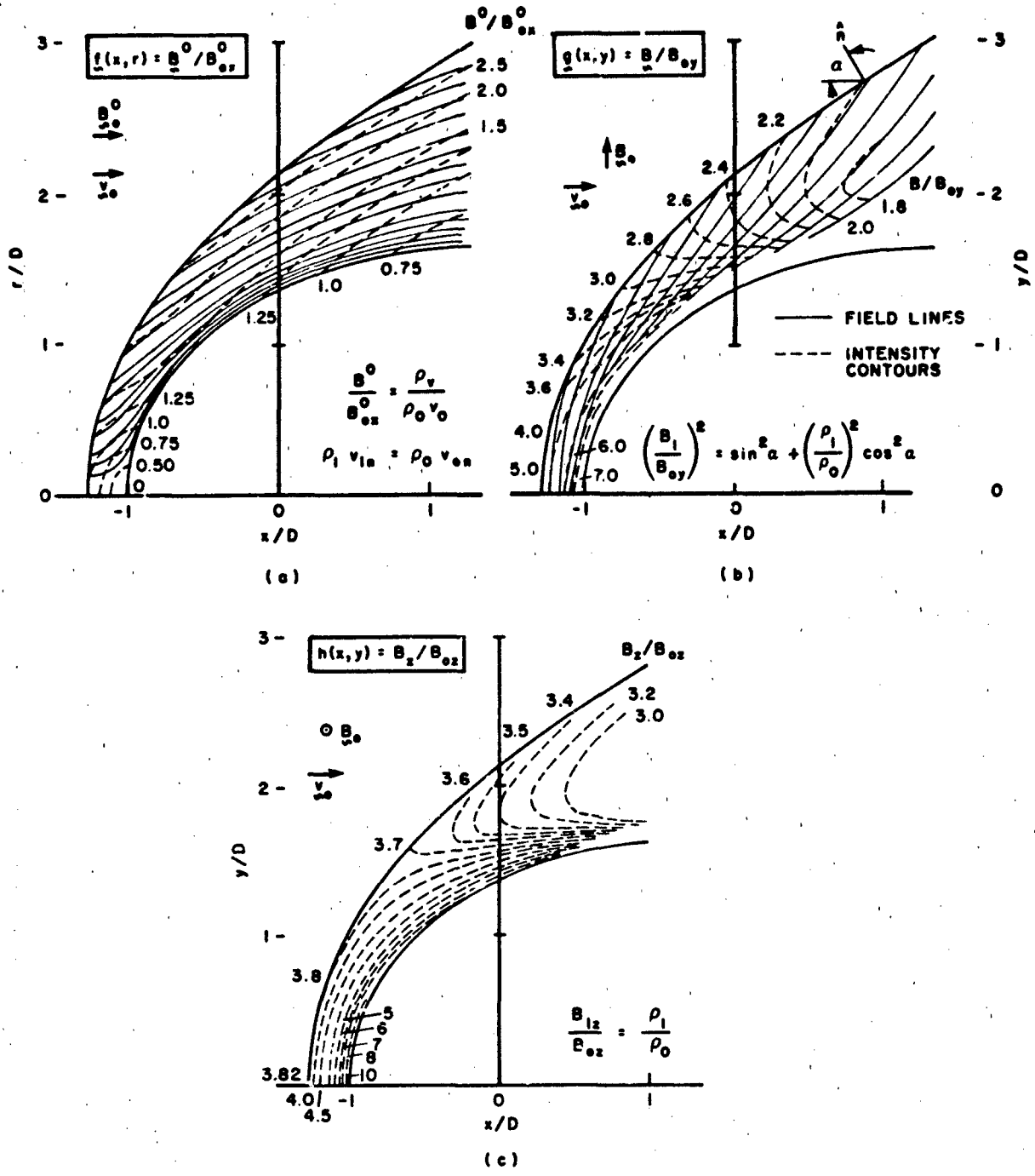


Figure 3-22. Magnetosheath magnetic fields for three orthogonal components of the interplanetary field calculated by assuming the field is frozen into the flow shown in Figure 3-21. The field directions are as shown in (a), (b), and (c). The jump relations at the shock are indicated and the dashed lines indicate contours of equal intensity ratio, whereas the solid lines follow the directions of the field lines [Aksne and Webster, 1970]. (Reprinted with permission from Pergamon Press, Ltd. © 1970.)

The magnetic field configuration is computed in this approximation by assuming the field is carried by the flow. The results are shown in Figure 3-22 [Alksne and Webster, 1970]. The incident field has been broken up into three orthogonal components with directions as shown on the figure. An arbitrary incident magnetic field will be described by an appropriate superposition of the three cases. The stream lines are given by the solid lines and the lines of equal magnetic field ratio are given by the dashed lines. The magnetic field lines become draped over the magnetosphere and tend to be aligned with the magnetopause direction.

Calculated values of magnetosheath electric fields are shown in Figure 3-23 [Alksne and Webster, 1970] but these have not been measured. Since there is no electric field associated with field aligned flow, electric fields for only two orthogonal magnetic field directions are shown, and the electric field due to an arbitrary magnetic field can be found by combining the two cases shown.

Measurements of the shape of the magnetosphere show some asymmetry that is not present in the results calculated from the high Mach number approximation. This asymmetry probably arises because of the the asymmetric magnetic field in the magnetosheath and the coupling between the fluid flow and the magnetic field. All of these effects will lead to magnetosheath pressure asymmetries which will distort the magnetosphere to some extent. In addition, the orientation of the earth's magnetic dipole relative to the direction of incidence of the solar wind must be taken into account.

The observed flow and averages of the magnetic fields in the magnetosheath agree satisfactorily with the magnetohydrodynamic models. The magnetosheath magnetic field is typically very disturbed but the vectors tend to be aligned with the magnetopause as shown in Figure 3-24a [Fairfield, 1967]. The average relative magnitudes of the magnetosheath field for 10 R_E intervals along the earth-sun line are shown in Figure 3-24b [Behannon and Fairfield, 1969]. The ratio of the magnitude of the magnetosheath field to the magnitude of the interplanetary field varies from 4 near the most sunward portion of the bow shock to less than 1 far down the flanks.

The variability of the magnetosheath fields is reflected in the power spectra as shown in Figure 3-25 [Fairfield, 1976]. The low frequency power levels of the variations of the field magnitudes differ by four orders of magnitude depending on the day. The September and November 1967 spectra are probably the most representative. Spectral peaks near 0.05 Hz and 0.07 Hz are typical but the exact frequency and relative magnitude are variable. The 31 July spectrum was taken on a day when the solar wind Mach number was low. At these times the power in magnetosheath variations is also quite low. The dashed lines with slopes of 1/f and $1/f^3$ are shown for comparison. The spectra typically steepen at frequencies above the proton gyrofrequency. At frequencies between 3.3×10^{-4} and 2.6×10^{-3} Hz, which are lower than those shown in the figure, the average slope is slightly steeper than 1/f. There is a tendency for the per-

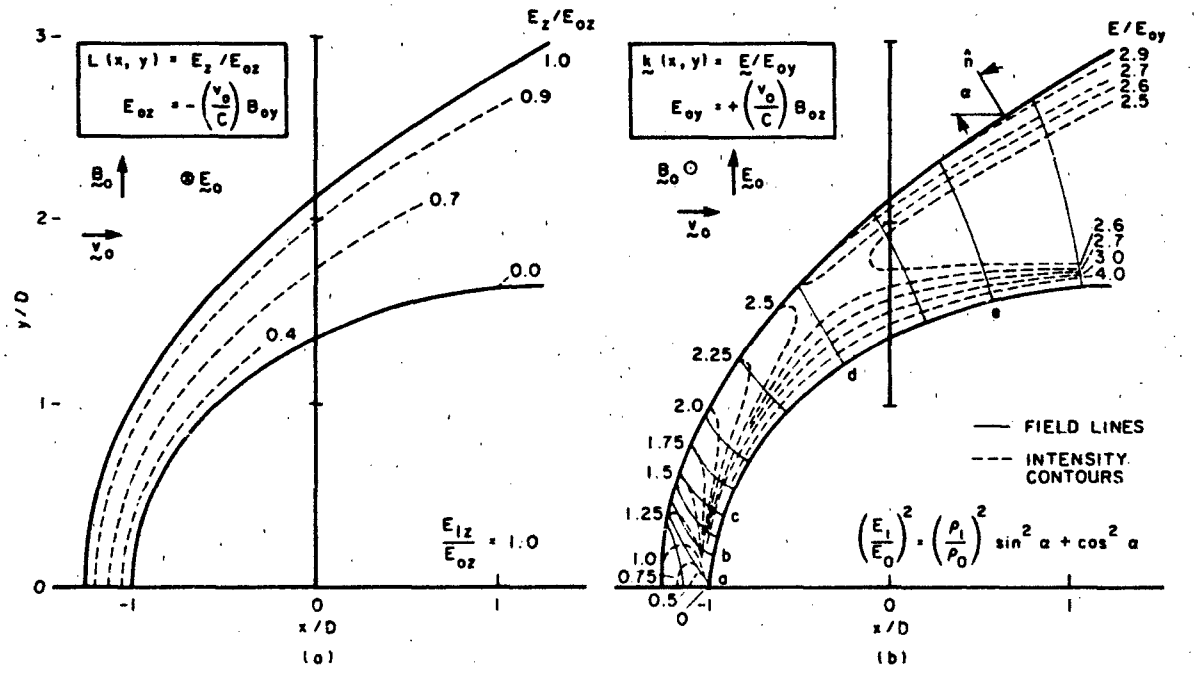
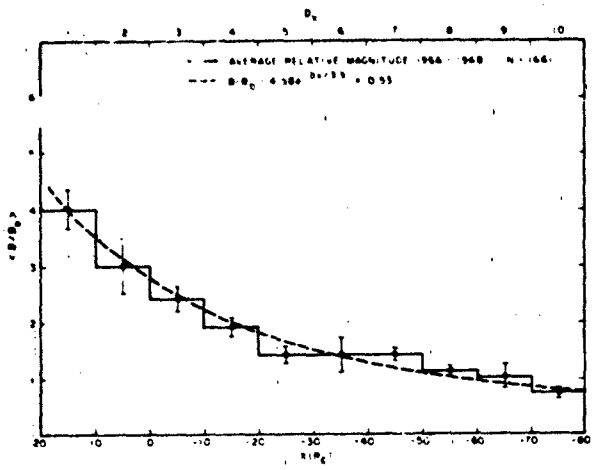
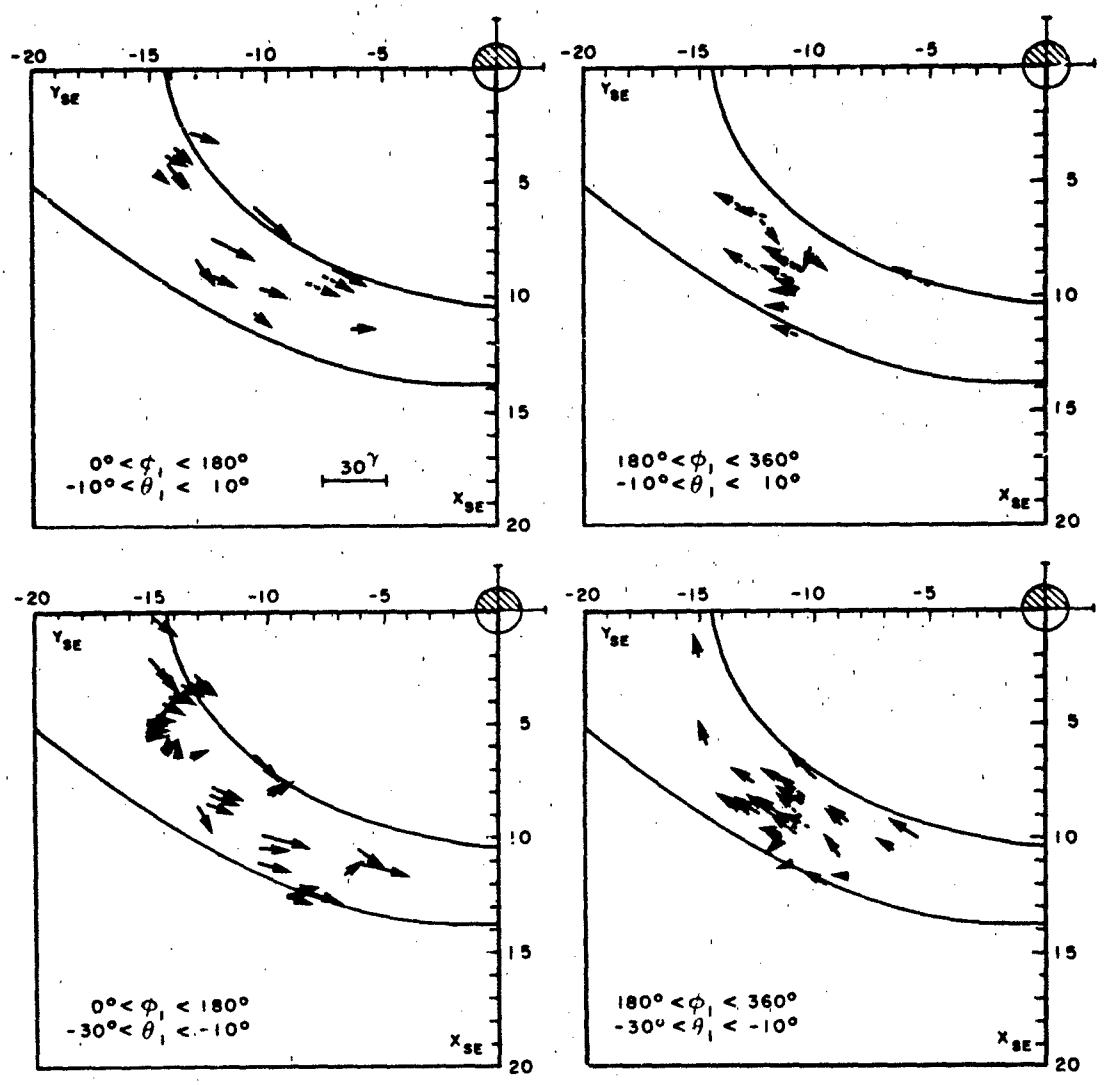


Figure 3-23. Magnetosheath electric fields calculated under the same assumptions as in Figure 3-22. The relations at the shock jump are indicated in each panel [Alksne and Webster, 1970]. (Reprinted with permission from Pergamon Press, Ltd. © 1970.)



(a)



(b)

Figure 3-24. Observations of the magnetosheath field: a) Averages of the magnetosheath magnetic field intensity normalized to the simultaneously measured interplanetary field. For each ΔX bin measurements are taken from all available values of y and z. The field is intense in the subsolar region and decreases to less than the interplanetary value in the distant flanks [Behannon and Fairfield, 1969] (Reprinted with permission of Pergamon Press, Ltd. © 1969.) b) Projections of vectors measured in the magnetosheath into the X, y plane for four different orientations of the interplanetary field, [Fairfield, 1967].

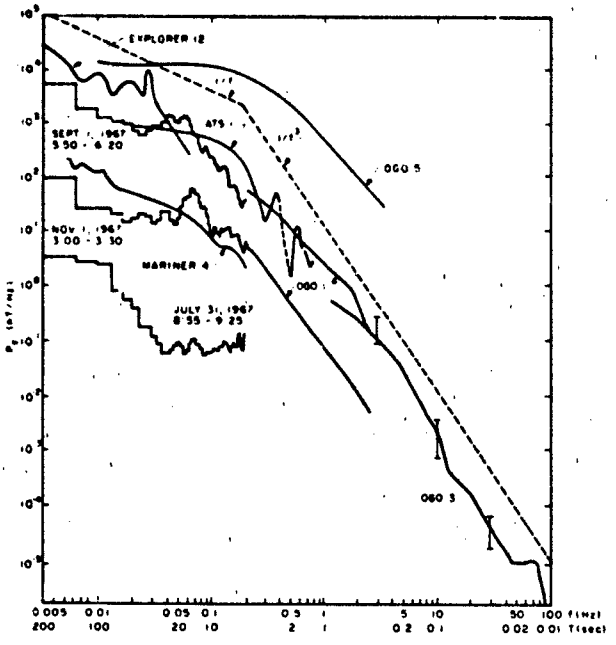


Figure 3-25. Composite of magnetosheath magnetic field spectra. The lower frequency data are magnetic field magnitude spectra while the highest frequency data are from fluctuations in the field components [Fairfield, 1976].

turbation vector to align itself with the shock surface for variations observed near the shock surface. In the inner portion of the magnetosheath, large amplitude waves occurring primarily in the field magnitude are observed. Rotational wave modes have been found to dominate over magnetoacoustic modes for wave periods greater than 10 minutes. There also appear to be plasma clouds with scales of several hundred to several thousand kilometers associated with weak fields that deviate from a direction tangent to the magnetopause. In addition to broad noise with a $1/f^3$ spectral shape there are often sporadic occurences of quasi-monochromatic bursts throughout the magnetosheath. These waves have a frequency of 50–200 Hz, durations of less than one to tens of seconds and amplitudes of tenths of nanoteslas.

Several sources have been suggested for magnetosheath field fluctuations. The waves and discontinuities in the solar wind are sources of some of the variability, as is the convection into the magnetosheath of waves produced at the bow shock or in the region of space just upstream from the bow shock. Waves incident on the bow shock will be transmitted across it and amplified and new waves of various types will be produced in the process. The details concerning the types of waves produced depend on the properties of the incident wave. The power in the variations and complexity of observed forms will be markedly increased by transmission. Interplanetary discontinuities will contribute to magnetosheath variability but they are not frequent enough in the solar wind to explain a significant portion of magnetosheath power. Processes taking place at the magnetopause or within the magnetosheath also contribute to observed waves in the magnetosheath.

3.5.2 Bow Shock

The flow of the solar wind past the earth is usually but not always highly supersonic and superalfvenic as far as the protons are concerned. However, both theory and observations imply that the treatment of the bow shock as if it were a hydrodynamic discontinuity is a gross oversimplification. Since the solar wind at earth is collisionless on a scale much larger than the magnetosphere, plasma properties are responsible for producing the bow shock discontinuity, which is a collisionless shock. The types of plasma phenomena involved will depend on the Mach number of the flow, the ratio (β) of the internal energy of the particles to the energy in the magnetic field, and the direction of the magnetic field relative to the shock surface. Figure 3-26 [Greenstadt et al., 1980] shows a conceptualization of collisionless shock phenomena at the earth's bow shock. The magnetic field is visualized as making an angle of about 45° with the solar wind, which is flowing from the direction of

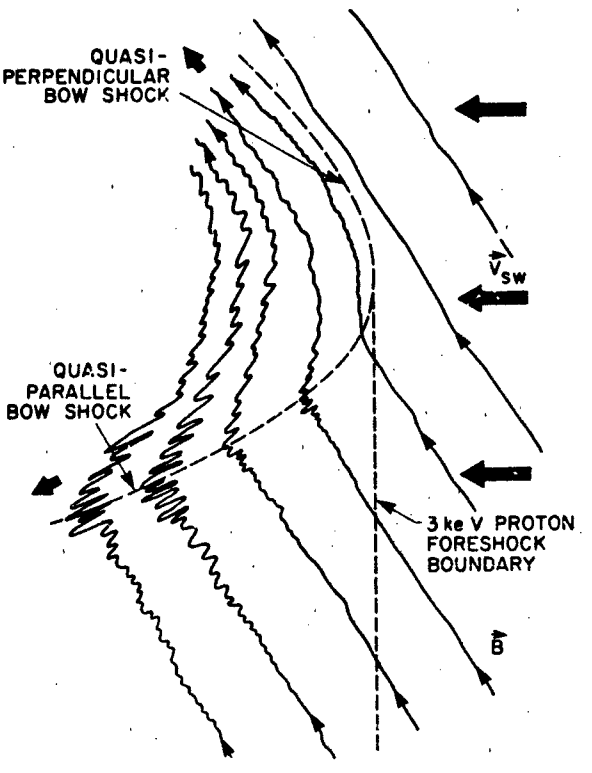


Figure 3-26. A schematic of the earth's bow shock showing the quasi perpendicular and quasiparallel regions. Perpendicular shocks most closely obey the classical magnetohydrodynamic equations but parallel shock region behavior is dominated by plasma effects [Greenstadt et al., 1980; Gary, 1981].

the sun, that is, the sun is beyond the lower right hand quadrant. Four regions of the shock surface are distinguished, the two extremes (strictly perpendicular and strictly parallel shocks) and the two intermediate regions (quasi-perpendicular and quasi-parallel shocks). A shock is called parallel when the upstream magnetic field is perpendicular to the plane of the shock and parallel to the upstream flow velocity. Conversely, perpendicular shocks have normals perpendicular to the direction of the magnetic field in the solar wind. The major part of the bow shock region has been divided into quasi-parallel and quasi-perpendicular regions based on empirical results. Quasi-perpendicular shocks have a monotonic, sawtooth, or wave-step magnetic profile and are sharply defined in space. Quasi-parallel shocks are thicker and do not show a clearly definable boundary between upstream and downstream parameters, either in the field or particle properties.

Figure 3-27 [Greenstadt and Fredericks, 1979] is a schematic of a cut in the plane containing the interplanetary field, showing the waves that appear in the parallel and quasi-parallel region in the vicinity of the classical position of the bow shock. The shocked magnetosheath plasma takes about 5 min to flow around the forward part of the magnetosheath so when parameters are averaged over periods of more than 5 min the shock should arrive at the shape given by steady state magnetohydrodynamic theory. The instantaneous shock is not in equilibrium. Figures 3-26 and 3-27 show schematically the waves in the vicinity of the boundary that propagate back into the solar wind and into the magnetosheath. The properties of both the quasi-perpendicular shock and the quasi-parallel shock depend on the range of β and M, where M is the Mach number of the flowing gas. Since the fast wave mode is the MHD wave that steepens into a shock, the Mach number for a perpendicular shock is given by $M = V/(C_A^2 + C_S^2)^{1/2}$ where C_A and C_S are the Alfvén and sound wave velocities. In Table 3-2 [Greenstadt and Fredericks, 1979] the characteristics of shock phenomena are given for various combinations of β and M. A value of $\beta \approx 1$ is the dividing line for the regimes of β parameter space and $M \approx 3$ is the dividing level for the Mach number. An estimate of the incidence of the specific ranges of β and M the solar wind can be found from Table 3-1. In the solar wind β varies from 0.5 to 4.6, and the ranges of sonic and Alfvénic Mach numbers are from 4 or 5 to over 10. However, solar winds with Alfvén numbers less than 1 have been observed on occasion.

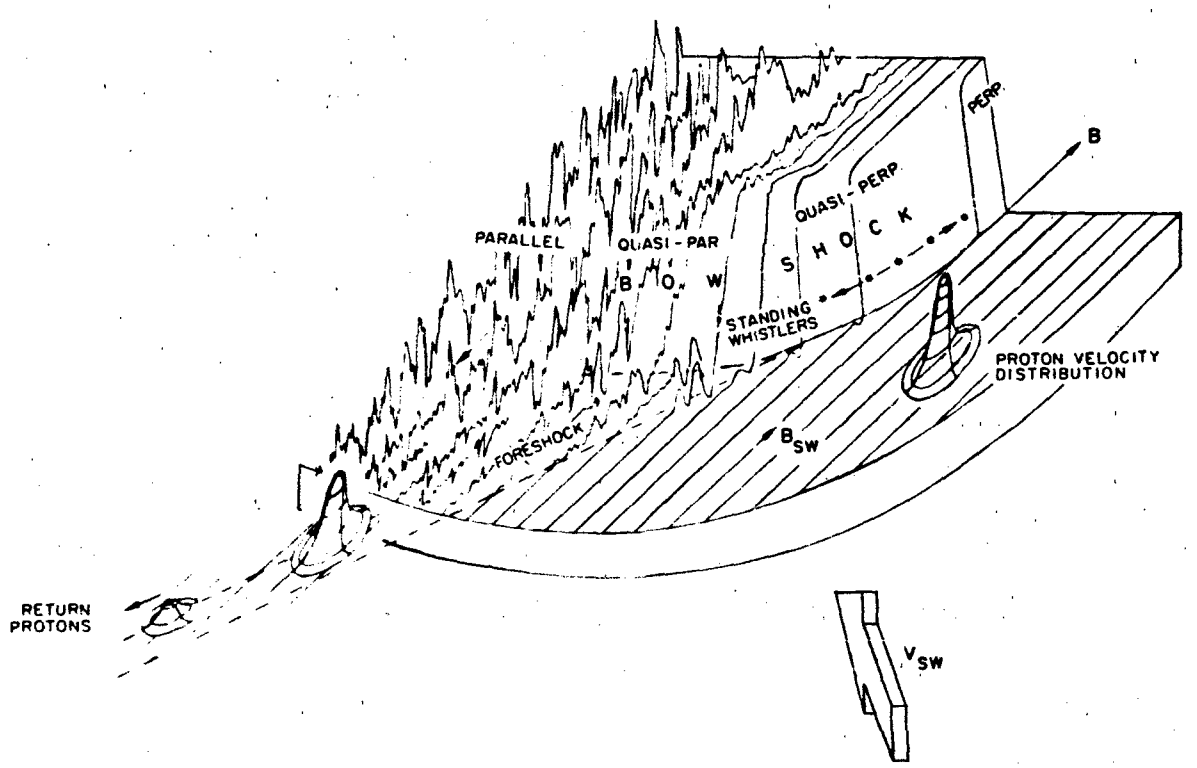


Figure 3-27. Conceptualization of collisionless shock phenomenology as manifested in the earth's curved bow shock. Unshocked interplanetary field direction B , is indicated on the foreground field "platform." Field magnitude is plotted vertically. The superimposed three-dimensional sketches represent solar wind proton thermal properties as number distributions in velocity space [Greenstadt and Fredericks, 1979].

Table 3-2. Macroscopic features of the bow shock by parameter class

Parameter Values	Plasma Conditions	Name of Structure	Quasi-perpendicular Features	Quasi-parallel Features
$\beta \ll$ $M \lesssim 3$	Cold plasma, low Mach number	Laminar	Clean field jump, sometimes with damped periodic waves, no turbulence. Relatively little proton temperature jump (T_{p2}/T_{p1}) $\approx$ 2 , maxwellian downstream distribution.	Multigradient field transition with embedded, nearly-periodic wavetrains. Upstream waves with strong periodic component, tens of seconds period, unknown plasma distributions.
$\beta < 1$, $M \gtrsim 3$	Cool plasma, high Mach number	Quasi-laminar	Clean field jump, downstream, nearly-periodic waves, little turbulence. Appreciable proton temperature jump, bimodal distribution, nonmaxwellian high energy tail downstream.	Multigradient field transition at least 2 R_i thick, with large amplitude pulses; extensive foreshock. Little change in solar wind streaming velocity; some heating with resulting non-maxwellian proton distributions differing from either solar wind or magnetosheath forms.
$\beta \approx 1$, $M \lesssim 3$	Warm plasma, low Mach number	Quasi-turbulent	Clean field jump, small scale turbulence. Little proton temperature rise, maxwellian downstream distribution.	No known example.
$\beta \approx 1$ $M \gtrsim 3$	Warm plasma, high Mach number	Turbulent	Irregular field fluctuations obscuring definite average field jump. Bimodal or multimodal proton distributions, nonmaxwellian downstream.	Multigradient, irregular field transition with large magnitude excursions, upstream waves. Unknown plasma distributions, but maxwellian deep in magnetosheath.
$\beta \gg 1$ $M > 3$	Hot, high velocity plasma	High-beta	Irregular field fluctuations of extremely high peak magnitude, possibly lowering β locally. Extensive precursor region with appreciable effect on approaching flow; proton distributions unknown.	

(From Greenstadt and Fredricks, [1979])

3.5.3 Upstream Waves and Particles

The solar wind-magnetosphere interaction generates substantial numbers of energetic particles with a high enough speed so that they can propagate back upstream guided by the ambient interplanetary field. The upstream particle region has a strong local time asymmetry with the major effect occurring usually on the dawnside on field lines connected to the quasi-parallel shock. This region of space upstream of the bow shock is sometimes called the foreshock and is characterized by many different types of particle distributions and wave modes. The energy flux of particles into the upstream region is found to be 10^{17}–10^{18} ergs, which is comparable to the peak energy released in a geomagnetic substorm. The presence of energetic heavy ions indicates that strong local acceleration is occurring. Observations of plasma waves, particles, and fields in the region has led to a partial understanding of some of the interrelationships among the phenomena [Tsurutani and Rodriguez, 1981].

The upstream energetic particle population has been divided into several types based on their observed velocity and phase distributions. Reflected ions are characterized by a beamlike distribution of solar wind ions that apparently have been reflected and energized at the bow shock. They have a streaming velocity 2 or 3 times the solar wind velocity and a density of about 0.1/cm^3. Their kinetic energy distribution typically maximizes at about 5 keV and has spread of 1–30 keV. They are detected on interplanetary field lines that connect to the bow shock and are nearly tangent to the shock surface. The sunward boundary of the region in which these particles are observed is defined by the field-aligned flow back into the solar wind plus $\mathbf{E} \times \mathbf{B}$ convection, where E is the solar wind electric field.

In addition to these beam-like ions there is a population of diffuse ions that has a nearly isotropic velocity distribution with a typical temperature of 4×10^7 K, an upstream flow speed of 0.8–1.2 times the solar wind velocity and a density that varies between 0.02 and 0.8/cm^3. The

energy density of the diffuse ions is comparable to that of the reflected ions. These particles are found downstream (antisunward) of the reflected ions, almost filling the entire foreshock region. An excellent correlation exists between diffuse ions and low frequency (0.01–1 Hz) MHD and ion-acoustic-like plasma waves. Recent observations indicate that the transition from distributions typical of reflected ions to that of diffuse ions is a smooth one and there are ions with intermediate properties in a region of space between the reflected and diffuse ion region.

Upstream electrons have energies of 0.5 to 100 keV and stream in a field aligned direction. The majority of events involve 1–2 keV particles with a peak flux of $10^4 cm^{-2} s^{-1} sr\ keV^{-1}$. They appear on interplanetary field lines newly connected to the bow shock and because of their high velocities can be found upstream of the ion foreshock boundary.

The region upstream from the earth's bow shock is also characterized by complex wave phenomena. There are MHD waves with an amplitude of 5 nT peak to valley and period of 10–60 s, which have a left-hand polarization in the spacecraft frame. The particles involved in the wave show a ± 15 km/s solar wind velocity variation as well as correlated density fluctuations indicating a magnetosonic component. There is a strong association between these waves and diffuse ions. Wave packets of whistler mode waves having intensities ≤ 1 nT and typically composed of 2 or 3 coherent cycles are observed. Long trains of more than 20 cycles have been seen on occasion. The emissions, though left hand polarized in the spacecraft frame have been shown to be Doppler-shifted right-hand emissions being convected back over the spacecraft. They propagate at small (~ 25°) angles to the magnetic field. The rest-frame frequencies have been deduced to be 0.1–40 times the proton gyrofrequency. They are related to the MHD waves and are detected during diffuse ion events. Whistler mode waves with frequencies 10–200 Hz are observed upstream of the bow shock. These emissions have a peak amplitude of 0.1 nT and occur at about 1/4 to 1/2 the electron gyrofrequency. It is believed that the most intense waves are associated with low-energy electron spikes coincident with intense diffuse ion fluxes.

Other waves that have been observed include small amplitude electrostatic waves in the 100 Hz–10 kHz range polarized primarily parallel to the ambient magnetic field. They have short wave lengths, $30 < \lambda < 215$ m, and a low relative energy density. Broadband impulsive electrostatic emissions with amplitudes proportional to f^{-2} for frequencies less than 3 kHz are observed to almost always accompany electron plasma oscillations. The electron plasma oscillations themselves have intensities up to 10 mV with wavelengths of 400–780 m and frequencies of 20–30 kHz. They are detected throughout the foreshock region and are driven by 0.2 to 1.5 keV electrons at the boundary of the electron foreshock. Low intensity electromagnetic emissions at twice the electron plasma frequency having a narrow bandwidth are detected upstream of the bow shock to distances of at least 240 R_E. The emission frequency is correlated with solar wind plasma density variations.

REFERENCES

Alksne, A.Y. and D.L. Webster, "Magnetic and Electric Fields in the Magnetosheath," *Planet. Space Sci.*, **18:** 1203, 1970.

Auer, R.D., "Magnetohydrodynamic Aspects of the Earth's Bow Shock 1. Equilibrium Bow Shock Position," *J. Geophys. Res.*, **79:** 5118, 1974.

Bame, S.J., J.R. Asbridge, W.C. Feldman, E.E. Fenimore and J.T. Gosling, "Solar Wind Heavy Ions from Flare-Heated Coronal Plasma," *Sol. Phys.* **62:** 179, 1979.

Barnes, A., "Hydromagnetic Waves and Turbulence in the Solar Wind," in *Solar System Plasma Physics*, edited by E.N. Parker, C.F. Kennel and L.J. Lanzerotti, North Holland, Amsterdam, 1979.

Behannon, K.W. and D.H. Fairfield, "Spatial Variation of the Magnetosheath Magnetic Field," *Planet. Space Sci.*, **17:** 1803, 1969.

Belcher, J.W. and L. Davis Jr., "Large Amplitude Alfvén Waves in the Interplanetary Medium," *J. Geophys. Res.*, **76:** 353, 1971.

Burlaga, L.F., "Micro-Scale Structures in the Interplanetary Medium," *Sol. Phys.*, **4:** 67, 1968.

Burlaga, L.F. and J.H. King, "Intense Interplanetary Magnetic Fields Observed by Geocentric Spacecraft During 1963–1975," *J. Geophys. Res.*, **84:** 6633, 1979.

Burlaga, L.F., E. Sittler, F. Mariani, and R. Schwenn, "Magnetic Loop Behind an Interplanetary Shock: Voyager, Helios and IMP8 Observations," *J. Geophys. Res.*, **86:** 6673, 1981.

Burlaga, L.F. and J.B. Turner, "Microscale Alfvén Waves in the Solar Wind at 1 AU," *J. Geophys. Res.*, **81:** 73, 1976.

Chapman, S., "Notes on the Solar Corona and Terrestrial Ionosphere," *Smithsonian Contrib. Astrophys.*, **2:** 1, 1957.

Coleman, P.J. Jr., "Turbulence, Viscosity and Dissipation in the Solar Wind Plasma," *Astrophys. J.*, **153:** 371, 1968.

Fairfield, D.G., "The Ordered Field of the Magnetosheath," *J. Geophys. Res.*, **72:** 5865, 1967.

Fairfield, D.H., "Magnetic Fields of the Magnetosheath," *Rev. Geophys. Space Phys.*, **14:** 117, 1976.

Feldman, W.C., J.R. Asbridge, S.J. Bame, and M.D. Montgomery, "Interplanetary Solar Wind Streams," *Rev. Geophys. Space Phys.*, **12:** 715, 1974.

Feldman, W.C., J.R. Asbridge, S.J. Bame, M.D. Montgomery and S.P. Gary, "Solar Wind Electrons," *J. Geophys. Res.*, **80:** 4181, 1975.

Feldman, W.C., J.R. Asbridge, S.J. Bame, and J.T. Gosling, "Plasma and Magnetic Fields from the Sun," in *The Solar Output and its Variations*, edited by Oran R. White, Colorado Associated University Press, Boulder, p. 351, 1977.

Feynman, J., "Solar Cycle and Long Term Changes in the Solar Wind," *Rev. Geophys. and Space Phys.*, **21:** 33, 1983.

Feynman, J. and N.U. Crooker, "The Solar Wind at the Turn of the Century," *Nature*, **275:** 626, 1978.

Feynman, J. and S.M. Silverman, "Auroral Changes During the 18th and 19th Centuries and Their Implications for the Solar Wind and the Long Term Variation of Sunspot Activity," *J. Geophys. Res.*, **85:** 2991, 1980.

Gary, S.P., "Microinstabilities Upstream of the Earth's Bow Shock: A Brief Review," *J. Geophys. Res.*, **86:** 4331, 1981.

Gosling, J.T., V.Z. Pizzo, and S.J. Bame, "Anomalously Low Proton Temperatures in the Solar Wind Following Interplanetary Shock Waves—Evidence for Magnetic Bottles?", *J. Geophys. Res.*, **78:** 2001, 1973.

Gosling, J.T., J.R. Asbridge, S.J. Bame, and W.C. Feldman, "Solar Wind Stream Interfaces," *J. Geophys. Res.*, **83:** 1401, 1978.

Gosling, J.T., J.R. Asbridge, S.J. Bame, W.C. Feldman and R.D. Zwicki, "Observations of Large Fluxes of He^+ in the Solar Wind Following an Interplanetary Shock," *J. Geophys. Res.*, **85:** 3431, 1980.

Greenstadt, E. W. and R.W. Fredericks, "Shock Systems in Collisionless Space Plasma," in *Solar System Plasma Physics III* edited by C.K. Kennel, L.J. Lanzerotti and E.N. Parker North Holland, Amsterdam, 1979.

Greenstadt, E.W., C.T. Russell and M. Hoppe, "Magnetic Field Orientation and Suprathermal Ion Streams in the Earth's Foreshock," *J. Geophys. Res.*, **85:** 3473, 1980.

Hedgecock, P.C., "Measurements of the Interplanetary Magnetic Field in Relation to the Modulation of Cosmic Rays," *Sol. Phys.*, **42:** 497, 1975.

Hirshberg, J., "The Transport of Flare Plasma from the Sun to the Earth," *Planet. Space Sci.*, **16**(3): 309–319, 1968.

Hirshberg, J., "The Interplanetary Field During the Rising Part of the Solar Cycle," *J. Geophys. Res.* **74:**5841, 1969.

Hirshberg, J., S.J. Bame, and D.E. Robbins, "Solar Flares and Solar Wind Helium Enrichments, July 1965–July 1967," *Sol. Phys.*, **23:** 467, 1972.

Hirshberg, J., Y Nakagawa, and R.E. Weilck, "Propagation of Sudden Disturbances Through a Non-Homogeneous Solar Wind." *J. Geophys. Res.*, **79:** 3726, 1974.

Hundhausen, A.J., *Coronal Expansion and Solar Wind*, Springer-Verlag, New York, 1972.

Neugebauer, M., "The Enhancement of Solar Wind Fluctuations at the Proton Gyroradius," *J. Geophys. Res.*, **80:** 998, 1975.

Neugebauer, M., "Observations of Solar-Wind Helium," *Fundamentals Cosmic Phys.*, **7:** 131, 1981.

Neugebauer, M. and C.W. Snyder, "Mariner 2 Observations of the Solar Wind," *J. Geophys. Res.*, **71:** 4469, 1966.

Neupert, W.M. and V. Pizzo, "Solar Coronal Holes as Sources of Recurrent Geomagnetic Disturbances," *J. Geophys. Res.*, **79:** 3701, 1974.

Ogilvie, K.W., M.A. Coplan, P. Bochsler, and J. Geiss "Abundance Ratios of $^3He^{++}/^4He^{++}$ in the Solar Wind," *J. Geophys. Res.*, **85:** 6021, 1980.

Parker, E.N., *Interplanetary Dynamical Process*, Interscience, New York, 1963.

Rickett, B.J., "Disturbances in the Solar Wind from IPS Measurements in August 1972," *Solar Phys.*, **43:** 237, 1975.

CHAPTER 3

Robbins, D.E., A.J. Hundhausen, and S.J. Bame, "Helium in the Solar Wind," *J. Geophys. Res.*, **75:** 1178, 1970.

Russell, C.T., "Geophysical Coordinate Transformations," *Cosmic Electrodynamics*, **2:** 184, 1971.

Schwenn, R., H. Rosenbauer, and K.H. Muhlhäuser, "Singly Ionized Helium in the Driven Gas of an Interplanetary Shock Wave," *Geophys. Res. Lett.*, **7:** 201, 1980.

Scudder, J. D. and S. Olbert, "A Theory of Local and Global Processes Which Affect Solar Wind Electrons 1., The Origin of Typical 1 AU Velocity Distribution Functions, Steady State Theory" *J. Geophys. Res.* **84:** 275, 1979.

Siscoe, G.L., "Evidence in the Auroral Record for Secular Solar Variability," *Rev. Geophys. Space Phys.*, **18:** 647, 1980.

Siscoe, G.L., L. Davis Jr., P.J. Coleman Jr., E.J. Smith, and D.E. Jones, "Power Spectra and Discontinuities of the Interplanetary Magnetic Field: Mariner 4," *J. Geophys. Res.*, **73,** 1968.

Siscoe, G.L., N.U. Crooker, and L. Christopher, "Solar Cycle Variation of the Interplanetary Magnetic Field," *Sol. Phys.* **56:**449, 1978.

Smith, E.J., "Interplanetary Magnetic Fields," *Rev. Geophys. Space Phys.*, **17:** 610, 1979.

Solodyna, C.V., J.W. Sari, and J.W. Belcher, "Plasma and Field Characteristics of Directional Discontinuities in the Interplanetary Medium," *J. Geophys. Res.*, **82:** 10, 1977.

Spreiter, J.R., A.L. Summers, and A.Y. Alksne, "Hydromagnetic Flow Around the Magnetosphere," *Planet. Space Sci.*, **14:** 223, 1966.

Spreiter, J.R. and A.Y. Alksne, "Plasma Flow Around the Magnetosphere," *Rev. Geophys. Space Phys.*, **7:** 11, 1969.

Suess, S.T., "The Solar Wind During the Maunder Minimum," *Planet. Space Sci.*, **27:** 1001, 1979.

Svalgaard, L., "Polar Cap Magnetic Variations and Their Relationship With the Interplanetary Magnetic Sector Structure.," *J. Geophys. Res.*, **78:** 2064, 1973.

Tsurutani, B.T. and E.J. Smith, "Interplanetary Discontinuities: Temporal Variations and the Radial Gradient from 1 to 8.5 AU," *J. Geopys. Res.*, **84:** 2773, 1979.

Tsurutani, B.T. and P. Rodriguez, "Upstream waves and particles: An Overview of ISEE Results," *J. Geophys. Res.*, **86:** 4319, 1981.

Wilcox, J.M. and N.F. Ness, "Quasi-Stationary Corotating Structure in the Interplanetary Medium," *J. Geophys. Res.*, **70:** 5793–5805, 1965.

Wilcox, J.M., L. Svalgaard, and P.C. Hedgecock, "Comparison of Inferred and Observed Interplanetary Magnetic Field, 1970–1972," *J. Geophys. Res.*, **80:** 3685, 1975.

Chapter 4

THE GEOMAGNETIC FIELD

D.J. Knecht and B.M. Shuman

The existence of the geomagnetic field has been recognized for a very long time. The usefulness of a magnet as a directional reference was probably known in China more than 1000 years ago and in Europe at least 800 years ago. As early as the 15th century, the earlier belief that a compass needle points true north was found to be incorrect and mariners and mapmakers took account of this. Recorded measurements of the magnetic declination (the deviation of the compass from true north) at various locations on the earth date back to the early 16th century, which also saw the discovery of the magnetic dip (the deviation of a compass needle from horizontal when unconstrained). Although experiments with magnets had been carried out since the 13th century, the concept that the earth itself is a magnet was not advanced until the end of the 16th century by Gilbert.

From these beginnings, geomagnetism as a branch of science was developed. It was first assumed that the magnetism of the earth was like that of a solid permanent magnet and was therefore expected to be constant in the absence of major geological changes, but this view was soon proved wrong; the secular variation (changes in the field over time intervals of years or centuries) was discovered in the 17th century. Transient variations of the field (geomagnetic disturbance) were observed during the 18th century, and geomagnetism was increasingly appreciated to be a dynamic phenomenon.

By the early 19th century, a large number of magnetic observatories had been established both in European countries and in the distant lands of their empires. Through coordinated measurements by many stations, the geographic dependence of some geomagnetic phenomena was discovered and the worldwide nature of major disturbances was established. The increasing volume and precision of accumulated data made discouragingly clear how complex were the phenomena being studied. Increasing international cooperation included investigations during the first International Polar Year (1882–1883). By this time, the correlation between the 11-year periodicities of sunspot occurrence and geomagnetic phenomena had been noted. Early in the 20th century the intimate connection between solar and geomagnetic phenomena was further established by the correlation of recurrent disturbances with the 27-day solar rotation and later by the correlation of magnetic storms with solar flares.

However, the most important connection, the fact that the geomagnetic field interacts with a continuous stream of solar plasma, was established only within the last twenty years. As a result of satellite investigations, these recent years have seen drastic revisions in many fundamental ideas concerning the configuration of current systems and the magnetic field above the surface of the earth.

But while satellite measurements have expanded our understanding of the space above us, the development of techniques for collecting and interpreting archaeological and geological data have led to some important discoveries about the earth below us: namely, that the continents have drifted thousands of miles from their original locations and that the entire geomagnetic field undergoes periodic reversals.

Today, geomagnetism encompasses two broad areas of theoretical study that are served by overlapping experimental data bases: the physics of the interior of the earth (which produces the steady and slowly varying field) and the physics of the magnetosphere and ionosphere (which produce the dynamic behavior of the field). The material in this chapter tends to neglect the physics in favor of a description of geomagnetic phenomena and their experimental observation, but this neglect is partly remedied in other chapters on the ionosphere and magnetosphere.

4.1 BASIC CONCEPTS

4.1.1 Units, Terminology, and Conventions

The geomagnetic field is characterized at any point by its direction and magnitude, which can be specified by two direction angles and the magnitude, the magnitude of three perpendicular components, or some other set of three independent parameters. Angles are commonly measured in degrees, minutes, and seconds. Prior to widespread adoption of mks units, the magnitude was usually given in units of oersted (magnetic intensity) or gauss (magnetic induction). Since the field is less than one oersted everywhere, the unit gamma was most useful; one gamma equals 10^{-5} oersted or 10^{-5} gauss and was used interchangeably for intensity and induction. Since the mks unit for field strength is very large

(one tesla = 10^4 gauss), the nanotesla (nT), which very conveniently equals one gamma, is now almost universally used.

Some of the angles and components commonly employed are shown in Figure 4-1. Standard terminology is as follows. The vector geomagnetic field is the vector **F**. Its magnitude F is called the total intensity or the total field.

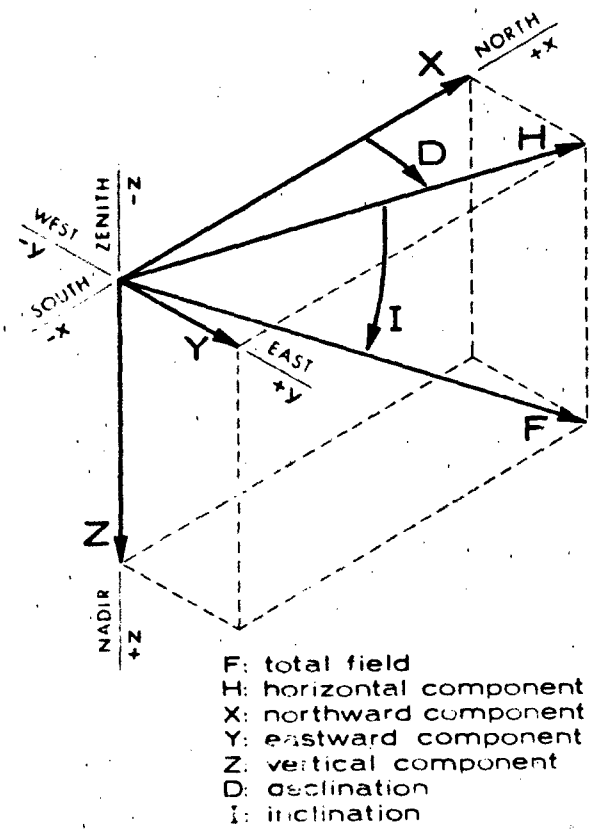


Figure 4-1. Definition and sign convention for the magnetic elements.

The magnitude H of the horizontal component vector **H** is called the horizontal intensity; the magnitude Z of the vertical component vector **Z** is called the vertical intensity. The northward, eastward, and downward components of the field are designated by the magnitudes X, Y, and Z, respectively, the Cartesian components of the field. The magnitude D of the angle between X and H is called the declination, the magnetic variation, or the variation of the compass. The magnitude I of the angle between **H** or **F** is called the inclination or the dip. The quantities F, H, X, Y, Z, D and I are called magnetic elements. The sets of elements used most commonly to specify the field are (H, D, Z); (F, I, D); and (X, Y, Z). The sign convention for each parameter is shown in the figure, all vectors and angles being positive as drawn.

4.1.2 Coordinate Systems

A number of coordinate systems are employed in the description of geomagnetic phenomena. Five of the most useful are the geographic, geomagnetic, geocentric solar-ecliptic, geocentric solar-magnetospheric, and solar-magnetic systems. They are shown in Figure 4-2 and are defined as follows.

A geographic coordinate system is one that is fixed with respect to the rotating earth and aligned with the axis of rotation. Most commonly used are the spherical polar coordinates r, θ, and ϕ, where r is geocentric distance, θ is colatitude (measured from the north geographic pole), and

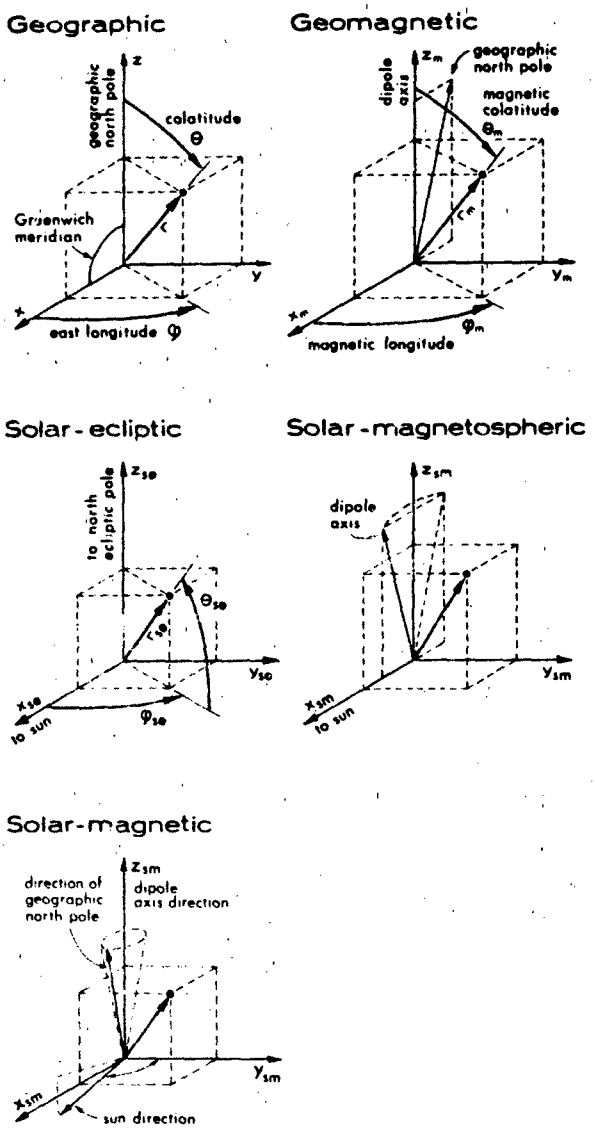


Figure 4-2. Several coordinate systems used in geomagnetism.

ϕ is east longitude (measured from the Greenwich meridian), with the earth assumed spherical. Sometimes altitude (above the surface of the earth) is specified in place of r, north or south latitude is specified in place of θ, and west longitude is specified in place of ϕ for values greater than 180 degrees. (Geodetic coordinates, which are defined relative to the nonspherical earth ellipsoid, must be used with care.)

The geomagnetic coordinate system is also a spherical polar system fixed relative to the earth, but the polar axis is the axis inclined 11.5 degrees to the axis of rotation, intersecting the earth surface at the point 78.5°N, 291.0°E which defines the geomagnetic north pole. This was at one time the axis of the best centered-dipole approximation to the field; current spherical-harmonic models of the main field would place the pole approximately 0.25 degrees farther north and 1.6 degrees farther west. Geomagnetic coordinates r, θ_m, and ϕ_m (and geomagnetic latitude and longitude) are defined by analogy with geographic coordinates, with ϕ_m (or geomagnetic longitude) measured from the American half of the great circle which passes through both the geographic and geomagnetic poles (that is, the zero-degree geomagnetic meridian coincides with 291.0°E geographic longitude over most of its length).

The corrected geomagnetic coordinate system is a refinement (of the geomagnetic coordinate system) that has proved useful in considering phenomena that involve propagation along lines of force of the geomagnetic field [Hakura, 1965]. It effectively provides a more accurate field-line connection from a point on (or near) the earth surface either to the equatorial plane or to its conjugate point than would be afforded by any simple dipole approximation of the geomagnetic field. Figure 4-3 shows how the corrected geomagnetic coordinates (latitude Φ_c and longitude Λ_c) are obtained for a point Q, at the earth surface, whose geomagnetic coordinates are Φ and Λ. Starting at Q, an "actual" field line (that is, one computed from an accurate higher-order model fitted to the actual geomagnetic field, as described in Section 4.6.1) is traced out to intersect the magnetic equatorial plane at the point A, which has polar coordinates L_c and Λ_c. From here a simple dipole field line is projected back toward the earth to a "landing point" Q_c. The corrected coordinates for point Q are the uncorrected coordinates of point Q_c (Φ_c and Λ_c). The value of Φ_c may be found from the dipole equation, $L_c = R \sec^2 \Phi_c$, where R is the radius of the earth. Assigning the corrected coordinates to the point Q permits an accurate description of phenomena in the actual field as if it were a simple dipole field; for example, points having the same coordinates in the northern and southern hemispheres will be actual conjugate points. A revised tabulation of corrected geomagnetic coordinates, using terms through order 7 in the spherical-harmonic model, has been published by Gustafsson [1970] for every 5° of geographic longitude and every 2° of geographic latitude.

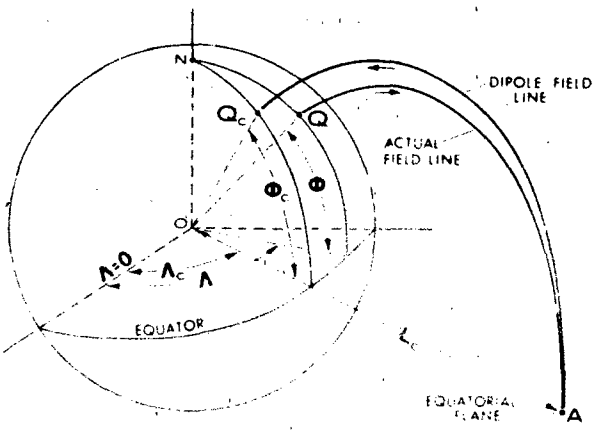


Figure 4-3. Method of finding corrected geomagnetic coordinates.

The geocentric solar-ecliptic (GSE) coordinate system is a right-handed Cartesian system with coordinates X_{se}, Y_{se}, and Z_{se}, and the center of the earth as origin. The positive X_{se} axis is directed toward the sun. The Z_{se} axis is directed toward the northern ecliptic pole, so both the X_{se} and Y_{se} axes lie in the ecliptic plane. This system therefore rotates slowly in space with the orbital period of the earth. In this system, field vectors are often resolved into two components, one lying in and the other perpendicular to the ecliptic plane; the direction of the former is specified by the angle ϕ between it and the X_{se} axis (positive counterclockwise when viewed from the northern pole). The direction of the total field is specified by ϕ and θ, where θ is the angle between the vector and the ecliptic plane (positive northward). This system is particularly useful for referencing data from interplanetary space, such as measurements of the undisturbed solar wind and the interplanetary magnetic field.

The geocentric solar magnetospheric (GSM) coordinate system is also a right-handed Cartesian system, with coordinates X_{sm}, Y_{sm}, and Z_{sm}, and its origin at the center of the earth. The positive X_{sm} axis is also directed toward the sun. It differs from the solar-ecliptic system in that the Z_{sm} axis lies in the plane containing both the X_{sm} axis and the geomagnetic dipole axis defined above. The system therefore not only rotates with the orbital period of the earth but also rocks back and forth through 23 degrees (a rotation about the X_{sm} axis) with a period of one day. This system is particularly useful for referencing data from distant regions of the magnetosphere, since time-dependent features that result from the conical motion of the dipole axis are, to a large extent, eliminated; that is, to a first approximation, the entire magnetosphere, in its main features, may be expected to rock back and forth in this way.

A related frame is the solar-magnetic (SM) coordinate system. In this system the Z axis is directed to the north

CHAPTER 4

along the geomagnetic dipole axis, and the Y axis is perpendicular to the earth-sun line toward the dusk side. The X axis is not always directed toward the sun in this system, but rocks back and forth through 11.5 degrees about the earth-sun line. This system differs from the GSM essentially by a rotation about the Y_{sm} axis.

In addition to strictly spatial coordinate systems, several so-called magnetic systems have been found useful in studying the motion of charged particles trapped in the magnetic field: these coordinates generally locate particles by reference to surfaces on which some magnetic parameter is constant, and since most particles are strongly controlled magnetically a great simplification of the data often results. Most widely used is the B-L coordinate system of McIlwain. As shown in Figure 4-4, surfaces of constant B (magnetic field intensity) are concentric, roughly ellipsoidal shells encircling the earth, while surfaces of constant L (a magnetic parameter) approximate the concentric shells generated by dipole field lines rotating with the earth. The mathematical definition of L arises from the equations of motion of particles in the field; to some degree of approximation, particles move to conserve three adiabatic invariants to which B and L are simply related. Since these coordinates are more useful to the study of trapped particles than to the study of the field itself, the reader is referred to standard texts on trapped-particle physics for a complete discussion.

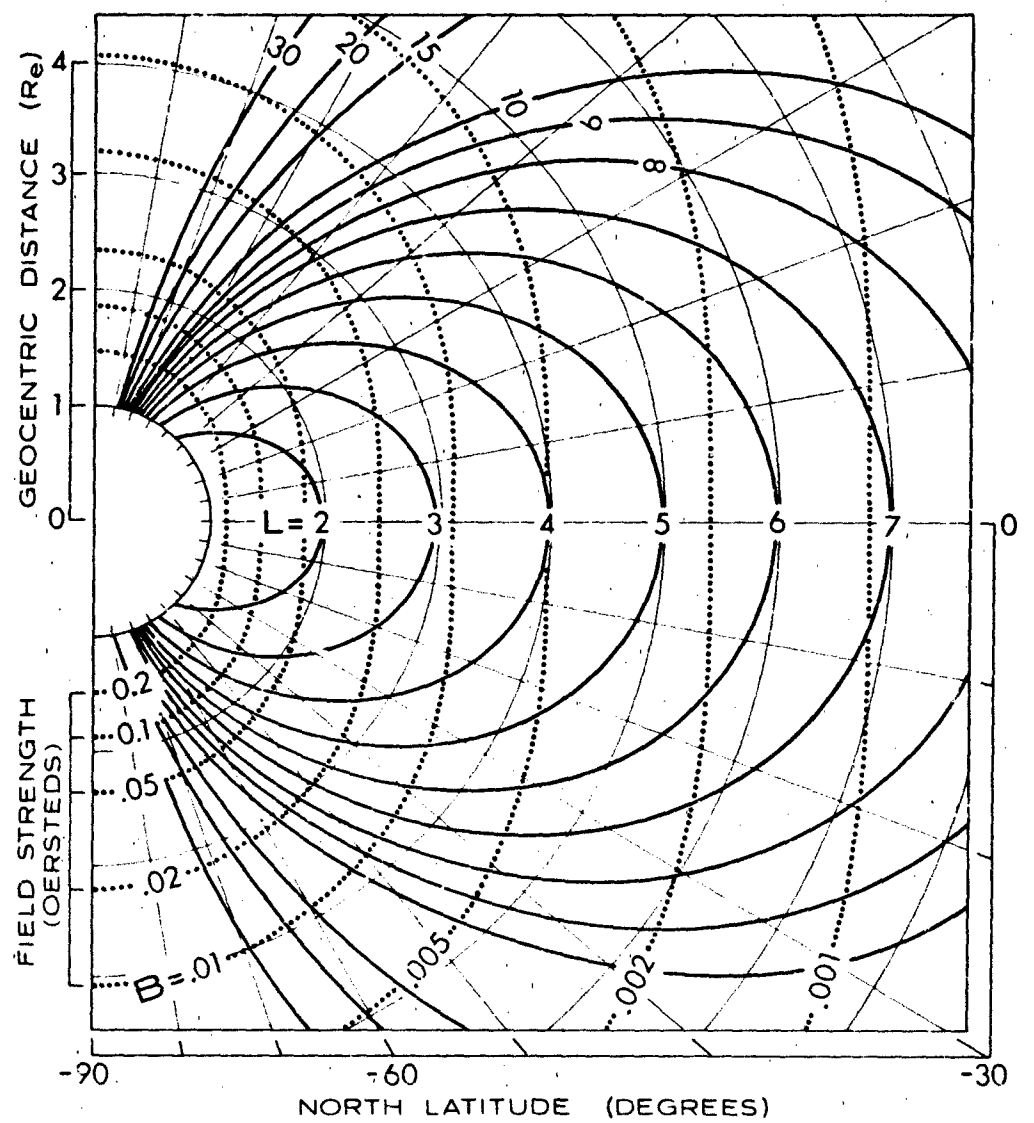


Figure 4-4. The B L coordinate system. The curves shown here are the intersection of a magnetic meridian plane with surfaces of constant B and constant L. (The difference between the actual field and a dipole field cannot be seen in a figure of this scale.)

4.1.3 Sources of the Geomagnetic Field

In considering a physical description of the field, a useful point of view to adopt is that of energy balance. A static field represents an energy density $B^2/8\pi$, and any change in the field implies a transfer of energy to or from the field. Understanding the field therefore implies identifying the energy sources and the causative physical mechanisms through which this energy generates (or is generated by) the field. Except in the case of permanent residual magnetism, a magnetic field is generated only by the macroscopic motion of electric charge, so the final step in any physical process affecting the field will involve electric currents, though the energy driving the currents may be drawn from various sources. At present, the terrestrial and extraterrestrial sources known to contribute appreciably to the geomagnetic field are the following:

1. Core motion. Convection motion of the conducting fluid core of the earth constitutes a self-exciting dynamo.
2. Crustal magnetization. Residual permanent magnetism exists in the crust of the earth.
3. Solar electromagnetic radiation. Atmospheric winds (produced by solar heating) move charged particles (produced by solar ionizing radiation); this constitutes an ionospheric current which generates a field.
4. Gravitation. The gravitational field of the sun and moon produces a tidal motion of air masses that generates a field in the same way as does the air motion from solar heating.
5. Solar corpuscular radiation and interplanetary field. A number of field contributions arise directly or indirectly from the interaction of the solar wind and its imbedded magnetic field with the main field of the earth. Some important effects are the compression of the main field by external plasma pressure, the intrusion of solar plasma into the main field, the heating of plasma already within the field, and the merging of magnetospheric and interplanetary fields.

There are a number of other obvious possible sources that, in fact, do not contribute appreciably; examples are the mantle of the earth and energetic cosmic rays.

4.1.4 The Steady Interior Field

A geometric-temporal description of the field is constructed from measurements made by observatories, stations, ships, rockets, and satellites, all of which are restricted in geometric coverage (geographic or spatial location) and temporal coverage (time period and frequency response). Observed phenomena tend to be classified accordingly. The traditional classification by frequency is very useful and is retained here.

The steady (nonvarying or dc) component of the field may be considered first. Although it is true that the entire field has been varying drastically over geological time scales, that portion which varies with periodicities greater than about a year is customarily considered to be the steady field, while the remainder is considered the variation field.

Most of the steady field arises from internal terrestrial sources (that is, below the surface of the earth, but excluding currents induced in the earth by external current systems) and is known as the main field. This field results primarily from convective motion of the core and is approximately of dipole configuration, having a strength at the surface of the earth of several times 10^4 nT. The dipole is centered close to the center of the earth, with its axis inclined about 11.5 degrees to the axis of rotation. About 10% of the main field, often termed the residual field, is nondipolar; it consists of both large-scale anomalies (up to thousands of kilometers), believed to be generated by eddy currents in the fluid core, and small-scale irregularities (down to 10 km) arising from residual magnetism in the crust. Changes in the main field (the so-called secular variation) are slow, with time constants of tens to thousands of years.

If the earth were in a perfect vacuum, its dipole field would extend outward without limit, merging smoothly with the fields of the sun and other planets in a simple additive fashion, the field strength declining inversely with the third power of geocentric distance. However, interplanetary space is not a vacuum but is filled with the ionized corona of the sun (the solar wind), which flows continuously outward past the planets. On a quiet day, near the earth, this plasma typically has a density of a few ions/cm^3 and a velocity of about 400 km/s. An important feature of the plasma is its high electrical conductivity. One result of applicable theory is that the magnetic field will be "frozen into" such a plasma; that is, the ions, electrons, and magnetic field will move together as a compressible fluid medium. When such a moving fluid encounters a stationary entity with which it can interact, such as the geomagnetic field, one or the other will be deflected, swept away, or otherwise modified by the collision. The total pressure of the solar wind is the sum of the pressure exerted by the momentum of the particles and the Maxwell pressure $B^2/8\pi$ of the frozen-in field. The geomagnetic field also contains highly conductive plasma, and this medium similarly sustains a pressure equal to the sum of the ambient-plasma and Maxwell pressures. When the pressures of the interplanetary medium and the geomagnetic field are compared, it is clear that at great distances the geomagnetic field will be swept away by the solar wind and that close to the earth the solar wind will be excluded by the field. At intermediate distances there must exist a region of interaction where the pressures are comparable and where rather complicated features can be expected. In the last decade, satellite experiments and theoretical development have discovered and clarified the principal features of this interaction. Some of them are illustrated in Figure 4-5. (The 23° tilt of the rotational axis and the 11.5° tilt of the dipole

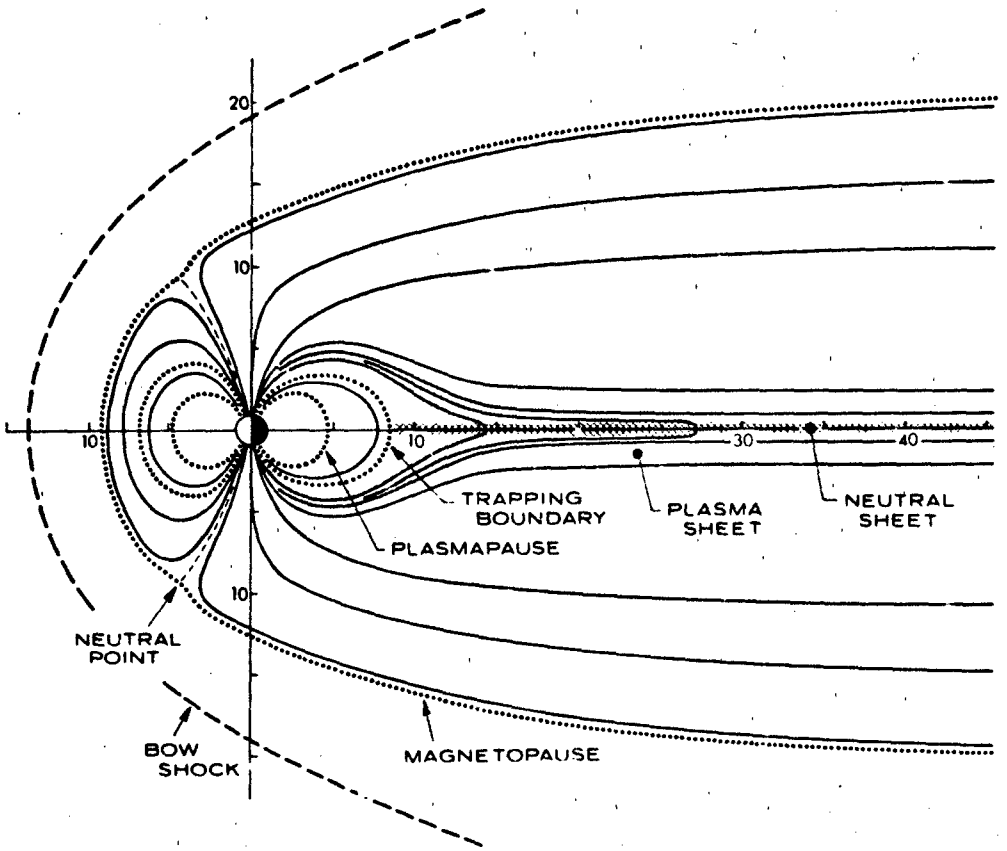


Figure 4-5. The general configuration of the magnetosphere, shown in a noon-meridian section.

axis have been neglected for simplicity.) The field of the earth extends to a geocentric distance of 10 R_E toward the sun (R_E being a unit of length equal to the radius of the earth) at which distance it terminates abruptly in a thin layer known as the magnetopause. The region interior to this layer is known as the magnetosphere. The region exterior to it contains the solar wind, which is "piled up", that is, compressed, deflected, heated, and made turbulent by the collision. These effects propagate some distance upstream, with the result that the wind is slowed over a distance of a few R_E. Since the velocity of the undisturbed wind is "supersonic", there exists a surface at which the velocity is just "sonic" and a stationary shock front, the bow shock, is created. The magnetopause is typically at about 10 R_E and the bow shock at about 14 R_E on the sun-earth line. The region between these is called the magnetosheath. Field lines from the high-latitude (polar cap) regions are swept back toward the night side and form a long geomagnetic tail. Although it might be expected that solar wind pressure perpendicular to the direction of its bulk flow should close the magnetosphere within a few tens of R_E behind the earth, that is not the case; the combined pressure of field and plasma within the tail is sufficient to prevent closing, and the tail has been observed at a distance of more than 1000 R_E behind the earth. It might also be expected that within the tail the north-polar field lines would be smoothly connected across the equatorial plane to the corresponding south-polar field lines, but this also does not happen; these field lines are drawn out into the tail, directed toward the earth above the plane and away from the earth below it. Thus, beyond a geocentric distance of about 10 R_E in the tail, the equatorial plane (neglecting tilts) is a sort of neutral sheet, across which there is a fairly abrupt field reversal, and the component perpendicular to the neutral sheet is very small.

A surface of discontinuity in the magnetic field implies the existence of a current flow in the surface, and the current pattern can be inferred from the field. On the sunward magnetopause the flow is characterized by an eastward current sheet (dawn-to-dusk) across the nose (subsolar point) of the magnetosphere. In the neutral sheet the flow is westward across the tail (also dawn-to-dusk), with return loops on the tailward part of the magnetopause. These currents are shown schematically in Figure 4-6. (The composite magnetopause current is not shown.)

The current system of the magnetopause acts to cancel the dipole field outside and enhance the field inside the

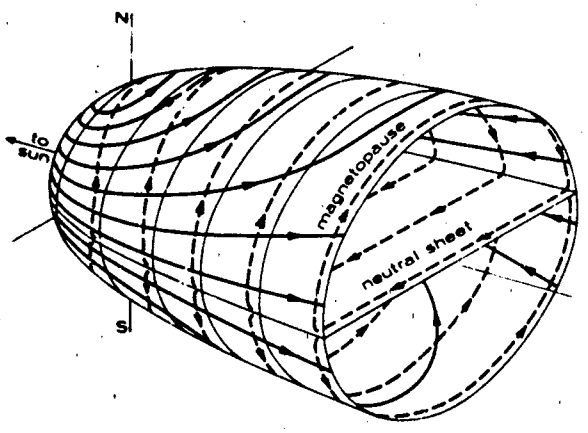


Figure 4-6. Flow patterns of the two principal current systems which determine the configuration of the magnetosphere [adapted after Axford, 1965]

magnetopause. This is equivalent to a compression of the geomagnetic field by the cavity to which it is confined. Because the cavity almost totally surrounds the earth, the field is compressed on all sides, but since the tail is open and the highest pressure is on the nose of the cavity, the compression is somewhat less on the night side. The compression results in an average increase of the equatorial surface field of about 0.1% (about 30 nT); just inside the magnetopause the increase is 100% (which is about 30 nT when the magnetopause is at 10 R_E but about 60 nT if it has been pushed in to 8 R_E). About a third of the surface increase results from diamagnetism in the solid earth.

The steady field then consists basically of the main field of the earth compressed by the cavity to which it is confined. In addition, most of time-varying field contributions discussed below are also (like the magnetopause current) likely to have an average dc value which may be thought of as part of the steady field. For example, convection of the outer magnetosphere and the flowing of a ring current are processes which continue even on the quietest days.

4.1.5 Quiet Variation External Fields

The earth with its core, atmosphere, and main field rotates in the interplanetary environment and moves along its orbit so that any point stationary in geographic coordinates experiences periodic variations in gravity force, solar illumination, and compression or other modification by solar wind effects. The field contributions that result from these motions vary diurnally and seasonally. Field contributions that vary this slowly and regularly and do not result from disturbances in the interplanetary environment are known as quiet variation fields. The analysis of experimental data to determine the quiet variation fields is of course difficult in the presence of magnetospheric disturbance, and criteria for separating field contributions have been somewhat arbitrary and subject to personal judgment. These fields were originally defined on the basis of data taken during a few of the quietest days per month. With better understanding of magnetospheric disturbance, improved measurements, and an awareness that quietness is only relative (any day being only more or less disturbed), it became more common to consider them in an idealized sense as being those variations that would exist if the earth were subjected only to an absolutely quiet external environment. More recently, evidence of a direct connection between quiet and disturbance variations has made this viewpoint less useful.

Quiet variation fields include several contributions. The solar quiet (Sq) variation field, which has a peak-to-peak amplitude of several tens of nanoteslas at most surface locations, is caused mostly by the generation of ionospheric currents by solar electromagnetic radiation. The L (lunar daily) variation field, which typically has an amplitude of a few nanoteslas at the surface, results from the generation of ionospheric currents by luni-solar atmospheric tides. Another contribution of a few nanoteslas at the surface results from the confinement of the main field by the solar wind, the compression being stronger on the day side than on the night side. Quiet variation fields are discussed in Section 4.4, which for both historical and practical reasons retains a ground-based perspective.

4.1.6 Disturbance External Fields

Variations in the geomagnetic field that do not have a simple periodicity and appear to result from changes in the interplanetary environment are called disturbed variation fields or geomagnetic disturbance and are denoted by D. The D field is that remaining after the steady and the quiet variation fields have been subtracted from the total. Large disturbances of relatively long duration, the behavior of which suggests some magnetospheric events as the cause, are termed geomagnetic storms. Except for some fluctuations attributable to irregular motion of the upper atmosphere, the sun is responsible for all disturbance effects recognized at present, and with only two exceptions, it is the solar wind with the frozen-in solar magnetic field that transmits the disturbance to the vicinity of the earth. The two exceptions are disturbances in which ionospheric conductivity is enhanced as the result of solar flares: polar cap absorption events (PCA), which result from low-energy protons from the flare, and solar-flare effects (SFE), which result from x-ray emissions from the flare.

Historically, the disturbance field has been studied by ground observations, with the hope of separating the observed surface field into components that could be explained in terms of current systems above the earth. A useful distinction has been the separation of the component that depends only on universal time (UT) from that which depends

on local time (LT); the former, usually called the storm-time variation and denoted by Dst, is, by definition, symmetric about the polar axis, while the latter, called the disturbance-daily variation and denoted by DS (or Ds), is asymmetric. Then D = Dst + DS. The component Dst was attributed to a ring current encircling the earth a few R_E above the equator, while DS was attributed to ionospheric currents generated by auroral particles precipitated from the ring current. Although better knowledge of the magnetosphere has made clear that ionospheric and magnetospheric current systems are intimately related, this separation is still useful. Other separations have been made or proposed, usually relating to a theoretical model of some postulated physical process. Many are currently in use in the literature but are likely to be revised as understanding improves.

Except as noted above, geomagnetic disturbance results from the interaction of the solar wind with the geomagnetic field. While some minimum level of disturbance may be expected to result from turbulence generated by instabilities in the flow of plasma around the magnetosphere even if solar wind properties were absolutely constant, most disturbance phenomena having characteristic times of minutes to days and observable in ground-based measurements of the field result from variations (often abrupt) in one or more of the solar wind parameters (for example, the density or velocity of the plasma or the direction or intensity of the interplanetary field). The largest disturbances of the magnetosphere are called magnetospheric storms and the corresponding disturbances of the geomagnetic field are called geomagnetic storms. While phenomena vary greatly from storm to storm it is possible to describe a typical or "classic" magnetic storm (see Section 4.5.1). Many other complex dynamic processes in the magnetosphere are manifested in magnetic-field disturbance; some of these are discussed in Chapter 8.

The dynamic behavior of the magnetosphere also includes oscillations, especially in accompaniment to slower magnetic disturbance, both because it is an elastic entity which can resonate and because a number of its dynamic processes generate oscillatory currents. Periodic and aperiodic field fluctuations with frequencies covering nearly eight decades (10^{-4} to 10^{4} Hz) are observed. In the lower frequency range (ULF up to 5 Hz) they are called geomagnetic pulsations. These are discussed in Section 4.5.2. Higher frequencies (ELF and VLF) are associated with the dynamics of ionospheric and magnetospheric plasmas.

4.2 MEASUREMENTS OF THE GEOMAGNETIC FIELD

Geomagnetic phenomena are studied experimentally through data obtained by ground stations, ships, aircraft, and space vehicles. This section discusses the instruments used for geomagnetic measurements and reviews the principal sources of such measurements.

4.2.1 Instrumentation

Instruments used over the past several hundred years to measure the intensity and direction of the magnetic field have been few in number and simple in principle, but during the past century they have been made very reliable and fairly precise. Although greater precision and sensitivity were needed earlier, major developments of new instruments came only in the past 35 years, partly because the older instruments were not adaptable for use on rockets and satellites. The principal instruments currently in use may be listed as follows.

Ground-based instruments exploit several physical principles. Several of the older instruments are based on the alignment or oscillation of a permanently magnetized needle in the field; these include the compass, dip circle, and magnetic theodolite, which measure D, I, and H, respectively, the three elements usually measured at observatories to determine the field. Several others rely on the induction of a voltage in a coil of wire. The coil may be rotated in the field as in the dip inductor, or may be fixed as in a large induction-coil magnetometer used to measure geomagnetic pulsations. Two magnetometers are based on the cancellation of a component of the geomagnetic field by the known field of an electromagnet; these are the H-magnetometer of Schuster and Smith and the Z-magnetometer of Dye. Of the newer instruments, several are based on atomic-resonance techniques; these are the proton precession (and proton vector), rubidium-vapor, and helium magnetometers. Another widely used instrument exploits the saturation characteristics of a ferromagnetic core; this is the fluxgate or saturable-core magnetometer. Most recently, a number of extremely sensitive instruments have been developed which utilize the quantum-mechanical behavior of Josephson junctions in a superconducting loop; these are known as SQUID magnetometers (for "superconducting quantum interference devices").

All of these magnetometers are in use for ground measurements. Satellite and rocket measurements rely mainly on the rubidium-vapor, induction-coil (often called searchcoil), and fluxgate magnetometers, which inherently have small size, low weight, modest power requirements, and an easily telemetered output. A brief description of several of the most important of these instruments follows.

Fixed induction coils of various sizes are used to measure rapid fluctuations in the field. To measure the vertical component, horizontal coils with diameters of nearly 10 km are laid out on the ground; for the other components, coils a few meters in diameter, but with many turns, are used. Also used for this purpose are much smaller coils which are wound around laminated mu-metal cores which concentrate magnetic flux for increased sensitivity. Since the quantity measured is the time rate of change of the field, the sensitivity is inherently proportional to the frequency of the fluctuation. A metal-core coil of 30 000 turns, having a diameter of 7 cm and a length of 2 m, can detect variations

of 0.001 nT at one Hz. A typical spacecraft searchcoil having a diameter of 2 cm and a length of 30 cm has a sensitivity 1000 times less.

The first and best developed of the atomic- or nuclear-resonance instruments is the proton precession magnetometer. The physical principle on which it depends follows. Individual protons in a hydrogenous material placed in a magnetic field have both a magnetic moment and an angular momentum, which coincide in direction; the field exerts on the proton a torque tending to align its moment with the field, but the existence of the angular momentum causes the common vector to precess about the field direction. Normally the precessing vectors are random in phase and produce no coherent signal, but if they are started with a common phase by suddenly releasing them after polarization by a strong field perpendicular to the field to be measured, they precess for some time in unison, producing at the precession frequency a signal which can be detected by a pickup coil surrounding the material. The precession (Larmor) frequency is directly proportional to the field, the constant of proportionality being $1/2\pi$ times the proton gyromagnetic ratio. This physical constant, known to an accuracy of better than one part in 10^5, has the value $26751.9 \times 10^5\ T^{-1}\ s^{-1}$ so the frequency for a field of 30 000 nT is 1.2773 kHz. In a typical instrument, the hydrogenous material is a fraction of a liter of water, alcohol, or n-heptane around which is wound a single coil, used first to produce the polarizing field of about 0.01 T and subsequently to detect the precession signal. After the sample is polarized, the coherence of the precession persists for a few seconds before being destroyed by thermal agitation. Several precautions and corrections are required, but the instrument is basically simple and reliable. Absolute measurements of the field can be made with an uncertainty as low as 0.1 nT. The sensitivity of the instrument can be increased to 0.01 nT by adding a microprocessor to process the precession signal. Versions for use in observatories, aircraft, ships, and rockets have been developed and a continuously self-oscillating version is under development.

The proton vector magnetometer combines the proton-precession magnetometer with two sets of Helmholtz coils arranged to null the H and Z components. To measure Z, **H** is first annulled by producing $-\mathbf{H}$ in the H coils. The null cannot be detected directly but is produced by using just half the current required to generate $-2\mathbf{H}$; the latter condition can be detected since the total intensity is then exactly the same as that with zero current in the coil. The current required to annul Z is then measured. First-order instrument errors, of which leveling alignment is most critical, can be corrected by appropriate checks with reversed coils. To keep the field gradient at the sensor low enough with moderate coil dimensions, a four-element Fenselau or Braunbeck coil array may replace the simpler Helmholtz coil. Since this instrument uses the proton precession magnetometer simply as a null detector, the precision of the measurement depends on the accuracy with which the generated fields are known; an uncertainty as low as about 0.3 nT is possible. This instrument is used to measure H and Z at many observatories.

In the last 25 years, a newer resonance instrument has been widely used. This is the alkali-vapor magnetometer, which relies on the Zeeman effect and the phenomenon of optical pumping. Any alkali vapor is suitable, but the rubidium isotopes 85 and 87 have been most used. The energy-level diagram for Rb-87, showing the Zeeman splitting in a magnetic field, is shown in Figure 4-7. When light having a wavelength of about 0.79476 μm is passed through a transparent cell filled with the vapor, resonance absorption and re-emission occurs involving transitions between the various Zeeman sublevels of the ground and first excited states. If the light is circularly polarized, the absorption transitions must have $\Delta m = +1$, so no transitions from the groundstate sublevel with $m = +2$ can occur. Eventually, all electrons are trapped in this substate and no further absorption can take place; the vapor becomes magnetically polarized and transparent. This process is called optical pumping. The polarization can be destroyed by impressing in a direction perpendicular to the ambient field a weak magnetic field oscillating at a frequency (the Larmor frequency) corresponding exactly to the Zeeman splitting (6.99 Hz/nT). Forbidden transitions between the various m-sublevels of the groundstate are induced, electrons trapped in the sublevel with $m = +2$ are redistributed to other sublevels, resonance absorption of the light is again possible, and the vapor is no longer transparent. The ambient magnetic field is determined by measuring the Larmor frequency (that is, the oscillating-field frequency that produces maximum light absorption). The simplest magnetometer consists of the vapor absorption cell surrounded by a coil to produce the Larmor-frequency field, an rf-excited vapor lamp with a filter to absorb all but the 0.79476 μm line, a circular polarizer between the lamp and the cell, and a photodetector to measure the intensity of transmitted light. The frequency of the impressed field may be adjusted manually for minimum transmission of light through the cell; there is a 20% change in absorption between the pumped and unpumped conditions. In more refined instruments, several corrections and improvements are incorporated and they are usually self-oscillating; that is, both the light intensity and the impressed field oscillate with the Larmor frequency, which is established using a feedback signal from the photodetector. The absolute accuracy of alkali-vapor magnetometers is limited by the inherent line width of the resonance (several nT for rubidium) and a further splitting of the Zeeman levels by second-order effects in the coupling of moments; the uncertainty in weak field regions such as the distant magnetosphere is negligible, but in strong-field regions, such as near the surface of the earth, it is seldom less than about 2 nT.

The helium magnetometer also depends on optical pumping. Its operation is similar to that of the alkali-vapor instrument, but since the groundstate of helium has zero

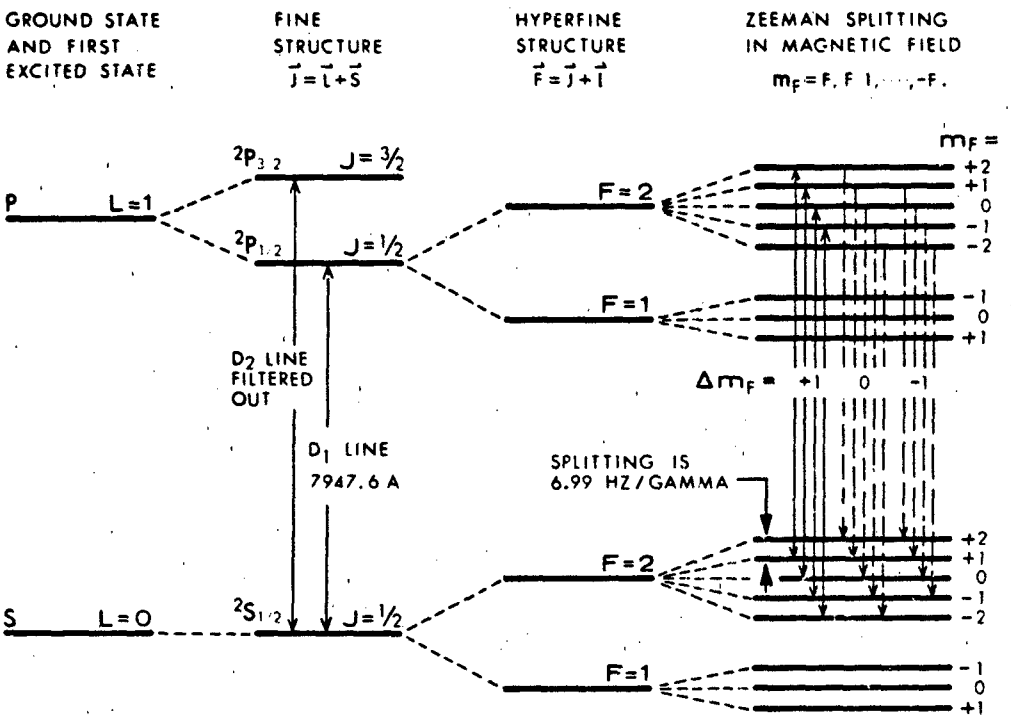


Figure 4-7. Energy-level diagram for Rubidium-87.

magnetic moment, electrons are trapped in a substate of the metastable 2^3P_1 state instead, being excited to the metastable state by an rf electric field. The inherent sensitivity is higher because of a greater Zeeman splitting (28 Hz/nT compared with 7 Hz/nT for Rb-87), but the line width of the resonance is much broader (about 100 nT).

The fluxgate magnetometer has been in use for nearly half a century, but its current usefulness results from extensive development in recent years. At present, it is probably the most widely used instrument in both ground and space measurements for geomagnetic and magnetospheric studies. In its simplest form, it consists of a highly permeable ferromagnetic core on which primary and secondary coils are wound. The primary winding is driven with a sinusoidal current that has an amplitude sufficient to drive the core material into saturation twice each cycle, thus changing the permeability of the material at a frequency twice that of the primary current. The flux in the core has two sources, the exciting field and the ambient external field. The former produces a large component at the driving frequency and, because of the changing permeability, other components at odd-harmonic frequencies. The latter produces a large component at the second-harmonic frequency and smaller components at higher even-harmonic frequencies, all due to the changing permeability; that is, the steady ambient field, which would otherwise produce no signal, is "gated" into the secondary winding by the changing permeability of the core. Since the exciting field produces only odd harmonics, the second harmonic is a measure of the ambient field. Its amplitude is proportional to the magnitude of the field component parallel to the core, and its phase indicates the sign. Several schemes may be used to eliminate the drive signal in the secondary winding. The primary and secondary coils can be wound on axes which are perpendicular, or if the core material is separated longitudinally into two halves, the primary winding can encircle the two halves in opposite senses. In either case the net excitation flux through the secondary can be made zero, and to the degree that this is achieved, only the desired even-harmonic signal is detected. A magnetometer of the latter type is shown schematically in Figure 4-8. Since a single sensor measures only one component, most applications use a three-sensor orthogonal array. It is also common to combine the instrument with coils, either external or wound in the sensor (like the calibration winding shown), to annul all or most of the ambient field. It has also been combined with a servomechanism to orient a three-sensor array such that two sensors read zero, the third measures the total field, and the servomechanism measures the field direction. This so-called orienting fluxgate was used extensively in world-wide magnetic mapping by aircraft. Since the fluxgate is an analog electronic instrument, a number of characteristics of the circuitry and core material limit its absolute accuracy; some instruments have an uncertainty of less than 1 nT. The sensitivity for both ground and space measurements is better than 0.1 nT. Since the excitation frequency may be as high

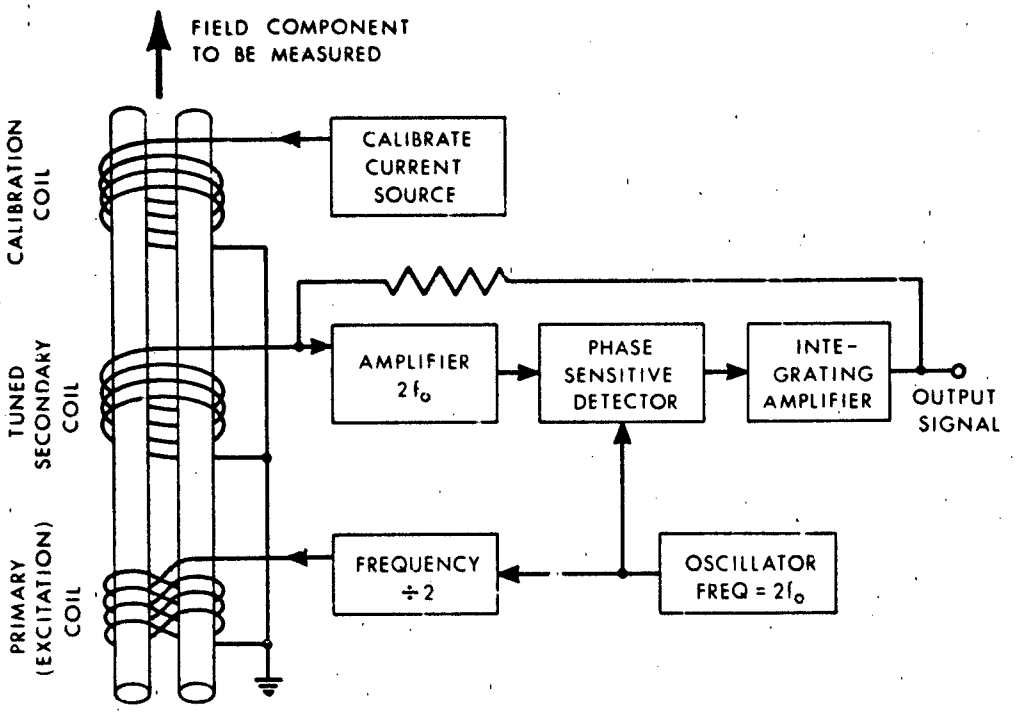


Figure 4-8. Schematic diagram of a typical fluxgate magnetometer.

as 10 kHz, field fluctuations with frequencies as high as 100 Hz or more can be measured.

The most recent and most sensitive instrument is the SQUID magnetometer, which operates in liquid helium at about 4 K. One version uses two Josephson junctions in a superconducting loop as shown in Figure 4-9. Each junction consists of two superconductors (the two halves of the loop) separated by a thin insulator across which a current flows because of the quantum-mechanical tunneling of electrons through the the insulator. Quantum mechanics requires that all electron pairs in the superconductor be in the same state, and therefore that a single wave function describe the entire loop. One result is that in a loop without junctions the magnetic flux enclosed by the loop cannot change (being "trapped") and is quantized in units of $h/2q_e$ (having the value 2.07×10^{-15} Wb), where h is Planck's constant and q_e is the charge of the electron. Adding the junctions to the loop results in a partial breakdown of the trapping; a non-integral multiple of the flux quantum can exist in the loop, but the behavior of the loop in attempting to maintain the trapping can be observed through its effect on the Josephson dc current flowing from A to B in the figure, which varies as shown with the applied magnetic field. Another version requiring only one junction in the loop observes this behavior by its effect on the impedance of a coil arranged to couple an externally generated rf magnetic field into the loop. In either case, the measured quantity exhibits an interference pattern, that is, a series of maxima and minima corresponding to the existence of integral and half-integral multiples of the flux quantum in the loop. The SQUID is not an absolute instrument; changes in the ambient magnetic field are measured by counting the number of maxima which occur in an output signal as the field changes. In practice

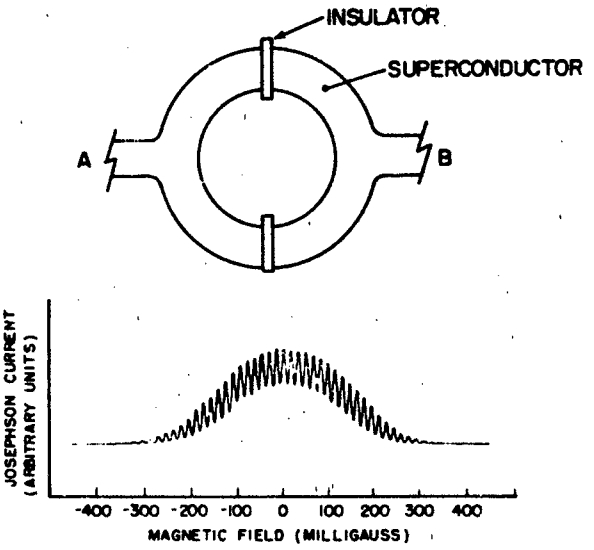


Figure 4-9. The basic geometry and Josephson-current response of a simple dc SQUID magnetometer.

both dc and rf instruments are equipped with feedback circuitry to keep the detector locked on a single maximum, and the amount of the feedback signal becomes the measure of the magnetic field. The SQUID is a vector instrument, since only the component perpendicular to the loop contributes. The sensitivity is extremely high; field changes as small as 10^{-5} nT can be observed. The main limitation is the need to keep the thermal energy kT very small compared with the energy of a flux quantum so the interference pattern is not obscured by thermal noise, and this limits the area of the loop to about 1 or 2 mm^2. The magnetometer inherently possesses excellent linearity, range, and frequency response. Its disadvantages are that it must be calibrated to some other standard, is susceptible to electromagnetic interference (which can cause loss of lock to the reference maximum and hence its calibration), and must be operated in a liquid-helium cryostat.

4.2.2 Ground Measurements

Since the early nineteenth century, the principal source of geomagnetic data has been a continually increasing number of magnetic observatories throughout the world. Recently, the advent of scientific satellites has greatly advanced the precision of global surveys and the study of magnetospheric phenomena, but this has not diminished the need for ground-based measurements.

By 1840 about 50 observatories were making coordinated measurements of the declination by hourly intervals. In recent years over 250 stations have been in continuous or intermittent operation. Of these, over 130 are formal observatories, most of which publish data on a regular basis. Others are repeat stations which are carefully marked and used periodically for standard measurements with portable instruments. Still others are special stations set up for a particular research problem. Their geographic distribution is shown in Figure 4-10, and a list of most stations in recent operation is given in Section 4.8.3. Almost all observatories measure three elements to define the vector field with an accuracy of about 1 nT and a time resolution of one minute. Many also derive K indices and make rapid-run magnetograms for better time resolution. Many of these data are available from the data centers listed in Section 4.8.1.

In recent years, a number of ground-station networks have been constructed to provide the particular geographic coverage needed for the study of ionospheric and magnetospheric dynamics. The largest number were in operation for the three-year International Magnetospheric Study (IMS) commencing in 1977, including five in North America, three in Europe, and (through international cooperation) one which circled the globe. Some are still in continuous or intermittent operation. Data from many of these are in digital form, have 10-second (or better) time resolution, and are available from the National Geophysical Data Center.

4.2.3 Satellite and Rocket Measurements

Sounding rockets carrying magnetometers have been used since 1949 to measure the intensity and height distribution of ionospheric current systems by measuring the change in the transverse magnetic-field component as the current layer was traversed and comparing it with that expected from an idealized infinite current sheet. The equatorial electrojet, the midlatitude Sq currents, and the auroral electrojet have all been studied in this manner using primarily proton-precession or alkali-vapor magnetometers but occasionally fluxgate or searchcoil instruments. Recently, attention has centered on the fine structure of field-aligned currents of the auroral oval, and a series of rocket-borne experiments has attempted to determine the configuration of the lines of force of the geomagnetic field by injecting energetic electron beams along the field line at rocket altitudes and detecting their return.

The use of satellites for magnetic measurements began in 1959 with the Sputnik 3 and Vanguard 3 spacecraft. Measurements since then have provided the drastically altered concept of the magnetosphere described in Section 4.1.4. In the last decade, deep-space satellites with highly elliptical orbits (Hawkeye 1; Heos 1 and 2; IMP H, I, and J; ISEE 1 and 2; OGO 5; and Prognoz 4, 5, 6, and 7) have carried fluxgate magnetometers to the boundary regions of the magnetosphere. Other satellites in elliptical orbits of lower apogee (Explorer 45, GEOS 1, and Jikiken) made measurements largely within the magnetosphere.

Low-altitude polar satellites are ideal for making measurements of the geomagnetic field to derive a spherical-harmonic model of the internal geomagnetic field, since they are nearly simultaneous and have a uniform global distribution. The OGO 6 satellite accomplished this during 1969–1970 using a rubidium-vapor magnetometer to obtain accurate scalar measurements of the field. These were improved on during 1979–1980 by the MAGSAT spacecraft, which obtained both scalar values, using a cesium-vapor magnetometer, and vector component values, using an automatic-field-offset fluxgate magnetometer. The scalar values have an accuracy of 1 nT; instrument and attitude uncertainties limit the accuracy of the vector values to 6 nT.

In the last decade, high-resolution fluxgate magnetometers on the polar-orbiting TRIAD and Air Force S3-2 satellites measured the distribution and intensity of field-aligned currents connecting the outer magnetosphere to the polar and auroral regions of the ionosphere. By applying spin-averaging techniques, data from the fluxgate magnetometers (carried for attitude determination on the ISIS 2, AE-C, and S3-3 satellites) were also used to measure field-aligned currents.

Magnetometers aboard geostationary satellites respond primarily to time variations in the geomagnetic field and have been useful as extraterrestrial observatories to monitor the fluctuations caused by magnetic storms. In the last decade, such spacecraft, all equipped with three-component

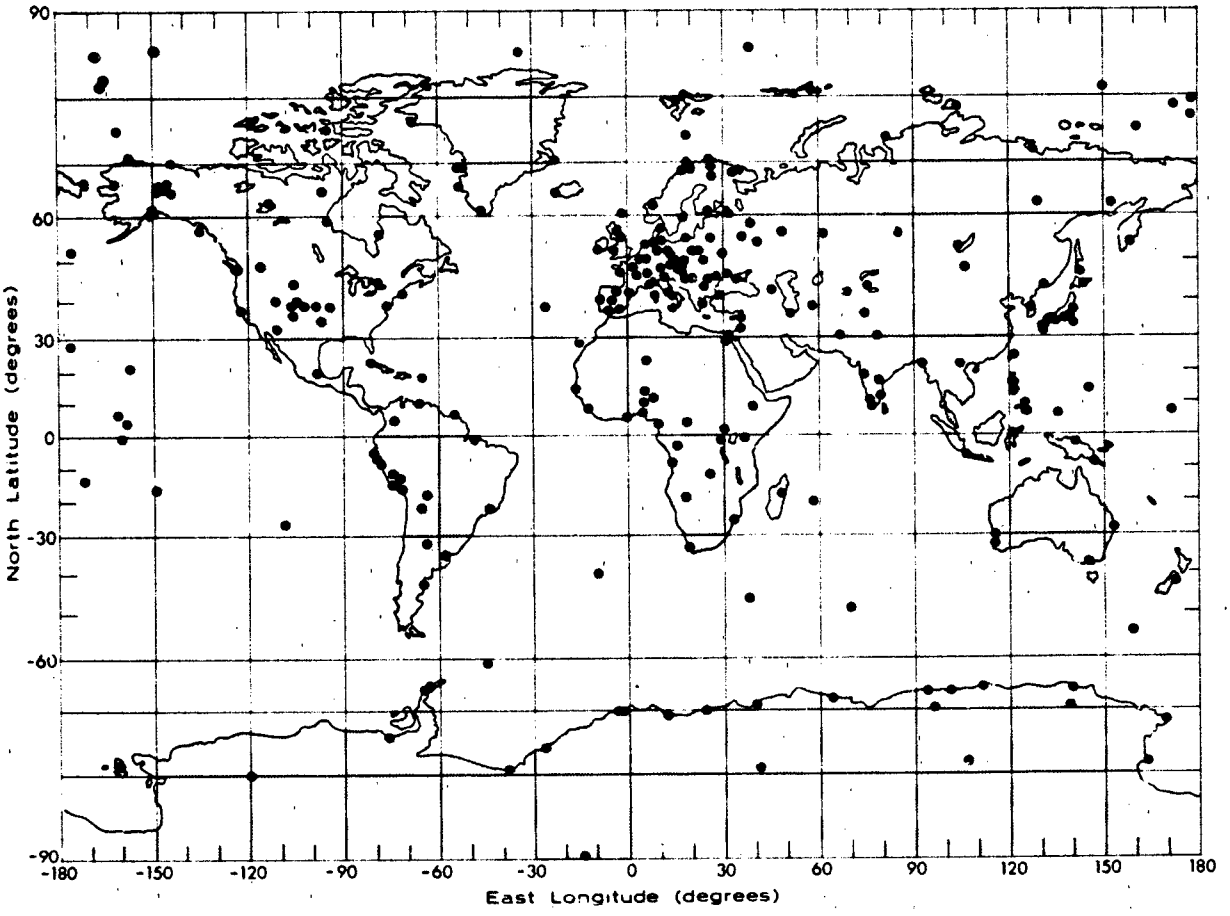


Figure 4-10. Geographical distribution of magnetic observatories.

fluxgate magnetometers, have included ATS 5 and 6; GEOS 2, GOES 1, 2, and 3; SMS 1 and 2; and (in an orbit slightly removed from the synchronous position) SCATHA.

4.3 THE MAIN FIELD

The steady field includes both the main field of terrestrial origin and the nonvarying components of external current systems. While only the former is discussed here, it should be remembered that the latter contributes to the surface field an amount which exceeds the uncertainty of present satellite survey measurements.

4.3.1 Basic Description

The detailed characteristics of the main field are most easily shown in world charts of the elements. Historically such charts were prepared by the hand fitting of curves to observatory and survey data. Now, however, experimental measurements and mathematical models (Section 4.6) are both sufficiently accurate that differences between them are not detectable on world charts of moderate scale, and charts are plotted from the models. Charts of the magnetic elements F, H, Z, and D are presented in Figures 4-11 through 4-14, respectively. These charts for epoch 1980.0 are plotted from the GSFC 9/80 model [Langel, 1982], but differences between all recent models and the actual field are too small to detect on these small-scale plots. Crustal anomalies, which seldom exceed 100 km in extent but are resolved by current models, are similarly unresolved in these plots.

An explanation of the source of the main field must be consistent with seismic and other geophysical data, and many proposed explanations have been discarded on the basis of convincing arguments. The most satisfactory theory is that the field is generated by a self-exciting dynamo system in which an emf generated by the motion of a conductor (molten iron) in a magnetic (excitation) field produces a current so oriented as to produce the excitation field. Although a detailed understanding of core circulations is still lacking, it is believed that the dipole part of the field results

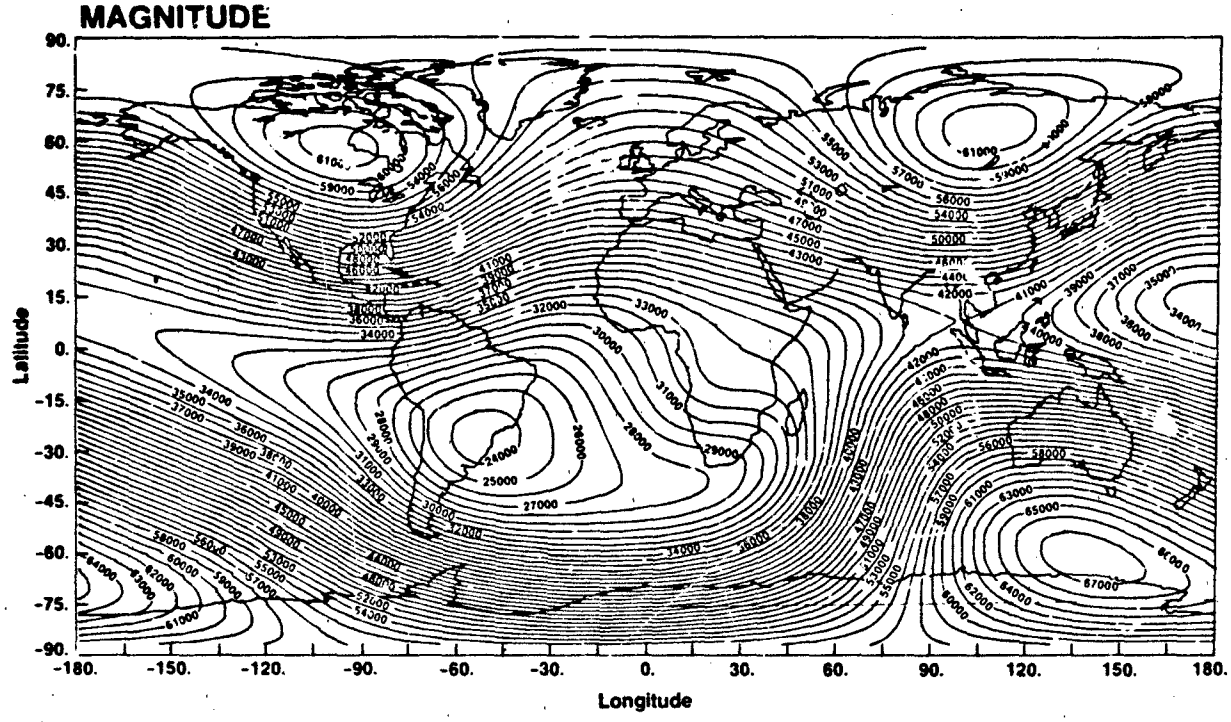


Figure 4-11. Contours of constant total field F at the surface of the earth from the model IGRF 1980.0.

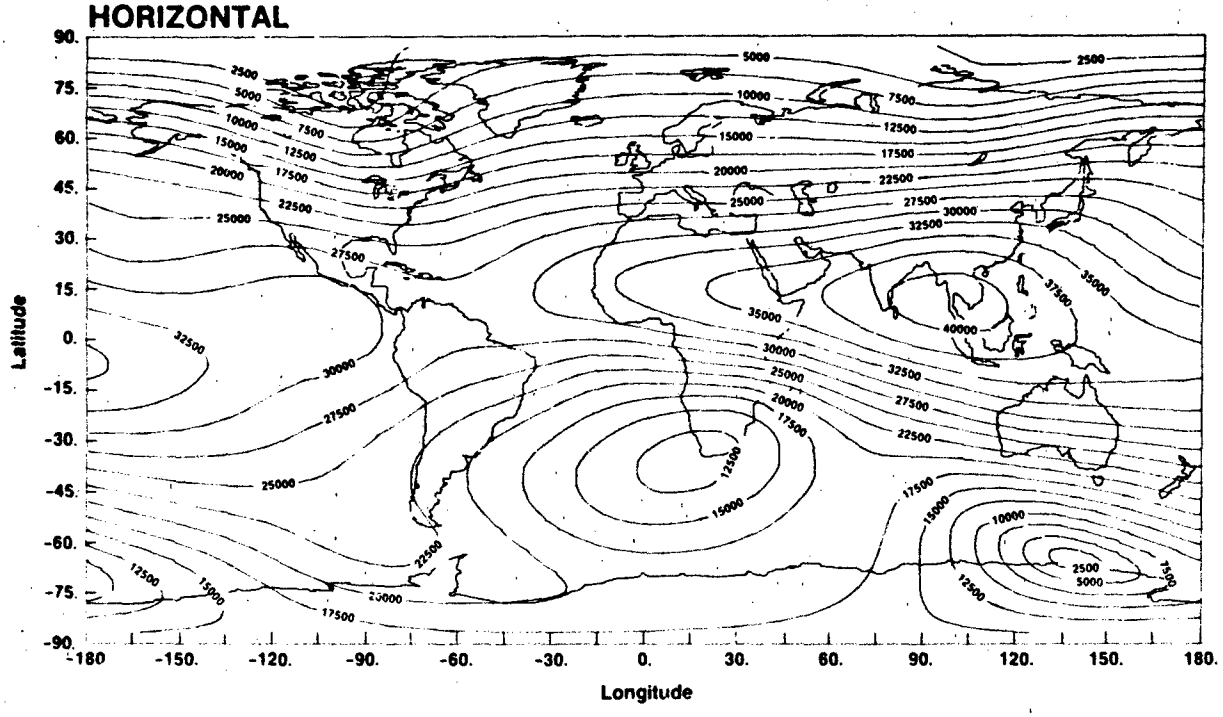


Figure 4-12. Contours of constant horizontal field H at the surface of the earth from the model IGRF 1980.0.

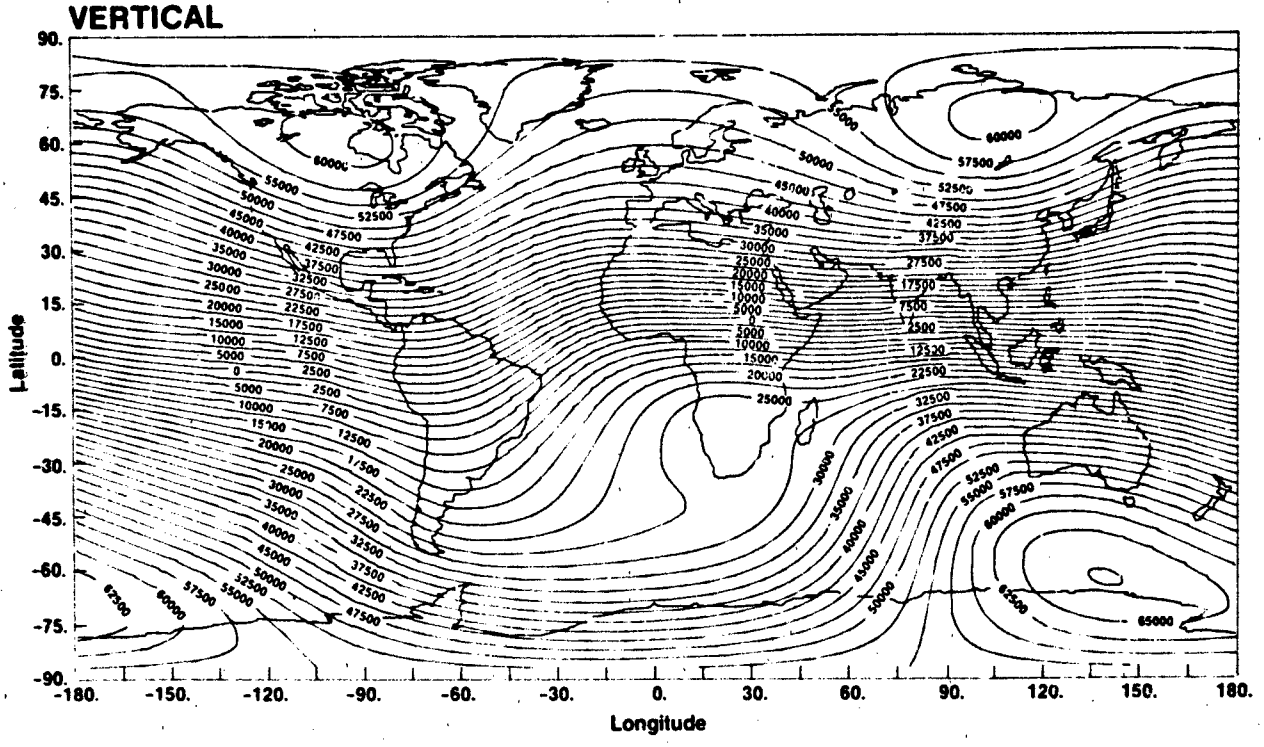


Figure 4-13. Contours of constant vertical field Z at the surface of the earth from the model IGRF 1980.0.

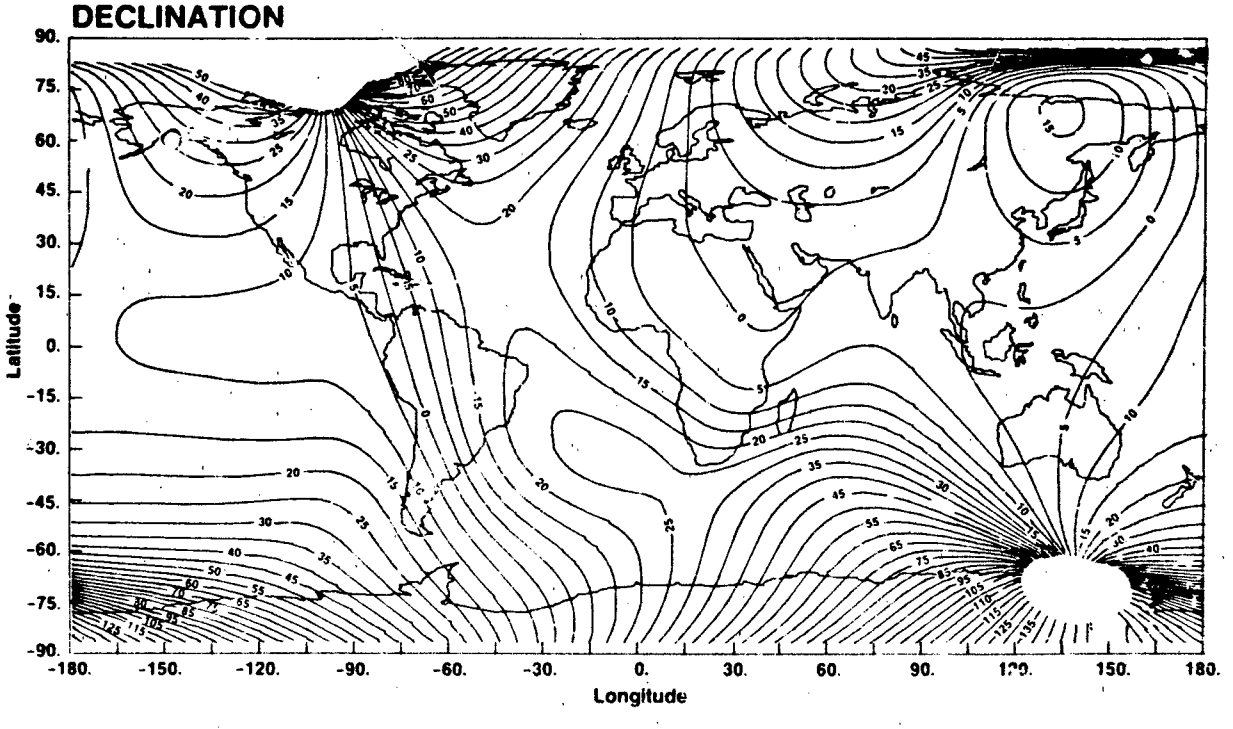


Figure 4-14. Contours of constant declination D at the surface of the earth from the model IGRF 1980.0.

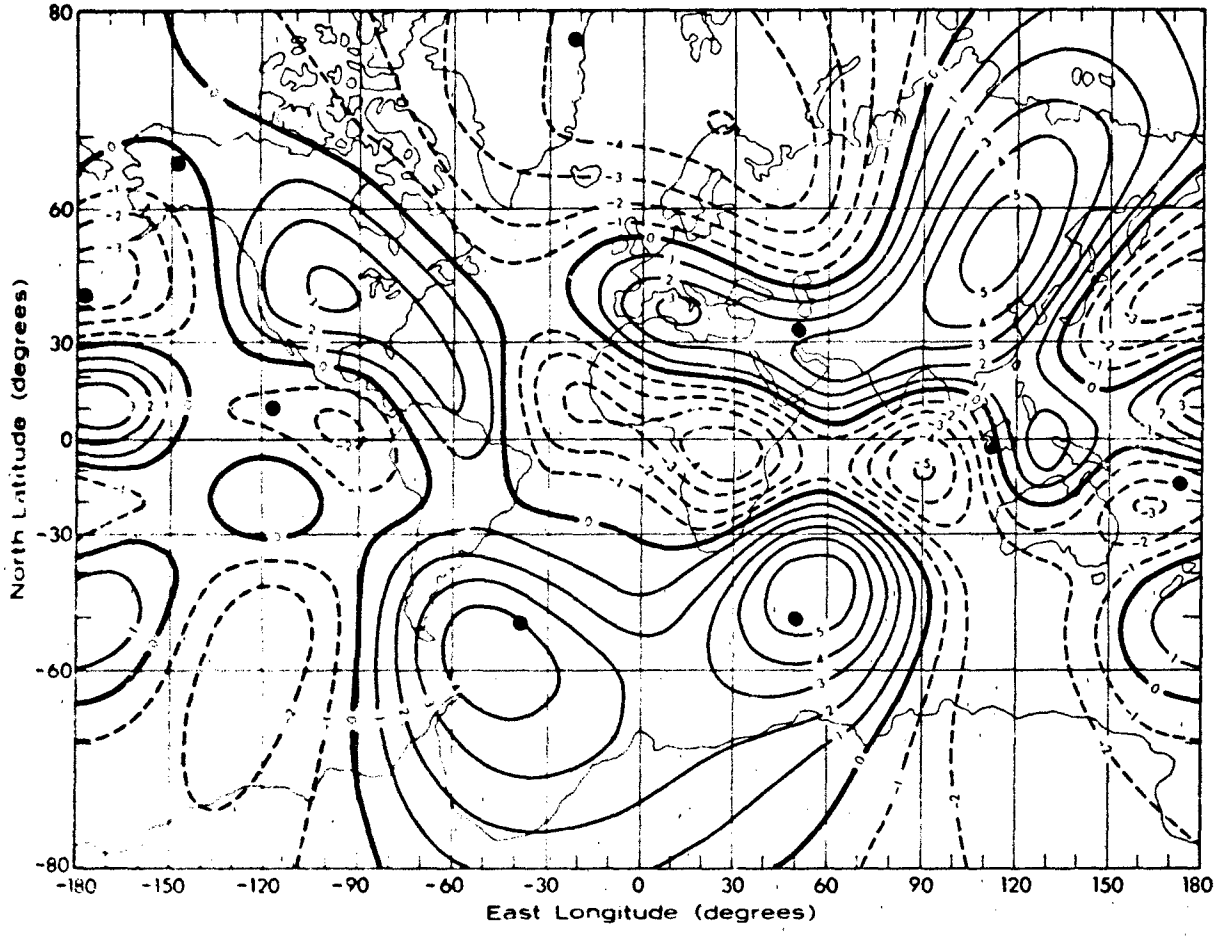


Figure 4-15. The vertical component of the nondipole field, extrapolated to the surface of the core (in oersteds). The locations of nine radial dipoles fitted to the field are indicated by solid circles [see Alldredge and Hurwitz, 1964].

from a major two-dimensional circulation while the nondipole regional anomalies arise from eddy circulations in the outer layer of the core. The regional anomalies can be fitted fairly well by assuming that there exist within the core about nine radial dipoles of various strengths that are equivalent to circulation vortices [Alldredge and Hurwitz, 1964]. These are shown in Figure 4-15. A new analysis of this type will be made of the MAGSAT data [Langel, 1980]. There are also local anomalies of limited extent caused by permanent magnetism in the crust. Because this source is so close to the surface, these fields sometimes exceed that arising from the core. The largest is the Kursk anomaly, 400 km south of Moscow, which is shown in Figure 4-16.

4.3.2 The Secular Variation

The secular variation refers to very slow changes in the main field, the time scale generally being measured in years. The variation during the past 400 years or so can be studied on the basis of historical records. For most of this period measurements separated by many years have been required, but current satellite surveys are precise enough to show changes occurring over an interval as short as a year. The variation has traditionally been found by comparing world magnetic charts for successive epochs; it is now derived by computer programs that fit mathematical models to the experimental data from different periods. Defined more narrowly, the secular variation refers to the time derivatives of the magnetic elements of the main field. The secular variation of the total field F is shown in Figure 4-17.

The secular variation observed over recent decades and centuries can be described as the sum of four components that can be described individually: a decrease in dipole strength, a westward drift of regional anomalies, a northward movement of the dipole, and residual nondrifting variations.

Over the entire period of recorded data, the dipole strength

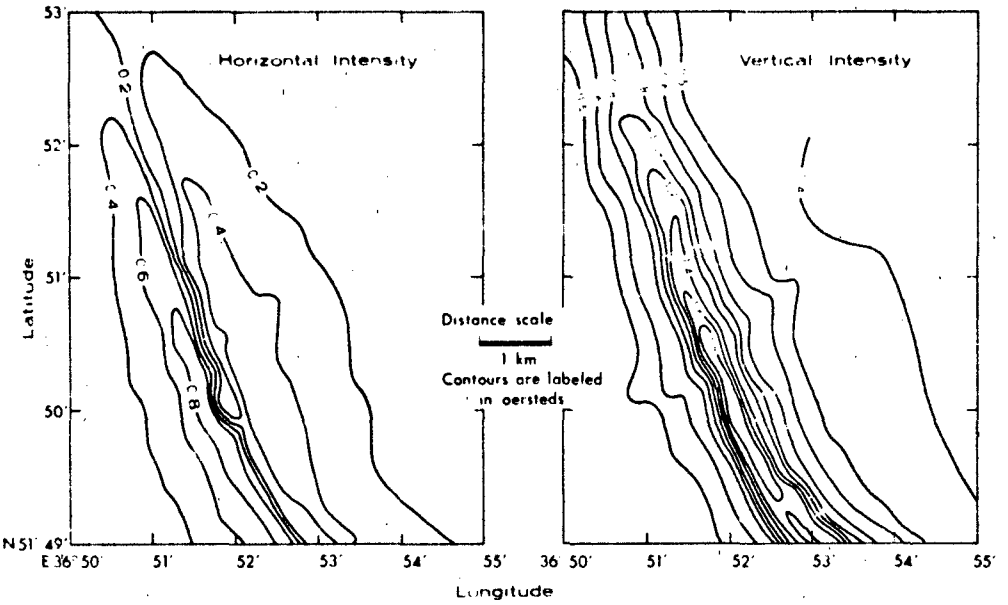


Figure 4-16. The northern portion of the Kursk anomaly. Isointensity contours of the horizontal and vertical field are shown in units of oersteds [after Chapman and Bartels, 1940].

has been decreasing at an average rate of about 0.05% per year (16 nT per year at the equator); data for the past 150 years are plotted in Figure 4-18.

It has long been observed that the major regional anomalies in the field appear to be moving westward, and mathematical analysis [Nagata, 1962] has confirmed that about 60% of the secular variation not attributable to dipole weakening can be accounted for by a westward drift of the nondipole field by about 0.2 degree per year. The cumulative drift over a period of 38 years is easily observed in the equatorial profiles of the vertical field shown in Figure 4-19.

A smaller part of the variation can be described as a northward movement of the dipole center with a velocity somewhat greater than 2 km per year. If this rate were to continue, the center would be outside the core after about 1500 years, so it is likely that the northward motion is merely the current phase of an axial oscillation of the currents that generate the dipole.

The remainder of the variation is relatively small (except in Antarctica). It seems to have about a dozen regional foci, but it has not been accurately measured.

4.3.3 Paleomagnetism

Paleomagnetism, the study of the geomagnetic field in times earlier than those for which recorded data exist, is based on the fact discovered more than a century ago that the natural remanent magnetism (NRM) of some rocks and archaeological samples is a measure of the geomagnetic field that existed at the time of their production. If the NRM is stable, its direction is the same as and its intensity is proportional to the field in which the sample was formed; however, only certain combinations of material, physical process, and conditions result in NRM that is stable enough for reliable results. The most reliable data result from thermoremanent magnetization, locked into the sample by cooling after formation at a higher temperature. The best archaeological samples are baked earths, such as kilns or hearths, from earlier civilizations, and the best geological samples include materials such as lava formed at high temperature. Very careful experimental techniques are required, but the validity of a great many paleomagnetic data and conclusions is well established.

The study of the secular change of geomagnetic-field intensity has been extended backward in time about 5000 years, as shown in Figure 4-20. Over the most recent 2000 years the intensity has tended to decrease at an average rate of about 10 nT per year, while for the preceding 2000 years an increase of similar magnitude is observed.

The study of the secular change of geomagnetic-field direction has been extended backward over both archaeological and geological time scales using data from rock samples. During the past several tens of thousands of years the direction appears to have made quasi-periodic oscillations ranging over several tens of degrees. However, measurements covering the past few millions of years have yielded the important discovery that the dipole axis has not wandered over the entire earth but has remained quite closely aligned

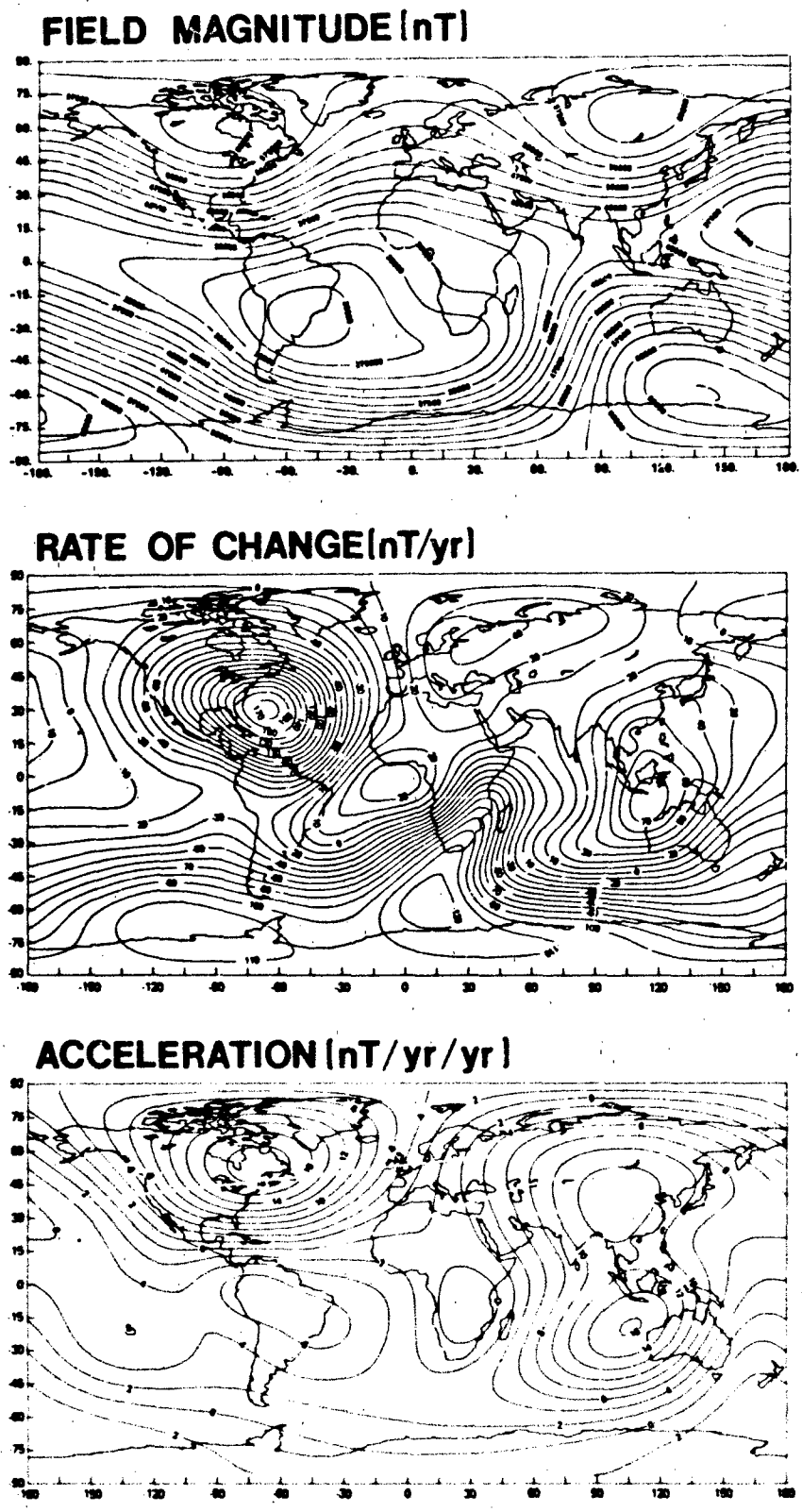


Figure 4-17. The secular variation of the total field for epoch 1980.0. Shown are the field itself (top panel), its first time derivative (center panel), and its second time derivative (bottom panel) [Langel, 1982].

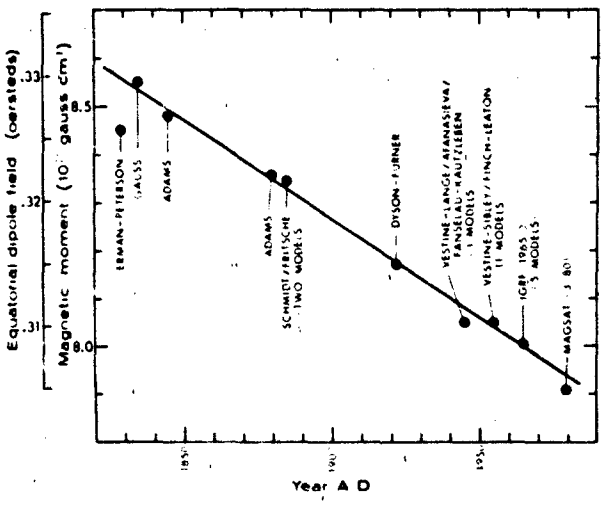


Figure 4-18. Change in equatorial field strength over the past 150 years [historical data from Vestine, 1962].

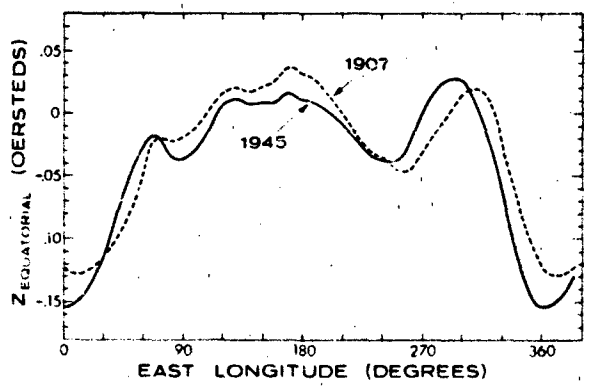


Figure 4-19. Profiles of the vertical equatorial field in the years 1907 and 1945, showing the cumulative westward drift of the non-dipole field over 38 years [after Bullard et al., 1950].

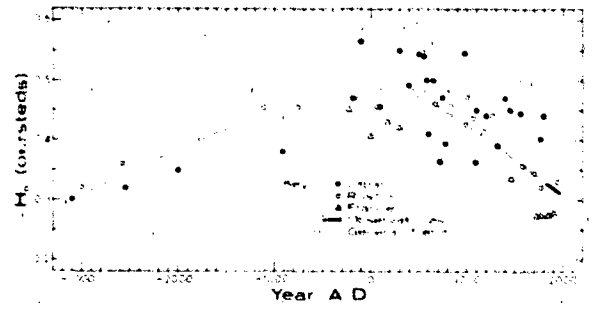


Figure 4-20 Equatorial field intensity in recent millenia, as deduced from measurements on archaeological samples and recent observatory data [after Nagata and Ozima, 1967]

with the axis of rotation. There seems generally to be a clockwise motion of the dipole axis about its mean position with a period of roughly 10 000 years suggesting a precession about the axis of rotation.

For earlier geological ages, paleomagnetic (as well as paleoclimatological) data obtained on any one continent tend to be consistent and yield a time history of the apparent motion of the magnetic poles; over hundreds of millions of years, such "virtual poles" seem to have moved systematically by many tens of degrees toward their present locations. However, there is a very large apparent disagreement between traces for different continents. The explanation is that during this time period the continents themselves have drifted large distances from their original locations.

Measurements covering about 500 million years have revealed that there exist reversely magnetized rocks, and careful study has established that these indicate that the field has periodically undergone complete reversals, the latest being only about one million years ago. It is not clear whether the present weakening of the dipole field represents the beginning of another reversal or a less drastic oscillation.

4.4 QUIET VARIATION FIELDS

A daily variation of the surface magnetic field was first noted in 1722. Although the ionosphere was not discovered until 1902, it was predicted as early as 1882 that the cause of the variation was electric current in a conducting layer of the atmosphere. Subsequent study has shown that the quiet variation includes a large effect of solar origin, a smaller effect of lunar origin, and a still smaller remainder due to other magnetospheric processes.

4.4.1 The Solar Quiet Daily Variation

The solar quiet daily variation (the Sq field) results principally from currents flowing in the E layer of the ionosphere. To a first approximation, this current system is stationary in nonrotating coordinates, and the field variation is observed on the ground as a function of local time because the earth rotates under the currents; therefore, it is similar for all observers at the same latitude, having the dependence shown in Figure 4-21. However, this is a poor approximation since the conductivities that determine the current pattern are controlled by the magnetic field, which is tilted, and the Sq variation is more nearly the same along contours of constant dip latitude than at constant geographic latitude. There is also a longitudinal dependence from such effects as the influence of ocean areas upon the strength of induced currents.

Both the conductivity, which permits this current to flow, and most of the electric field, which powers the current, are produced by solar electromagnetic radiation. The

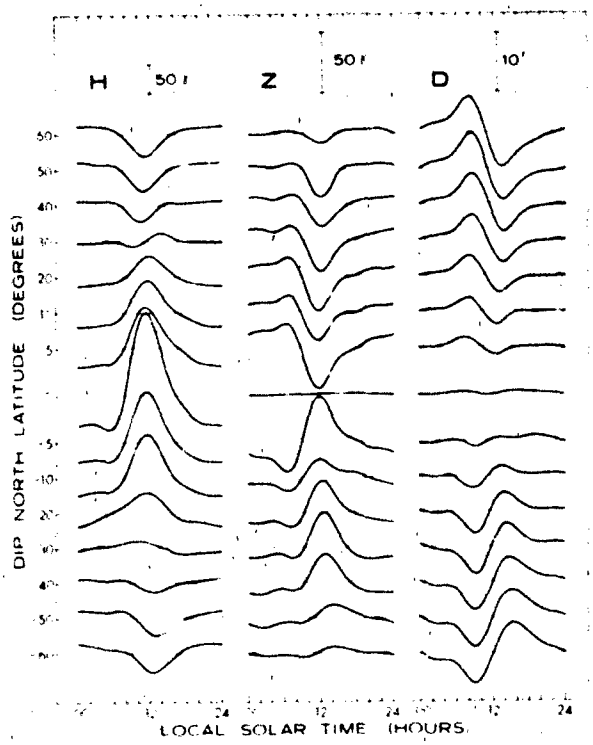


Figure 4-21 Worldwide average of the solar quiet variation near the equinoxes at solar maximum (March, April, September, and October, 1958) [after Matsushita, 1967]

major part of the electric field appears to be generated in the manner of a dynamo by high-speed tidal winds produced by solar heating of the atmosphere, resulting in a variation termed $Sq^{(v)}$. However, it is also clear that part of the electric field, particularly the high-latitude part, originates in the magnetosphere and is communicated to the ionosphere by field-aligned currents resulting in a variation termed $Sq^{(p)}$ [Matsushita, 1975]. Figure 4-22 shows the $Sq^{(v)}$ and total Sq ($Sq^{(v)} + Sq^{(p)}$) ionospheric current systems inferred from a global array of measurements. Induced earth currents (not shown) contribute roughly one-quarter to one-third of the total Sq field.

One notable feature is a concentration of current at the magnetic dip equator. This so-called equatorial electrojet is, in fact, only a few hundred kilometers wide, more concentrated than can be reproduced in the figure by the spherical-harmonic model used to compute the currents. The electrojet exists because of a special circumstance. The fact that the field at the dip equator is exactly horizontal creates a narrow belt of high conductivity in the following way. An electric field impressed perpendicular to a magnetic field (here eastward and northward, respectively) would normally produce a Hall current flowing perpendicular to both (here vertically). However, in this case the conductive medium is bounded in the vertical direction, the Hall current is inhibited, and a polarization results. It is found experimentally and theoretically that the polarization enhances the effective conductivity in the direction of the electric field (Cowling conductivity). At all other points, even slightly off the dip equator, the conductivity along the slightly tilted field lines is sufficient to allow the polarization to leak off partially, and the Cowling conductivity is much less enhanced.

Both the strength and pattern of the Sq variation show a dependence on longitude, season, year, and solar cycle [Matsushita and Maeda, 1965a]. The dependence on season is strong, with the current vortex in a given hemisphere becoming more intense during its local summer. The dependence on solar cycle is also strong; while E-layer ionization increases by 50% from solar minimum to solar maximum, the Sq variation increases by about 100%, presumably because the wind speed also increases. The Sq variation is enhanced and diminished by the changes in solar radiation produced by solar flares and eclipses, respectively.

The Sq field also exhibits large changes from day to day, but the reasons for this are not yet well understood. The current evidence that the Sq current system is partially driven by magnetospheric processes associated with disturbance phenomena blurs the traditional distinction between quiet variation and disturbance variation fields.

4.4.2 The Lunar Daily Variation

The lunar daily variation (the L field) is generated in the same manner as the Sq field, except that the responsible winds are produced by luni-solar gravitational tides and there is currently no evidence for any contribution to the electric field by magnetospheric processes. The dominant behavior is a semidiurnal variation; the amplitude is about an order of magnitude smaller than the Sq amplitude. As in the case of Sq, about 30% of the L field is produced by induced earth currents. Figure 4-23 shows the average L variation in the elements H, Z, and D near the time of an equinox and for a mean lunar age. Figure 4-24 shows the inferred ionospheric (but not induced) currents. A lunar equatorial electrojet is a principal feature, existing (and being more intense than shown) for the same reasons given in the case of the Sq variation.

The dependence of L on several parameters is consistent with expectations [Matsushita and Maeda, 1965b]. There is a seasonal dependence, as for Sq, and also a dependence on lunar age. The solar-cycle dependence on L is smaller than that of Sq, the variation being about 30% (instead of 100%) greater at solar maximum than at solar minimum because the increased activity increases only the conductivities and not the tidal wind speeds. The longitudinal dependence, if it exists, is too small to have been established to date.

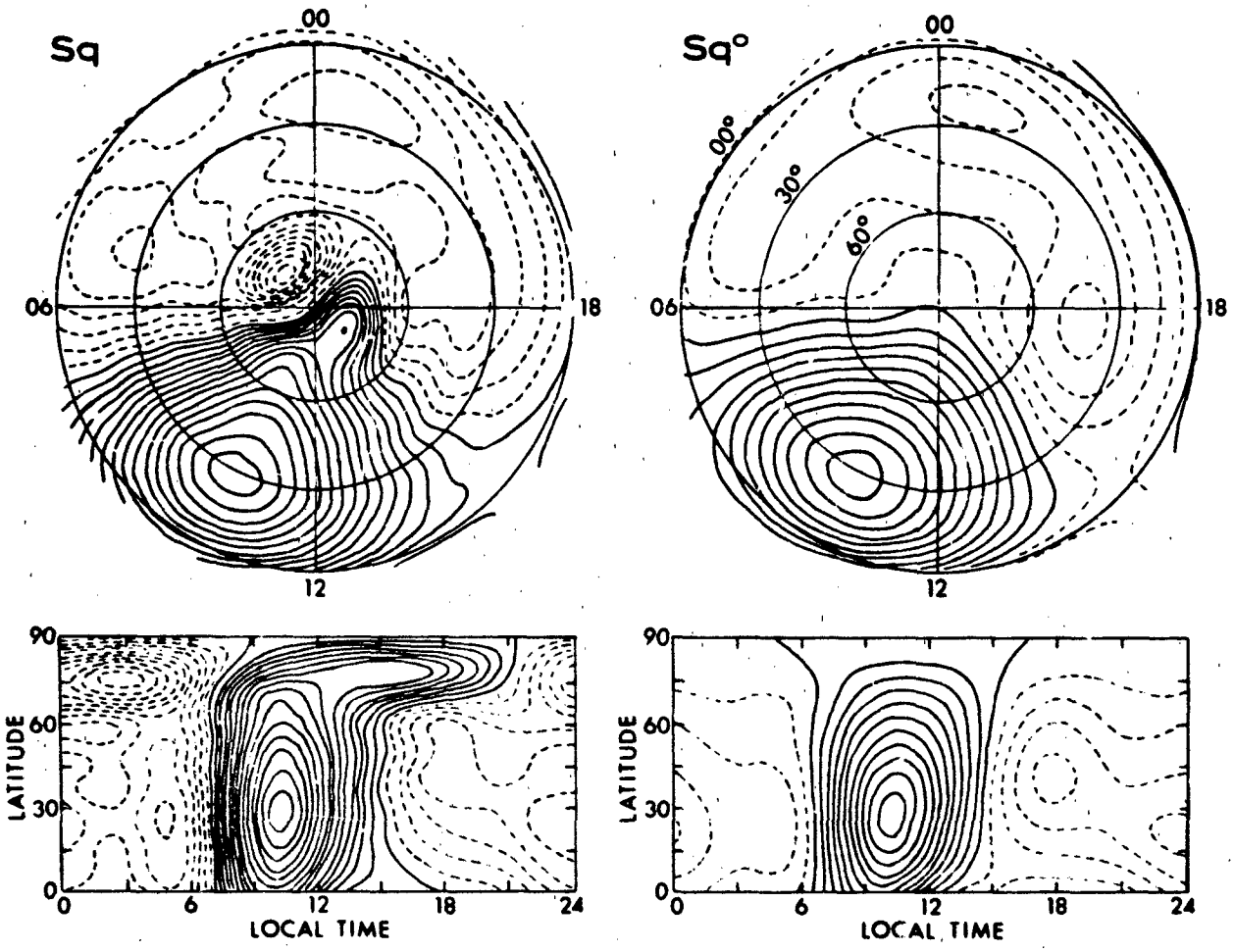


Figure 4-22. Ionospheric currents inferred from the observed Sq variation. The right panels include only the tide-produced currents (Sq^q). The left panels include the convection-produced polar currents ($Sq = Sq^{||} + Sq^q$). The current between adjacent solid (broken) contours is 10 000 amperes counterclockwise (clockwise) [Matsushita, 1975].

4.4.3 Magnetospheric Daily Variation

Except for the contribution to the high-latitude Sq field, the daily variation in the surface field that results from the rotation of the earth within its magnetosphere has not been observed directly since it is similar to and certainly smaller than the Sq and L variations. The dayside-nightside difference in compression of the field by a quiet solar wind results in a surface diurnal variation computed to be about 3 nT, and the surface diurnal components of other magnetospheric fields (such as a quiet-time ring current) that could arise from asymmetric geometries are probably negligible. In the outer magnetosphere, of course, such diurnal effects are large, but measurements made there are usually referenced to a coordinate system that does not rotate with the earth; the effect of the rotating earth is then a somewhat different problem.

4.5 DISTURBANCE FIELDS

4.5.1 Geomagnetic Storms and Substorms

The term geomagnetic storm refers to the geomagnetic effects of a magnetospheric storm, which, broadly defined, is any large prolonged disturbance of the magnetosphere by variations in the solar wind. These storms, observed in recordings (magnetograms) of the surface magnetic field, exhibit great variability and complexity, reflecting the complexity of solar phenomena. However, a classic storm, the features of which are frequently observed, can be described as follows. It includes two energizing parts and a subsequent recovery.

The first part consists of a sudden commencement and an initial phase. These result from a change in compression

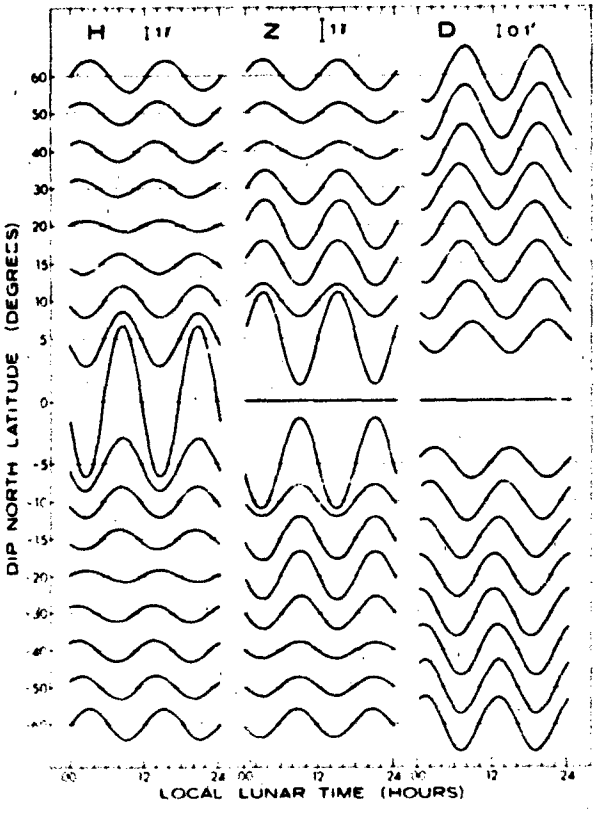


Figure 4-23. Lunar semidiurnal variation near the time of an equinox, derived from data covering the period 1841–1962 [after Matsushita and Maeda, 1965b].

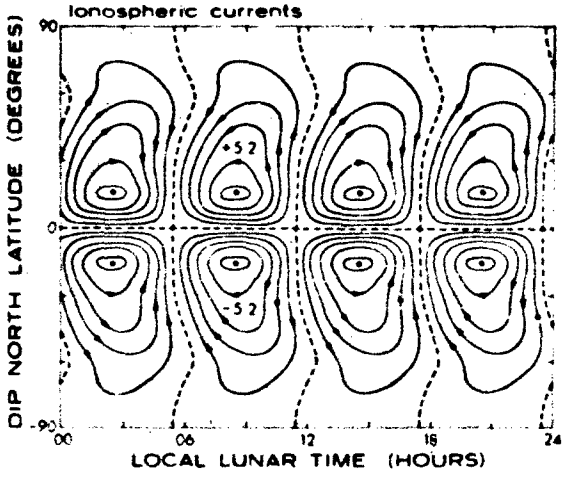


Figure 4-24. Ionospheric currents inferred from the observed L variation. Current between adjacent contours is 1000 amperes, and each dot indicates a vortex center with the total current in thousands of amperes [after Matsushita and Maeda, 1965b].

of the magnetosphere following the passage of a discontinuity such as a shock front propagating in the solar wind and correlate well with the pressure exerted by the bulk flow. The sudden commencement (SC) is seen at low-latitude observatories as an impulsive increase in H, typically having a rise time of one to six minutes and an amplitude of several tens of nanoteslas and observed over the entire earth with a spread in arrival time of less than a minute. Depending on location and the particular storm, it may be positive, negative, double-valued or absent. The rise time corresponds to the time required for the discontinuity to reach all points of the magnetopause and be transmitted to the ground as a hydromagnetic wave. When not followed by the later phases of a storm, this phenomenon is called a sudden impulse (SI). The initial phase typically lasts two to eight hours, during which the field remains compressed by the increased solar-wind pressure following the discontinuity.

The second part is the main phase. It results from an inflation of the magnetosphere by a ring current and is best correlated with a previous southward turning of the interplanetary field, which permits energy to be extracted from the solar wind by a merging of the interplanetary and geomagnetic fields at the magnetopause. It is seen at low latitudes as a rapid decrease in the field to values which are below the prestorm level, often by more than 100 nT and infrequently by more than 1000 nT. It develops over a period of a few hours to a day and is characterized by noise (large fluctuations with a broad risetime spectrum) and an asymmetry in local time (earliest development in the late-afternoon sector). Since solar wind discontinuities usually involve changes in both pressure and field direction, storms typically show both compression and inflation effects, but this is not always the case, and storms without sudden commencements or storms which fail to develop a main phase are not uncommon.

The final part is the recovery phase. It consists of a quiet increase of the field toward the prestorm level with a characteristic time which is typically about one day but sometimes much longer; the recovery is often faster at first than later. Recovery results from a decrease in the ring-current plasma when the source is terminated and the existing plasma is lost by various mechanisms.

The term magnetospheric substorm, denotes a process by which energy extracted from the solar wind and stored in the magnetosphere is dissipated. It is so named because the main phase of a large magnetic storm often appears to be the superposition of many substorms, each of which contributes particles to the main-phase ring current. The intermittent and impulsive nature of the substorms accounts for the characteristic noise of the main phase. The substorm is the principal instability of the magnetosphere and very common, it is often observed almost daily and is seldom absent for many days. The process takes place near local

midnight and is manifested in auroral and geomagnetic observations by the auroral breakup and the geomagnetic bay. It is now clear that it involves a short-circuiting through the auroral-zone ionosphere (via field-aligned currents) of the cross-tail neutral-sheet current. This permits closed field lines that had been distended by the cross-tail current to relax earthward into a more dipolar shape, carrying with them particles for the main-phase ring current. Substorms appear to be self-limiting, lasting for less than an hour; larger amounts of energy are often dissipated in a sequence of substorms or one substorm with multiple onsets.

Prior to the most recent decade, many studies of geomagnetic disturbance sought to explain various phenomena by constructing "equivalent" current systems confined to the ionosphere. More recently, there has emerged a much better understanding of the importance and general nature of field-aligned currents in the magnetosphere and the intimate connection between ionosphere and magnetosphere. Chapter 8 discusses some of the magnetospheric-ionospheric processes and current systems, including substorms, that manifest themselves in geomagnetic phenomena and that traditionally have been studied from that perspective.

4.5.2 Geomagnetic Pulsations

Variations of the geomagnetic field having periods from less than one to several hundreds of seconds are observed both on the ground and in the magnetosphere. They are ultra-low-frequency (ULF) waves with frequencies below the ion gyrofrequency that propagate as hydromagnetic waves in the magnetosphere. They have commonly been called micropulsations, geomagnetic pulsations, or simply pulsations. Wave amplitudes range from tenths to several hundred nanoteslas, with the largest amplitudes usually occurring in the longer-period waves at high latitudes. Simultaneous periodic variations in particle precipitation, auroral intensity, electric fields, x ray bursts, and particle flux are often observed. The pulsations have been classified into two principal types: Pc (continuous) pulsations, which often have very sinusoidal waveforms, and Pi (irregular) pulsations. More detailed classification is discussed by Jacobs [1970]; further subdivisions and characteristics of the waves are given in Figure 4-25 [Saito, 1978]. The shortest-period ULF waves are thought to result from cyclotron instabilities with charged particles in the magnetosphere. An overview of these waves can be found in Jacobs [1970] and Nishida [1978]. A description of the large variety of higher-frequency (outside the ULF band) plasma waves in the magnetosphere is beyond the present scope; however, useful summaries have been given by Jacobs [1970] and Shawhan [1979]. Several collections of papers presented for review at meetings have been published [Southwood, 1980; Orr, 1981; IAGA, 1982]. Other useful reviews have been given by Saito [1969], Orr [1973], Lanzerotti and Southwood [1979], Shawhan [1979], Southwood [1979], Hughes [1982], Singer [1982], Southwood and Hughes [1983], and Hughes [1983].

Dungey [1954] suggested that the long-period ULF pulsations observed on the ground were hydromagnetic waves resonating on geomagnetic field lines. In the idealized case, field lines can be considered fixed at both ends in a perfectly conducting ionosphere, and harmonic standing waves can exist on flux tubes. In space and on the ground, magnetic observations of spatial variations in wave frequency and polarization characteristics have supported this picture. An approximate expression for calculating the resonance period is given by the so-called WKB or time-of-flight approximation

$$T = \frac{2}{\eta} \int \frac{ds}{v_A(s)} \propto L^4 \left[\sum_i n_i m_i \right]^{1/2} \qquad (4.1)$$

where the Alfvén speed is given by $v_A = B/(\mu_0 n_i m_i)^{1/2}$, the integration is carried out between conjugate ionospheres, η is the harmonic number, ds is an element of length along a field line, n_i and m_i are the number density and mass of species i in the plasma, and L is the McIlwain L-shell parameter. This method is least accurate for the fundamental mode but more accurate calculations of eigenperiods have been carried out using the wave equations for low-frequency hydromagnetic waves in a nonuniform magnetic field (for example, a dipole field). The full hydromagnetic wave equations are coupled and have not been solved in general. However, with certain simplifying assumptions (such as considering magnetic perturbations strictly transverse to the background field) they have been solved for poloidal-mode (radial field displacement) and toroidal-mode (azimuthal field displacement) waves. A particularly useful table for calculating eigenperiods of the uncoupled poloidal and toroidal mode waves is given by Cummings et al. [1969]. Solutions in an arbitrary field geometry have been developed by Singer et al. [1981]. The theoretical development of hydromagnetic waves has been discussed more fully by Southwood and Hughes [1982].

Several mechanisms have been suggested as sources of magnetic pulsations; however, it is usually difficult to link a particular wave observation to a specific source definitively. For waves external to the magnetosphere, sources include the Kelvin-Helmholtz instability (wind-over-water) at the magnetopause boundary, quasi-parallel wave excitation where waves upstream of the bow shock penetrate into the magnetosphere, and sudden impulses due to solar wind discontinuities encountering the magnetopause. For waves internal to the magnetosphere, sources include magnetic substorms, real-space or velocity-space gradients in the magnetospheric plasma distribution, ionospheric conductivity discontinuities, and unsteady large-scale plasma convection. Many dayside Pc pulsations probably originate

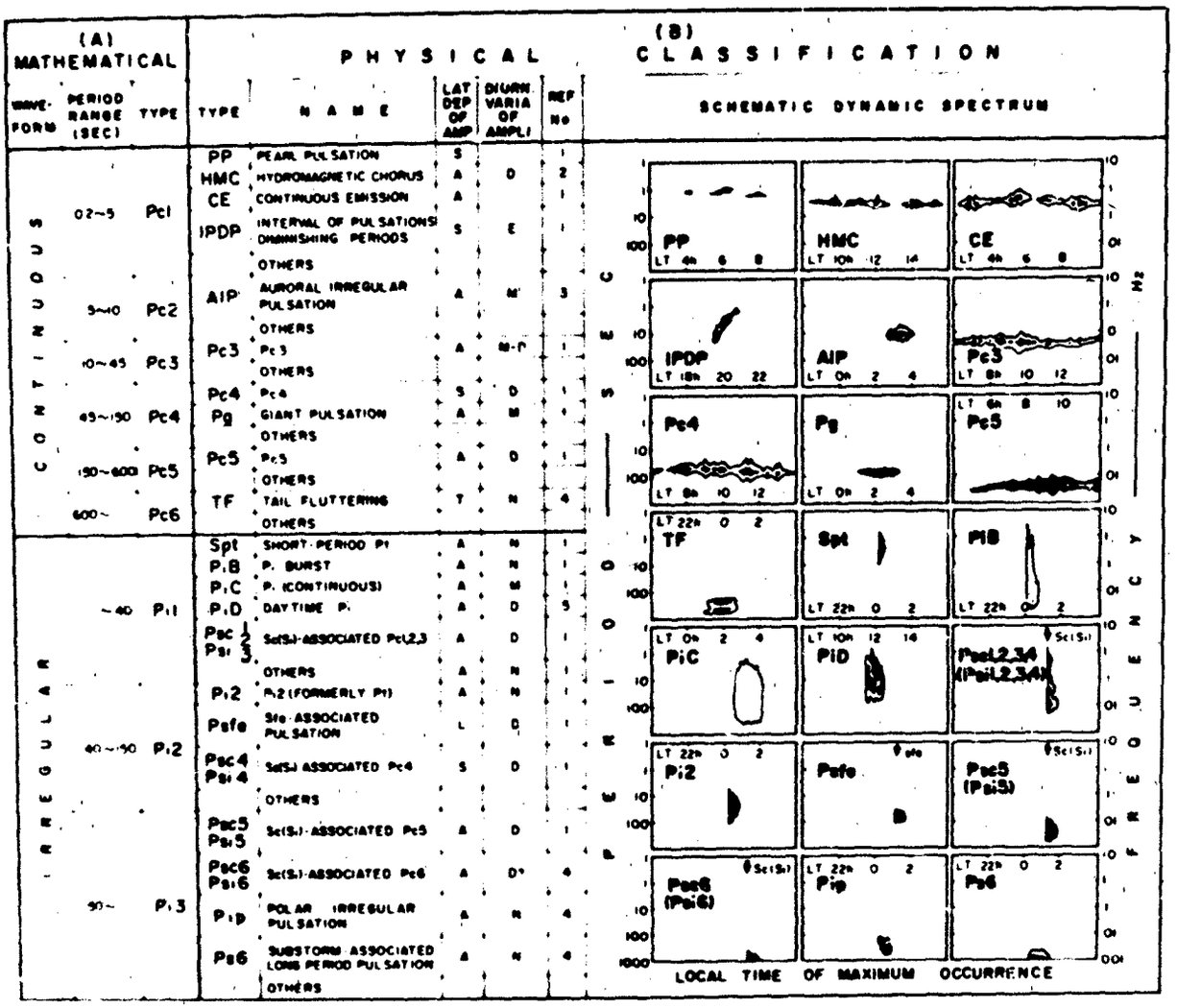


Figure 4-25. Classification of geomagnetic pulsations [Saito, 1978].

as surface waves generated by the Kelvin-Helmholtz instability that evanescently decay inside the magnetosphere where they couple to transverse-mode standing Alfvén waves on resonant field lines [Southwood, 1974; Chen and Hasegawa, 1974]. Pi pulsations are observed primarily on the nightside of the earth and have been clearly linked to the onset of magnetospheric substorms [Southwood and Stuart, 1980].

The ionosphere influences long-period ULF waves in the magnetosphere as well as modifying the signal observed on the ground from that which would be observed in space. Wave energy is lost in the ionosphere through Joule heating by Pedersen currents, and significant amounts of energy may be deposited in the high-latitude ionosphere by waves [Greenwald and Walker, 1980]. The ionospheric Pedersen currents shield the incident magnetic field from the ground, and the signals observed on the ground are due to Hall currents in the ionosphere. A consequence of this is that the polarization of the horizontal wave observed on the ground is rotated by 90° from that which would be observed in space. In addition, the ionosphere acts as a spatial filter and screens short-scale horizontal signal variations $\lesssim 120$ km (approximate height of the E region) from the ground.

The reflection coefficient of an Alfvén wave by the ionosphere is

$$R = \frac{1 - \mu_0 v_A \Sigma_p}{1 + \mu_0 v_A \Sigma_p} \qquad (4.2)$$

where Σ_p is the height-integrated Pedersen conductivity in the ionosphere, and v_A is the Alfvén velocity along the

oscillating flux tube. The ratio of the effective field-line conductance $(\mu_o v_A)^{-1}$ to the ionospheric conductance is one control on wave damping in addition to wave-particle interactions in the magnetosphere. For example, on the nightside of the earth where $\Sigma_p \sim (\mu_o v_A)^{-1}$, the wave may damp out after only a few cycles, forming Pi2-type pulsations, whereas on the dayside, where $\Sigma_p \gg (\mu_o v_A)^{-1}$, the reflectance is high and waves can exist for many cycles as Pc pulsations. The amplitude of the wave signals on the ground is also controlled by the ionospheric conductivity and the horizontal wave number of the waves. An expression relating the ground field strength b_g to that in the magnetosphere b_m is

$$\frac{b_g}{b_m} \propto e^{kh\frac{\Sigma_P}{\Sigma_H}} \tag{4.3}$$

where h is the height of the E region, k is the horizontal wave number, and Σ_H and Σ_p are the height-integrated Pedersen and Hall conductivities, respectively.

Magnetic pulsations contain information about their source and the regions through which they propagate. Accordingly, attempts have been made to use ground observations of pulsations as a diagnostic tool for inferring properties of the solar-wind and magnetospheric environments. Ground-pulsation characteristics such as occurrence rate, amplitude, and frequency have been correlated with various combinations of the solar wind velocity and the interplanetary magnetic-field strength and its component values, parameters which are important for generating pulsations in the solar wind or at the magnetopause boundary. Although this research is still in its infancy, it may eventually be possible to monitor solar wind properties from ground-based pulsation measurements, a much easier and less costly technique than *in situ* satellite measurements. Magnetospheric plasma density is another property that would be useful to measure by ground-based observations, and attempts have been made, using the notion of field-line resonance. As can be seen from Equation (4.1), measurement of pulsation period permits determination of plasma mass density along a flux tube, providing the standing-wave harmonic and field geometry are known. This problem is further complicated by the actual distribution of mass on the field line and it does not distinguish plasma species. Nevertheless, it has been used to give estimates of mass density on flux tubes. Another use of pulsations as diagnostics of the magnetospheric environment is that of determining substorm onsets using Pi2 pulsations. It appears that each substorm onset may be associated with a Pi2, although the converse is probably not true. The Pi2 signal is particularly useful because it begins abruptly and can often be observed over a large portion of the earth. In space the signal propagates faster than the plasma injected at substorm onset, and at some distance from the injection source the Pi2 can be used to foretell the arrival of energetic particles.

4.6 QUANTITATIVE MAGNETIC-FIELD MODELS

4.6.1 The Internal Field

A first approximation to the geomagnetic field near the surface of the earth is an earth-centered dipole with its axis tilted to intersect the earth at 78.5°N, 291.0°E, the geomagnetic north pole; the other intersection at 78.5°S, 111.0°E is the geomagnetic south pole. (These values were adopted internationally to define the geomagnetic coordinate system; however, the dipole terms of the first International Geomagnetic Reference Field (epoch 1965.0), also adopted internationally, place the pole at 78.56°N, 290.24°E, and this axis is also sometimes used.) In spherical coordinates, r, θ, and φ, with r measured from the center of the earth and θ measured from the dipole axis (geomagnetic colatitude), the dipole field has the vector components

$$B_r = -\frac{M}{r^3} 2\cos\theta$$

$$B_\theta = -\frac{M}{r^3}\sin\theta \tag{4.4}$$

$$B_\phi = 0.$$

The total intensity is then

$$B = -\frac{M}{r^3}[3\cos\theta + 1]^{1/2} \tag{4.5}$$

where M is the dipole moment of the earth (about 7.9×10^{15} T/m^3). The equation of a field line is obtained by equating its slope to the slope of the field vector:

$$\frac{r d\theta}{dr} = \frac{B_\theta}{B_r} = \frac{\tan\theta}{2}$$

Integration yields

$$r = r_0 \sin^2\theta \tag{4.6}$$

where the constant of integration r_0 is the geocentric distance at which the field line crosses the geomagnetic equator. A magnetic shell, which is the surface of revolution generated by rotating a field line about the dipole axis, is a useful concept in studying charged-particle motion. In the B-L coordinate system, trapped particles which move adiabatically to conserve three invariants of the motion are confined to a shell of constant L. The B-L coordinate system is, roughly speaking, a transformation of the actual distorted field of the earth into an equivalent dipole field (employing a multipole expansion of the former, as discussed below).

and the shells of constant L transform into the simple magnetic shells of the dipole field, with $L = r_0$.

The centered dipole is a poor approximation to the actual field. It yields an equatorial field intensity of 30 800 nT, while the actual intensity ranges between values of 41 000 and 24 000 nT, more than 25% above and below the calculated value. Much of this discrepancy can be removed by taking the dipole to be eccentric, a fairly good fit being obtained by displacing the center of the dipole 436 km toward 15.6°N, 150.9°E and inclining the dipole axis to intersect the spherical earth at 81.0°N, 275.3°E and 75.0°S, 120.4°E [Parkinson and Cleary, 1958]. However, discrepancies of about 10% remain, which is not good enough for most purposes. Since much better models are readily available for machine computation, the eccentric dipole is seldom used.

Any field derivable from a potential function can be expressed in terms of a multipole expansion of the potential. The coefficients of the various terms can then be adjusted by a least-squares method to give the best fit to the measured field. The methodology for such analyses is well developed, and sets of coefficients along with computer programs to calculate the field and a number of related parameters are readily available. In simplest form, the analysis is as follows. The magnetic scalar potential can be written as a spherical-harmonic expansion:

$$V = a \sum_{n=1}^{x} \sum_{m=0}^{n} P_n^m(\cos\theta) \times \left[\left(\frac{a}{r}\right)^{n+1} (g_n^m \cos m\phi + h_n^m \sin m\phi) + \left(\frac{a}{r}\right)^{n} (A_n^m \cos m\phi + B_n^m \sin m\phi) \right], \quad (4.7)$$

where r, θ, and φ are the geographical polar coordinates of radial distance, colatitude, and east longitude, and a is the radius of the earth ($a = R_1$). The functions $P_n^m(\cos\theta)$ are the Schmidt functions:

$$P_n^m(\cos\theta) = \left[\frac{\epsilon_m (n-m)!}{(n+m)!}\right]^{1/2} \times \left[\frac{(1-\cos^2\theta)^{m/2}}{2^n n!} \frac{d^{n+m}}{d(\cos\theta)^{n+m}} (\cos^2\theta - 1)^n\right]$$
$$\epsilon_m = 2 \text{ if } m > 0$$
$$\epsilon_m = 1 \text{ if } m = 0. \quad (4.8)$$

The second quantity in brackets is the associated Legendre function $P_{n,m}(\cos\theta)$. Its numerical multiplier makes the Schmidt function partially normalized. (Its mean-square value integrated over the sphere is not unity but $(2n+1)^{-1/2}$). The coefficients g_n^m, h_n^m, A_n^m, and B_n^m appropriate to this function, referred to as "Schmidt coefficients", are now quite universally used in quoting results of analyses, although earlier analyses have used other normalizations, particularly the Gaussian, in which the numerical multiplier is replaced by $[2^n n!(n-m)!/(2n)!]$.

In the potential, those terms containing g_n^m and h_n^m arise from sources internal to the earth, while those containing A_n^m and B_n^m arise from external currents; the potential function is valid in the space above the surface and below the external current system. The field is given by

$$\mathbf{B} = -\nabla V. \quad (4.9)$$

The northward, eastward, and downward components of the field are thus

$$X = \frac{1}{r}\frac{\partial V}{\partial \theta}$$
$$Y = -\frac{1}{r\sin\theta}\frac{\partial V}{\partial \phi} \quad (4.10)$$
$$Z = \frac{\partial V}{\partial r}.$$

The construction of a model consists of adjusting the coefficients g_n^m, h_n^m, A_n^m, and B_n^m to fit experimental measurements taken over a network covering the entire earth, using a least-squares procedure to minimize some quantity such as the weighted sum of the squares of the differences between calculated and measured X, Y, and Z values. As the degree of the multipole terms increases, the magnitude of their effect decreases, and it is pointless to retain more terms than are warranted by the accuracy of the experimental measurements.

Twenty years ago all models were based on surface measurements. Since analyses were unable to identify contributions from external sources, the coefficients A_n^m and B_n^m were usually assumed to be zero. Although measurements were not simultaneous, the existing knowledge of the secular variation was used to adjust them to a common epoch to which the resulting model applied. The significance of coefficients of order greater than about six was dubious.

Since then field models of much greater precision have been developed on the basis of data obtained by low-orbiting polar satellites that can provide nearly simultaneous coverage of the entire earth. A first major advance was made possible by three spacecraft of the NASA POGO series, which carried accurate total-intensity magnetometers. A second advance has now been afforded by the MAGSAT satellite, launched in late 1979, which made accurate three-component measurements of the vector field. The increased precision of these recent data require the retention of higher-

order coefficients in the multipole expansion; the latest models typically include terms to order 13. At satellite altitude (e.g., 500 km), these models specify the field to an accuracy of about 10 nT and clearly resolve regional anomalies resulting from crustal magnetism. Also discernible in the data are the contributions of the external field, including that of the magnetic-storm ring current. Since satellite surveys now span about 15 years, the secular variation can be increasingly well determined. Current models derived from data spanning such a time interval include the time dependence by specifying the coefficients for a reference epoch and also their first and second time derivatives.

Until recently, substantial discrepancies existed between models derived from different data bases. The value of a standard reference field was recognized, and beginning with epoch 1965.0 such a model, the International Geomagnetic Reference Field (IGRF), has been adopted. IGRF 1965.0 was an average model; for any specific purpose another model was likely to provide greater accuracy, but the IGRF served well as a common reference. This difficulty no longer exists; IGRF 1980.0 is the most accurate model available. Its time-derivative coefficients permit its use until 1985 when the next IGRF is scheduled. Retrospective studies have also resulted in a Definitive Geomagnetic Reference Field, consisting of models at five-year intervals (DGRF 1965, DGRF 1970, and DGRF 1975) with linear interpolation of coefficients between them. The DGRF is not expected to be revised further, whereas the IGRF may be (becoming DGRF 1980). A continuous bridge from the DGRF 1975 to the IGRF 1980 is made by the PGRF 1975 ("P" for "provisional"), a linear interpolation between the two. This pattern is expected to be followed for further updates.

The coefficients for the three DGRF models and IGRF 1980 have been published in *EoS* [IAGA, 1981] and elsewhere; those for IGRF 1980 are listed in Table 4-1. These and other field models can be provided by World Data Center A for Rockets and Satellites (Section 4.8.1) in the form of coefficients and computer programs to generate field values.

4.6.2 The External Field

The external field is that portion of the geomagnetic field that results from currents flowing above the surface of the earth. Even at the surface and at low altitudes, where the contribution of the external field is a very small percentage, measurements are now sufficiently precise that it is included in the spherical-harmonic models. As the strength of the internal field weakens with increasing distance from the earth, the external field becomes relatively more important. However, even with external-field coefficients, the spherical-harmonic models described above lose their usefulness beyond about two or three earth radii. The reason is that the geometry of the external field does not lend itself to spherical-harmonic analysis. Therefore, a separate model of the external field must be constructed, and contributions from the internal and external models can then be added to specify the total field at any location in the magnetosphere.

Large-scale current systems, often only partially understood, are associated with many complex magnetospheric processes. However, three major current systems appear to explain the basic configuration of the magnetosphere: (1) the magnetopause current, which creates the magnetopause and flows in it, confining the geomagnetic field to the magnetosphere; (2) the neutral-sheet current, which is driven by the dynamo-generated potential across the magnetosphere and causes polar-cap field lines to be drawn out away from the earth into the tail; and (3) the ring current, which consists of plasma drifting around the earth in the region where field lines have a more dipolar shape. Several useful models have been developed by Olson and Pfitzer [1974, 1977], Tsyganenko [1976], and others to compute the magnetic field produced by these currents. One of the Olson-Pfitzer models includes the tilt of the earth dipole; the other permits varying the strengths of the three currents independently.

External-field models describe a field configuration that is much less stable than the internal field. In studying some magnetospheric process, the model can be used to describe the general field configuration while the process is modeled separately. However, the currents involved in the process may be intimately associated with the model currents. An example is the substorm, in which part of the neutral-sheet current is diverted along field lines and through the ionosphere causing an abrupt change in the configuration of the magnetospheric magnetic field. The development of self-consistent models for such processes is just beginning.

4.7 GEOMAGNETIC ACTIVITY INDICES

The magnetograms and averaged data supplied by magnetic observatories are too detailed for many purposes, and it has long been found useful to compute some numerical parameters that indicate the level of general magnetic activity or of specific types of disturbance. The International Association of Geomagnetism and Aeronomy (IAGA) has formally adopted 19 of more than 30 such indices currently in use. A brief description is given here of several of the most useful indices; a complete review has been given by Mayaud [1980]. The availability and sources of indices are discussed in Section 4.8.5.

4.7.1 General Activity Indices K, Ks, Kp, and Km

The K index, a measure of the irregular variations of standard magnetograms, is an indicator of the general level of disturbance at a given observatory. The index is defined

CHAPTER 4

Table 4-1. Spherical-harmonic coefficients of the IGRF 1980.0.

n	m	g	h	dg/dt	dh/dt	n	m	g	h	dg/dt	dh/dt
1	0	−29988		22.4		8	0	20		0.8	
1	1	−1957	5606	11.3	−15.9	8	1	7	7	−0.2	−0.1
						8	2	1	−18	−0.3	−0.7
2	0	−1997		−18.3		8	3	−11	4	0.3	0.0
2	1	3028	−2129	3.2	−12.7	8	4	−7	−22	−0.8	−0.8
2	2	1662	−199	7.0	−25.2	8	5	4	9	−0.2	0.2
						8	6	3	16	0.7	0.2
3	0	1279		0.0		8	7	7	−13	−0.3	−1.1
3	1	−2181	−335	−6.5	0.2	8	8	−1	−15	1.2	0.8
3	2	1251	271	−0.7	2.7						
3	3	833	−252	1.0	−7.9	9	0	6			
						9	1	11	−21		
4	0	938		−1.4		9	2	2	16		
4	1	783	212	−1.4	4.6	9	3	−12	9		
4	2	398	−257	−8.2	1.6	9	4	9	−5		
4	3	−419	53	−1.8	2.9	9	5	−3	−7		
4	4	199	−298	−5.0	0.4	9	6	−1	9		
						9	7	7	10		
5	0	−219		1.5		9	8	1	−6		
5	1	357	46	0.4	1.8	9	9	−5	2		
5	2	261	149	−0.8	−0.4						
5	3	−74	−150	−3.3	0.0	10	0	−3			
5	4	−162	−78	0.2	1.3	10	1	−4	1		
5	5	−48	92	1.4	2.1	10	2	2	1		
						10	3	−5	2		
6	0	49		0.4		10	4	−2	5		
6	1	65	−15	0.0	−0.5	10	5	5	−4		
6	2	42	93	3.4	−1.4	10	6	3	−1		
6	3	−192	71	0.8	0.0	10	7	1	−2		
6	4	4	−43	0.8	−1.6	10	8	2	4		
6	5	14	−2	0.3	0.5	10	9	3	−1		
6	6	−108	17	−0.1	0.0	10	10	0	−6		
7	0	70		−1.0							
7	1	−59	−83	−0.8	−0.4						
7	2	2	−28	0.4	0.4						
7	3	20	−5	0.5	0.2						
7	4	−13	16	1.6	1.4						
7	5	1	18	0.1	−0.5						
7	6	11	−23	0.1	−0.1						
7	7	−2	−10	0.0	1.1						

for each 3-hour interval on the basis of the largest value R of the 3-hour ranges in X, Y, D, or H, where the range is the difference between the highest and lowest deviations from the regular daily variation. The K value for a given value of R is found from a table in which the location of the station has been taken into account to permit comparison between K values from different stations; the index reflects mainly auroral-zone activity, so stations nearer the auroral zone are much more sensitive. A two-letter subscript identifies the station. The K_{FR} index, from Fredericksburg, Virginia, has sometimes been used as a standard measure of activity for the United States. K values are integers ranging from 0 through 9; the calibration for six representative stations is given in Table 4-2; it may be noted that the K scale is roughly logarithmic with range.

The Kp index ("p" for "planetary") is probably the most widely used of all the indices. It was intended as a measure of the worldwide average level of activity; however, it is very sensitive to certain auroral-zone activity and insensitive to some other types of disturbance. It is based on the K

Table 4-2. Definition of the K scale for six representative observatories

Name of Observatory	Geomagnetic Lat.	Geomagnetic Long.	Lower Limit of R in nT for the K Value to be 0	1	2	3	4	5	6	7	8	9
Godhavn	79.9	32.5	0	15	30	60	120	210	360	600	1000	1500
Sitka	60.0	275.3	0	10	20	40	80	140	240	400	660	1000
Huancayo	−0.6	353.8	0	6	12	24	48	85	145	240	400	600
Fredericksburg	49.6	349.8	0	5	10	20	40	70	120	200	330	500
Tucson	40.4	312.2	0	4	8	16	30	50	85	140	230	350
Honolulu	21.1	266.5	0	3	6	12	24	40	70	120	200	300

Table 4-3. Value of the ap index for a given value of Kp.

If Kp =	0o	0+	1−	1o	1+	2−	2o	2+	3−	3o	3+	4−	4o	4+
then ap =	0	2	3	4	5	6	7	9	12	15	18	22	27	32
If Kp =	5−	5o	5+	6−	6o	6+	7−	7o	7+	8−	8o	8+	9−	9o
then ap =	39	49	56	67	80	94	111	132	154	179	207	236	300	400

indices from 12 stations between geomagnetic latitudes 48° and 63°, selected for good longitude coverage. Values of K are first used to find the Ks index ("s" for "standardized") from tables which remove the characteristic seasonal behavior at the station. The Ks index ranges continuously from 0.0 to 9.0 but is quoted in thirds of an integer by use of the symbols −, o, and +, such that the interval, for example, from 3.5 to 4.5 includes the Ks values 4−, 4o, and 4+. Ks can therefore assume the 28 values 0o, 0+, 1−, 1o, 1+, . . . , 9−, 9o. The Kp value for each 3-hour interval is the average of Ks from the 12 stations. The Kp index is published in two forms: numerical tables and the "musical-note" diagram reproduced in Figure 4-26; in the diagram each successive solar rotation is plotted under the previous one to expose 27-day recurrences of activity.

A newer index Km ("m" for "mondial", meaning global) has recently been developed as an improvement on Kp. It is an average of two component indices Kn and Ks (not to be confused with the standardized K index Ks) derived from northern- and southern-hemisphere stations, respectively. Some of its advantage lies in a simpler and more direct conversion of the data, but the major improvement comes from a much better geographical distribution of stations.

4.7.2 Related Indices ap, Ap, ak, Ak, am, and aa

Because the 3-hour K-type indices are defined with a roughly logarithmic scale, they are not suitable for simple averaging to obtain a daily index (though this is not uncommonly done). To convert to a roughly linear scale (that is, reconversion to an equivalent range), the ap index is defined from Kp by use of Table 4-3, and the daily index Ap is then defined to be the average of the eight 3-hour ap indices for the day.

In the same way, the single-station indices ak and Ak are found from K by use of Table 4-4. In this form, these indices are normalized for station location (because K is),

Table 4-4. Value of ak for a given value of K.

If K =	0	1	2	3	4	5	6	7	8	9
then ak =	0	3	7	15	27	48	80	140	240	400

and comparison between stations is convenient. To remove the normalization, these indices are sometimes multiplied by a calibration factor f, given by $f = R_9/250$, where R_9 is the lower limit of R for K = 9. Indices so adjusted are quoted in units of nanoteslas. Thus at Fredericksburg, f = (500 nT)/250 = 2 nT. As an example of the complete circle, when R = 27 nT, K = 3, and ak = 15 (dimensionless) or ak = 30 nT.

The am index is similarly related to Km (there are also indices an and as corresponding to Kn and Ks), but in this case the a-type indices are derived directly from the data and the corresponding K-type index follows; that is, the awkward conversion-reconversion to and from a logarithmic scale is avoided. The am index is probably, in many respects, the best worldwide index currently available.

The aa index is similar to the am index but is derived from only two observatories situated at approximately antipodal locations in England and Australia. Although its development in the early 1970s was motivated by the availability of records from two old observatories (Greenwich and Melbourne) and has resulted in an index which spans 115 years (1868 to the present), it has also proved to be an

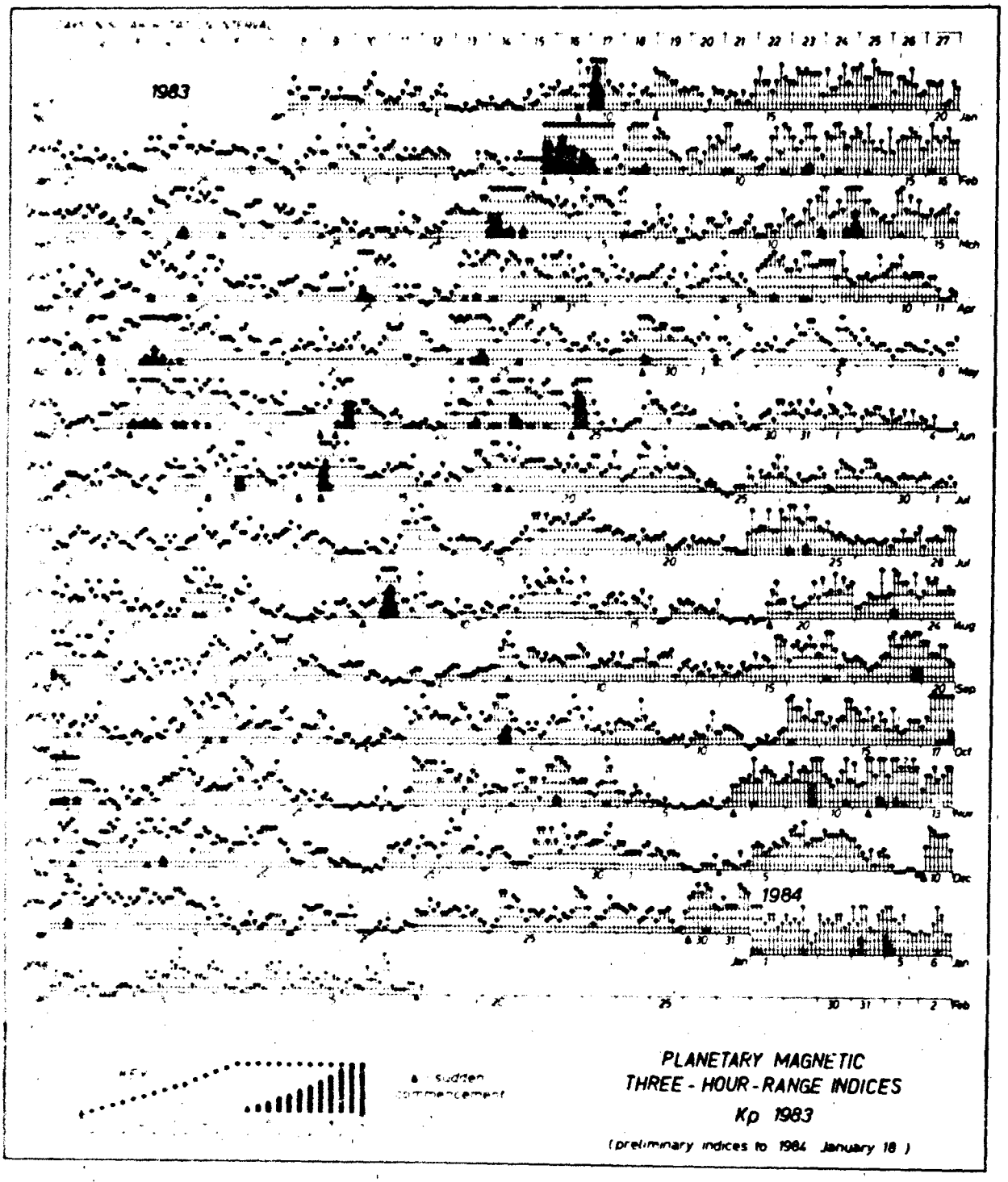


Figure 4-26. Musical note diagram of Kp values for the year 1983, produced by the International Service of Geomagnetic Indices

Table 4-5. Scale for finding Cp from the daily sum of ap. For each Cp interval the value listed is the upper limit of the daily-sum values in the interval.

If sum<	22	34	44	55	66	78	90	104	120	139	164	190	228
then Cp=	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2
If sum<	273	320	379	453	561	729	1119	1399	1699	1999	2399	3199	—
then Cp=	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5

easily obtainable but very good approximation to the am index.

4.7.3 Character Indices Cp, Ci, and C9

The Cp index, called the daily planetary character figure, is a number that ranges continuously from 0.0 to 2.5 and is quoted to tenths of an integer. It is found, using Table 4-5, from the daily sum (the sum of the eight 3-hour values) of the ap index. It restores the quasilogarithmic relationship originally introduced by the Kp scale and later removed by the ap scale. The peculiarity of its scale results from the intention that it should replace and be equivalent to a valuable older index, the Ci index, called the international daily character figure, available for all days since 1884. Cp is more reliable and objective, but usually differs from Ci by less than 0.2. Ci is obtained for each Greenwich day as the arithmetic mean of the C index values from a number of observatories; at each observatory the C index, called the daily magnetic character index, is reported as 0, 1, or 2 according to whether the UT day is quiet, moderately disturbed, or highly disturbed, respectively.

The C9 index, which can be found from Cp or Ci by use of Table 4-6, is sometimes used in their place especially for graphing, since it uses the more convenient single-digit zero-to-nine scale. Figure 4-27 shows a standard plot of the C9 index that, like the Kp musical-note diagram, shows recurrent 27-day activity.

4.7.4 Auroral-Electrojet Indices AE, AU, AL, and Ao

The AE index called the auroral electrojet activity index is derived from magnetograms from a number of auroral-zone stations distributed in longitude. It is intended to measure the strength of the auroral electrojet tabulated for hourly and (to some extent) 2.5-minute intervals. Obtained graphically from a single combined plot of the deviations at all stations of the element H from its normal quiet-time value, AE is the height of the envelope (that is, the difference) between the curves AU and AL (for "upper" and "lower") drawn through the maximum and minimum excursions of the deviations, respectively. The values of AU and AL are also used as indices, and another index Ao is defined to be the mean deviation, the curve midway between AU and AL.

4.7.5 Ring-Current Index Dst

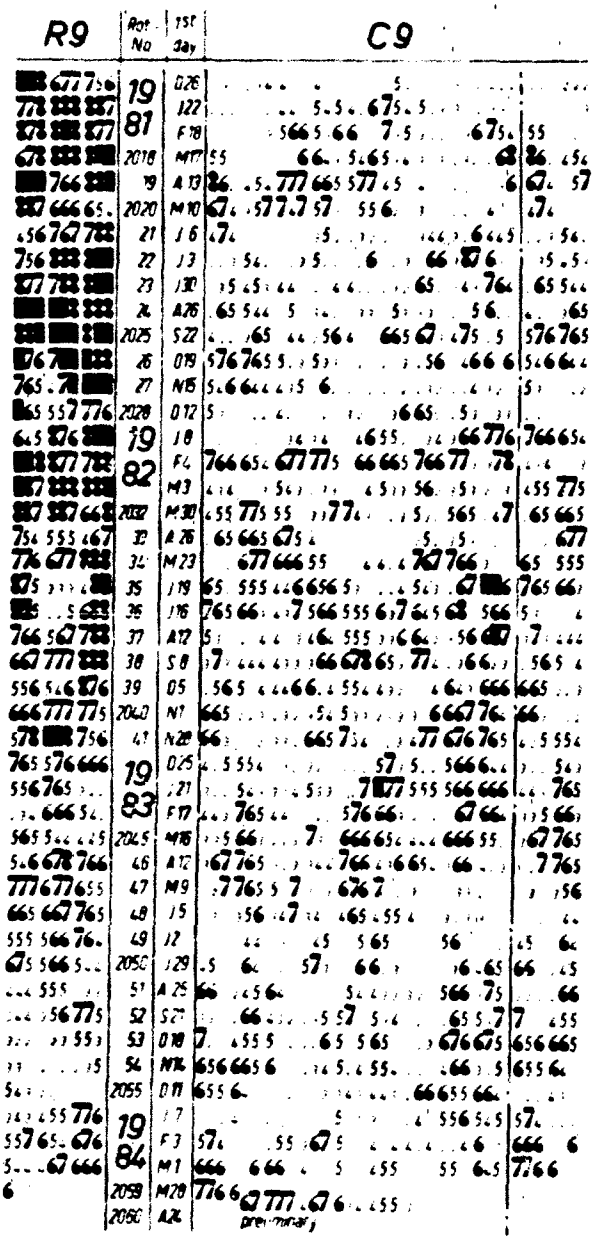
The most widely used and available index of low-latitude activity is the hourly Dst index. It is the magnitude of the normalized horizontal component of the Dst field (axially symmetric disturbance), as determined from the data obtained by four low-latitude observatories distributed in longitude. The Dst index was designed as a measure of the magnetospheric ring current of magnetic storms; therefore, high-latitude and equatorial stations are avoided to minimize the effects of auroral and equatorial electrojets. The derivation was described by Sugiura [1964], who since the IGY has been the principal source of this index.

4.7.6 Other Indices

The Q index is an index which provides better time resolution than those based on the 3-hour K indices, but its availability is limited. It is a measure of high-latitude magnetic activity for each 15-minute interval, assigned by a number of observatories above 58 degrees geomagnetic latitude. It is computed, from the more disturbed of the elements X and Y, as the sum S of the absolute values of the maximum positive and maximum negative deviations from the normal (quiet) curve, with the provision that if the "negative" ("positive") deviation does not indeed become negative (positive), it is considered to be zero. The conversion

Table 4-6. Scale for finding C9 from Cp or Ci.

If Cp or Ci =	0.0 or 0.1	0.2 or 0.3	0.4 or 0.5	0.6 or 0.7	0.8 or 0.9	1.0 or 1.1	1.2 thru 1.4	1.5 thru 1.8	1.9	2.0 thru 2.5
then C9 =	0	1	2	3	4	5	6	7	8	9



DAILY GEOMAGNETIC CHARACTER FIGURES C9 AND 3-DAY MEAN SUNSPOT NUMBERS R9 (after Bartels)

Figure 4-27 Plot of the C9 index and sunspot number R9 for the period 1981–1984.

scale for obtaining Q from S, applicable to all stations, is given in Table 4-7.

The indices U, u, u_1, and Δ are all derived from the difference between consecutive daily means of some element of the field. The U figure is the absolute magnitude of the difference between the daily means of the horizontal component for the given and preceding days. The u figure is a reduction of U to an equivalent equatorial value by the relation $u = U/(\sin\theta_m \cos\psi)$, where θ_m is geomagnetic colatitude and ψ is the angle between H and the magnetic meridian. The u_1 figure is obtained from u by a nonlinear conversion intended to reduce the influence of large disturbances. These indices have been useful historically for monthly and annual changes, but since they were intended to measure ring-current effects, the Dst index has largely superseded them.

The W measure applies to quiet variation fields rather than disturbance phenomena, being derived from the amplitude of the Sq variation. It is a measure of solar radiation effects and the intensity of the equatorial electrojet [Bartels, 1946].

4.8 SOURCES OF GEOMAGNETIC DATA

The following paragraphs provide some specific information on the availability of geomagnetic data. In recent years, the increasing emphasis on global studies of magnetospheric phenomena and the correlation of satellite and ground-based measurements have encouraged the extensive collection of such data at several data centers where they are made readily available. For data that are not archived there, the centers can often provide a referral to an alternative source.

4.8.1 Data Centers

Two national data centers in the United States have extensive holdings of geomagnetic data. The National Geophysical Data Center (NGDC) is operated by NOAA (Environmental Data and Information Service, NOAA, D63, 325 Broadway, Boulder, CO 80303, phone 303-497-6323). The NGDC is a recent successor to the National Geophysical and Solar-Terrestrial Data Center (NGSDC) and also serves internationally as the World Data Center A for Solar-Terrestrial Physics (WDC-A for STP). The National Space Science Data Center (NSSDC) is operated by NASA (Goddard Space Flight Center, Code 601, Greenbelt, MD 20771, phone 301-344-6695) and also serves as the World Data Center A for Rockets and Satellites. NGDC has magnetograms (as plots and computer tapes) and magnetic indices (in tabular, computer tape, and chart form). Both centers have space measurements (mostly nonduplicative) of mag

Table 4-7 Scale for finding Q from S. The listed value of S (in nT) is the upper limit for the corresponding value of Q

If S <	10	20	40	80	140	240	400	660	1000	1500	2000	—
then Q =	0	1	2	3	4	5	6	7	8	9	10	11

netospheric magnetic fields, electric fields, and ULF VLF phenomena. Both centers issue catalogs of their holdings and publications [Allen et al., 1982; National Space Science Data Center, 1981 and 1982].

Subcenters of the World Data Centers are also relevant to geomagnetism. They are organized into WDC-A (subcenters in the United States, including the two mentioned above), WDC-B (subcenters in the USSR, including one for solar-terrestrial physics at IZMIRAN, Moscow), and WDC-C (twenty subcenters in Europe, Japan, and India). WDC-A for STP now includes in its catalog the geomagnetic holdings of WDC-B2 (Moscow) and WDC-C2 (Kyoto).

4.8.2 International Organizations

The International Council of Scientific Unions (ICSU) includes 18 unions, which are nongovernmental organizations dedicated to scientific progress in various disciplines. Geomagnetism is a concern of the International Union of Geodesy and Geophysics (IUGG), which in the past has initiated such international collaborations as the International Geophysical Year and the International Magnetospheric Study. Within the IUGG a number of associations have been formed; geomagnetism falls within the concerns of the International Association of Geomagnetism and Aeronomy (IAGA). This association was originally the Section de Magnetisme et Electricite Terrestres (1919) and later the International Association of Terrestrial Magnetism and Electricity (1930) before taking its present name in 1954.

The IAGA has historically been a principal source of publications of geomagnetic data, including its Bulletins 12 series (covering 1940 through 1969) and Bulletins 32 series (covering 1970 through the present) of geomagnetic data, indices, rapid variations, and special intervals. The International Service of Geomagnetic Indices (ISGI) operates under supervision of the IAGA. Its services include the regular distribution of the Kp and C9 tables and charts from its Gottingen office (Institut für Geophysik, 34 Gottingen, W. Germany). Publications of both the IAGA and the ISGI are available from the IUGG Publications Office (39 ter rue Gay-Lussac, 75005 Paris, France). A catalog is available.

4.8.3 Magnetic Stations

Nearly 400 magnetic stations, data from which are held in the World Data Centers, are listed in Table 4-8, which is a condensation (omitting two sets of corrected coordinates) of Table II of the WDC-A catalog. Stations are listed in order of increasing uncorrected geomagnetic colatitude.

4.8.4 Magnetic Field Models

A number of magnetic-field models and computer programs for their use are now available from the National Space Science Data Center [see NSSDC, 1981, p. 55]. For many uses where the highest precision is not required hard-copy charts and tables may be more convenient. The IGRF 1975.0 field is available in this form. Barraclough and Fabiano [1977] present 14 charts (elements X, Y, Z, H, D, I, F and the secular variation in terms of their first time derivatives) and tables which list values for every five degrees of longitude and every two degrees of latitude. Publications resulting from the most current work with MAGSAT and other data can be expected from cooperative efforts at both NASA and the United States Geological Survey [Langel, 1980].

4.8.5 Activity Indices and Charts

At present, an excellent collection of indices has been assembled at the NGDC (Section 4.8.1). The widely used tables and charts furnished by the ISGI, as well as the publications of the IAGA, may be obtained through the IUGG Publications Office (see Section 4.8.2). A variety of sources current through 1971 was listed by Knecht [1972].

The periods for which various indices have been computed and published vary greatly, often depending on the resources available to the workers who have the capability. Even some of the currently most useful indices, such as AE and Dst, have been delayed. An enquiry as to the current coverage of any index (beyond that listed in the catalog) may be directed to the NGDC.

ACKNOWLEDGMENTS

The section on magnetic pulsations was written by H.J. Singer. Reference materials, recent data, and/or helpful comment were generously provided by R.H. Langel, J.C. Cain, and N.W. Peddie on current magnetic-field models; S. Matsushita on recent work on the Sq variation; T. Saito on the classification of pulsations; and J.H. Allen on the activities and holdings of the WDC-A for STP. The authors are very grateful for these contributions.

CHAPTER 4

Table 4-8 List of geomagnetic stations from which the WDC-A has data holdings, listed in order of increasing geomagnetic colatitude. Successive columns are observatory name, international three-letter code, and uncorrected geomagnetic colatitude and east longitude

[illegible]	[illegible]	1.75	12.66	RIO GRANDE	RIO	25.43	260.96	SEDDIN	SED	37.87 98.11
[illegible] AFB	[illegible]	[illegible]	9.81	NORTHWAY	NWA	25.62	295.52	GOTTINGEN	GTT	37.99 94.82
ALERT	ALE	4.02	359.17	MALIYE KARNAFULY	MKL	25.62	149.82	NIEMEGK	NGK	38.01 97.67
RESOLUTE BAY	RES	6.82	232.87	MELVILLE AFB	MLV	25.63	13.72	DOURBES	DOU	38.30 88.88
NORTH POLE [illegible]	[illegible]	6.41	[illegible]	KOTZEBUE	KOT	25.99	243.55	MANHAY	MAB	38.32 90.04
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	9.28	152.90	ISLAND LAKE	ISL	26.16	326.66	CASPER	CSR	38.42 316.25
[illegible]	[illegible]	10.65	58.65	HEALY	HEA	26.19	258.17	MINSK	MNK	38.70 111.42
[illegible] BAY	[illegible]	10.69	268.34	[illegible]	[illegible]	26.53	120.77	UGUIT	UGT	38.72 153.19
PELLY BAY	PEL	11.87	322.55	BELLY ISLAND	BEY	26.60	155.68	COLLMBERG	CLL	38.78 97.61
[illegible]	[illegible]	11.92	143.85	NORWAY HOUSE	NHO	26.62	322.51	YAKUTSK	YAK	38.79 194.94
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	12.19	311.67	[illegible]	[illegible]	26.65	126.55	VAL JOYEUX	VJ	39.01 89.70
JOHNSON POINT	JOH	12.04	272.79	CAPE SCHMIDT	CPS	26.89	228.72	MAGADAN	MGD	39.02 211.57
[illegible] BAY	[illegible]	13.18	237.31	DEERING ISLAND	DIK	26.90	162.52	UKHTA	UKH	39.21 153.04
[illegible]	[illegible]	13.64	208.33	LOVOZERO	LOZ	27.21	127.74	P. TUNGUSKA	POD	39.23 165.40
[illegible]	[illegible]	14.17	30.74	KANDALAKSHA	KND	27.56	125.02	MOSCOW	MOS	39.28 121.51
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible] BAY	[illegible]	14.57	199.00	[illegible]	[illegible]	27.75	80.64	SWIDER	SWI	39.61 105.69
[illegible]	[illegible]	14.60	82.63	LERWICK	LER	27.81	258.60	BELSK	BEL	39.80 105.13
[illegible]	[illegible]	14.60	267.87	TALKEETNA	TLK	27.97	238.48	NANTES	NTS	39.82 81.39
[illegible]	[illegible]	14.60	131.33	CAPE WELLEN	CWE	27.99	101.06	CHAMBON-LA-FORET	CLF	39.83 89.58
NEW [illegible]	[illegible]	14.60	274.22	DOMBAS	DOB	27.99	902.97	ANDREWS AFB	AMS	39.95 352.15
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	15.44	131.69	KHARASAVEY	KHS	28.21	192.72	CHELTENHAM	CLH	40.06 352.24
CAPE PARRY	CPY	16.00	271.96	TAMSUY	TMA	28.28	155.77	PRUHONICE	PRU	40.36 99.43
[illegible] LAKE	[illegible]	16.17	317.73	ANCHORAGE	ANC	28.78	259.73	CARROLLTON	CAX	40.39 332.14
[illegible]	[illegible]	16.49	326.72	AMEDNMA	AMD	29.27	147.69	FREDERICKSBURG	FRD	40.58 351.58
NORTH POLE 6	[illegible]	17.17	174.54	TIKIE BAY	TIK	29.33	192.38	BELOIT	BLT	40.75 326.63
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible] INLET	[illegible]	17.19	324.72	SEYAHA	SEY	29.69	195.52	KAZAN	KZN	40.76 131.37
NORTH POLE 12	[illegible]	17.50	219.58	SITKA	SIT	29.75	277.08	BOULDER	BOU	40.96 318.24
NORTH POLE 15	[illegible]	18.00	204.65	WHITESHELL	WHS	30.29	327.44	BOROV	BOV	41.12 97.51
TIKHAYA BAY	TAH	18.48	153.60	WINNIPEG	WNP	30.61	325.12	BALDWIN	HAL	41.14 330.29
HEISS ISLAND	HIS	18.68	156.47	GLENLEA	GLL	30.61	325.16	BURLINGTON	BRT	41.41 321.97
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
NORTH POLE 8	[illegible]	18.68	223.65	ARKHANGELSK	ARK	31.16	129.25	FURSTENFELDBRUCK	FUR	41.45 94.48
[illegible]	[illegible]	19.18	38.56	CAPE KENNEDY	CKA	31.34	359.54	ARTI	ARS	41.52 139.57
[illegible]	[illegible]	19.00	124.35	NORILSK	NOR	31.40	169.68	SVERDLOVSK	SVD	41.54 141.70
[illegible]	[illegible]	19.03	325.51	LORING AFB	LOR	31.77	2.97	CLIMAX	CLI	41.85 317.35
NARSSARSSUAQ	NAQ	19.22	58.70	ST JOHNS	STJ	31.77	22.95	LEADVILLE	LDV	41.93 317.26
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	19.29	267.10	ENDERBY	END	31.77	106.22	CHATONNAYE	CTE	42.01 89.50
[illegible]	[illegible]	19.13	37.24	ESKDALEMUIR	ESK	31.85	88.05	TEVRIZ	TEV	42.02 150.97
[illegible] ISLAND	[illegible]	19.31	254.77	LVOV	LVV	32.15	106.77	LVOV	LVV	42.19 106.96
[illegible]	[illegible]	20.10	72.13	NURMIJARVI	NUR	32.32	119.53	PRICE	PCU	42.23 312.07
NEWMAN WELLS	NOW	20.67	279.83	NYDA	NDA	33.14	154.34	WIEN-KOBENZL	WIK	42.33 99.34
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	20.81	292.35	OTTAWA	OTT	33.34	353.43	KIEV	KIV	42.62 113.25
[illegible]	[illegible]	20.81	292.38	STONYHURST	STO	33.36	83.87	NAGYCENK	NCK	43.02 99.44
[illegible]	[illegible]	21.12	242.60	VALENTIA	VAL	33.68	74.70	ADAK	ADA	43.04 241.60
[illegible]	[illegible]	21.28	325.30	LENINGRAD	LNN	33.92	118.29	HURBANOVO	HRB	43.08 100.92
NORTH POLE 10	[illegible]	21.37	211.09	SLUTZK	SLU	34.15	117.94	O'GYALLA	OGY	43.08 100.92
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
ARCTIC VILLAGE	AVI	21.77	256.17	RUDE SKOV	RSV	34.39	99.58	BUDAKESZI	BUZ	43.55 101.45
[illegible]	[illegible]	22.39	325.46	[illegible]	[illegible]	34.58	99.60	TIHANY	THY	43.95 100.25
[illegible]	[illegible]	22.56	113.41	NEWPORT	NEW	34.74	301.85	TOMSK	TMK	44.04 160.64
[illegible]	[illegible]	22.31	288.15	AGINCOURT	AGN	35.07	348.84	TULSA	TUL	44.06 330.06
CAPE CHELANIA	CCH	22.77	157.50	VICTORIA	VIC	35.63	294.73	LOGRONO	LGR	44.25 78.46
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	23.04	117.42	HARTLAND	HAD	35.69	80.20	MONTE CAPELLINO	MCP	44.46 90.71
FORT [illegible]	[illegible]	23.07	258.85	[illegible]	[illegible]	35.71	95.19	GENOVA	GEN	44.46 90.71
[illegible]	[illegible]	23.22	325.64	[illegible] HAVEN	WHV	35.75	94.14	ROBURENT	ROB	44.49 89.56
[illegible]	[illegible]	23.62	121.06	[illegible]	[illegible]	36.03	152.03	CASTELLACCIO	CAO	44.57 90.64
GREAT WHALE R.	[illegible]	23.64	349.46	GREENWICH	GRW	36.08	85.02	SAO MIGUEL	SMG	44.70 52.30
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	23.53	133.18	WITTEVEEN	WIT	36.12	92.32	PETROPAVLOVSK	PET	45.02 219.87
[illegible] LAKE	[illegible]	23.84	316.90	ABINGER	ABN	36.29	84.47	POLA	POL	45.11 99.59
[illegible]	[illegible]	23.77	316.30	WESTON	WES	36.30	358.86	COIMBRA	COI	45.29 71.59
[illegible]	[illegible]	24.15	325.60	DE BILT	DRN	36.51	90.48	ESPANOLA	EPN	45.32 318.33
[illegible]	[illegible]	24.91	116.38	SREDNIKAN	SRE	36.63	211.76	NOVOSIBIRSK	NVS	45.33 158.78
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
FORT SMITH	FSM	24.32	299.97	HEL	HLP	36.79	104.74	CASTLE ROCK	CRC	46.36 300.28
[illegible]	[illegible]	25.06	259.14	KALININGRAD	KNG	37.04	106.45	EBRO	EBR	46.41 80.91
[illegible] AFB	[illegible]	25.09	293.03	BOROK	BOK	37.12	124.15	TOLEDO	TOL	46.44 75.95
PETSAMO	PTS	25.21	126.47	BOCHUM	BOC	37.50	92.15	ODESSA	ODE	46.51 112.23
[illegible]	[illegible]	25.43	122.68	POTSDAM	POT	37.79	98.21	GROCKA	GCK	46.69 102.08

Table 4-8. (Continued)

DALLAS	DAL	47.05	329.44	FUQUENE	FUQ	75.35	356.55	PORT MORESBY	PMG	108.28	219.50
L'AQUILA	AQU	47.41	94.14	PARAMARIBO	PAR	75.27	15.79	TSUMEB	TSU	108.46	84.24
SUBIACO	SUA	47.68	107.22	FREETOWN	FTN	75.54	59.15	PILAR	PIL	110.42	9.93
LA MADDALENA	LMD	47.75	89.92	LUNPING	LNP	76.11	190.90	TANANARIVE	TAN	113.89	113.96
MT WILSON	MWC	48.64	309.48	UJJAIN	UJJ	76.51	148.22	LAS ACACIAS	LAS	113.93	11.99
SIMFEROPOL	SIM	48.97	114.45	ZARIA	ZAR	76.73	80.47	MAURITIUS	MRT	116.75	125.38
IRKUTSK	IRT	49.01	175.69	KONTAGORA	KNT	77.05	78.15	PLAISANCE	PLS	117.11	124.04
SAN FERNANDO	SFS	49.14	72.64	WAKE ISLAND	WKE	77.18	234.47	MAPUTO	LMM	117.92	97.51
CAPRI	CPI	49.35	98.98	CHAPA	CPA	78.92	174.47	TRELEW	TRW	121.97	4.97
PANAGYURISHTE	PAG	49.57	104.54	HONG KONG	HKC	79.09	184.35	CAPE GOOD HOPE	CGH	122.97	81.41
TUCSON	TUC	49.49	319.44	KUMASI	KUM	79.35	70.16	CAPE TOWN	CTO	122.99	81.39
KARAGANDA	KSU	49.66	149.76	IBADAN	IBD	79.66	76.03	HERMANUS	HER	123.59	81.99
ALMERIA	ALM	49.68	76.58	TATUOCA	TTN	80.70	22.19	BRISBANE	BRS	125.44	228.32
ISTANBUL-KANDILLI	IST	51.70	108.69	ALIBAG	ABG	80.76	144.58	WATHEROO	WAT	131.54	187.28
KHABAROVSK	KHB	51.92	201.24	HYDERABAD	HYB	82.39	150.24	MAGALLANES	MGS	131.88	359.61
AVINOFES	AVI	52.20	70.45	TALARA	TAL	83.91	349.08	GNANGARA	GNA	133.02	187.40
NOVOKAZALINSK	NKK	52.41	132.55	FUSHUN	FUS	84.27	52.09	CANBERRA	CAN	133.66	275.74
TUZHNO SAKHALINS	YSS	52.82	207.32	PALMYRA ISLAND	PAT	84.56	267.21	SOUTH GEORGIA	SGE	134.51	27.19
ULAN BATOR	URA	53.16	177.63	MOJA	MOJ	84.58	79.34	TOOLANGI	TOO	136.34	222.32
TBILISI	TFS	53.47	123.24	BAGUIO	BAG	84.70	190.60	AMBERLEY	AML	137.34	255.80
PENDELI	PEG	53.64	102.91	ADDIS ABABA	AAE	84.92	110.51	EYREWELL	EYR	137.67	253.52
KZIL-AGACH	KZA	54.59	153.76	RANGPI	RNG	85.46	89.97	CHRISTCHURCH	CHR	137.72	253.84
CANARIAS	TEN	55.34	59.99	CHICLAYO	CCL	85.63	350.65	MARION ISLAND	MRN	139.24	95.90
SANTA CRUZ	SCI	55.34	59.99	GUAM	GUA	85.73	214.26	LAUDER	LAU	139.72	251.36
MEMAMBETSU	MMB	55.70	209.70	FANNING	FAN	85.97	270.24	ORCADAS DEL SUR	ORC	140.25	19.31
HAVANA	HVN	55.97	346.94	ANTIPOLO	ANO	86.50	191.22	PORT ALFRED	CZT	141.51	110.57
ALMA ATA	AAA	56.53	151.88	MANILA	MAN	86.53	191.04	GONZALES VIDELA	GVD	143.57	5.41
VLADIVOSTOK	VLA	56.90	199.56	MINTINUPA	MUT	86.73	191.09	ARGENTINE ISLAND	AIA	145.93	4.35
TASHKENT	TKT	57.64	145.19	BANGKOK	BKK	87.46	171.23	CAMPBELL ISLAND	CAI	146.97	294.44
ASHKHABAD	ASH	59.56	134.30	CHIMBOTE	CHB	87.89	351.91	PORT-AUX-FRANCAIS	KGL	147.43	129.83
KSARA	KSA	60.00	112.96	MAJURO	MJR	88.35	241.05	MACQUARIE ISLAND	MCQ	150.75	244.61
TEKELYUCAN	TLO	60.14	329.62	ANNAMALAINAGAR	ANN	88.61	150.45	HEARD ISLAND	HIS	151.48	131.91
WIELDES	WDS	60.58	5.41	KODAIKANAL	KOD	89.40	148.45	SANAE	SNA	153.95	45.30
SAN JUAN	SJG	63.59	4.65	JARVIS ISLAND	JRV	90.30	270.45	EIGHTS	EGS	153.96	356.08
TEHRAN	TEH	60.75	127.73	ETTAYAPURAM	ETT	90.57	148.84	NORWAY STATION	NWS	154.11	49.09
MIZUSAWA	MIZ	60.84	207.69	CEBU	CEB	90.68	194.08	LAZAREV	LZV	155.98	56.69
MEJZIMA	MJI	61.23	185.75	HUANCAYO	HUA	90.77	355.13	HALLEY BAY	HBA	156.06	29.25
OHAGAWA	OHA	61.39	208.12	TRIVANDRUM	TRD	91.08	147.75	NOVOLAZAREVSKAYA	NVL	156.51	54.94
MOOSTIN	MOO	61.71	111.65	CUZCO	CUS	92.21	358.49	ROI BAUDOUIN	RBD	158.33	64.67
NITSANIM	NGM	61.81	111.24	MOGADISCIO	MGC	92.90	115.68	SYOWA STATION	SYO	159.96	79.38
AMATSIA	AMT	62.04	111.49	KOROR	KOR	92.95	204.73	MOLODEZHNAYA	MOL	159.97	57.36
MEIWAN	MEW	62.42	207.69	BINZA	BIN	93.45	84.91	BYRD STATION	BYR	160.62	336.71
MISALLAT	MLT	63.28	107.20	DAVAO	DAV	93.81	195.90	MAWSON	MAW	163.31	105.08
PAKHERA	PAK	63.69	207.25	LWIRO	LWI	94.01	98.60	LITTLE AMERICA	LAA	163.95	312.23
TORVO	TOR	64.21	206.90	YAUCA	YAU	94.24	355.88	LENINGRADSKAYA	LEN	164.33	261.97
KANIYAN	KNZ	64.68	207.19	NAIROBI	NAI	94.63	106.47	CAPE HALLETT	HLL	164.47	278.80
TAMANRASSET	TAM	64.92	81.84	CHARACATO	ARE	95.15	358.97	DUMONT DURVILLE	DRV	165.33	232.38
MIYAKO	MIY	65.02	202.81	LA PAZ	LPB	95.19	2.22	MIRNY	MIR	167.02	149.66
CANCHOU	LZH	65.17	174.87	PENANG	PNG	95.68	170.81	OASIS	OAS	167.37	153.53
MIDWAY	MDY	65.59	241.75	LUANDA	LUA	97.49	91.96	PLATEAU	PTU	167.53	93.79
SIMOSATO	SSO	66.69	207.74	DAR ES SALAAM	DRS	100.44	108.90	WILKES	WIL	167.57	181.35
HACHIJO	HJO	66.81	207.29	LA QUIACA	LQA	100.92	4.54	CHARCOT	CTX	167.94	235.93
HATIZYO	HTY	66.86	207.57	VASSOURAS	VSS	102.22	25.21	SOUTH POLE	SPA	168.68	0.00
AGO	AGO	67.74	199.16	HOLLANDIA	HNA	102.25	211.70	SCOTT BASE	SBA	168.80	294.35
QUETTA	QUE	68.40	140.97	RIO DE JANEIRO	RDJ	102.77	25.60	MCMURDO	MCM	168.82	294.31
MOHO-BILU	MHB	68.68	267.97	SAN JOSE	SJO	102.86	23.07	PIONERSKAYA	PIO	170.33	149.88
MEANO	MAO	69.28	56.81	ELIZABETHVILLE	ELI	102.94	95.37	VOSTOK	VOS	170.41	111.76
COSTA RICA	CRO	69.93	345.69	KARAVIA	KVA	102.94	95.37				
KANOYA	KNY	69.20	199.36	TANITI	TAH	105.11	284.04				
SARMANLA	SAM	69.44	150.35	PAMATAI	PPT	105.12	284.09				
[illegible]	[illegible]	69.80	190.80	APIA	API	105.35	261.58				
ZIKAWEI	ZKW	69.88	185.27	KUYPER	KUY	107.14	127.02				
SHESHAN	SHS	70.01	194.65	TANGERANG	TNG	107.47	176.01				
CHICHIJIMA	CHI	72.47	210.19	BATAVIA	BTV	107.49	177.12				
JAIPUR	JAI	72.63	148.68	EASTER ISLAND	EIC	109.19	323.98				

CHAPTER 4

REFERENCES

Alldredge, L.R. and L. Hurwitz, "Radial Dipoles as the Sources of the Earth's Main Magnetic field," *J. Geophys. Res.*, **69:** 2631, 1964.

Allen, J.H., C.C. Abston, Z.R. Kharin, and N. Papitashvili, International Catalog of Geomagnetic Data, *Report UAG-86*, World Data Center A for Solar-Terrestrial Physics, Boulder, CO, 1982.

Axford, W.I., "Tail of the Magnetosphere," *J. Geophys. Res.*, **70:** 1231, 1965.

Barraclough, D.R. and E.B. Fabiano (eds) Grid Values and Charts for the IGRF 1975.0, *IAGA Bulletin No. 38*, 1977. (Available not from IUGG but as NTIS No. PB-276 63x.)

Bartels, J., "Geomagnetic Data on Variations of Solar Radiation: Part I—Wave Radiation," *Terr. Magn. Atmos. Elec.*, **5:** 181, 1946.

Bullard, E.C., C. Freedman, H. Gellman, and J. Nixon, "The Westward Drift of the Earth's Magnetic Field," *Phil. Trans. Roy. Soc.*, **243A:** 67, 1950.

Chapman, S. and J. Bartels, *Geomagnetism*, Oxford University Press, London, 1940.

Chen, L. and A. Hasegawa, "A Theory of Long-Period Magnetic Pulsations, 1, Steady State Excitation of Field Line Resonance," *J. Geophys. Res.*, **79:** 1024, 1974.

Cummings, W.D., R.J. O'Sullivan, and P.J. Coleman, Jr., "Standing Alfven Waves in the Magnetosphere," *J. Geophys. Res.*, **74:** 778, 1969.

Dungey, J.W., "Electrodynamics of the Outer Atmosphere," *Sci. Rep. 69*, Ionos. Res. Lab., Pennsylvania State University, 1954. (Also partially contained in Dungey, J.W., "The Structure of the Exosphere, or Adventures in Velocity Space," *Geophysics, The Earth's Environment*, edited by DeWitt, Hubblot and Lebeau, Gordon and Breach, New York, 1963).

Greenwald, R.A. and A.D.M. Walker, "Energetics of Long Period Resonant Hydromagnetic Waves," *Geophys. Res. Lett.*, **7:** 745, 1980.

Gustafsson, G., "A Revised Corrected Geomagnetic Coordinate System," *Ark. Geofys.*, **5:** 595, 1970.

Hakura, Y., "Tables and Maps of Geomagnetic Coordinates Corrected by the Higher Order Spherical Harmonic Terms," *Rep. Ionosph. Space Res. Japan*, **19:** 121, 1965.

Hughes, W.J., "Pulsation Research During the IMS," *Rev. Geophys. Space Phys.*, **20:** 641, 1982.

Hughes, W.J., "Hydromagnetic Waves in the Magnetosphere", submitted to U.S. National Report to 18th General Assembly of IUGG Hamburg, Aug. 1983.

IAGA Division I Working Group 1, International Geomagnetic Reference Fields: DGRF 1965, DGRF 1970, DGRF 1975, and IGRF 1980, *Eos Trans AGU* **62:** 1169, 1981.

IAGA, "Review Papers from the ULF Symposium held at the 4th IAGA Scientific Assembly, Edinburgh, 14 August 1981," *Planet. Space Sci.*, **30:** 1199–1258, 1982.

Jacobs, J.A., *Geomagnetic Micropulsations*, Springer-Verlag, New York, 1970.

Knecht, D.J., "The Geomagnetic Field," *AFCRL-72-0570*, AD 759837, 1972.

Langel, R.A., "Magsat Scientific Investigations" *NASA Technical Memorandum 80698*, May, 1980.

Langel, R.A., private communication, 1982.

Lanzerotti, L.J. and D.J. Southwood, "Hydromagnetic Waves," in *Solar System Plasma Physics*, III, 109, edited by E.N. Parker, C.F. Kennel, L.J. Lanzerotti, North Holland, Amsterdam, 1979.

Matsushita, S., "Solar Quiet and Lunar Daily Variation Fields," in *Physics of Geomagnetic Phenomena*, edited by S. Matsushita and W.H. Campbell, Academic Press, New York, 1967.

Matsushita, S., "Morphology of Slowly Varying Geomagnetic External Fields—A Review," *Phys. Earth Plan. Int.*, **10:** 299, 1975.

Matsushita, S. and H. Maeda, "On the Geomagnetic Solar Quiet Daily Variation Field During the IGY," *J. Geophys. Res.*, **70:** 2535, 1965a.

Matsushita, S. and H. Maeda, "On the Geomagnetic Lunar Daily Variation Field," *J. Geophys. Res.*, **70:** 2559, 1965b.

Mayaud, P.N., *Derivation, Meaning, and Use of Geomagnetic Indices*, (Geophysical Monograph 22) AGU, Washington, D.C., 1980.

Nagata, T., "Two Main Aspects of Geomagnetic Secular Variation—Westward Drift and Nondrifting Components" in *Benedum Earth Magnetism Symposium* edited by T. Nagata, University of Pittsburgh Press, 1962.

Nagata, T. and M. Ozima, "Paleomagnetism," in *Physics of Geomagnetic Phenomena* edited by S. Matsushita and W.H. Campbell, Academic Press, New York, 1967.

National Space Science Data Center, NSSDC Data Listing, *NSSDC Report 81-11*, Goddard Space Flight Center, Greenbelt, Md., September 1981.

National Space Science Data Center, Data Catalog Series for Space Science and Applications Flight Missions, *NSSDC Report 82-21* (Vol. 1A) and *NSSDC Report 82-22* (Vol. 2A), Goddard Space Flight Center, Greenbelt, Md., September 1982.

Nishida, A., *Geomagnetic Diagnosis of the Magnetosphere*, Springer-Verlag, New York, 1978.

Olson, W.P. and K.A. Pfitzer, "A Quantitative Model of the Magnetospheric Magnetic Field," *J. Geophys. Res.*, **79:** 3739, 1974.

Olson, W.P., and K.A. Pfitzer, "Magnetospheric Magnetic Field Modeling," McDonnell-Douglas Astronautics Co. Rept., 1977.

Orr, D., "Magnetic Pulsations Within the Magnetosphere: A Review," *J. Atmos. Terr. Phys.*, **35:** 1, 1973.

Orr, D., (ed.), "Geomagnetic Pulsations: Papers from 7th Annual E.G.S. Meeting," *J. Atmos. Terr. Phys.*, **43**(9), 1981.

Parkinson, W.D. and J. Cleary, "The Eccentric Geomagnetic Dipole," *Geophys. J. Roy. Astron. Soc.*, **1:** 346, 1958.

Saito, T., "Geomagnetic Pulsations," *Space Sci. Rev.*, **10:** 319, 1969.

Saito, T., "Long-period Irregular Magnetic Pulsations, Pi 3," *Space Sci. Rev.*, **21:** 427, 1978.

Shawhan, S.D., "Magnetospheric Plasma Wave Research 1975–1978," *Rev. Geophys. Space Phys.* **17:** 705, 1979.

Singer, H.J., "Multisatellite Observations of Resonant Hydromagnetic Waves," *Planet. Space Sci.*, **30:** 1209, 1982.

Singer, H.J., D.J. Southwood, R.J. Walker, and M.G. Kivelson, "Alfvén Wave Resonances in a Realistic Magnetospheric Magnetic Field Geometry," *J. Geophys. Res.*, **86:** 4589, 1981.

Southwood, D.J., "Some Features of Field Line Resonances in the Magnetosphere," *Planet. Space Sci.*, **22:** 483, 1974.

Southwood, D.J., "Waves, Wave Particle Interactions and Magnetic Oscillations," *IAGA Report*, International Association of Geomagnetism and Aeronomy, Dec. 1979.

Southwood, D.J. (ed.), "ULF Pulsations in the Magnetosphere," *J. Geomagn. Geoelec.*, **32**, Suppl. II, 1980.

Southwood, D.J. and W.J. Hughes, "Theory of Hydromagnetic Waves in the Magnetosphere," *Space Sci. Rev.*, **35:** 301, 1983.

Southwood, D.J. and W.F. Stuart, "Pulsations at the Substorm Onset," *Dynamics of the Magnetosphere*, edited by S.-I. Akasofu, D. Reidel, Dordrecht, Holland 1980.

Sugiura, M., "Hourly Values of Equatorial Dst for the IGY," *Annals of the Int. Geophys. Year*, **35:** 49, 1964.

Tsyganenko, H.A., "A Model of the Cis-Lunar Magnetospheric Field," *Ann. Geophys.*, **32:** 1, 1976.

Vestine, E.H., "Influence of the Earth's Core upon the Rate of the Earth's Rotation," in *Benedum Earth Magnetism Symposium*, edited by T. Nagata, University of Pittsburgh Press, Pittsburgh, 1962.

Chapter 5

THE RADIATION BELTS

W.N. Spjeldvik and P.L. Rothwell

Interaction of the solar wind flow with the earth's magnetic field gives rise to a cavity in the interplanetary medium known as the earth's magnetosphere. Within this cavity there exists a limited region where the motion of energetic particles is confined by the earth's magnetic field. This region comprises the earth's *radiation belts* as depicted in Figure 5-1. The radiation belt region contains electrons, protons, helium, carbon, oxygen, and other ions with energies from less than 1 keV to hundreds of MeV. Particles below 200 keV energy represent the principal corpuscular energy density and form the extraterrestrial ring current. Confinement (or trapping) of these particles results from the dipolar-like topology of the geomagnetic field which is characterized by magnetic field lines that converge at high latitudes towards the poles resulting in a relative minimum magnetic field strength region in the vicinity of the geomagnetic equator. Figure 5-2 illustrates principal aspects of a charged particle trajectory in magnetic mirror field geometry. The magnetic force ($\mathbf{F} \propto \mathbf{V} \times \mathbf{B}$) deflects the particle velocity vector $\mathbf{V}$ so that the particle spirals around the magnetic field $\mathbf{B}$. The convergence of the magnetic lines of force causes a tightening of the spiral angle and eventually a reflection of the particle from the high magnetic field region (mirror point); for this reason the earth's magnetic field is capable of confining charged particles. A detailed account of single particle motion in magnetic fields is found in Roederer [1970].

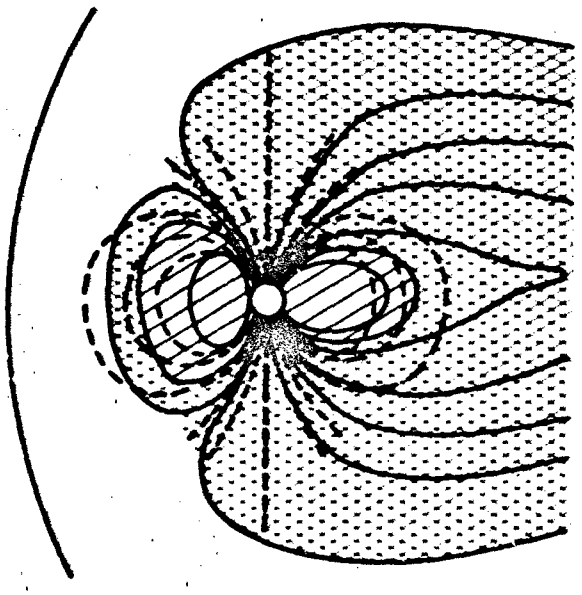


Figure 5-1. Cross-section of the earth's magnetosphere in the noon-midnight meridian showing relative locations (lightly shaded regions) of the earth's radiation belts in the overall magnetospheric topology.

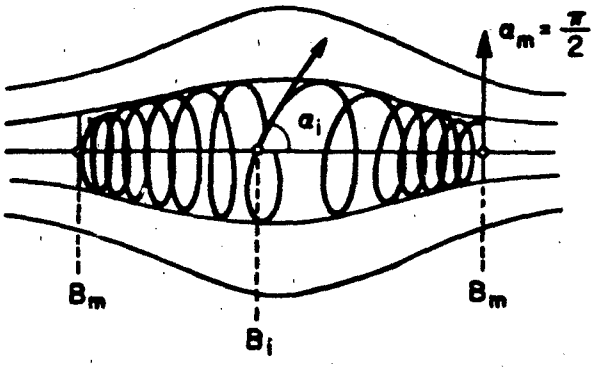


Figure 5-2. A charged particle trajectory in a magnetic "bottle". Conservation of the first adiabatic invariant can cause the spiraling particle to be reflected where the magnetic field is stronger. This causes the particle to be trapped by the magnetic field.

To a fair approximation, the earth's magnetic field in the radiation belt region can be described in terms of a magnetic dipole located near the center of the earth. The dipole moment is $M = 0.312\ G\ R_I^3$, and the dipole is directed so that the magnetic south pole on the earth's surface is located in northern Greenland (geographic coordinates: 78.5° N, 291° E); on the earth, the northern end of the compass needle points to this location. The spatial distribution of the dipolar magnetic field strength is

$$B = B_I \left(\frac{R}{R_I}\right)^{-3} \frac{\left[4 - 3\cos^2\lambda\right]^{1/2}}{\cos^6\lambda} \qquad (5.1)$$

where R is the radial distance measured from the center of the earth, R_E is the radius of the earth, $B_E = 0.312$ G is the equatorial field at $R = R_E$ and λ is magnetic latitude. A detailed account of the earth's magnetic field and its variability is found in Chapman and Bartels [1951]. Since the geomagnetic field is inhomogeneous, a radiation belt particle experiences varying magnetic field strengths over its trajectory. Field variations on a length scale of the order of the particle's gyroradius cause a net drift across the magnetic field in the azimuthal direction around the earth as illustrated in Figure 5-3. This is a direct result of the field strength being greater closer to the earth causing

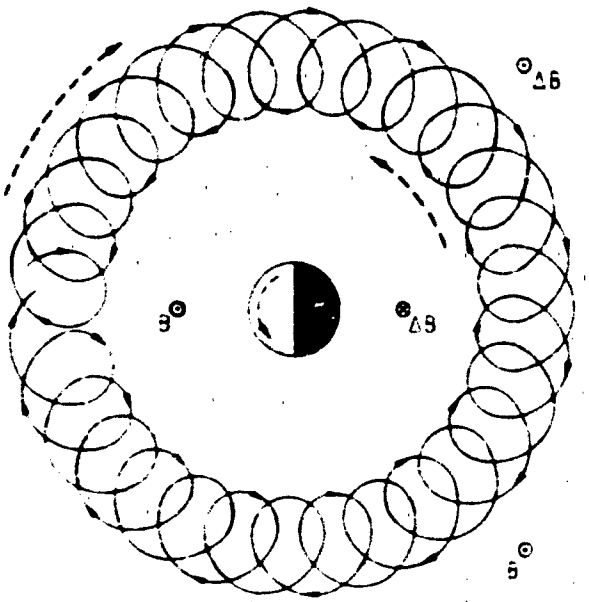


Figure 5-3. Schematic representation of the gyration and azimuthal drift (solid curve) of an equatorially mirroring proton with associated current patterns (dashed curves) [Shulz and Lanzerotti, 1974].

the particle orbital radius of curvature to be less there. The direction of the magnetic force depends on the sign of the particle charge: electrons drift eastward and positive ions drift westward. Thus the energetic trapped particles are spread out in a belt-like configuration around the earth, forming the *radiation belts*.

Radiation belt electrons move at very high speeds. For example, electrons with a kinetic energy of 500 keV move at 85% the speed of light. Radiation belt electrons must thus be studied using relativistic theory. Ions, because they are substantially heavier, generally move at subrelativistic velocities; at 500 keV the proton speed is 3% of the speed of light, while the heavier ions are even slower at the same energy.

The composition and flux intensities of the earth's radiation belts are determined by the strength of the sources, internal transport processes, and loss mechanisms. The outer boundary of the radiation belt trapping region occurs at the point where the magnetic field is no longer able to maintain stable trapping, and at low altitudes the earth's atmosphere forms an effective boundary for radiation belt particles. During geomagnetically quiet conditions, the radiation belt region extends from the top of the atmosphere along dipole field lines to an equatorial radial distance of at least 7 earth radii. Energetic ions and electrons that encounter the dense atmosphere collide with the atmospheric constituents and are readily lost from the radiation belts. Particles with mirror points well inside the atmosphere (nominally below ~100 km altitude) are said to be within the atmospheric (bounce) *loss cone*. Kinematically, the angle between the velocity vector of such particles and the magnetic field direction (the pitch angle) at the equator is below a certain value, known as the loss cone angle. Figure 5-4 exemplifies this.

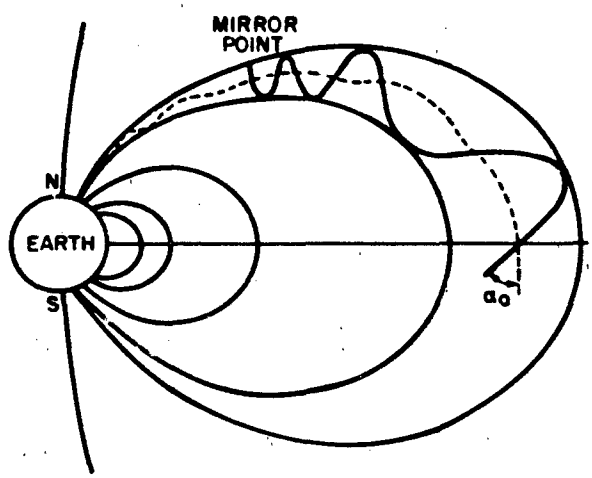


Figure 5-4. Illustration of magnetic mirroring in a dipolar magnetic field. The single particle trajectory shown in the solid line is for a particle outside the atmospheric bounce loss cone and the dashed line represents the trajectory of a particle inside the loss cone. The latter particle will encounter the denser parts of the earth's atmosphere (mirror point height nominally below 100 km) and will thus precipitate from the radiation belts.

The earth's radiation environment is best studied with combined experimental and theoretical means. On one hand, it is impossible to encompass the entire magnetosphere by experimental techniques or even to measure all the physical parameters that may have bearing on the dynamical phenomena; on the other hand, this environment is so complex that there can be little hope of theoretically predicting the total radiation belt behavior solely from a set of mathematical postulates. For example, the governing diffusion equations describing the trapped radiation phenomena may be known in analytic form,

but the transport coefficients that enter into them must be empirically determined.

This chapter gives an account of these and other applied concepts. The current state of knowledge of the geomagnetically trapped radiation is described both from the theoretical perspective and from direct observations. We demonstrate how this knowledge is used to construct physical models of the radiation belts. Empirical radiation belt models based on data compiled from many spacecraft and a brief survey of man's interaction with geospace are also presented.

5.1 THEORETICAL PRELIMINARIES

The difficult mathematical problem of the motion of energetic charged particles in a dipolar magnetic field was extensively studied during the first half of the twentieth century. A general analytic solution to the equation of motion was never found, and in most cases particle orbit tracing had to be done numerically. The interested reader is referred to Störmer [1955]. Physical approximations that lead to great simplification have, however, been found. This is known as the *adiabatic theory* for trapped particles [Alfvén and Fälthammer, 1963], and the earth's radiation belts have now been successfully described in terms of adiabatic invariants and their perturbation.

5.1.1 Single Particle Motion

An ensemble of ions and electrons moving in space constitutes a plasma that can exhibit many modes of collective as well as single particle behavior. In the presence of electric ($\mathbf{E}$) and magnetic ($\mathbf{B}$) fields these particles are subject to the electromagnetic Lorentz force, $\mathbf{F} = q(\mathbf{E} + \mathbf{V} \times \mathbf{B})$, where q and $\mathbf{V}$ are the particle charge and velocity vector respectively. For ions $q = Ze$, where Z is the ionic charge state and e is the unit charge; for electrons $q = -e$. This force controls the particle motion, and collectively the ensemble of charged particles can modify the fields through induction, charge separation, and electrical currents formed by differential ion and electron motion. When the latter effects are negligible, the particles move independently of each other and the *single particle motion* approximation is applicable.

For geomagnetically trapped particles there exist three quasi-periodic motions: gyro motion around the magnetic field lines, bounce motion between the conjugate mirrorpoints, and drift motion around the earth. The fundamental physics in this approximation is described in detail by Alfvén and Fälthammar [1963]. One should note that the frequencies associated with each of these periodic motions are such that $f_{gyro} >> f_{bounce} >> f_{drift}$. For this reason the three types of motions are largely uncoupled. Figure 5-5 illustrates numerical values of

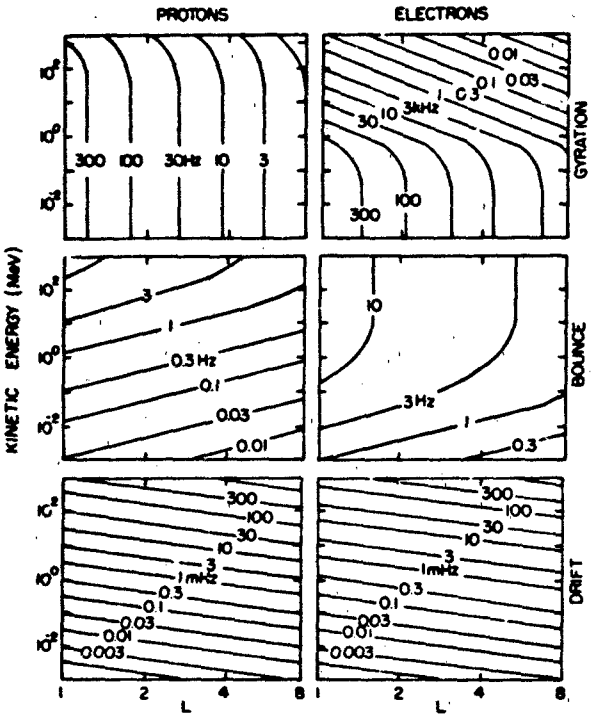


Figure 5-5. The gyration, bounce, and drift frequencies for equatorially mirroring particles in a dipole field as function of L-shell for different particle energies [Shulz and Lanzerotti, 1974].

these fundamental particle motion frequencies for protons and electrons in the earth's radiation belts [Schulz and Lanzerotti, 1974].

5.1.2 Adiabatic Invariants

In general, the motion of charged particles is such that momentum and energy can be transferred between the different particles, and between the particles and the fields that influence their motion. Therefore, it is not always possible to identify constants of motion. However, under certain conditions these energy and momentum exchanges are very small, and it is possible to identify specific quantities that remain virtually unchanged with the particle motion. These are called *adiabatic invariants*.

Associated with each of the three quasi-periodic modes of motion is an adiabatic invariant related to the Hamilton-Jacobi action variable:

$$J_i = \oint_i \left[\mathbf{P} + \frac{q}{c}\mathbf{A}\right] \cdot d\boldsymbol{\ell} \qquad (5.2)$$

where $d\boldsymbol{\ell}$ is a vector line element along the path of integration. Here $\mathbf{P}$ is particle momentum and $\mathbf{A}$ is the magnetic vector potential (that is, $\mathbf{B} = \nabla \times \mathbf{A}$). The integration

is extended over the particle orbit for gyro motion, bounce motion, and azimuthal drift motion (for i = 1, 2, 3 respectively). If the particle's trajectory closed exactly on itself, then the action variables J_i would be absolute constants of motion. Finite spatial and temporal variations in **B** prevent perfect closure, and thus the J_is are at best *approximate* constants.

5.1.2.1 First Adiabatic Invariant. J_1 is obtained by evaluating the integral in Equation (5.2) over the particle gyro motion only, that is, over the particle orbit projection in a plane perpendicular to **B**. Using subscripts $\parallel$ and $\perp$ to denote directions parallel and perpendicular to **B**, and by virtue of Stokes' theorem [Schultz and Lanzerotti, 1974] one derives

$$J_1 = 2\pi\rho_g P_\perp + \frac{q}{c}\pi\rho_g^2 B . \qquad (5.3)$$

where $B = |\mathbf{B}|$, p is particle momentum and $\rho_g = P_\perp/|q|B$ is the particle gyro (or cyclotron) radius. From this, one defines the first adiabatic invariant

$$\mu = \frac{P_\perp^2}{2m_o B} = \frac{P^2 \sin^2\alpha}{2m_o B} , \qquad (5.4)$$

also known as the relativistic magnetic moment. Here m_o is the particle rest mass, and α = arc sin $(\mathbf{B}\cdot\mathbf{P}/PB)$ is the pitch angle illustrated in Figure 5-6.

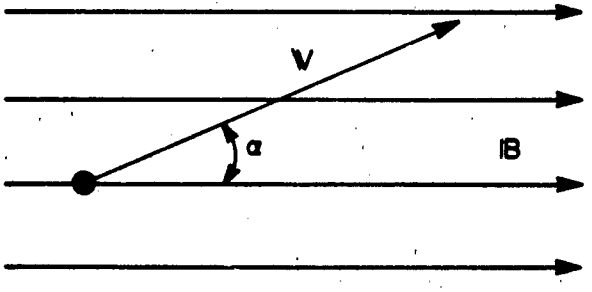


Figure 5-6. A particle with its velocity vector inclined to the magnetic field at an angle α. This angle is called the particle pitch angle.

For non-relativistic particles,

$$\mu = \frac{\epsilon_\perp}{B} = \frac{\frac{1}{2} m_o V_\perp^2}{B} \qquad (5.5a)$$

where $\epsilon_\perp$ is the particle kinetic energy associated with the directions perpendicular to the local magnetic field direction, while for relativistic particles

$$\mu = \frac{P_\perp^2}{2m_o B} = \gamma^2 \left[\frac{\frac{1}{2} m_o V_\perp^2}{B}\right] \qquad (5.5b)$$

where the relativistic factor $\gamma = 1/\sqrt{1-v^2/c^2}$. μ is an approximate constant of motion when both of the following conditions are fulfilled:

— The *spatial* scale of B-field variation is much larger than the particle gyroradius

$$\frac{B}{|\nabla B|} >> \rho_g = P_\perp/B|q| = \gamma m_o V \sin\alpha/B|q| \qquad (5.6)$$

— The time scale of change of the B-field is much larger than the particle gyroperiod

$$T >> \tau_g = 2\pi\rho_g/v_\perp = 2\pi\gamma m_o/B|q| . \qquad (5.7)$$

5.1.2.2 Second Adiabatic Invariant. J_2 is obtained by evaluating the integral in Equation (5.2) over the bounce trajectory and averaged over the gyro motion, or equivalently along the magnetic field line (guiding center field line) around which the particle gyrates, and thereby defining the second adiabatic invariant:

$$J = \frac{1}{2} J_2 = \frac{1}{2}\oint \mathbf{P}\cdot d\ell = \int_{-\ell m}^{+\ell m} P_\parallel d\ell \qquad (5.8)$$

where $d\ell$ is an element of length along that field line segment and ℓm is the curvelinear distance of the mirrorpoints from the equator measured along the guiding center magnetic field line. Since equatorially mirroring particles ($\alpha_o = \pi/2$) do not have any bounce motion, it follows that J=0 for such particles.

Provided the particle mirrorpoints are above the dense atmosphere, J will remain an approximate constant when the time scale of B-field variation is much larger than the particle bounce time between the conjugate mirrorpoints

$$T >> \tau_B = \int_{\ell m}^{+\ell m} d\ell / V_\parallel(\ell) . \qquad (5.9)$$

Constancy of the first adiabatic invariant μ implies that

$$\frac{\sin^2\alpha_o}{B_o} = \frac{\sin^2\alpha}{B} = \frac{1}{B_m} = \text{constant} \qquad (5.10)$$

over the bounce motion between the mirror points. Here the subscript zero denotes equatorial quantities and B_m is the magnetic field induction at one of the mirror points

(where $\alpha = \pi/2$). Equation (5.10) is known as the mirror equation. Using (5.10) one finds

$$\tau_B = \frac{m_o\gamma}{P} \int_{\ell m}^{+\ell m} d\ell \left(1 - \frac{B}{B_o} \sin^2\alpha_o\right)^{-1/2}$$

$$= \frac{m_o\gamma}{P} \int_{\ell m}^{+\ell m} d\ell \left(1 - \frac{B}{B_m}\right)^{-1/2}, \qquad (5.11)$$

and in dipolar coordinates (Equation 5.11) becomes

$$\tau_B = \frac{2m_o\gamma}{P} T(\alpha_o), \qquad (5.12)$$

where $T(\alpha_o)$ is the bounce time integral given by

$$T(\alpha_o) = \int_0^{\lambda_m(\alpha_o)} \frac{\cos\lambda\,[4-3\cos^2\lambda]\,d\lambda}{\left\{1 - \frac{\sin^2\alpha_o}{\cos^6\lambda}\,[4-3\cos^2\lambda]^{1/2}\right\}^{1/2}}, \qquad (5.13)$$

and $\lambda_m(\alpha_o)$ is the magnetic latitude of the mirrorpoint which depends on the equatorial pitch angle α_o.

To a fair approximation

$$T(\alpha_o) \approx 1.30-0.56\,\sin\alpha_o \quad \text{[Hamlin et al., 1961]}, \qquad (5.14)$$

or alternatively

$$T(\alpha_o) \approx 1.3802-0.3198\,(\sin\alpha_o + [\sin\alpha_o]^{1/2})$$

[Schulz and Lanzerotti, 1974]. (5.15)

Other approximations are given by Davidson [1973]. Gradients of $T(\alpha_o)$ should, however, not be derived from such approximations. From Equations (5.1) and (5.10) it follows that

$$B_m = \frac{B_o}{\sin^2\alpha_o} = \frac{B_E}{L^3\sin^2\alpha_o}, \qquad (5.16)$$

where in a dipolar magnetic field $L = (R/R_E)$ is the equatorial distance of a given field line, and

$$\sin^2\alpha_o = \frac{\cos^6\lambda_m(\alpha_o)}{[4-3\cos^2\lambda_m(\alpha_o)]^{1/2}}. \qquad (5.17)$$

Although Equation (5.17) cannot be solved explicitly for $\lambda_m(\alpha_o)$, a numerical solution is easily obtained, or one may approximate as in Hamlin et al. [1961]

$$\cos\lambda_m(\alpha_o) \approx [\sin\alpha_o]^{1/4}. \qquad (5.18)$$

It should be emphasized that using a dipolar magnetic field representation explicitly disregards any azimuthal asymmetries of the geomagnetic field. Such asymmetries do exist and become significant beyond L=5. Under such conditions a different magnetic field representation should be used, and this is outlined in Section 5.6.1.4.

5.1.2.3 Third Adiabatic Invariant. J_3 is obtained by evaluating the integral in Equation (5.2) over the particle drift motion around the earth, and averaged over gyro and bounce motion

$$J_3 = \frac{q}{c}\Phi = \frac{q}{c}\oint \mathbf{A}\cdot d\mathbf{l} \qquad (5.19)$$

where Φ is the magnetic flux enclosed by the azimuthal drift orbit and l is linear azimuthal distance. Using Stokes' theorem yields

$$\Phi = \oint \mathbf{A}\cdot d\mathbf{l} = \int_S \mathbf{B}\cdot d\mathbf{S}, \qquad (5.20)$$

where S is a surface bounded by the azimuthal drift path. In a dipolar magnetic field one calculates [Roederer, 1970]

$$\Phi = -\frac{2\pi B_E R_E^2}{L}, \qquad (5.21)$$

where L is the McIlwain [1961] L-parameter.

J_3 will remain approximately constant when the time scale of B-field change is much longer than the azimuthal drift time $\tau_d = \oint_{drift} ds/V_d$ around the earth. Determination of the azimuthal drift velocity is discussed in the next section.

5.1.3 Particle Drift Motion. In a uniform magnetic field, charged particles execute a spiral motion such that the angle between the particle velocity vector and the magnetic field direction (the pitch angle) remains constant. When the magnetic field lines converge, the particle will respond to an effective net magnetic force from

higher to lower magnetic field strengths. This is illustrated in Figure 5-7. The physical reason for this force is

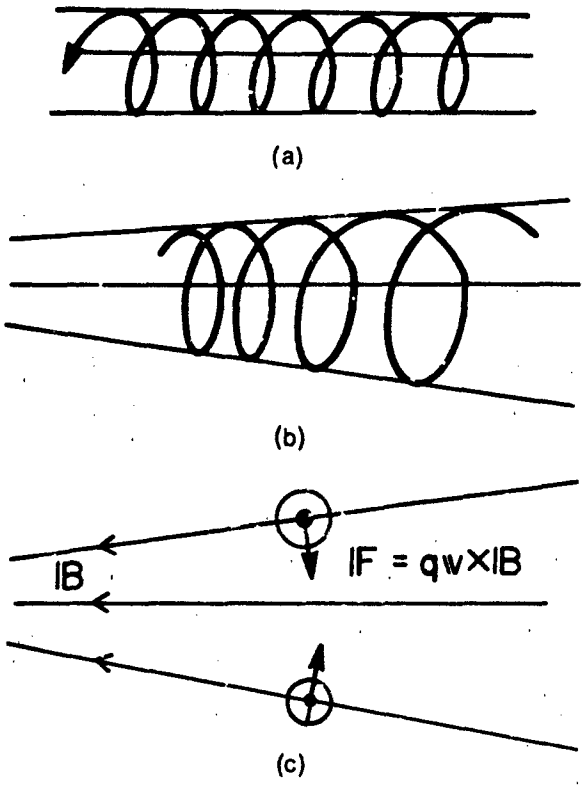


Figure 5-7. (a) Particle motion in a uniform magnetic field (uniform spiral motion).
(b) The tightening of spiral motion in a converging magnetic field.
(c) Illustration of magnetic force with gyroaveraged net component in the ∇B direction in a converging magnetic field.

that the particle gyro motion produces an elementary current (which may be interpreted as a magnetic dipole current loop). For each such loop the effective current is

$$i_p = \frac{dq}{dt} = \frac{qP_\perp}{2\pi\rho_g m_o \gamma}, \qquad (5.22)$$

where ρ_g is the mean gyroradius over the loop. The magnetic moment of a current loop enclosing an area A is

$$M = i_p A = \frac{P_\perp^2}{2m_o \gamma B} \qquad (5.23)$$

where $A = \pi\rho_g^2$, which is the particle magnetic moment itself. The particle will therefore, averaged over its gyro motion, be subject to a net force $F = -M\nabla_\parallel |B|$ in the direction along the field lines away from the higher field region.

In general, the magnetic field may also have an intensity gradient across the field lines. This is illustrated in Figure 5-8. Charged particles moving in such a magnetic

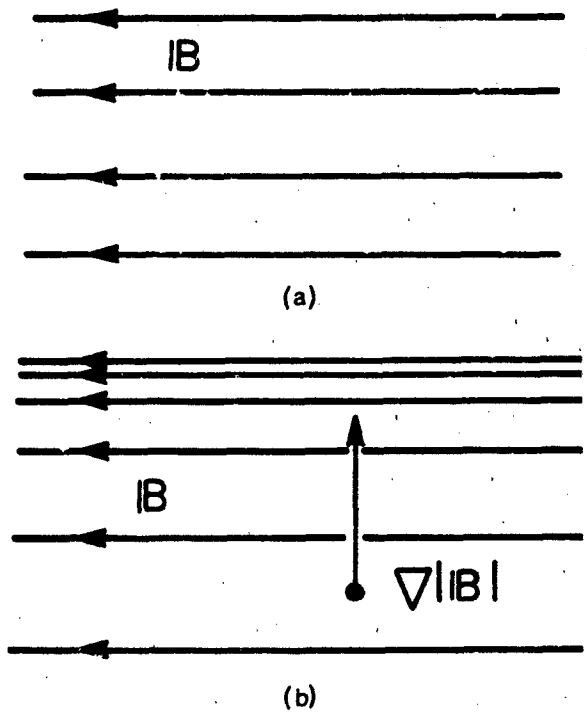


Figure 5-8. (a) A uniform B-field where field lines are represented as evenly spaced.
(b) A magnetic field with increasing strength perpendicular to B, represented as denser field lines with increasing B.

field will have a smaller gyroradius in the higher field region and a larger gyroradius in the lower field region. As a consequence, there will be a net drift velocity perpendicular to the magnetic field direction. In this figure a positive charge would drift into the paper and a negative charge out of the paper. Defining the angular gyrofrequency

$$\Omega = \frac{|q|\,B}{m_o \gamma c}, \qquad (5.24)$$

one can express the instantaneous vector gyroradius as

$$\boldsymbol{\rho}_g = \frac{\mathbf{P}\times\mathbf{B}}{m_o\gamma\Omega B} = c\,\frac{\mathbf{P}\times\mathbf{B}}{qB^2}, \qquad (5.25)$$

and the drift velocity is then the time rate of change of ρ_g

$$\mathbf{V}_d = \frac{d\boldsymbol{\rho}_g}{dt} = \frac{c}{qB^2}\left(\frac{d\mathbf{P}}{dt}\times\mathbf{B}\right) \qquad (5.26)$$

where it is assumed that the magnetic field is constant in time. With $\mathbf{F} = \frac{d\mathbf{p}}{dt}$ being the net force due to the cross-B gradient, one obtains the "gradient drift"

$$\mathbf{V}_{gd} = \frac{c}{qB^3}\left(M\nabla_\perp|\mathbf{B}| \times \mathbf{B}\right),$$

$$= \frac{Mc}{qB^2}\left(\nabla_\perp|\mathbf{B}| \times \mathbf{B}\right), \qquad (5.27a)$$

$$= \frac{cp^2}{2m_0q\gamma B^3}\left(\nabla_\perp|\mathbf{B}| \times \mathbf{B}\right),$$

which non-relativistically is just

$$\mathbf{V}_{gd} = \frac{c\epsilon}{qB^3}\left(\mathbf{B} \times \nabla_\perp|\mathbf{B}|\right). \qquad (5.27b)$$

The earth's magnetic field is also curved (that is, the dipolar-like field lines form loops from pole to pole), and the field line radius of curvature is given by

$$R_c = \frac{R_0}{3}\cos\lambda\,\frac{(1+3\sin^2\lambda)^{3/2}}{1+\sin^2\lambda}, \qquad (5.28)$$

where the individual field lines are described by the dipole relation

$$R = R_0\cos^2\lambda \qquad (5.29)$$

with $R_0 = R_E L$. Thus a charged particle moving in that field will experience a centrifugal force

$$\mathbf{F}_c = \frac{p^2}{m_0\gamma R_c}\mathbf{n} \qquad (5.30)$$

where $\mathbf{n}$ is a unit vector in the direction away from the instantaneous field line center of curvature. This causes a drift velocity

$$\mathbf{V}_{cd} = \frac{c}{qB^2}\left(\frac{p^2}{m_0\gamma R_c}\mathbf{n} \times \mathbf{B}\right). \qquad (5.31)$$

and in the absence of significant plasma currents $\nabla \times \mathbf{B} = 0$ and $\nabla_\perp B = \frac{B}{R_c}\mathbf{n}$ [Roederer, 1970], and one can write

$$\mathbf{V}_{cd} = \frac{cp^2}{m_0\gamma qB^3}\left(\mathbf{B} \times \nabla_\perp|\mathbf{B}|\right), \qquad (5.32)$$

which non-relativistically becomes

$$\mathbf{V}_{cd} = \frac{2c\epsilon_\parallel}{qB^3}\left(\mathbf{B} \times \nabla_\perp|\mathbf{B}|\right). \qquad (5.33)$$

Although the effect of the earth's gravitational field is rather small compared to other forces on radiation belt particles, it can easily be included:

$$\mathbf{V}_{grav} = \frac{cm_0\gamma}{qB^2}\left(\mathbf{g} \times \mathbf{B}\right) \qquad (5.34)$$

where $\mathbf{g}$ is the vector gravitational acceleration.

The effect of a weak, externally imposed electric field is also easily taken into account:

$$\mathbf{V}_{ed} = \frac{c}{qB^2}\left(q\mathbf{E} \times \mathbf{B}\right) = c\,\frac{\mathbf{E} \times \mathbf{B}}{B^2}. \qquad (5.35)$$

The electric field drift is *independent* of particle charge and mass as long as either is non-zero. Thus, under the influence of an electrostatic field, ions and electrons drift together (plasma flow), while under the influence of an inhomogeneous magnetic field, oppositely charged particles drift in opposite directions (causing current flow).

The total particle drift velocity is then the superposition of the contributing drifts:

$$\mathbf{V}_d = c\,\frac{p_\perp^2 + 2p_\parallel^2}{2m_0\gamma qB^3}\left(\mathbf{B} \times \nabla_\perp|\mathbf{B}|\right) \qquad (5.36)$$

$$+ \frac{c}{B^2}\left(\mathbf{E} \times \mathbf{B}\right) + c\,\frac{m_0\gamma}{qB^2}\left(\mathbf{g} \times \mathbf{B}\right)$$

$$\mathbf{V}_d = \frac{\epsilon_\perp + 2\epsilon_\parallel}{qB^3}\left(\mathbf{B} \times \nabla_\perp|\mathbf{B}|\right)$$

$$+ \frac{C}{B^2}\left(\mathbf{E} \times \mathbf{B}\right) + \frac{m_0\gamma C}{qB^2}\left(\mathbf{g} \times \mathbf{B}\right). \qquad (5.37)$$

In the dipolar magnetic field representation an approximate formula for the drift period is given by Davidson [1977]

$$T_d \approx \frac{1.43K_t}{LE\gamma(V/c)^2(1+0.42\sin\alpha_0)}, \qquad (5.38)$$

where $K_t = 1.0308 \times 10^4$ seconds for electrons, $K_t = 5.655$ seconds for protons, and $K_t = \frac{4\pi Z_i^2 B_E R_E^2}{3M_i C^3}$ for ions of mass M_i and charge state Z_i.

The cartoon in Figure 5-9 illustrates the principal drift effects associated with the different drift mechanisms. In Equations (5.36) and (5.37) the terms are listed in order of their importance in the radiation belts. Above ~10 keV the magnetic gradient curvature drift is generally strongest, and static electric field and gravity effects are usually neglected in radiation belt studies. The gradient-curvature drift carries energetic electrons towards the east and ions to the west. Thus there will be a net westward electrical current encircling the earth. This is the *extraterrestrial ring current*. These findings are summarized in Figure 5-10.

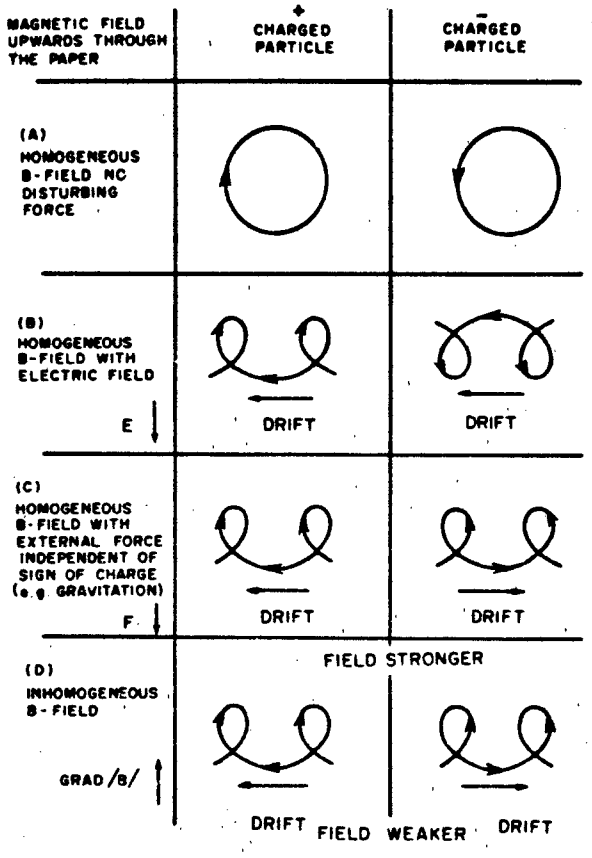


Figure 5-9. Summary illustration of drift of both positive and negative charged particles in a) a uniform B-field, no external force; b) a uniform B-field with a perpendicular electric field; c) a uniform B-field with an external force which is independent of electric charge such as gravity; and d) a B-field with a gradient [Alfven and Falthammar, 1963].

5.2 TRAPPED RADIATION SOURCES

Precisely where the radiation belt particles come from and how they are accelerated to energies in the keV and MeV range are still areas of research for which a comprehensive answer is not yet available.

5.2.1 Qualitative Description

A number of sources are considered responsible, and the effectiveness of each probably also varies with time:

1. Particles from the sun, including solar wind particles and energetic solar particle emissions, possibly via magnetotail storage. For a description of solar cosmic rays see Chapter 6.
2. Particles from the earth's ionosphere, including the polar wind flow into the magnetotail, particles flowing up the magnetic field lines to form the plasmasphere and particles accelerated out of the auroral ionosphere (Chapters 8 and 9).
3. Cosmic ray albedo neutron decay (CRAND) within the trapping region.
4. Particles arriving at the earth having been accelerated in interplanetary shock waves or in the magnetospheres of other planets.
5. Low energy components of galactic cosmic rays (Chapter 6).
6. *In situ* acceleration of pre-existing low energy trapped particles within the radiation belts.

The solar wind flows past the earth's magnetosphere virtually at all times (Chapter 3). Some of these particles may find their way through the outer regions of the magnetosphere to the stable trapping region [Hovestadt et al., 1978]; this process may be particularly effective during periods of southward heliospheric magnetic field. Direct transient injections of solar energetic particles probably also occur, particularly in conjunction with magnetic storms.

Ionospheric particles diffusing out of the polar ionosphere (polar wind) escape into the magnetotail region from which some may become energized and injected into the trapping region. Auroral electric fields are intermittent and can have a significant component parallel to the magnetic field, and ions and electrons from the topside auroral ionosphere can be accelerated to multi-keV energies. The wave fields associated with plasma waves may also cause particle acceleration. This could be a source of H^+, He^+, O^+, and electrons provided other processes act to trap the particles.

Cosmic rays impacting the earth's atmosphere undergo nuclear reactions, and a flux of neutrons escapes from the top of the atmosphere. Free neutrons are unstable and decay into proton, electron and neutrino triplets on a time scale of ~1000 seconds. If the electrically charged decay products find themselves within the radiation belts, they will immediately be subject to the magnetic force and may become trapped. Empirically, this is an important source of multi-MeV protons in the innermost part of the inner radiation zone ($L \lesssim 1.5$). There is little direct information about the efficiency of direct extraterrestrial energetic particle trapping in the radiation belts. One may surmise, however, that time variability of the geomagnetic field is needed for trapping to occur or that incident extraterrestrial energetic ions in low charge

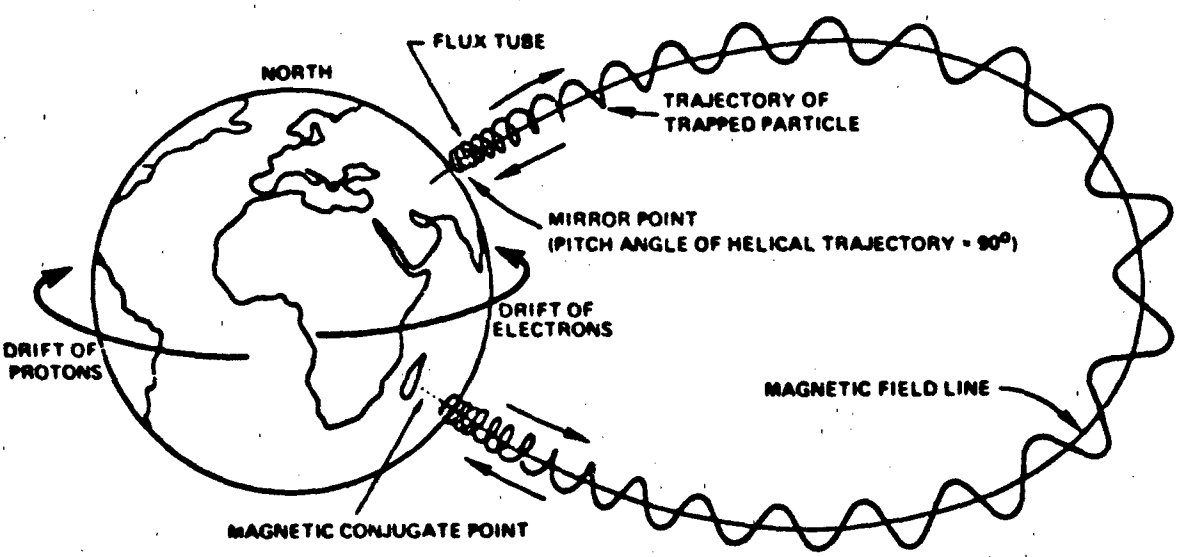


Figure 5-10. A descriptive drawing of the three types of motion of particles trapped in the earth's magnetic field

states (for example C^+ and O^+) may charge exchange to higher charge states (for example, C^{6+} and O^{8+}) within the magnetosphere so that their gyro radii become small enough ($\rho_g = P_\perp / qB$) for trapping.

The plasma sheet in the earth's magnetotail is considered an important reservoir for radiation belt particles. However, the plasma sheet particles themselves are likely to be a mixture of particles from several of the prime sources mentioned above. During magnetospheric substorms plasma sheet particles convect inward toward the earth, and in the process can be accelerated and may become trapped in the radiation belts. Unfortunately, little is known about the specific details of the time dependent trapping process and it is not yet possible to make a quantitative evaluation of the strength and characteristics of this source.

Current research also points to the earth's ionosphere as an important contributor to the lower energy particle population below a few tens of keV, perhaps with a roughly equal contribution from solar wind particles. In contrast, the high energy particles above ~500 keV appear to have an extraterrestrial source. The former conclusion is derived from observations of dominant oxygen fluxes at times, while the latter stems from observed carbon-to-oxygen ratios of order unity. For the very important intermediate energy range where most of the radiation belt energy density is found, there is no experimental result indicating the source.

A simple theory that seeks to explain observed storm-time enhancements of radiation belt particle fluxes at tens and hundreds of keV energies has been advocated [Lyons and Williams, 1980]. Particles existing in the outer radiation zone may suddenly, during the storm main phase, be subject to an electric field that transports them towards lower L-shells on a time scale that preserves μ and J but violates the constancy of Φ. For equatorially mirroring particles, a radial displacement from L=5 to L=3 increases the particle energy by a factor of ~5. Furthermore, if the particle spectrum follows a power law distribution $j(E) = (E\ E_0)^{\ u}$ with u = 3, for example, then the apparent particle flux enhancement seen at fixed energies will be a factor of 125. Seen at a fixed L-shell, the observable flux increase will be even greater if the pre-storm radial distribution falls off toward lower L-shells (as may be the case at lower radiation belt energies), and smaller if this flux gradient is negative. Given the knowledge of the storm perturbation electric field (magnitude, direction, azimuthal extent, and duration) together with observations of the pre-storm radiation belt structure, this *in situ* "source" is in principle assessable.

5.2.2 Simplifying Assumptions

For most of the radiation belt source mechanisms accurate quantitative information is still lacking.

Ideally, one would like to know

1. the source strength for different particle species as a function of energy and pitch angle,
2. the effective source locations within and on the boundaries of the trapping region, and
3. the source strength as function of the different geophysical conditions during quiet and disturbed times.

Unfortunately, contemporary research has not yet yielded quantitative answers to these requirements. Without this information, how can we understand and model the earth's radiation belts?

CHAPTER 5

For quiet time conditions one can solve the steady state radiation belt transport equations for the interior of the radiation belts subject to suitable outer zone *boundary conditions* on the trapped fluxes. This amounts to the assumption that the radiation belt source is capable of supplying particles to the outer radiation zone boundary at a rate sufficient to offset losses within the trapping region. The existence of long-term approximate stability of the radiation belts as a whole during extended quiet periods supports this contention, and fortunately trapped flux observations from geostationary spacecraft, such as ATS-6, of the outer zone flux levels at L = 6.6 makes this a feasible solution.

This is not a satisfactory situation for magnetic storms and other disturbances. As a consequence most radiation belt modeling has been done for steady state, quiet time conditions. Time dependent radiation belt modeling would require time dependent boundary conditions averaged over local time. It is conceivable that data from several geostaticnary satellites may be used for this purpose, but this has not yet been done.

5.3 TRANSPORT PROCESSES IN THE RADIATION BELTS

As we have seen, in the static geomagnetic field, radiation belt particles execute the three periodic motions: gyration around the magnetic field lines, bounce motion between mirror points, and azimuthal drift around the earth. The latter type of periodicity is caused by the gradient-curvature drift motion. Effects of gravity and electric fields cause departure from this simple picture, but both forces are primarily important at low energies typically below a few tens of keV.

5.3.1 Convection

Magnetospheric convection results from externally imposed electric fields. The solar wind flows past the earth at a velocity in the range 200-600 km sec. This implies a "convection" electric field $\mathbf{E} = \mathbf{V}_c \times \mathbf{B}$ across the earth's magnetosphere directed from dawn to dusk. Combined with the electric field induced by the rotation of the earth a characteristic magnetospheric convection pattern is set up (Axford, 1964; Roederer, 1970). The radial corotational electric field is induced by the earth's magnetic dipole field corotating with the earth. Figure 5-11 shows model electric equipotential lines around the earth due to these electric fields in a time-independent situation. Particles of quite low energies ($\lesssim$ 1 keV) are primarily controlled by the electric field drift, and their motion approximately follows the equipotential lines as indicated by the arrows in Figure 5-11. Notice the topologically distinct regions: near the earth where the convective motion follows oval paths around the earth, and at greater distances where the drift paths are open to the magnetopause. Departures from this overall configuration, of course, take place during disturbed conditions when the imposed "convection" electric field is time variable and the low energy particle distributions are not in equilibrium [Harel et al., 1981 a & b].

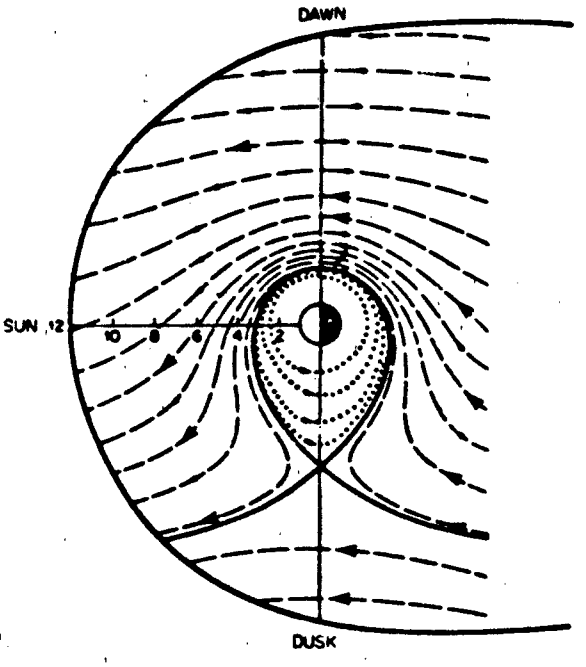


Figure 5-11. Equipotential contours for an electric field in the earth's equatorial plane (dashed lines). These are also drift paths for very low energy particles. The electric field is a superposition of a corotational E-field due to the rotation of the earth and its imbedded magnetic field and a uniform dawn-dusk electric field. The separatrix (solid curve) is the low energy particle Alfvén layer which separates the open and closed drift paths [Kivelson, 1979]. The outer oval represents the magnetopause in the equatorial plane.

5.3.2 Stochastic Processes: Phenomenology

At higher energies, particularly above a few tens of keV, the dominant drift of trapped particles is due to the gradient and curvature effects of the geomagnetic field. The prime drift motion is therefore circular (with gyro and bounce motion superimposed) around the earth. Departures from this pattern are due to the *fluctuating* nature of the geoelectric and geomagnetic fields induced by variations in the solar wind flow and internal magnetospheric processes. Radiation belt particles are also subject to interactions with plasma waves and suffer collisions with exospheric neutral atoms and low energy plasma particles. Common to these processes is their randomness in occurrence, and their effects are described by stochastic analysis that can be reduced to diffusion theory: radial diffusion and pitch angle diffusion. The physical ideas are illustrated in Figure 5-12. Radial diffusion transports radiation belt particles across the dipolar-like magnetic field lines in the radial direction, and pitch

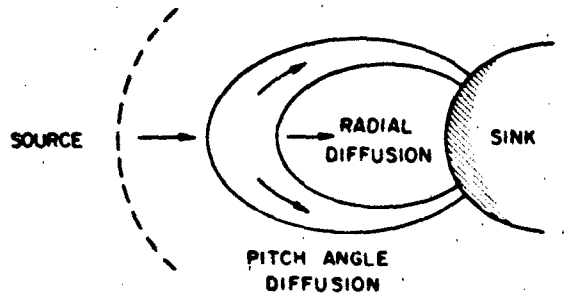


Figure 5-12. A conceptual representation of pitch angle and radial diffusion in the earth's radiation belts. Diffusion occurs in either direction, but in most cases there is a net diffusion flux in the indicated direction towards the earth's atmosphere.

angle diffusion alters the particle pitch angle (or equivalently, the mirrorpoint location). In both cases the earth's atmosphere is a sink: for radial diffusion by transport to very low L-shells, and for pitch angle diffusion by lowering the mirrorpoints into the atmosphere. In addition to diffusive processes, energy degradation from collisions with exospheric particles also occurs.

5.3.3 Effects of Field Fluctuations

Adiabatic invariants are useful substitutes for particle constants of motion. It should be kept in mind, however, that the earth's magnetic field is never perfectly static. Field fluctuations are associated with micropulsations [Jacobs, 1970], magnetospheric substorms [Akasofu, 1968], geomagnetic storms and other phenomena, and field oscillations vary considerably in magnitude, frequencies, and principal location. The adiabatic invariants are said to be violated when electric or magnetic variations take place near or above the adiabatic motion frequency in question. Even slow field variations may violate the third adiabatic invariant Φ ($T\sim$ minutes to hours) while μ-violation requires wave-like fluctuations on a time scale of ~milliseconds. Macroscopically, the earth's radiation belts are subject to field fluctuations that occur at quasi-random times. Their effects are best described by stochastic methods [Chandrasekhar, 1965] that treat the mean deviations in the adiabatic invariants and the associated distribution function.

Define a particle distribution function $f = f(\mu, J, \Phi; t)$ such that the number of particles located within a parameter space volume element given by $\mu - \frac{1}{2}d\mu$ to $\mu + \frac{1}{2}d\mu$, $J - \frac{1}{2}dJ$ to $J + \frac{1}{2}dJ$, $\Phi - \frac{1}{2}d\Phi$ to $\Phi + \frac{1}{2}d\Phi$ is given by

$$dN = f(\mu, J, \Phi; t)\, d\mu\, dJ\, d\Phi$$

at a time t. The function f is thus the particle density in this parameter space spanned by the adiabatic invariants.

Let $P(\mu, J, \Phi; \Delta\mu, \Delta J, \Delta\Phi)$ be the probability that a mean change $\Delta\mu$, ΔJ, and $\Delta\Phi$ takes place in the adiabatic invariants per unit time ($|\Delta\mu| << \mu$, $|\Delta J| << J$, and $|\Delta\Phi| << \Phi$). The distribution function $f(\mu, J, \Phi; t)$ averaged over gyrophase, bounce phase, and azimuthal drift phase is then governed by the relation

$$f(\mu, J, \Phi; t) = \iiint d(\Delta\mu)\, d(\Delta J)\, d(\Delta\Phi)\; f(\mu - \Delta\mu, J - \Delta J, \Phi - \Delta\Phi; t - \Delta t)\; P(\mu - \Delta\mu, J - \Delta J, \Phi - \Delta\Phi; \Delta\mu, \Delta J, \Delta\Phi), \tag{5.39}$$

and one expands f and P in Taylor series around the unperturbed quantities to arrive at

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial \mu}\left(\frac{\langle \Delta\mu \rangle}{\Delta t} f\right) - \frac{\partial}{\partial J}\left(\frac{\langle \Delta J \rangle}{\Delta t} f\right) - \frac{\partial}{\partial \Phi}\left(\frac{\langle \Delta\Phi \rangle}{\Delta t} f\right) + \frac{\partial^2}{\partial \mu^2}\left(\frac{\langle (\Delta\mu)^2 \rangle}{2\Delta t} f\right) + \frac{\partial^2}{\partial J^2}\left(\frac{\langle (\Delta J)^2 \rangle}{2\Delta t} f\right) + \frac{\partial^2}{\partial \Phi^2}\left(\frac{\langle (\Delta\Phi)^2 \rangle}{2\Delta t} f\right) + \frac{\partial^2}{\partial \mu \partial J}\left(\frac{\langle \Delta\mu \Delta J \rangle}{\Delta t} f\right) + \frac{\partial^2}{\partial \mu \partial \Phi}\left(\frac{\langle \Delta\mu \Delta\Phi \rangle}{\Delta t} f\right) + \frac{\partial^2}{\partial J \partial \Phi}\left(\frac{\langle \Delta J \Delta\Phi \rangle}{\Delta t} f\right) + \text{higher order terms,} \tag{5.40}$$

where the stochastic coefficients are defined by

$$\langle \Delta i \rangle = \iiint d(\Delta\mu)\, d(\Delta J)\, d(\Delta\Phi)\; P(\mu, J, \Phi; \Delta\mu, \Delta J, \Delta\Phi)\, \Delta i$$

$$\langle \Delta i \Delta j \rangle = \iiint d(\Delta\mu)\, d(\Delta J)\, d(\Delta\Phi)\; P(\mu, J, \Phi; \Delta\mu, \Delta J, \Delta\Phi)\, \Delta i \Delta j \tag{5.41}$$

with i and j being μ, J or Φ in all permutations.

The transport equation (5.41) can be greatly simplified by recognizing that violation of one adiabatic invariant is almost always uncorrelated with the process violating another. In that case all cross coefficients must vanish:

$$\langle \Delta\mu \Delta J \rangle = \langle \Delta\mu \Delta\Phi \rangle = \langle \Delta J \Delta\Phi \rangle = 0. \tag{5.42}$$

Furthermore, $\langle \Delta i \rangle$ and $\frac{1}{2}\langle (\Delta i)^2 \rangle$ ($i = \mu, J, \Phi$) are related. In the absence of external sources and losses, diffusion would proceed to transport particles from overabundant regions of parameter space until all gradients in the distribution function had vanished, and for each diffusion mode

$$\langle \Delta i \rangle \quad \frac{\partial}{\partial i} \frac{\langle (\Delta i)^2 \rangle}{2} = 0 \quad (i = \mu, J, \Phi). \tag{5.43}$$

Equation (5.41) then simplifies to

$$\left(\frac{\partial f}{\partial t}\right)_{transp} = \sum_i \frac{\partial}{\partial i}\left(D_{ii} \frac{\partial f}{\partial i}\right) \quad (i = \mu, J, \Phi), \tag{5.44}$$

where

$$D_{ii} = \frac{\langle (\Delta i)^2 \rangle}{2\Delta t} \quad (i = \mu, J, \Phi). \tag{5.45}$$

Equation (5.44) is the *diffusion equation* describing the earth's radiation belts. It is valid whenever the perturbations are small (but accumulative). Transformation to other variables, such as ϕ_1, ϕ_2, ϕ_3 is facilitated by the Jacobian $G = G(\mu, J, \Phi, \phi_1, \phi_2, \phi_3)$ such that

$$\left(\frac{\partial F}{\partial t}\right)_{transp} = \sum_j \frac{1}{G} \frac{\partial}{\partial \phi_j} \cdot \left(D_{\phi_j \phi_j} G \frac{\partial F}{\partial \phi_j}\right) \quad (j = 1, 2, 3), \tag{5.46}$$

where

$$D_{\phi_j \phi_j} = D_{ii}\left(\frac{\partial \phi_j}{\partial i}\right)^2 \quad (i = \mu, J, \Phi), \tag{5.47}$$

and

$$F = F(\phi_1, \phi_2, \phi_3; t). \tag{5.48}$$

5.3.4 Radial Diffusion

Field fluctuations on a time scale comparable to the azimuthal drift time around the earth can violate the third adiabatic invariant Φ but preserve the μ and J invariants. This is the limit of pure radial diffusion. To study this process requires knowledge of the diffusion coefficient, sources, and losses. It is convenient to use the dipole L-shell coordinates defined through

$$\Phi = 2\pi R_E^2 \frac{B_E}{L}. \tag{5.49}$$

where $B_E \approx 0.312$ G is the equatorial B-value at the surface of the earth where $L = 1$.

In this case the Jacobian coordinate transformation is just

$$G(L; \Phi) = \frac{\partial \Phi}{\partial L} = -2\pi R_E^2 B_E \frac{1}{L^2}, \tag{5.50}$$

and the pure radial diffusion equation becomes

$$\frac{\partial f}{\partial t} = L^2 \frac{\partial}{\partial L}\left[D_{LL} L^{-2} \frac{\partial f}{\partial L}\right] + \underline{S} - \underline{L}, \tag{5.51}$$

where $\underline{S}$ and $\underline{L}$ represent particle source and loss functions. D_{LL} is the pure radial diffusion coefficient at constant μ and J values. It has been estimated that

$$D_{LL} \simeq D_{LL}(\alpha_o) = D_{LL}\left(\frac{\pi}{2}\right)\bar{d}(\alpha_o) \tag{5.52}$$

where $\bar{d}(\alpha_o)$ is a function only of the particle equatorial pitch angle and $D_{LL} = D_{LL}\left(\frac{\pi}{2}\right)$ is the radial diffusion coefficient for equatorially mirroring $\left(\alpha_o = \frac{\pi}{2}\right)$ particles where $\bar{d}\left(\frac{\pi}{2}\right) = 1$. For a description of $\bar{d}(\alpha_o)$ see Schulz [1975a].

Both geomagnetic and geoelectric field fluctuations contribute to D_{LL}. For geomagnetic fluctuations it can be shown that

$$D_{LL}^{(M)}(\mu, L) = \omega_d^2 \bar{P}^{(M)}(\omega_d) L^{10} \tag{5.53}$$

where $\bar{P}^{(M)}$ is the fluctuation power spectral density evaluated at the azimuthal angular drift frequency $\omega_d = \frac{V_d}{LR_E} = \frac{2\pi}{t_d}$. Empirically $\bar{P}^{(M)}(\omega) \propto \omega^{-r}$ where the value of r most often is $r = 2 \pm 1$. When $r = 2$, one obtains the very simple expression

$$D_{LL}^{(M)} = K^{(M)} L^{10} \tag{5.54}$$

where $K^{(M)}$ is a factor dependent on the fluctuation magnitudes. For geoelectric field fluctuations it can be shown that

$$D_{LL}^{(E)}(\mu, L) = \frac{c^2}{8B^2} \sum_{n=0}^{\infty} \bar{P}_n^{(E)}(\omega_d) L^6 \tag{5.55}$$

where $\bar{P}_n^{(E)}$ is the n-th spatial Fourier component of the electric field power spectral decomposition. For details, see Fälthammar [1968], and Cornwall [1968].

The actual calculation of $D_{LL}^{(E)}$ is fairly complicated and the reader is referred to the research literature. However, for typical substorm conditions Cornwall [1972] derived the simple relation

$$D_{LL}^{(E)} = K^{(E)} \frac{L^{10}}{L^4 + (\mu_M / Z_1)^2}, \qquad (5.56)$$

where μ_M is the magnetic moment in MeV G and Z_1 is the particle charge state number.

Both $K^{(M)}$ and $K^{(E)}$ are dependent on the geophysical activity. Likely values of $K^{(M)}$ fall in the range 2×10^{-10} to 2×10^{-8} (L-shells)^2 per day and $K^{(E)}$ may be found in the range 10^{-6} to 10^{-4} (L-shells)^2 per day. The total radial diffusion coefficient is then

$$D_{LL} = D_{LL}^{(M)} + D_{LL}^{(E)}. \qquad (5.57)$$

It should be emphasized that the relations in Equations (5.54) and (5.56) represent simplified considerations that may not always be realized. If, for example, $\bar{P}^{(M)}$ or $\bar{P}^{(E)}$ do not follow ω^{-2} dependences, then $D_{LL}^{(M)}$ will involve dependence on μ, and $D_{LL}^{(E)}$ may have a different form.

5.3.5 Pitch Angle Diffusion

The presence of plasma and electromagnetic waves in the radiation belts implies fast low-amplitude field fluctuations, and some of these waves (such as the ELF whistler mode) can violate the first adiabatic invariant μ. These fast fluctuations will, in principle, also violate Φ and J; however, for these adiabatic invariants the effect is likely to be at least in part averaged out. It is convenient to convert from μ to equatorial pitch angle coordinates

$$\mu = \frac{p^2 L^3 \sin^2\alpha_o}{B_E} \qquad (5.58)$$

whereby the applicable Jacobian is

$$G(\alpha_o; \mu) = \frac{\partial \mu}{\partial \alpha_o} = \frac{p^2 L^3 \sin 2\alpha_o}{B_E} \qquad (5.59)$$

at constant particle momentum. In the pure pitch angle diffusion limit $|\mathbf{p}|$ is unchanged and only the particle *direction* of motion changes.

The pure pitch angle diffusion equation becomes

$$\frac{\partial f}{\partial t} = \frac{1}{\sin 2\alpha_o T(\alpha_o)} \frac{\partial}{\partial \mu} \left[\bar{D}_{\alpha_o \alpha_o} \sin 2\alpha_o T(\alpha_o) \frac{\partial f}{\partial \alpha_o} \right] + \underline{S}^* - \underline{L}^* \qquad (5.60)$$

where $\underline{S}^*$ and $\underline{L}^*$ are the source and loss functions appropriate for the pitch angle diffusion process. $\bar{D}_{\alpha_o \alpha_o}$ is the bounce averaged pure pitch angle diffusion coefficient related to the local pitch angle diffusion coefficient $D_{\alpha\alpha}$ by

$$\bar{D}_{\alpha_o \alpha_o} = \frac{1}{\tau_b} \int_0^{\tau_b} D_{\alpha\alpha} \left(\frac{\partial \alpha_o}{\partial \alpha} \right)^2 dt. \qquad (5.61)$$

The actual calculation of $\bar{D}_{\alpha_o \alpha_o}$ from observed plasma wave distributions in the magnetosphere is quite complicated. The reader is referred to the research literature [Lyons et al., 1971, 1972; Retterer et al. 1983 and references therein].

In general, pitch angle diffusion can come about by violation of μ only (that is change in $p_\perp$) by violation of J only (change in $p_\parallel$), or by a combined violation of both μ and J. The pure pitch angle diffusion limit is ideally realized when both μ and J are violated so that the ratio of $p_\perp / p_\parallel$ changes while $|\mathbf{p}|$ remains almost constant. When this happens, there is essentially no energy exchange between waves and particles. Physically, this can mean that the principal interaction is between the particle and the magnetic field of the wave.

5.3.6 Energy Diffusion

Particles can become energized when interacting with waves (wave damping) or can lose energy to the waves (wave instability). One may write a pure energy diffusion equation in the form

$$\frac{\partial f}{\partial t} = G^\dagger \frac{\partial}{\partial \epsilon} \left[D_{\epsilon\epsilon} G^\dagger \frac{\partial f}{\partial \epsilon} \right] + S^\dagger - L^\dagger. \qquad (5.62)$$

However, such an equation has not been used much in radiation belt physics, since almost inevitably α_o also changes. The more general case of coupled energy and angular diffusion and the associated diffusion coefficients needs further research. For information relating to such coupled processes, see Schulz and Lanzerotti [1974].

5.4 LOSS MECHANISMS

Energetic particles residing in the radiation belts are subject to collisional interactions with coexisting particle populations. The most important of these are the earth's main atmosphere, the atomic hydrogen exosphere, and the plasmasphere. Near the earth (that is, at very low L-shells or for small equatorial pitch angles) such collisions constitute a dominant energetic particle loss mechanism. But even in the central parts of the radiation belts Coulomb collisions and charge exchange can be quite significant. Pitch angle scattering of particles into the atmospheric bounce loss cone (where particle-particle collisions are dominant) is also of great significance, especially for radiation belt electrons.

5.4.1 Exosphere

The terrestrial exosphere (or geocorona) is a continuation of the atmosphere to great altitudes where collisions are infrequent and the constituents follow ballistic trajectories. The principal constituent is thought to be atomic hydrogen with a density ranging from $\sim 10^4$ atoms cm^{-3} at 10^3 km altitude to $\sim 10^2$ atoms cm^{-3} at 3×10^4 km altitude. Table 5-1 gives the mean atomic

Table 5-1 Number density of neutral hydrogen at the equator for an exospheric temperature of 950 K and for average geomagnetic conditions [Tinsley, 1976]

T = 950 K			
L-Shell	[H] (1 cm^{-3})	L-Shell	[H] (1 cm^{-3})
1.1	15000.	1.5	16000.
2.0	3700.	2.5	1500.
3.0	800.	3.5	470.
4.0	300.	4.5	210.
5.0	148.	5.5	120.
6.0	98.	6.5	83.

hydrogen number density [H] as function of L-shell at the equator [for example, Tinsley, 1976] and it is thought to be an average representation for an exospheric temperature of ~950 K. The exospheric temperatures and densities will of course change with solar and geomagnetic activity.

Also overlapping the radiation belts is the terrestrial plasmasphere consisting of thermal ions and electrons and contained within an L-shell range roughly below about L-4 to 6 by the effect of the corotational electric field of the earth. Figure 5-11 shows the electric equipotential lines in the equatorial plane formed by the combination of the corotational electric field and the solar wind induced dawn-dusk "convection" electric field. The separatrix between closed (around the earth) and open equipotential lines is related to the static plasmapause, although in a dynamic, time variable situation no simple relation between the two exists. Empirically, there is a much higher density of cold (1000 - 10000 K) plasmaspheric particles below the plasmapause than beyond it. Based on data deduced from ducted VLF wave propagation experiments, Cornwall [1972] estimated the following average plasmaspheric particle densities:

$$[e] = 250\,(L - L_0)^K \text{ cm}^{-3} \text{ for } L \leq L_0$$
$$[e] = 13\,(L - L_0)^K \text{ cm}^{-3} \text{ for } L > L_0 \qquad (5.63)$$

where $L_0 \approx 4.1$ and $K \approx 4.64$.

It is well known that the plasmasphere deviates frequently and strongly from this functional form. It is nevertheless thought that this formula represents average long-term cold plasma densities appropriate for long-term steady state radiation belts studies. Further improvement, including storm and substorm variability, will be needed when time dependent radiation belt models are developed.

5.4.2 Coulomb Collisons

Coulomb collisions are inelastic interactions between charged particles. A radiation belt particle "colliding" with an exospheric neutral hydrogen atom, for example, will interact with the internal atomic electric field whenever the impact parameter is less than the atomic radius, or with the electric field from a thermal (plasma) proton or electron out to the particle's Debye shielding distance. The encounter will result in energy transfer from the energetic (incident) particle, and in deflection (angular scattering) of both particles. Changes in ionization states of either or both particles can also occur. On the average, angular scattering is important for radiation belt electrons, but usually not so important for the much heavier radiation belt ions. The collisional scattering process is qualitatively illustrated in Figure 5-13.

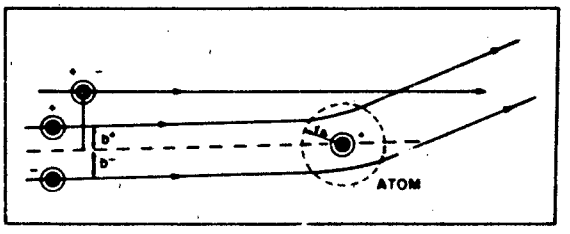


Figure 5-13. Deflection of a positive and negative particle in the internal atomic electric field of the target atom. b' and b⁻ are impact parameters and r_A is the atomic radius. If the impact parameter is greater than r_A, no interaction occurs. For plasma particles r_A is replaced by the Debye shielding distance λ_D.

Radiation belt particles are much more energetic than exospheric particles and energy is transferred from the energetic ion to the atomic bound electron(s) or plasma electrons. The changes in the energy spectrum and directional characteristics of particles traversing a material substance whose effective thickness is greater than the collision mean free path is a difficult mathematical problem. Great simplification is, however, obtained when the differential collision cross sections are small enough to overwhelmingly favor very small energy losses and directional changes in each collision. Fortunately, this is true for Coulomb collisions in which the classical Rutherford cross section is valid. See Rossi and Olbert [1970] or Jackson [1975] for details.

One finds that the average energy loss rate for an

energetic particle passing through a gas of atomic particles is given by

$$\frac{d\epsilon}{dt} = -4Z_i^2 Z_t [n] r_e (m_e c^3 \beta) F(\beta, Z_t) \quad (5.64)$$

where Z_i is the net charge state number of the incident particle, Z_t is the nuclear charge number of the target gas atoms, m_e is the electron mass, r_e is the classical electron radius, c is the speed of light, $\beta = V/c$ and $F(\beta, Z_t)$ is a slowly varying function determined quantum mechanically by

$$F(\beta, Z_t) = 2.9 + \ln \left[\pi^2 m_e^2 c^4 (1-\beta^2)^{1/2} I^2(Z_t)\right]$$

for energetic electrons, (5.65)

and

$$F(\beta, Z_t) = 2\beta^2 + \ln \left[4 m_e^2 c^4 \beta^4 (1-\beta^2)^2 I^2(Z_t)\right]$$

for energetic ions (5.66)

where $I(Z_t) = 13.5\ Z_t$ (in eV) is an approximate value of the ionization potential for the gas atoms [Rossi and Olbert, 1970].

Let $f = f(\mu, J, \Phi; t)$ denote the distribution function for equatorial radiation belt particles. The changes due to stochastic energy loss from the energetic particles can be described by

$$\left(\frac{\partial f}{\partial t}\right)_{Coulomb} = \frac{\partial}{\partial \mu}(<\Delta\mu\ \Delta t> f) - \frac{\partial}{\partial J}(<\Delta J\ \Delta t> f), \quad (5.67)$$

and the stochastic time average $<\Delta\mu\ \Delta t>$ and $<\Delta J\ \Delta t>$ are then simply given by

$$<\Delta\mu\ \Delta t> = \left(\frac{\partial \mu}{\partial \epsilon}\right)\left(\frac{d\epsilon}{dt}\right) \quad (5.68)$$

and

$$<\Delta J\ \Delta t> = \left(\frac{\partial J}{\partial \epsilon}\right)\left(\frac{d\epsilon}{dt}\right). \quad (5.69)$$

Cornwall [1972] noted that it is possible to treat the slowly varying logarithmic term in the expression for $\frac{d\epsilon}{dt}$ as an approximate constant. By including contributions from energetic particle plasma electron collisions as well as from energetic particle exospheric atom collisions, one derives the expression valid for equatorially mirroring ($\alpha_{eq} = \pi/2$) particles:

$$\left(\frac{\partial f}{\partial t}\right)_{Coulomb} = G_C(L)\, \mu^{1/2} \left(\frac{\partial f}{\partial \mu}\right) \quad (5.70)$$

The Coulomb collision loss factor is given by

$$G_C(L) = (50\sqrt{2}\ \pi e^4 \sqrt{m_H}\ L^{9/2}/(m_e B_E^{3/2}))\,[H]\, Z_i^2 \quad (5.71)$$

where $m_H = 1.67 \times 10^{-24}$ grams is the hydrogen atom mass and $e = (4.80286 \pm 0.00009) \times 10^{-10}$ esu is the unit charge. Z_i is the ionic charge state number for incident energetic ions [Cornwall, 1972; Spjeldvik, 1977]. For a fuller treatment of the collisional process the reader is referred to the treatise by Mott and Massey [1952] or Rossi and Olbert [1970].

5.4.3 Charge Exchange

Collisional encounters may also change the ionization state of the colliding particles. This may entail ionization of the "target" atom and or alteration of the net ionic charge of the incident particle. The latter is of importance for radiation belt ions since almost all of the physical processes depend directly on their charge state. The simplest of the charge exchange reactions is the one that neutralizes radiation belt protons (H):

$$\underline{H}^+ + H \rightarrow \underline{H} + H^+ \quad (5.72)$$

where underlying denotes energetic particles. The above reaction is an example of the electron capture process in which the incident proton picks up (or captures) the orbital electron from the thermal hydrogen atom, which then becomes a low energy proton. Once neutralized, the incident proton (now fast neutral hydrogen) is no longer subject to the magnetic deflecting force and escapes from the trapping region. Macroscopically these events occur at random, and fast neutral atoms thus exit the radiation belts in all directions. Since the speed of these particles greatly exceeds the earth's gravitational escape speed, many disappear to outer space. Some of these fast neutral atoms move towards the earth where they produce secondary interactions upon entering the atmosphere.

In a similar manner, other singly charged ions can be neutralized in such collisions:

$$\begin{aligned} \underline{He}^+ + H &\rightarrow \underline{He} + H^+ \\ \underline{C}^+ + H &\rightarrow \underline{C} + H^+ \\ \underline{O}^+ + H &\rightarrow \underline{O} + H^+ . \end{aligned} \quad (5.73)$$

These newly generated thermal hydrogen ions (protons) make a contribution to the earth's plasma envelope; however, their rate of formation is probably less than the rate at which such ions are supplied from the topside

ionosphere. One may note that the total charge is conserved in the radiation belts under the charge exchange reactions.

There is also another class of ion charge exchange; this results from electron stripping reactions by which the energetic ions (necessarily heavier than hydrogen) lose one or more of the remaining bound electrons. For example, there is a finite probability that an ion (say O^+) in a collision with a thermal hydrogen atom loses several of its electrons:

$$\underline{O}^{+} + H \rightarrow \underline{O}^{5+} + H^{+} + 4\underline{e}^{-} \qquad (5.74)$$

In this example four electrons are stripped off. Whether the fast O^+ ion actually loses one or more of its electrons or captures one from the hydrogen atom strongly depends on the energy of the incident ion and the details of the collision. Since this type of reaction preserves a non-zero ionic state, the ion remains trapped within the radiation belts. For this reason one distinguishes between "internal" charge exchange reactions such as the stripping reactions:

$$\begin{aligned} \underline{O}^{2+} + H &\rightarrow \underline{O}^{3+} + H + \underline{e}^{-} \\ \underline{O}^{3+} + H &\rightarrow \underline{O}^{7+} + H + 4\underline{e}^{-} \end{aligned} \qquad (5.75)$$

with their reverse reactions such as

$$\begin{aligned} \underline{O}^{8+} + H &\rightarrow \underline{O}^{7+} + H^{+} \\ \underline{O}^{7+} + H &\rightarrow \underline{O}^{6+} + H^{+} \\ \underline{O}^{2+} + H &\rightarrow \underline{O}^{+} + H^{+} \end{aligned} \qquad (5.76)$$

on one hand, and reactions that produce immediate particle loss (lowest charge state to neutral) noted above on the other. Notice also that to become neutralized, an O^{8+} ion requires a minimum of eight separate collisions with hydrogen atoms (since only one electron may be captured in each collision). This should be contrasted with electron stripping reactions in which the multiple charge state changes occur in a single collisional encounter. The principal features of the charge exchange chemistry are illustrated in Figure 5-14.

The probability that a given charge exchange process actually takes place in a collision is expressed as a reaction cross section. Such cross sections derive from laboratory studies of collision processes; however, for many ions the pertinent cross sections have not yet been measured over a sufficiently large range of particle energies. Figure 5-15 shows the cross sections for the charge exchange of protons incident on atomic hydrogen from a compilation by Spjeldvik [1977]. Notice that the charge exchange cross section is quite high ($\sigma_{10}^{P} > 10^{-15}$ cm^2) below about 20 keV, but falls off sharply with higher energies.

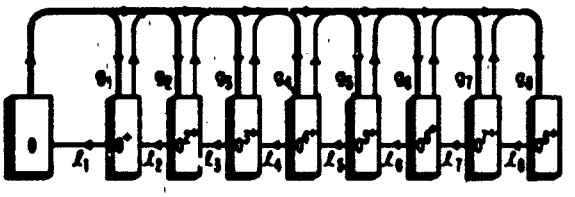


Figure 5-14. Flow diagram for energetic oxygen ion charge exchange. This figure illustrates that while multiple charge state increases are possible in a single encounter, only single step charge state reductions can occur in a hydrogen atom gas. Oxygen ions in lower charge states may be further ionized (through single or multiple electron stripping) in single collisional encounters with the exospheric gas atoms. The ions may lose net charge in collisions by capturing the bound electron from a thermal hydrogen atom. To become neutralized, a fully ionized oxygen ion requires a minimum of eight separate collisions.

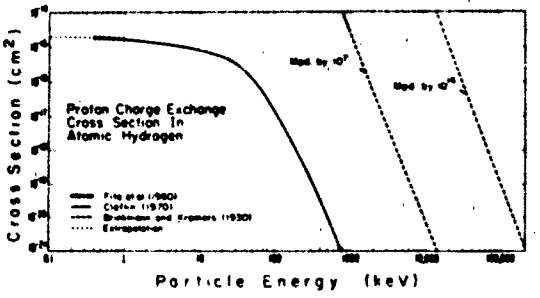


Figure 5-15. Proton charge exchange cross section as function of proton energy. It is worth noting that this cross section is high ($\sigma 10^{P} \sim 10^{-15}$ cm^2) below a few tens of keV causing short proton lifetimes at those energies, and it falls off rapidly towards higher energies. The values between 0.4 and 1 keV are from the experimental work of Fite et al. [1960], the data from 1 keV are due to a compilation by Claflin [1970], and above 1000 keV the theoretical results from Brinkmann and Kramers [1930] have been used [Spjeldvik, 1977].

For ions heavier than protons, multiple potential charge states are available. For helium ions, one must consider not only the cross section for the neutralization reaction but charge state changes: state 1 → state 2 and state 2 → state 1 as well. Thus, for helium there are three important cross sections to be included. A point worth noting is that the process transforming He^+ to He^{2+} dominates over the charge state reducing reactions in the high energy part of the radiation belts, essentially above 800 keV. This is explicitly depicted in Figure 5-16 which shows the three charge exchange cross sections for radiation belt helium ions.

Charge exchange cross sections for the heavier ions such as carbon and oxygen have been measured over a small fraction of the radiation belt energy range. The situation is particularly severe for carbon ions where

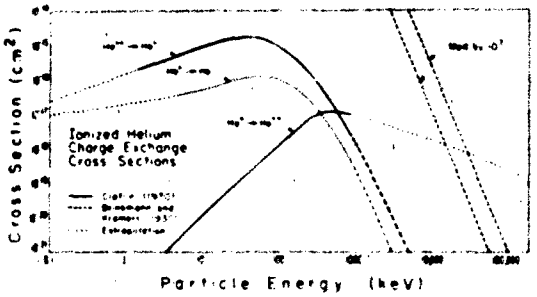


Figure 5-16. Helium ion charge exchange cross section as function of energy. Solid lines show the values compiled by Claflin [1970] from many experiments, and the dotted lines are extrapolations. The dashed lines are computed from the theoretical work of Brinkmann and Kramers [1930]. The three pertinent charge exchange cross sections are: $He^+ \rightarrow He$ (ion neutralization), $He^{2+} \rightarrow He^+$ and $He^{2+} \rightarrow He^+$. The extrapolated curve above 1000 keV (for $He^+ \rightarrow He^{2+}$) is progressively unreliable towards higher energies [Spjeldvik and Fritz, 1978b].

almost no measurements have been made. For oxygen ions there exist a number of measurements, but unfortunately the laboratory work used particles other than atomic hydrogen as targets. Figure 5-17 shows estimated

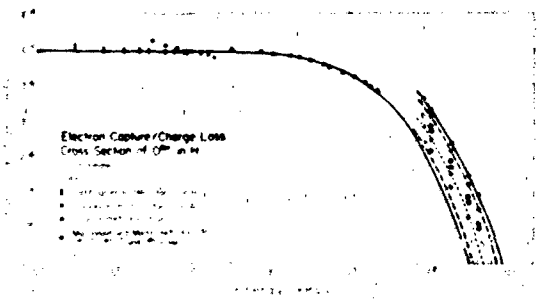


Figure 5-17. Charge exchange cross sections for (atomic) oxygen ions in an atomic hydrogen gas: Charge loss Electron capture cross sections $O^+ \rightarrow O^{i+}$ estimated from sparse available data. Stebbings et al. [1960] give data with H-targets up to 8 keV. Lo et al. [1971] give data with O-targets and MacDonald and Martin [1971] give data with He-targets. Also shown are results with rarefied air-targets [Solov'ev et al., 1972]. The cross sections for reactions with $i>1$ have not been measured below ~ 10 MeV [Spjeldvik and Fritz, 1978a].

ion charge loss (electron capture) cross sections based on a compilation of a number of measurements using H, air and O as target particles [Spjeldvik and Fritz, 1978a]. Estimates of the charge gain (electron loss) cross sections have been made from limited data available. Examples are illustrated in Figure 5-18. For the remaining reactions between the multiple charge states, no laboratory measurements have been reported, and one must use crude estimates [Spjeldvik and Fritz, 1978a].

The mathematical description of the effects of the

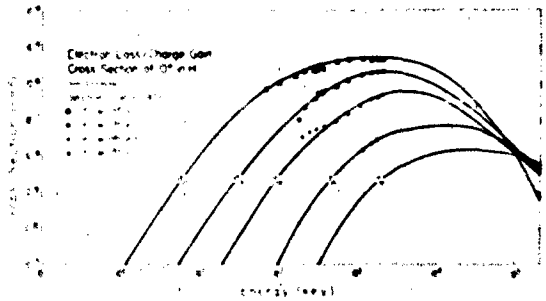


Figure 5-18. Charge exchange cross sections for (atomic) oxygen ions in an atomic hydrogen gas: Charge gain Electron loss cross sections $O^+ \rightarrow O^{(i-1)+} + j\,e^-$. The data shown are from Lo et al., [1971] for O-targets. The solid lines are crude estimates made by Spjeldvik and Fritz [1978a]. More reliable cross sections for oxygen ions and other heavy ions need to be established.

charge exchange processes on radiation belt ion distribution functions is through loss and gain terms. For radiation belt protons one has the expression

$$\left(\frac{\partial f}{\partial t}\right)_{\text{Charge Exchange}} = -\Lambda_{10} f \qquad (5.77)$$

since the loss rate is proportional to the number of protons present (or more precisely, the distribution function f), and also proportional to the factor $\Lambda_{ij} = \langle V\sigma_{ij}[H]\rangle$ where V is the ion speed and σ_{ij} the charge exchange cross section for transformation state i → state j. To describe the heavier ions where more than one charge state applies, one considers a distribution function for each charge state. For helium ions, let f_1 and f_2 be the distribution functions for He^+ and He^{2+} ions respectively. The charge exchange processes are then represented by

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{Charge Exchange}} = -\Lambda_{10} f_1 + \Lambda_{21} f_2 - \Lambda_{12} f_1 \qquad (5.78)$$

$$\left(\frac{\partial f_2}{\partial t}\right)_{\text{Charge Exchange}} = -\Lambda_{21} f_2 + \Lambda_{12} f_1 \qquad (5.79)$$

and this provides a coupling between the distribution functions.

In general, for an ion species with s available charge states one obtains charge exchange expressions of the form

$$\left(\frac{\partial f_i}{\partial t}\right)_{\text{Charge Exchange}} = \sum_{j=0}^{s} \Lambda_{ij} f_i + \sum_{j=1}^{s} \Lambda_{ji} f_j \quad \text{for } i \neq j, \qquad (5.80)$$

and when the exosphere predominantly consists of

atomic hydrogen (as in the case of the earth) the first summation contains contributions from j = i − 1 only.

5.4.4 Wave-Particle Interactions

Plasma waves play an important role in radiation belt physics. Different conditions for wave propagation, growth and decay in the magnetosphere have led to delineation into numerous *wave modes*. For a detailed classification of these modes see Stix [1962] or Kennel et al. [1979]. In a previous section, the radiation belt transport equation was given; here we shall outline some of the concepts concerning the interactions.

Angular scattering can result from interactions with electromagnetic waves. As an example, consider gyro (or cyclotron) resonance: a radiation belt particle spiraling around a magnetic line of force (the guiding center locus field line) will gyrate at a rate determined by the magnetic field strength, mass, and charge of the particle. An electromagnetic wave (with electric and magnetic wave vectors) also propagating along that field line will rotate according to its wave frequency. When both the sense of rotation and the rotation frequency match for both wave and particle, the particle will be subject to an essentially constant "wave" field for the duration of the encounter. The particle can exchange energy with the wave through the electrical interaction and or deflection can occur through the magnetic interaction. Higher order resonances can also take place, for example, if the rotation rates differ by a factor of two. When the particle-wave interaction is primarily via the magnetic wave vector (as in the case of the radiation belt electron whistler mode hiss wave interaction) the result is primarily angular scattering; this is the pure pitch angle scattering limit. Other resonances including bounce resonance and drift resonance can also be important.

These waves in the radiation belts may remain in the area where they were generated (very small group velocity) or propagate afar (large group velocity). They have frequency and wave length, and there are different polarization properties. Propagation properties are determined by a dispersion relation [Stix, 1962] that in part depends on the density of the plasma in which the wave exists and the geomagnetic field. Stability or instability of the waves is frequently determined by the energetic particles with which the waves can interact.

Plasma waves are said to be unstable to growth if interactions with the charged particles transfer energy to the waves; if the transfer is from the waves to the particles the waves are said to be damped. If plasma waves interact with radiation belt particles in such a way as to have their principal interactions locally, then it is possible to self-consistently give a theoretical treatment of both wave and particle properties. On the other hand, if the waves have significant spatial propagation so that they, for example, gain energy from one particle population and propagate to interact further with another particle population elsewhere, then the latter process is termed "parasitic". Both types of processes are of major significance in the radiation belts.

5.4.4.1 Pitch Angle Scattering Into the Loss Cone. Interactions between energetic particles and plasma waves can significantly affect the trapped particle population's energy and angular distribution. If the pitch angle is altered so that the particle finds itself within the atmospheric bounce loss cone, it will have a high probability of becoming lost upon entering the atmosphere. For ions this probability is virtually 100% while for energetic electrons a certain fraction of the precipitated particles are backscattered up into the magnetosphere only to encounter the conjugate hemisphere during the succeeding bounce motion.

The directional change in the angular scattering process due to plasma wave or collisional interactions is random: to lowest order it is just as likely for a single particle to be scattered from lower to higher pitch angles as in the opposite direction. However, when the particles are *anisotropically* distributed in pitch angle (for example, with $\frac{\partial f}{\partial \alpha_0} > 0$), the number of particles scattered from the higher particle density region towards the lower density region is greater than the number scattered the other way. As a consequence, the stochastic process is biased by the particle pitch angle distribution and a net *diffusion flux* occurs. This pitch angle diffusion flux is just

$$F_{diff} = \overline{D}_{\alpha_0\alpha_0} \sin 2\alpha_0 \, T(\alpha_0) \, \frac{\partial f}{\partial \alpha_0} \qquad (5.81)$$

so that the diffusive transport term in (5.45) may be written as simply

$$\left(\frac{\partial f}{\partial t}\right)_{diff} = \frac{1}{\sin 2\alpha_0 T(\alpha_0)} \, \frac{\partial F_{diff}}{\partial \alpha_0} \, . \qquad (5.82)$$

Atmospheric particle losses within the bounce loss cone generally cause $f(\alpha_0 < \alpha_{OLC})$ to be near zero (except under conditions of extremely strong scattering). This generally favors a positive pitch angle anisotropy (loss cone distribution), and particles can be lost to the atmosphere at any L-shell due to the pitch angle scattering process. Of course, $\partial f / \partial \alpha_0$ may be negative under certain conditions, and thus net reverse diffusion can take place. Examples are an atmospheric *source cone* in the auroral zone as a consequence of auroral electric fields, effects of L-shell splitting due to the deviations from the dipolar azimuthal symmetry of the magentic field [Roederer, 1970], or particle injections during disturbed times.

Analysis of resonant wave-particle interactions violating the first adiabatic invariant have generally considered waves with frequencies near the gyro frequency of the energetic particles. The condition for resonance between waves of angular frequency ω and particles at the gyrofrequency Ω is given by

$$\omega - k_{\parallel} v_{\parallel} = n\,\Omega;\ n = 0, \pm 1, \pm 2, \pm 3, \ldots \text{ [Lyons, 1979]}, \tag{5.83}$$

where $k_{\parallel}$ and $v_{\parallel}$ are the parallel (to **B**) wave vector and particle velocity respectively. For a given wave mode there is a dispersion relation linking ω and **k**. The *cyclotron harmonic resonances* have the (Doppler shifted) wave frequency equal to a harmonic ($n = \pm 1$, $n = \pm 2$, $n = \pm 3$, ...) of the particle gyrofrequency, and the classical *Landau resonance* ($n = 0$) has the wave parallel phase velocity $v_{\parallel} = \omega / k_{\parallel}$. All these resonances result in particle diffusion in both energy and equatorial pitch angle α_o.

The Landau resonance results in diffusion solely in $v_{\parallel}$, conserving $v_{\perp}$. Thus the pitch angle change $\Delta\alpha$ at a given location along the particle trajectory is related to the parallel velocity change $\Delta v_{\parallel}$ by

$$\Delta\alpha = -\frac{\sin^2\alpha}{v_{\perp}}\,\Delta v_{\parallel}. \tag{5.84}$$

The cyclotron harmonic resonance can likewise produce pitch angle changes resulting from the energy exchange, or the interaction can be primarily with the magnetic wave vector causing pitch angle scattering more directly with little energy exchange. The latter process is particularly important for electrons.

In general, the study of wave-particle interactions requires an extensive mathematical treatment, and the necessary derivations and analysis are beyond the scope of the present chapter. The interested reader is referred to a number of works on the subject [Stix, 1962; Sagdeev and Galeev, 1969; Lerche, 1968; Kennel and Engelman, 1966; Lyons et al., 1971, 1972; Retterer et al., 1983].

5.4.4.2 Scattering of Energetic Electrons. Pitch angle scattering is particularly important for radiation belt electrons. Here, we bypass the extensive mathematical treatment found in Lyons et al. [1971, 1972]. The different resonances are illustrated in Figure 5-19 which depicts the regions of cyclotron resonance in velocity space resulting from waves distributed over a band of parallel wave vectors $\Delta k_{\parallel}$. Resonance at each cyclotron harmonic occurs over a band of parallel velocities $\Delta v_{\parallel}$; for simplicity, relativistic effects for electrons >500 keV are not included. No cyclotron harmonic interaction occurs for $v < v_{\parallel min}$, and the value of $v_{\parallel min}$ depends on the actual $k_{\parallel}$ band over which the wave energy is distributed. This is locally true where the geomagnetic field may be considered uniform on the scale of the gyroradius.

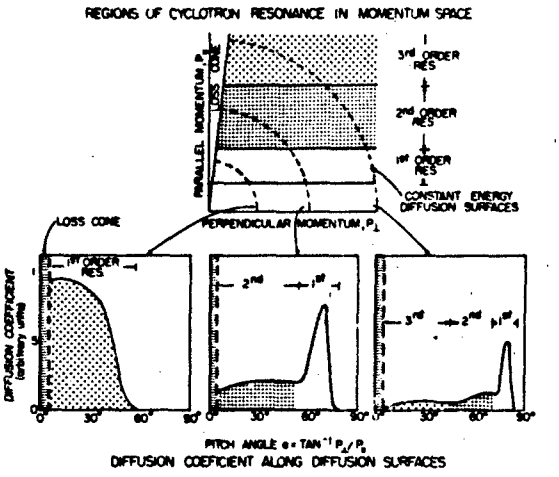


Figure 5-19. Radiation belt electron – ELF whistler mode wave interactions: Regions of cyclotron resonance in momentum space. Upper panel shows the effects of pure pitch angle diffusion (conserving electron energy) for the resonance with plasmaspheric ELF whistler mode turbulence. Lower panels show the effect upon progressively more energetic electrons, and that for the most energetic electrons the high order resonances become increasingly important as the electrons diffuse in pitch angle towards the atmospheric bounce loss cone. The resulting pitch angle diffusion coefficient thus becomes a strong function of pitch angle [Lyons et al., 1972].

The earth's magnetic field is, however, quite inhomogeneous when considered in its totality, and the wave energy is, in general, unevenly distributed over the space of the radiation belts. During the particle bounce motion, as the particles move away from the geomagnetic equator along its trajectory, the increasing magnetic field strength causes both particle pitch angle and parallel velocity to change. The increase in the local pitch angle as a particle moves away from the equatorial plane (see Equation 5.10) implies that cyclotron resonance can occur for all equatorial parallel particle energies greater than a minimum value $E_{\parallel min}$.

Figure 5-20 illustrates radiation belt electron and ELF whistler mode wave propagation. The significant wave-particle interactions for energetic electrons are shown. Note that the wave propagation does not necessarily follow magnetic field lines, and therefore the particles may interact with waves generated over a significant volume of the magnetosphere. For comparison, the inner and outer radiation zones (for electrons) are indicated.

Satellite measurements have shown that a band of whistler mode waves centered around a few hundred hertz exists essentially continually within the plasmasphere. Because of the persistence and audio frequency range, this type of wave is called *hiss*. Ray tracing studies have shown that the plasmaspheric hiss can readily propagate across the geomagnetic field lines [Lyons and Thorne, 1970] and thus fill a great volume of the inner

magnetosphere with wave energy. These waves are believed to be generated in the outer regions of the plasmasphere, and within the plasmasphere the ELF hiss turbulence is the dominant wave component that interacts with radiation belt electrons. Waves that may influence trapped particles can also be generated by earth-based radio sources.

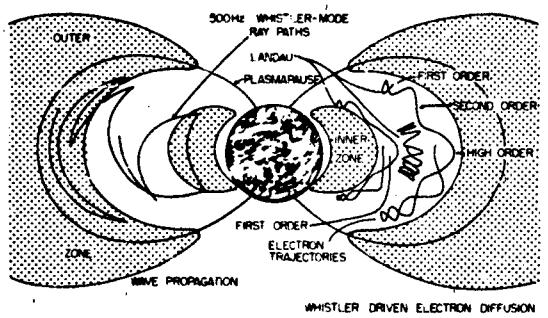


Figure 5-20. Spatial illustration of radiation belt electron wave-particle interactions. Characteristic locations for the inner and outer electron radiation zones are shaded, and the magnetic field line approximating the average plasmapause location is also indicated. Left side: 0.5 kHz whistler-mode ray paths in the radiation belts showing internal reflection and cross-L propagation characteristics [from Lyons and Thorne, 1970]. Right side: Typical energetic radiation belt electron trajectories indicating the spiral motion between the magnetic mirror points. The spatial regions where the Landau and different cyclotron resonances are most effective are noted [Lyons et al. 1972].

Based on typical observed wave characteristics, Lyons et al. [1972] calculated the pitch angle diffusion coefficient for both cyclotron and Landau resonances shown in Figure 5-21. They also computed lifetimes for energetic electrons subject to this wave-particle interaction process. An example of these lifetimes for average modeled wave parameters and a normalized wave amplitude of 35 mγ is shown in Figure 5-22. These lifetimes take on great significance in the modeling of the radiation belt electron structure and will be discussed in more detail in the modeling section.

5.4.4.3 Limit On Radiation Belt Particle Fluxes. Plasma waves generated by radiation belt particles locally will have their growth rates in part controlled by the intensity of the trapped energetic flux of those particles in resonance with the waves. While radiation belt electrons within the plasmasphere are controlled by the parasitic type of interaction mechanisms, it has been found that the radiation belt particle fluxes beyond the plasmasphere can be effectively limited by the self-generating wave mechanism. In the following a few principal aspects of the radiation belt saturation process are outlined; for a more detailed mathematical treatment the reader is referred to Kennel and Petschek [1966] or Schulz [1975b].

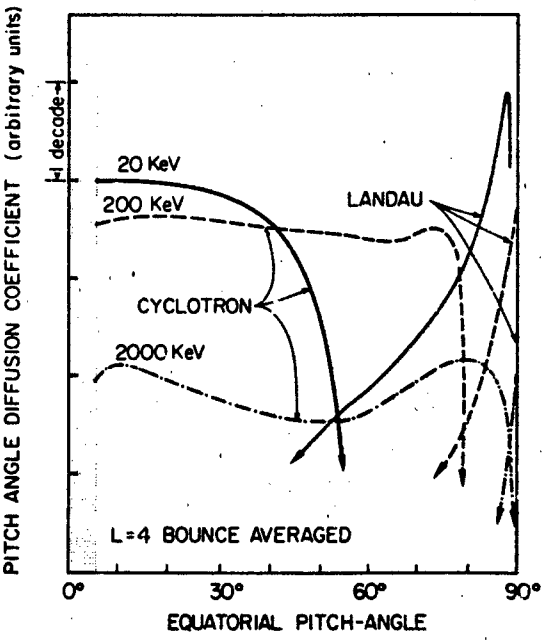


Figure 5-21. The bounce averaged electron pitch angle diffusion coefficient $D_{\alpha_0\alpha_0}$ calculated for all cyclotron harmonic resonances and the Landau resonance. Note that at each energy there is a region of very low pitch angle diffusion (for 20 keV electrons, near 87°); this "bottleneck" in the pitch angle diffusion coefficient is the cause of the "bumps" in the actual electron pitch angle distributions giving rise to the so-called bell-shaped distributions [Lyons et al., 1972].

It has been demonstrated that there is a limiting value I′ that the radiation belt integral omni-directional flux I cannot exceed without provoking a cyclotron wave instability. Schulz and Lanzerotti [1974] estimated this limit to be about $I' = 10^{11} L^{-4}$ particles cm^{-3} sec^{-1} integrated over all energies and pitch angles. The linear wave growth rate is γ_g, and therefore the growth rate for wave energy is $2\gamma_g$. An incipient wave undergoes a partial reflection (reflection coefficient $R \leq 1$) upon traveling a distance $d \sim L R_E$; the remaining fraction, 1-R, of the wave is lost from the radiation belts. The time interval between wave reflections is $T = L R_E / V_g$ where $V_g = (|d\omega/dk|)$ is the group wave velocity. The condition for marginal stability is that the waves on the average do not grow further in time; that is, the decrease in the wave amplitude upon reflection is restored in one traversal between bounces:

$$R \exp\left[2\gamma_g L R_E / V_g\right] = 1. \qquad (5.85)$$

This defines the marginal growth rate to be

$$\gamma_g = (V_g / 2LR_E) |\ln R|. \qquad (5.86)$$

If I exceeds I′ (that is, if γ exceeds γ_g) the consequence is a net growth of wave energy, and the stronger waves

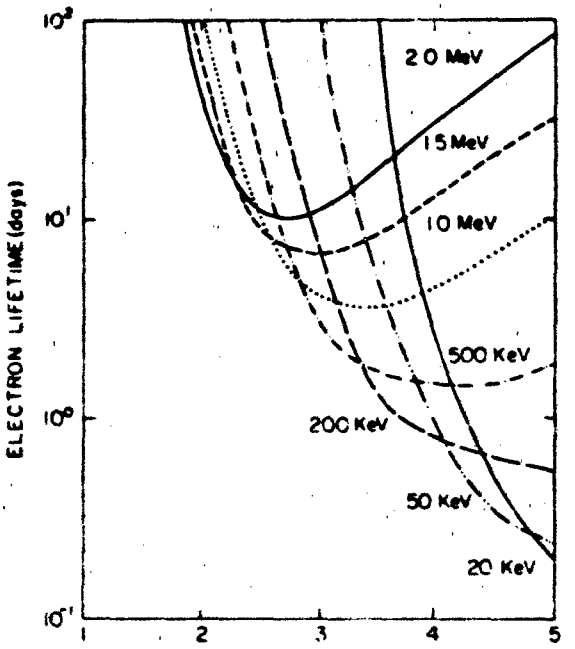


Figure 5-22. Radiation belt electron precipitation lifetimes due to electron pitch angle scattering into the atmospheric bounce loss cone. These lifetimes are given for an average whistler wave amplitude of $B_{w_{rms}} = 35$ mγ. Notice that there is generally a minimum in these lifetimes at each energy, and that this minimum is displaced towards lower L-shells with higher electron energy. At 2000 keV this minimum occurs at L = 2.7 while at 500 keV it is substantially broader and occurs at L = 4.2. This minimum electron lifetime is an important factor for the electron "slot" location defining the separation between the two radiation belts.

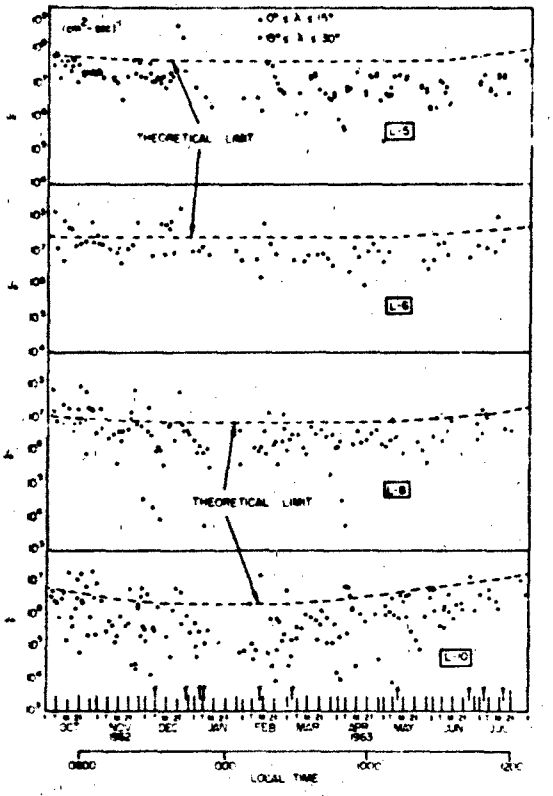


Figure 5-23. Comparison of the predicted whistler self-excitation limit with E>40 keV electron data recorded with the Explorer 14 spacecraft [Kennel and Petschek, 1966].

scatter the particles in pitch angle resulting in precipitation into the atmosphere and reduction in the particle flux until I no longer exceeds I'. If an external particle *injection source* is strong enough to more than offset this maximum loss rate then the radiation belt particle fluxes will increase beyond the stably trapped flux limit I'; this may indeed happen for impulsive sources associated with disturbed conditions. The aftermath of such an injection will, however, be a rapid reduction to the stably trapped flux intensity level.

Figure 5-23 shows a comparison of the theoretical radiation belt flux limit with data [Kennel and Petschek, 1966]. Notice that most of the time the outer radiation zone particle fluxes are just below the theoretical characteristic flux limit.

5.4.4.4 Strong Diffusion Limit. There is, however, another important limit to consider, namely the upper limit on the rate at which radiation belt particles can become lost to the atmosphere by scattering into the bounce loss cone. Consider the effect of very rapid pitch angle scattering. The particle pitch angle distribution will then become essentially isotropic; that is, the flux per differential solid angle interval will be the same at all pitch angles, even within the loss cone. Within the loss cone ($\alpha_0 \leq \alpha_{OLC}$) the particles will be lost from the radiation belts on a time scale τ_b, and for the *strong diffusion limit* to be reached, the scattering mechanism must be faster than this. The solid angle of the loss cone is just $\Omega_{LC} = 2\pi(1 - \cos\alpha_{OLC})$ and the maximum particle flux loss rate δ is then

$$\delta = \frac{\Omega_{LC}}{\tau_b} . \qquad (5.87)$$

The only exception to this concept occurs when the loss cone becomes a source cone due to the auroral electric field parallel to the magnetic field, and a source cone distribution exiting one hemisphere can become lost through direct precipitation into the conjugate hemisphere.

5.5 THEORETICAL RADIATION BELT MODELS

Two types of radiation belt models now exist: empirical models based on compiled observations and theoretical models derived from our knowledge of radiation belt

physics. The ideal empirical models represent the standards with which the theoretical models must be compared, and they are also useful for engineering purposes. However, they give only a smoothed statistical picture of the time period when the data were obtained, and the data collection process is subject to experimental errors and misinterpretations. Theoretical models, on the other hand, can be used to simulate and predict radiation belt behavior under a variety of conditions: quiet times, disturbed times, or magnetic storms. They can be used to study ionic species and charge state distributions for which no experimental information is available. Such models are only as valid as the physical approximation they are based on, however. In the following sections simple theoretical models are outlined.

5.5.1 Quiet Time, Steady State Models

Studies of quiet geomagnetic conditions using theoretical radiation belt models serve as an important test of our current understanding of the trapped radiation environment. It is assumed that under quiet time conditions the radiation belts can be described in terms of an equilibrium balance between sources, internal transport, and losses. Symbolically one writes

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{source} + \left(\frac{\partial f}{\partial t}\right)_{transport} + \left(\frac{\partial f}{\partial t}\right)_{loss} = 0 \quad (5.88)$$

The sources are considered to be located on the boundary of the trapping region (and thus describable via a boundary condition on the trapped flux for each species) as well as in the interior of the trapping region (such as the CRAND-source). The transport mechanisms are radial diffusion and pitch angle diffusion. Particle losses are caused by charge exchange ion neutralization, energy degradation, and pitch angle diffusion into the bounce loss cone. Particles diffusing outward may encounter the magnetopause and become lost from the magnetosphere. Radiation belt theoretical models constitute the combination of these processes.

5.5.1.1 Formulation of Boundary Conditions. For steady state conditions one needs in general boundary conditions on the particle distribution function f:

(1) At $L = L_{max}$ (outer boundary) one specifies the particles energy spectra and angular distributions, which defines f at the outer boundary.

(2) At $L = L_{min} = 1$ (inner boundary formed by the surface of the earth) the particle distribution function vanishes ($f = 0$).

(3) At $\alpha_0 = \pi/2$: $\partial f/\partial\alpha_0 = 0$.

(4) At $\alpha_0 = 0$ or π : $\partial f/\partial\alpha_0 = 0$.

(5) At $P = P_{max}$ (or $E = E_{max}$): $f = 0$ where $E_{max} \approx 100$ MeV.

(6) By symmetry $f(\alpha_0) \equiv f(\pi - \alpha_0)$ so that one needs to consider only the range $0 \leq \alpha_0 \leq \pi/2$.

The complete radiation belt modeling can be simplified in a way that retains much of the essential physics but greatly simplifies the mathematics, namely separating radial diffusion from pitch angle diffusion. One treats radial diffusion for $\alpha_0 = \pi/2$ particles only, and with given $f(\alpha_0 = \pi/2)$ the pitch angle diffusion process at fixed L-shells only. This approach is not strictly correct when there is significant interaction between the pitch angle and radial diffusion modes, but it reduces an almost unmanageable numerical problem to cases where solutions can be found. The results should therefore be treated with some caution.

5.5.1.2 Electron Model. Radiation belt electrons are described by the diffusion equation

$$\frac{\partial f}{\partial t} = L^2 \frac{\partial}{\partial L}\left[D_{LL} L^{-2} \frac{\partial f}{\partial L}\right] + \frac{1}{\sin 2\alpha_0 T(\alpha_0)} \frac{\partial}{\partial \alpha_0}\left[D_{\alpha_0\alpha_0} \sin 2\alpha_0 T(\alpha_0) \frac{\partial f}{\partial \alpha_0}\right] + G_c(L)\mu^{1/2} \frac{\partial f}{\partial \mu}, \quad (5.89)$$

where $G_c(L)$ is the Coulomb loss term for electrons. Lyons and Thorne [1973] separated radial and pitch angle diffusion by substituting for Equation (5.71) the two equations

$$\frac{\partial f}{\partial t} = L^2 \frac{\partial}{\partial L}\left[D_{LL} L^{-2} \frac{\partial f}{\partial L}\right] - \frac{f}{\tau_w} - \frac{f}{\tau_{cc}} \quad (5.90)$$

and

$$\frac{\partial f}{\partial t} = \frac{1}{\sin 2\alpha_0 T(\alpha_0)} \frac{\partial}{\partial \alpha_0}\left[D_{\alpha_0\alpha_0} \sin 2\alpha_0 T(\alpha_0) \frac{\partial f}{\partial \alpha_0}\right] \quad (5.91)$$

They solved Equation (5.91) under an assumed angular distribution preserving exponential decay conditions and obtained the lowest normal mode pitch angle distribution and electron precipitation lifetimes τ_w. Their results are shown in Figures 5-24 and 5-25 respectively. The precipitation lifetimes were calculated assuming a mean EL1 Whistler mode wave amplitude of 35nT. For other wave amplitudes, these lifetimes scale as $(B_w/35m\gamma)^{-2}$.

Using these characteristic electron scattering precipitation lifetimes it is possible to solve the steady state radial diffusion Equation (5.90) for equatorially mirroring radiation belt electrons. The results are shown in Figure 5-25.

A simplification in this work was *a priori* approximation of the Coulomb collision "loss time" which was defined as $\tau_{cc} \approx \epsilon(d\epsilon/dt)^{-1}$ where $d\epsilon/dt$ is given by Equation (5.64); this is a rather crude approximation for the last term in (5.90), since τ_{cc} really depends on the distribution function itself as given by Equation (5.66).

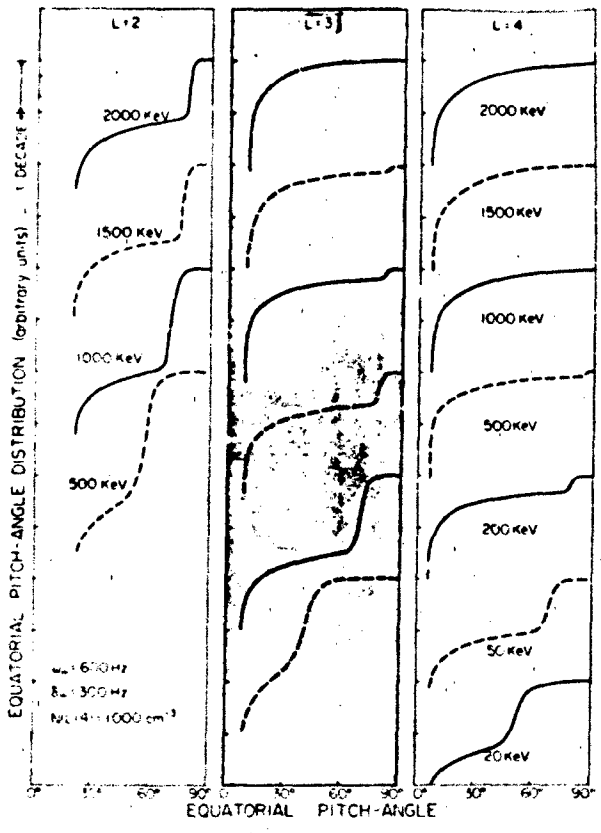


Figure 5-24. Predicted radiation belt electron pitch-angle distributions [Lyons and Thorne, 1973]. The example shown was calculated using a Gaussian ELF whistler mode hiss turbulence frequency distribution peaked at 600 Hz and with a bandwidth of 300 Hz.

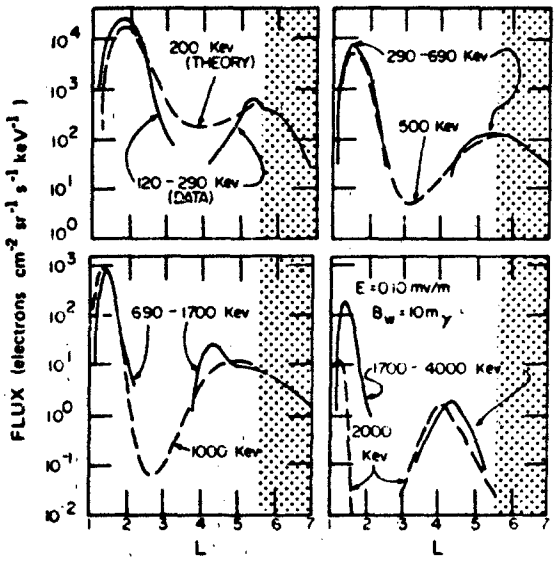


Figure 5-25. Theoretical and observed radiation belt electron radial flux distributions at the geomagnetic equator. The theoretical results (dashed curves) are calculated from radial diffusion of energetic electrons subject to pitch-angle scattering due to the (resonant) ELF whistler mode turbulence (hiss) within the plasmasphere. The data (solid lines) are taken from the OGO 5 spacecraft. Notice how well the theory simulates the location and width of the observed radiation belt separation (slot) region over a wide range of electron energies leaving essentially only the finer details of the equilibrium distribution for future investigation [Lyons and Thorne, 1973].

Much of the essential radiation belt electron physics is retained, however. Notice how well the electron slot region that separates the radiation zone into two belts is reproduced compared with actual data from the OGO-5 spacecraft, both in width and L-shell location for different energies. Perfect agreement in all details should not be expected since the theory treats average conditions based on parametrized wave characteristics, while the data are for a specific time when the recordings were made.

From such calculations we have learned that for energetic electrons, wave-particle interactions are the cause of the division into *two distinct belts*, since electrons in the "slot" region are lost to the atmosphere at a fast rate. Those few electrons that survive the cross-L transport to arrive in the inner radiation zone became very stable since the time scale for wave particle scattering is very long there.

5.5.1.3 Proton Model. Empirically, radiation belt protons are not distributed into two radiation zones. At constant proton energy there is generally a single peak in the proton flux radial distribution and this peak is located on lower L-shells with higher proton energy. No specific wave mode has yet been identified that interacts in a dominant way with >100 keV protons below the average plasmapause location. On the other hand, protons (like all ions) are subject to the charge exchange process.

Neglecting pitch angle scattering, the radial diffusion equation for equatorially mirroring ($\alpha_0 = \pi/2$) protons may be written

$$\left(\frac{\partial f}{\partial t}\right) = L^2 \partial/\partial L\left[D_{LL} L^{-2}(\partial f/\partial L)\right] - \Lambda_{10} f + G(L)\,\mu^{-1/2}(\partial f/\partial \mu) \tag{5.92}$$

where $\Lambda_{10} = \langle \sigma_{10}^{p} [H] v \rangle$ is the charge exchange loss frequency per unit distribution function f. To simulate the steady state radiation belts (that is $\frac{\partial}{\partial t} \to 0$) this equation can likewise be solved when suitable boundary conditions are imposed. The boundary conditions are established by using a measured outer zone proton flux spectrum, and by covering an energy range sufficiently large that $f(L,\mu) \approx 0$ at $\mu > \mu_{max}$ where μ_{max} corresponds to the upper limit of the radiation belt trapping

energies considered. Figure 5-26 shows the computed

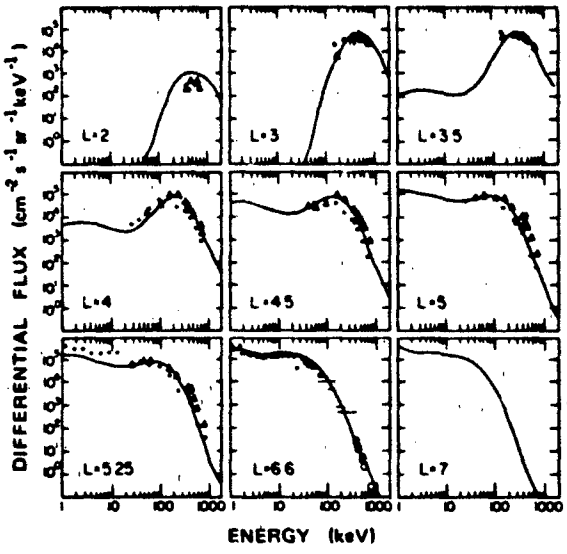


Figure 5-26. Radiation belt proton energy spectra obtained from a theoretical calculation based on proton radial diffusion subject to Coulomb collisions and charge exchange losses. No wave-particle interactions were considered. Boundary conditions were imposed at L-6.6 with data from the ATS-6 spacecraft, and the computed spectra at lower L-shells are compared with available equatorial data from Explorer 45. The results of the theoretical calculations below a few tens of keV energies may be inaccurate since convection processes may dominate over diffusion at these low energies [Spjeldvik, 1977].

radiation belt proton spectra obtained with this method, and data are shown for comparison (taken from Explorer 45 below L = 5.3). Notice that the spectral maximum found at L<5 generally is displaced towards higher energies with lower L-shells. The flux values computed below a few tens of keV are below radiation belt energies (convective processes also operate in that range) and should not be considered reliable. The comparison with data is quite good, however, giving support to the usefulness of the radial diffusion theory above at least 100 keV.

A complementary view of the theoretical radiation belt proton fluxes is given in Figure 5-27 which depicts the computed radial distributions for different proton energies. Qualitatively similar to the observations, the radial flux maxima are displaced towards lower L-shells with higher proton energy. On L-shells well beyond the flux radial peak location, the characteristic time scales of diffusive transport are substantially shorter than that of the loss processes. This situation is called diffusive equilibrium. There is also an important connection with the magnitude of the radial diffusion coefficient. As D_{LL} increases, the diffusive equilibrium radial range extends towards lower L-shells, and the radial peak location is found at lower L. Also, the absolute magnitude of the flux radial peak increases as D_{LL} increases. An enhanced

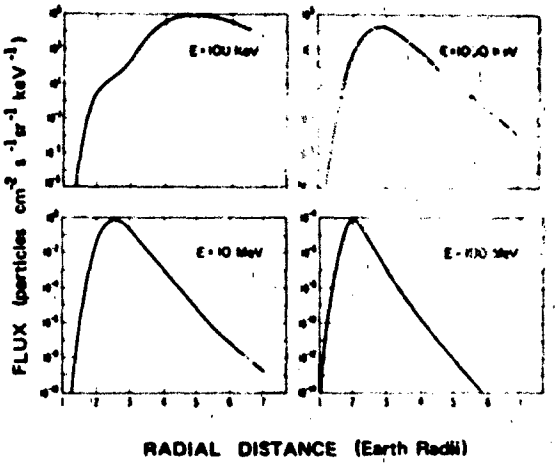


Figure 5-27. Theoretical proton radial distributions calculated from proton radial diffusion subject to Coulomb collision energy losses and charge exchange loss. Notice there is no division (or slot) within the trapping region. The curves now show that the radial flux maximum is displaced towards lower L-shells with higher proton energy. For details about the numerical calculations see Spjeldvik [1977].

D_{LL} causes a substantial increase in the trapped fluxes on L-shells below the peak flux location. Proton models that treat other than equatorially mirroring pitch angles in a comprehensive manner have not yet been developed. The reason is partially that $\bar{D}_{\alpha_o\alpha_o}$ is not know for protons, and partially that solving the simultaneous proton radial and pitch angle diffusion problem subject to the simultaneous μ-variations is mathematically difficult owing to the great inhomogeneity in the coefficients.

5.5.1.4 Heavy Ion Models. Besides electrons and protons, the earth's radiation belts contain appreciable fluxes of energetic helium, carbon, oxygen, and other ions. Theoretical modeling of these ion populations can be carried out by a generalization of the proton model. One must consider a distribution function for each charge state of a given radiation belt species, as well as the coupling between charge states imposed by the charge exchange chemistry. For *helium ions* one may write the coupled equations for equatorially mirroring ions:

$$\left(\frac{\partial f_1}{\partial t}\right) = L^2\partial/\partial L\left[D_{LL_1}L^{-2}(\partial f_1/\partial L)\right] + G_1(L)\,\mu^{1/2}(\partial f_1/\partial\mu) - \Lambda_{10}f_1 + \Lambda_{21}f_2 - \Lambda_{12}f_1 \quad (5.93)$$

and

$$\frac{\partial f_2}{\partial t} = L^2\partial/\partial L\left[D_{LL_2}L^{-2}(\partial f_2/\partial L)\right] + G_2(L)\,\mu^{1/2}(\partial f_2/\partial\mu) - \Lambda_{21}f_2 + \Lambda_{21}f_1 \quad (5.94)$$

where the subscripts indicate the ion charge state.

These equations have also been solved numerically [Cornwall, 1972; Spjeldvik and Fritz, 1978a], and some of the findings are illustrated in Figure 5-28. Notice that

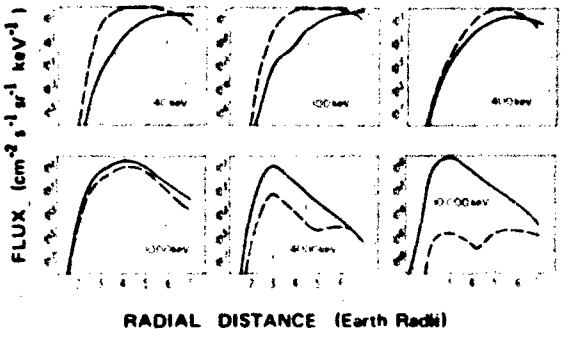


Figure 5-28 Theoretical helium ion radial distributions calculated from radial diffusion theory with Coulomb collisions and charge exchange. Solid curves show equatorially mirroring fluxes of He^{++} while the dashed curves depict He^{+}. Notice that He^{+} is dominant below ~ 1 MeV while He^{++} is dominant at higher energies [Spjeldvik and Fritz, 1978].

there is a transition from charge state 1 (He^{+}) to the state 2 (He^{2+}) at energies around 1 MeV, so that the lower charge state of helium ions is most abundant at the lower energies and the higher charge state at the higher energies. The spectral features computed for helium ions below ~ 1 MeV stem from the energy dependence of the radial diffusion coefficient D_{LL} and the loss rates. These have yet to be verified or refuted by observations.

For even heavier ions a multitude of charge states are available to the radiation belt heavy ions: for carbon ions there are six positive charge states, and for oxygen ions there are eight. In general, for a given ion charge state of an elemental species with s available states, the radial diffusion equation for equatorially mirroring ions may be written

$$\frac{\partial f_i}{\partial t} = L^2 \partial/\partial L \left[D_{LL_i} L^{-2} (\partial f_i/\partial L) \right] + G_i(L) \mu^{-1/2} (\partial f_i/\partial \mu) + \sum_{j=0}^{s} \Lambda_{ij} f_i + \sum_{j=1}^{s} \Lambda_{ji} f_i \quad \text{for } i \neq j \tag{5.95}$$

where Λ_{ij} denotes charge state transformation from state i to state j, so that $\Lambda_{ij} = 0$ for $i = j$ [Spjeldvik, 1979].

The radial diffusion equation has been solved for equatorially mirroring radiation belt oxygen ions [Spjeldvik and Fritz, 1978b], and one of the findings is that oxygen ions can be more numerous than protons at multi MeV energies. The oxygen ion charge state distribution is predicted to vary from a dominance of O^{+} ions at energies below ~ 100 keV to successively higher charge states at the higher energies; for example at 4 MeV the fourth charge state should be most abundant. A significant finding, illustrated in Figure 5-29, is that the radiation belt charge state distribution in the interior of the radiation belts ($L < 5$) is almost completely independent of the charge state distribution of the outer radiation zone source. In fact, this figure shows that regardless of the source being ionospheric (source charge state 1) or solar (source charge state 6 dominant) a steady state radiation belt distribution over ion charge states evolves that is independent of the source characteristics in the interior of the trapping region.

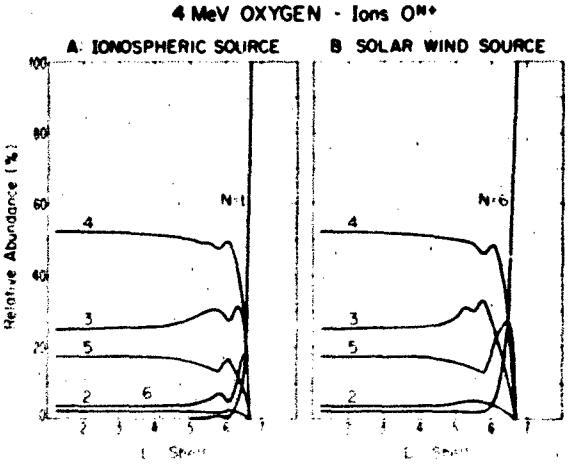


Figure 5-29 Relative charge state distribution for radiation belt oxygen ions. Left panel: an ionospheric source of O^{+} ions is assumed; Right panel: a solar (wind) source of O^{6+} ions is assumed. One finds the steady state radiation belt charge state distribution in the interior of the trapping region becomes largely independent of the source charge state characteristics.

The findings concerning charge states of energetic radiation belt ions may be summarized as follows:

1. Higher charge states are crucially important for the overall structure of the heavy ion component of the earth's radiation belts above ~ 100 keV.
2. Charge state redistribution processes are of major importance throughout much of the radiation belts and at all energies.
3. Radiation belt charge state distribution becomes largely independent of source charge state characteristics because of the frequent "internal" charge exchange.
4. Relative charge state distribution is to a large extent independent of the diffusive transport rate in much of the inner magnetosphere.

The last two conclusions are valid below a charge state redistribution zone adjoining the particle injection region in the outer radiation zone.

5.5.1.5 Theoretical Radiation Belt Ionic Composition. Because of experimental difficulties, only very limited information is available concerning the actual composition of radiation belt ion fluxes. For this reason we present a theoretical prediction of radiation belt protons, helium and oxygen ion fluxes at L = 3.25 (in the center of the radiation belts). This is shown in Figure 5-30. It must

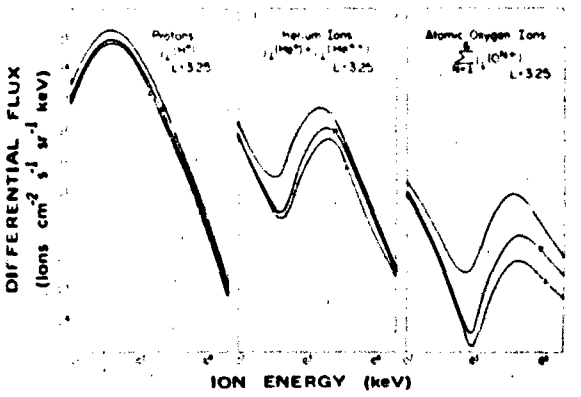


Figure 5-30. Theoretical model of the radiation belt ion fluxes. Comparison of theoretically predicted energy spectra of proton, helium and oxygen fluxes at L = 3.25 for different diffusion coefficients [Spjeldvik, 1979].

be kept in mind that these predictions have yet to be verified or refuted by experimental observation, particularly in the 100–1000 keV range. The prediction is that protons should be the dominant ion species from 0.1 to 1 MeV, helium ions should dominate at 1 to several MeV per ion, and at higher energies oxygen (and possibly also carbon) should dominate. These comparisons are done at *equal total ion energy*. If comparison is made at *equal energy per nucleon*, then protons would be the dominant constituent at almost all radiation belt energies beyond a few tens of keV.

5.5.2 Geomagnetic Storms

Most radiation belt modeling deals with equilibrium conditions appropriate for quiet time conditions. There is at the present time no fully comprehensive radiation belt model for the effects of geomagnetic storms and other disturbances. The reason is partially the incomplete knowledge of the radiation belt particle source mechanisms, and in part that disturbed time modeling is difficult because of the changes in magnetic field and time varying electric fields that must be also incorporated. Research is being conducted in this area, but no definite models are available as of this writing.

Specific types of disturbances and associated time variability have, however, been analyzed. Among these are studies of the post-geomagnetic storm decay of energetic particle fluxes following the (yet not fully explained) injection events. Lyons and Thorne [1973] have demonstrated that the radiation belt electron particle scattering lifetime can explain quantitatively the restoration of the radiation belt slot region after it becomes filled in during the storm injection process. Spjeldvik and Thorne [1975] subsequently demonstrated that the precipitating electron flux into the middle latitude ionosphere following such storms causes enhanced D-region ionization of sufficient magnitude to explain VLF radio wave phase anomalies observed at such times. Spjeldvik and Lyons [1980] have suggested a simplified prediction model for these effects.

There is reason to think that magnetic storms differ considerably from one another, not only by the magnitude of the ring current storm index D_{st} but also in the injection characteristics of energetic ion fluxes. Once injected, radiation belt particles become subject to the normal processes in the trapping region discussed earlier. For specific magnetic storms, the post-storm decay of protons and heavy ions has also been studied, and fair agreement between predictions and observations has been reported in limited energy ranges where the data were available [Spjeldvik and Fritz, 1981a & b].

5.5.2.1 The Variability of the Parameters. Within the framework of radiation belt diffusion theory, the injection of particles is described by a (time variable) source term $S(L, P, \alpha_0; t)$. The boundary conditions may also be time variable, reflecting dynamic conditions on the outer edge of the stable trapping region $f(L_{max}, P, \alpha_0; t)$, and the transport coefficients, D_{LL} and $\bar{D}_{\alpha_0\alpha_0}$, will also be time variable reflecting the geomagnetic activity level. Based on limited data, Lanzerotti et al. [1978] estimated that the radial diffusion coefficient might vary with the K_p geomagnetic index as

$$D_{LL}^{(M)} = 10^{\xi} L^{10} \qquad (5.96)$$

with

$$\xi = (9.6 \; 0.07 \sum_{1\ 2\ day} K_p). \qquad (5.97)$$

It is not yet known how $D_{LL}^{(E)}$ and $\bar{D}_{\alpha_0\alpha_0}$ might vary with geomagnetic conditions. It is known however that the exospheric neutral density varies with activity (because of the heating of the upper atmosphere), and the plasmaspheric densities certainly vary strongly. A first approach may be to solve the appropriate transport equations using perturbation theory for the different variables. Some geomagnetic conditions may, however, be too drastically altered to be treated as perturbations, so caution is in order. At the present time there are many unknowns, and specific models have yet to be developed.

5.5.2.2 Magnetic Topology Variations. During the early phase of magnetic storms the earth's magnetic field

becomes compressed on the dayside. It has been observed that the subsolar magnetopause can be pushed inward from an average location of $\sim 10 R_E$ to $\sim 5 R_E$ during large storms. This implies a large B-field change. Depending on the rapidness of the field change, the particle population may respond adiabatically or non-adiabatically. However, most magnetic storms do not cause such a large perturbation of the magnetic field. Increase in the lower energy (~10–500keV) radiation belt ion fluxes produces an enhanced *ring current* (see Section 5.7) that depresses the earth's magnetic field earthward of the enhanced particle flux region, and causes adiabatic deceleration of the trapped radiation belt particle fluxes. These effects must also be incorporated into storm-time radiation belt prediction models.

On the other hand, geoelectric fields penetrating into the trapping region during disturbed conditions [Harel et al., 1981a, b] can cause cross-L, non-diffusive transport and thus adiabatic acceleration of the particle population [Lyons and Williams, 1980]. The relative influence of these processes depends on particle energy.

5.6 EMPIRICAL RADIATION MODELS

Since the discovery of the earth's radiation belts, the population of trapped particles has been measured with ever improving instrumentation. The early Geiger counters flown on the first few spacecraft had little or no particle identification capabilities. Subsequent instruments used foil techniques, solid state detectors, magnetic and or electric deflection techniques and electronic signal discrimination. This led to a clear separation of electrons and ions (which were then assumed to be solely protons). During the later years of space exploration the presence of a multitude of different ion species was established. The purpose of this section is to provide a brief overview of existing radiation belt data, as a source guide for those who require crude numerical estimates of the radiation environment. It is emphasized that the older empirical "proton" model in reality represents contributions from many ion species, and that sometimes the heavy ion contribution may be dominant.

5.6.1 Data Acquisition and Processing

Empirical radiation belt models are compiled by NASA National Space Science Data Center, Goddard Space Flight Center in Maryland [Vette et al., 1979 and references therein]. These models represent a systematic effort to compile many years of data containing a large number of disparate satellite observations into a few key models. These observations were separated in space and time, and made with highly varying instrumentation so that subjective judgments were necessary regarding data quality. The complexity of the task is appreciated if one considers the volume of space to be covered and the time variation in the particle fluxes. Most of the data used in the NASA models were obtained in the 1960's and early 1970's. For example, the present AP-8 model for protons is a combination of 94 different instrument energy-detector channels from instruments flown on 24 satellites [Sawyer and Vette, 1976]. The combined data sets were smoothed to obtain omnidirectional flux distributions in the B, L parameter space. The fluxes were integrated over all pitch angles and, therefore, directional information was not utilized. Other limitations in the data sets are variations in instrument response and the lack of local time dependence considerations. Also note that the flux models are compiled from a data base obtained over a brief time period in comparison with other geophysical and solar time scales. Long term dynamical changes in the radiation belts are not well understood so that extrapolation to different epochs must be done cautiously. For example, much higher fluxes than the models give have recently been observed. One example of very long term effects is the adiabatic energization of inner belt protons by the secular variation of the earth's magnetic field [Schulz and Paulikas, 1972].

5.6.1.1 Protons (Ions). It has become customary to refer to energetic protons located below $L \sim 2$ as inner belt or inner zone protons. This is a misnomer, however, since the proton fluxes normally have only one radial flux peak. Sawyer and Vette [1976] have completed an extensive model of the trapped "proton" environment out to $L = 6.6$ for energies above 100 keV. The measured ions were labeled "protons", although no actual ion identification was made. The model is composed of two parts, AP8MIN and AP8MAX, which correspond to observations made during the 1960-1970 sunspot minimum and maximum conditions, respectively. The difference between these two models is thought to result in part from differences in upper atmosphere heating during solar active periods such that the trapped particle collision rate (due to the exospheric expansion) was increased. AP8MAX differs from AP8MIN for altitudes less than about 1000 km and for L values less than 2.9. Steep spatial gradients in the ion fluxes at lower altitudes are difficult to determine accurately.

An equatorial profile of the ion (proton) fluxes as given by AP8MIN is shown in Figure 5-31. Note the presence of >400 MeV protons for $L < 2$. This is consistent with higher energy protons being produced by the decay of neutrons produced in the atmosphere by cosmic rays (the CRAND source). The lower energy fluxes can arise from inward radial diffusion as discussed in the theoretical modeling section.

The AP8 proton models include no data after 1970 and very little data above 100 MeV energies [Vette et al., 1978]. Although the "proton" belt is considered far more static than the electron belts, significant enhancement

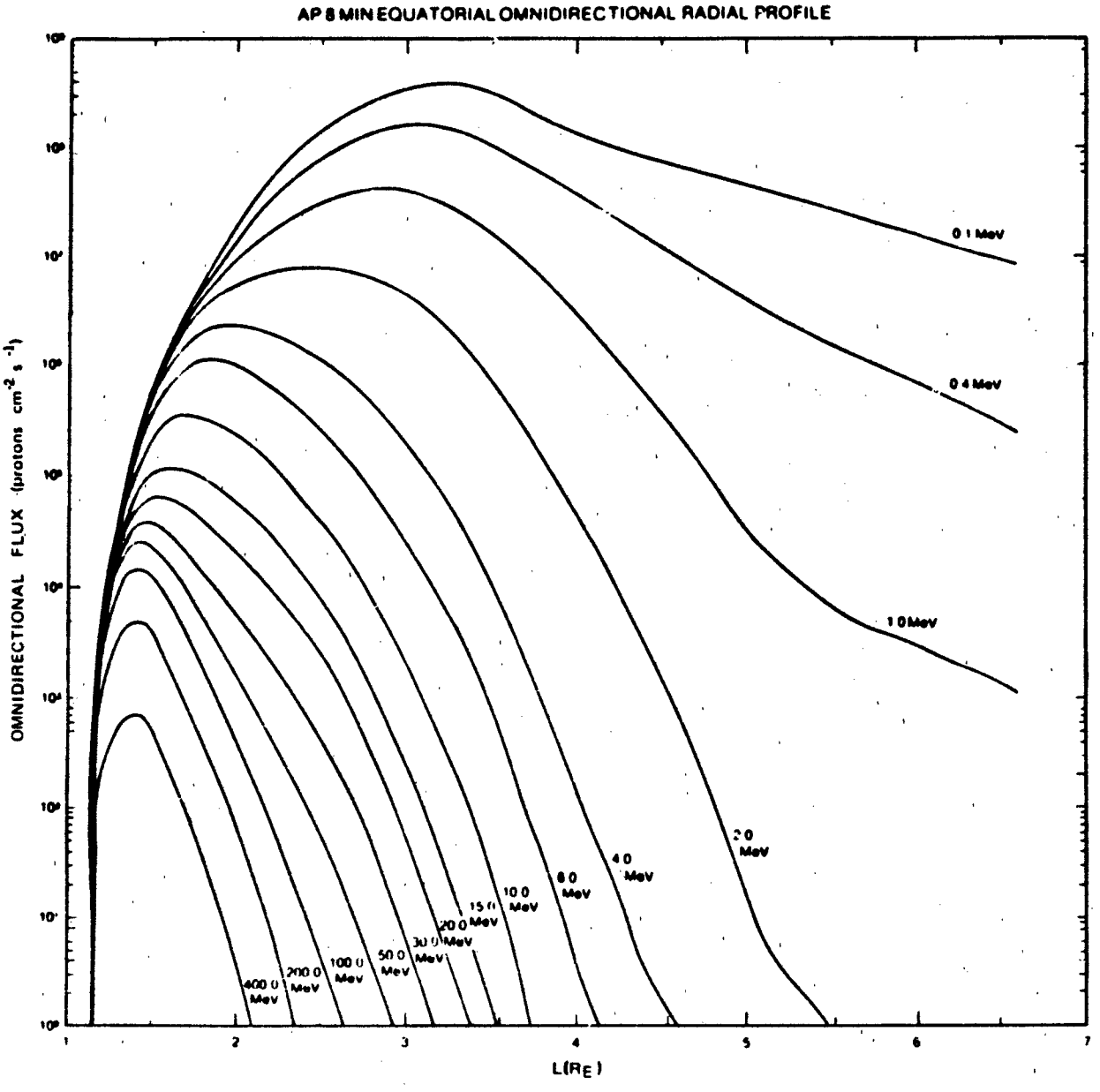
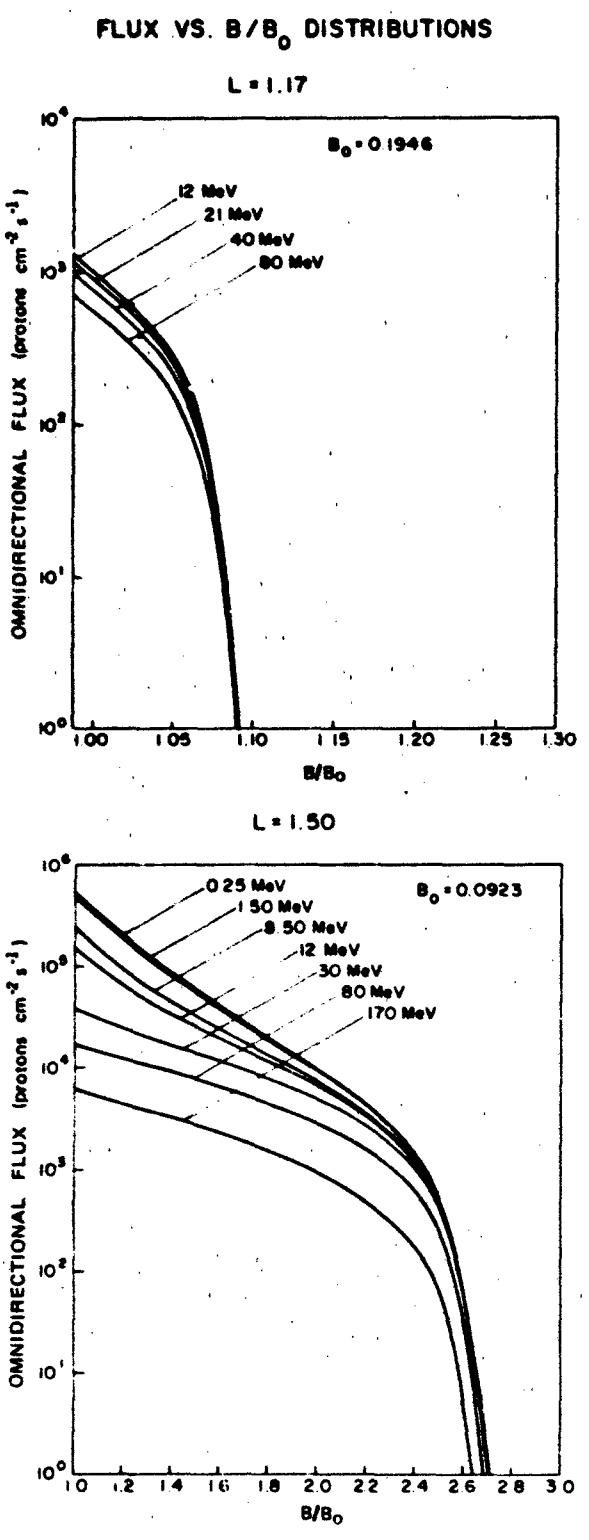


Figure 5-31. Radial distribution of AP8MIN omnidirectional fluxes of protons in the equatorial plane with energies above threshold values between 0.1 and 400.0 MeV [Sawyer and Vette, 1976].

and depletions have been observed. A secondary equatorial proton peak (40-110 MeV) of unknown origin was observed by McIlwain [1965]. This peak moved from L = 2.25 to L = 2.1 earth radii between January 1963 and January 1965. During the large May 1967 magnetic storm 2.2 - 8.2 MeV proton fluxes were observed to increase more than a factor of 10 at L = 2.2 [Bostrom et al., 1971]. Lower energy (0.26-0.65 MeV) protons were observed to be preferentially enhanced during the same storm at L = 3.0 [Rothwell and Katz, 1973]. Beyond L = 5 earth radii, order-of-magnitude fluctuations occur on time scales as short as 10 minutes [Sawyer and Vette, 1976]. Very large increases in MeV *heavy* ion fluxes at L = 2 5 were observed during the August 1972 magnetic storm event [Spjeldvik and Fritz, 1981a,b,c]. For this reason engineering applications should consider large deviations from the mean models.

Off-equatorially mirroring ions intercept the upper

atmosphere at higher equatorial pitch angles on lower L-shells. For example, Figure 5-32 shows the B/B_o distribution for three different L-values just above the



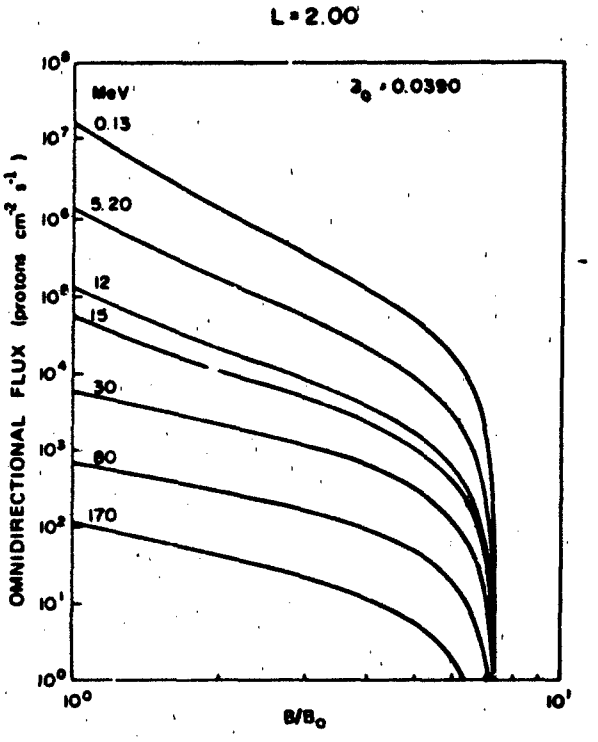


Figure 5-32. Empirical inner zone radiation belt proton (ion) flux vs B/B_o for three representative L-shells, L=1.17, 1.50, and 2.0, and several representative energies as given by the AP8MIN proton model [Sawyer and Vette, 1976].

atmosphere. At L = 1.17 the equatorial pitch angle distributions have very steep loss cone gradients such that the particle flux vanishes at $\alpha_{oc} \leq 73°$. At L = 1.50 this "cut off" is $\alpha_{oc} \sim 37°$ and for L = 2.00, $\alpha_{oc} \sim 21°$. These cutoffs come about because the atmospheric bounce loss cone is wide at the lower L-shells as seen in equatorial pitch angle.

The South Atlantic Magnetic Anomaly has a controlling effect on the inner zone particle fluxes in the vicinity of the loss cone. This anomaly arises from the earth's magnetic field being less intense at a latitude longitude region located near the coast of Brazil. Trapped particles at these low L-shells will encounter their lowest mirroring altitude (H_{min}) and thus the densest atmosphere in this longitude region. Figure 5-33 shows omnidirectional proton flux contours in protons cm^{-2} sec^{-1} MeV^{-1} at 750 km altitude. These contours were derived from 5 to 7 MeV (dashed lines) and 28 to 45 MeV (solid lines) data channels on the AFGL particle identifier instrument flown on AF Satellite 72-1 in 1972. The data show that the location of the proton peak flux in the South Atlantic is dependent on energy. The 5 to 7 MeV peak is located around L = 2 while the 28 to 45 MeV peak is located around L = 1.3.

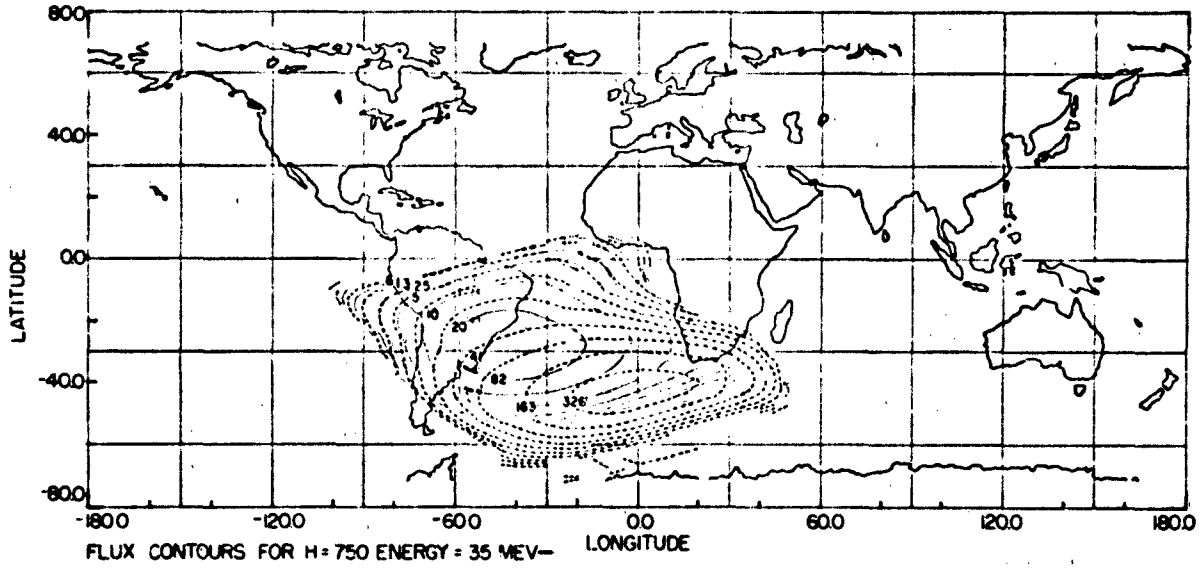


Figure 5-33. Proton isointensity flux contours as measured in the South Atlantic anomaly at an altitude of 750 km. The solid lines depict 28-45 MeV proton (ion) fluxes and the dashed lines 5-7 MeV proton fluxes. The flux units are particles/(cm^2-s-MeV).

The azimuthal drift around the earth of radiation belt particles through the South Atlantic anomaly region produces a "windshield wiper" effect. Electrons present in the enhanced loss cone created by the locally low B-field region are lost. These electrons are subsequently replenished by pitch angle diffusion during their drift around the earth outside the anomaly.

Solar cycle variations and high altitude nuclear detonations that modify the atmospheric composition significantly affect the trapped particle populations. A 15-year time study of inner belt 55 MeV protons concluded that the observed flux variations were consistent with expected atmospheric loss processes [Parsignault et al., 1981].

5.6.1.2 Heavy Ions. The abundance of trapped heavy ions give clues to the origin of the radiation belt particles. At high energies they also constitute a hazard to operational space systems. Over the last decade data have been acquired to sketch the spectral and angular dependence of the helium and carbon-nitrogen-oxygen (CNO) fluxes.

Helium Ions: Helium ions are sometimes referred to as alpha particles, although in a strict sense an alpha particle is only the totally ionized state (He^{2+}) of a helium ion. Figure 5-34 shows average equatorial helium ion spectra over a range of L-shells during quiet-time conditions [Fritz and Spjeldvik, 1979]. Notice that these spectra are very steep at $L \sim 4$, become much harder with lower L-shells, and are almost flat at $L \sim 3$. This is expected since the lower energy helium ions are preferentially lost through the ion exchange mechanism and Coulomb collision energy degradation, as the particles diffuse radially inward. The observed ratio between helium ion and proton fluxes simultaneously observed are given in Figure 5-35. Care must be exercised whether the ratio is

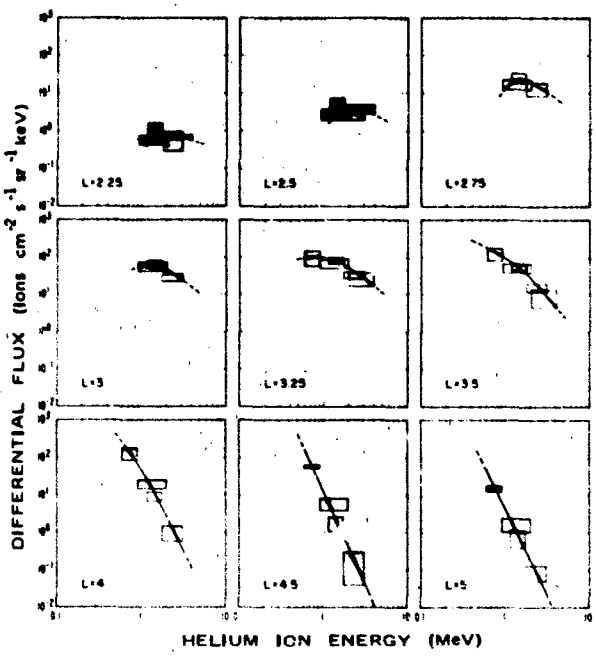


Figure 5-34. Energy spectra of equatorial radiation belt helium ions deduced from mass ion observations on Explorer 45 during the geomagnetically quiet period June 1-15, 1972. The data are given at L = 2.25, 2.5, 2.75, 3, 3.25, 3.5, 4, 4.5, and 5 [Fritz and Spjeldvik, 1979].

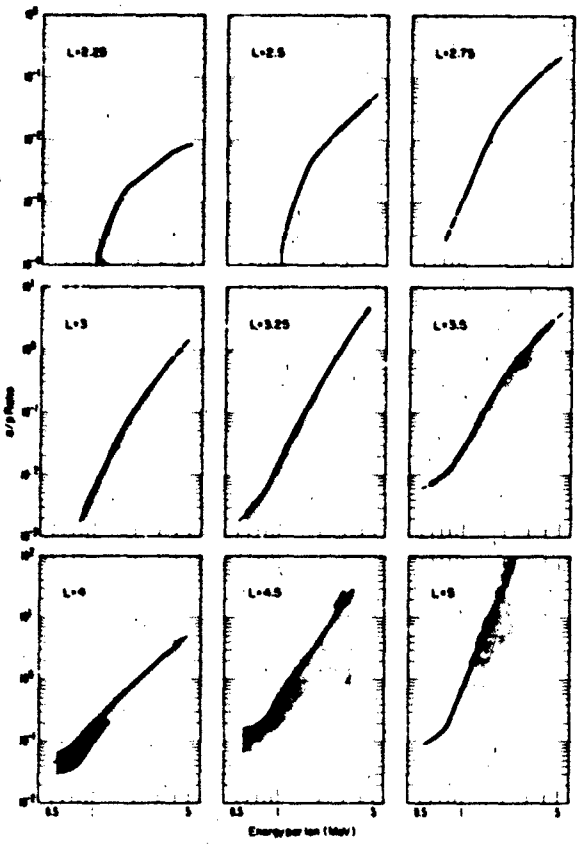


Figure 5-35. Quiet time He/p ion flux ratios in the equatorial radiation belts of the earth deduced from Explorer 45 observations during June 1-15, 1972. The ratios are calculated from flux observations at equal energy per ion. The shaded areas depict the experimental uncertainty in the data [Fritz and Spjeldvik, 1979].

defined at the same total ion energy or (as in this figure) at the same energy per nucleon. At higher total ion energies the heavy ion flux may be dominant over the proton flux at some L-shells (Figure 5-35) while at the same energy per nucleon proton fluxes usually dominate.

The equatorial helium ion pitch angle distribution is generally more anisotropic than the comparable proton pitch angle distribution. For L-values between 2.5 and 4.5 the proton fluxes most often vary as $\sin^n\alpha_0$ where $n = 4 \pm 2$ and α_0 is the equatorial pitch angle while for helium ions $n \approx 10 \pm 4$. Beyond $L \sim 5$ the helium fluxes are quite dynamic and characteristic quiet-time values are difficult to define. During magnetic storms the fluxes of energetic helium ions can increase by orders of magnitude in the heart of the radiation belts ($L \sim 2$-5), and the relative abundance of the different ion species can vary. At higher L-shells substorm effects can be significant. For example, during the 18 June 1974 substorm, helium ions were more numerous than protons at geosynchronous altitude for $E \geq 800$ keV ion [Fritz and Wilken, 1976]. Blake and Fennell, [1981] also have noted that a strictly sinusoidal pitch-angle distribution did not completely describe the 98-240 keV nucleon ion fluxes at the geosynchronous altitude ($L \sim 6.6$).

Carbon-Nitrogen-Oxygen (CNO): Trapped oxygen ions can at times be much more numerous than protons particularly at $L \geq 5$, when compared at equal total ion energy. Figure 5-36 shows the radial flux profiles of carbon and oxygen ions as measured by the ISEE 1 spacecraft [Hovestadt et al., 1978]. At equal total ion energies in the MeV range the carbon to oxygen flux ratio is of the order of 0.5, and at equal energy per nucleon the carbon flux can dominate. This indicates an extraterrestrial source for the very energetic trapped heavy ions since the ionospheric C/O-ratio is $< 10^{-5}$ [Blake, 1973]. The CNO flux pitch angle anisotropy is even more pronounced than that of helium, having a value of the anisotropy n-index typically between 12 and 16. At the higher total energies (>800 keV) and at geosynchronous altitudes the CNO flux has been measured to be higher than that of protons and helium.

Very Heavy Ions: Ions heavier than oxygen are also present in the earth's radiation belts in small quantities. It has been demonstrated that substantial injection of ions with nuclear charge $Z \geq 9$ can take place during some magnetic storms, and many orders of magnitude flux intensity enhancements at $\epsilon \geq 10$ MeV have been observed lasting for many months [Spjeldvik and Fritz, 1981c]. During such disturbed conditions the trapped fluxes of other ions (He, O, etc.) can also be greatly enhanced [Spjeldvik and Fritz, 1981a,b].

5.6.1.3 Trapped Electrons. Empirical flux models have been developed that describe the inner and outer electron radiation belts. The National Space Science Data Center at NASA Goddard Space Flight Center has developed two inner belt models: AE-6 [Teague et al., 1976] for sunspot maximum, and AE-5 epoch 1975 [Teague and Vette, 1974] for sunspot minimum conditions. There is also an outer belt electron model called AE-4. An updated outer belt model, AE-7, is now under development. In this section a brief description of the models is given and typical electron flux versus L-shell profiles, energy spectra, and pitch angle distributions are quoted.

Inner Belt Electrons: The empirical data AE-5 model is based on flux data from five satellites, OGO 1, OGO 3, 1963-38C, OV3-3 and Explorer 26 [Teague and Vette, 1972]. This data base covered the period from December 1964 to December 1967, representing a transition from solar (sunspot) minimum towards maximum conditions. During this period the time-averaged Zurich sunspot number R_z ranged in value from about 10 (December 1964) to around 100 (December 1967). In constructing the AE-5 model it was assumed that the total inner belt electron flux is composed of four components: 1) quiet day flux at solar minimum, 2) quiet day flux at other

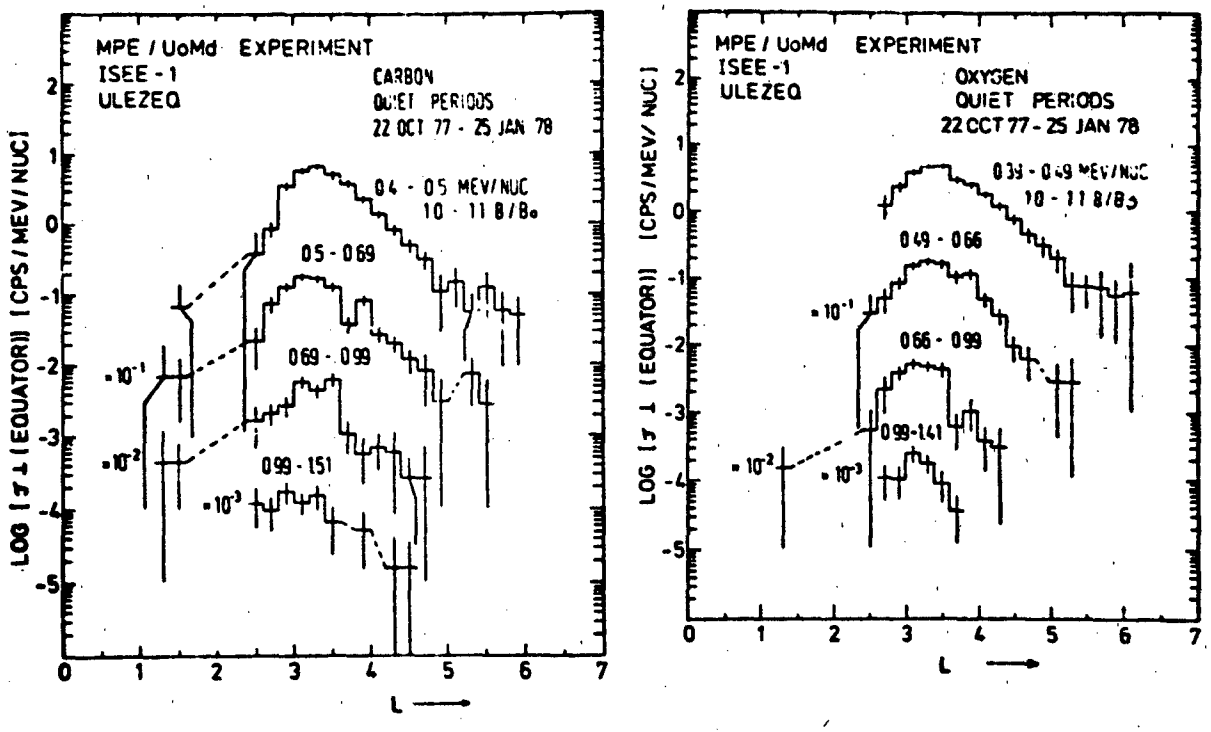


Figure 5-36. Equatorially trapped carbon and oxygen ion fluxes measured from October 1977-January 1978 during quiet times. Count rates may be approximately converted to omnidirectional flux units: ions (cm^2-s keV) by multiplying by 50 ΔE where ΔE is the appropriate energy passband in keV [Hovestadt et al., 1981].

times during the solar cycle, 3) storm time flux and 4) residual flux from the 1962 high altitude Starfish nuclear explosion. By late 1967 the Starfish generated energetic electron flux at $\epsilon < 1$ MeV had decayed to insignificant levels. It should be noted that for energies $\epsilon > 690$ keV the available data for the inner belt AE-5 model were quite limited. For example, within the observing range of the instruments the monthly averaged OGO 1 and OGO 3 data showed non-zero counts in this energy range only when the Starfish fluxes were still present or during storm times.

The quiet day solar cycle variation was defined by taking the ratio of the omnidirectional flux measured from solar minimum to a standard reference epoch (chosen as October 1967). Insufficient data necessitated the assumption of B-field (and therefore particle pitch angle distribution) independence for the geospace solar cycle variation. Also, the presence of Starfish residual energetic electrons restricted the analysis to later times when these man made radiation belt electron fluxes had decayed away. It should be emphasized that the solar cycle variation has been determined for only one specific cycle and that it may not apply to others.

The effect of magnetic storms on the time averaged inner radiation belt electron flux depends on 1) the frequency at which magnetic storms occur, 2) the magnitude of the storm time electron flux enhancement over quiet time values, and 3) the duration and characteristics of the magnetic storms. The inner-belt electron flux increases are infrequent but substantial and long-lasting, so that a model storm effect was not extractable from this NASA data base. However, the average long term impact of magnetic storms was estimated by forming the ratio R_s which is the average electron flux (June 1966 - December 1967) divided by the quiet-time electron flux (October 1967). The results are shown in Figure 5-37 and reflect a pronounced flux peaking at 1 MeV. Of course, the largest relative storm time energetic electron flux enhancements are found in the electron "slot-region" located at $L \sim 2$ - 4, depending on energy.

An inner belt electron model for solar minimum was derived from the model discussed above [Teague and Vette, 1974]. This is called the *AE-5 Epoch 1975 Projected Model* and was constructed by removing the estimated temporal variations. For example, the Starfish residual energetic electron flux component ($L < 1.6\ R_E$, $\epsilon > 700$ keV) was removed, using the residual Starfish electron model of Teague and Stassinopolous [1972].

Similarly a *solar maximum model, AE-6*, was constructed using the AE-5 model at solar maximum values (epoch 1967) and with the estimated Starfish residual (background) energetic electron fluxes also subtracted out. This model is called *AE-6 Epoch 1980*.

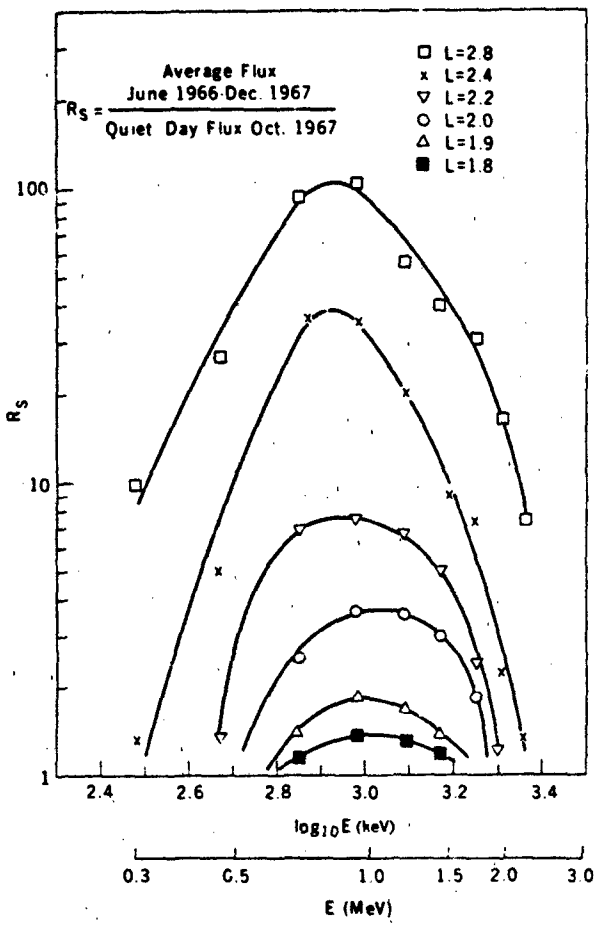


Figure 5-37. Estimate of the relative average effect of substorms on inner belt electron fluxes (Teague and Vette, 1972). It is cautioned that different disturbances can have different effects.

The electron models were verified by comparing them with additional data sets from the spacecraft OV3-3, OV1-13, OV1-19, OGO 5, OGO 1, OGO 3, 1963-038C, and OGO 4 [Teague et al., 1979]. These additional data were also obtained before 1970 and substantiate the model mean values for that particular epoch. Figure 5-38 shows a comparison between these data sets and the AE-6 and AE-5 1967 models. This is the quiet time energy spectra of equatorially mirroring electrons as measured at L = 1.5. The OGO 1 day 300 1964 results are evidently seriously contaminated by the Starfish detonation residual energetic electrons. Otherwise, these results show fair agreement between the AE-6 electron model and the different data. Figure 5-39 shows comparison of the model and measured equatorial pitch angle distributions at L = 1.4. The AE-5 1967 (solar maximum) model and the AE-5 1975 (solar minimum) model bracket the data within a factor ± 2 to 3, except at very low equatorial pitch angles.

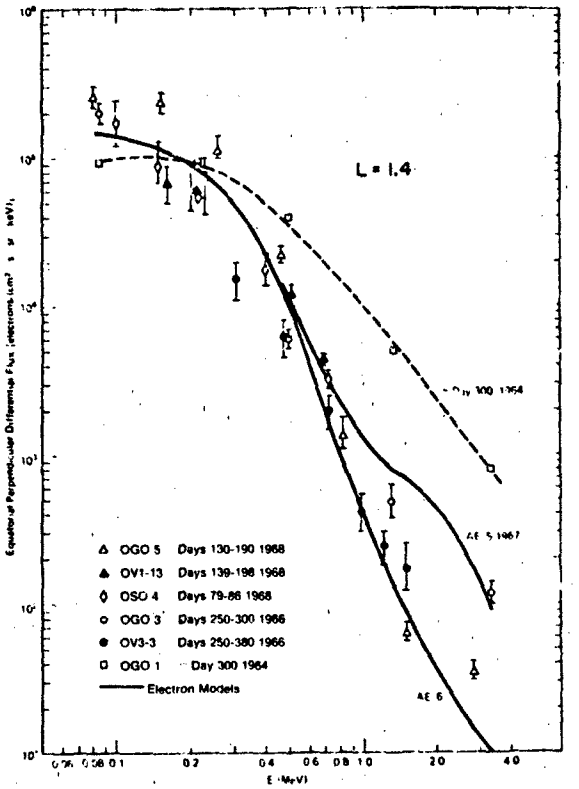


Figure 5-38. Empirical radiation belt electron fluxes at L = 1.4: Energy spectra comparison of the AE-5 1967 and AE-6 inner belt electron models with experimental data taken from the listed satellites for the times shown [Teague et al., 1979].

The quiet day inner belt electron fluxes for $E < 690$ keV at $1.3 < L < 2.4$ can be represented by an analytic formula based on the empirical data [Teague and Vette, 1972]. The energetic electron flux is parametrized as follows:

$$j(\alpha_0, L, \epsilon) = A(\alpha_0, L)\, \epsilon \exp\left[-\epsilon/\epsilon_0(\alpha_0, L)\right] \qquad (5.98)$$

where L is the McIlwain L-shell parameter, α_0 is the equatorial pitch angle and the parameters, $A(\alpha_0, L)$, $\epsilon_0(\alpha_0, L)$, are related to the equatorial values ($\alpha_0 = 90°$) by the following empirical expressions:

$$A(\alpha_0, L) = A(90°, L)\, \frac{\sin^m(\alpha_0 - \alpha_{OLC})}{\sin^m(\phi - \alpha_{OLC})} \qquad \phi > \alpha_0 \geq \alpha_c$$

$$= A(90°, L) \qquad 90° \geq \alpha_0 \geq \phi \qquad (5.99)$$

$$\epsilon_0(\alpha_0, L) = \epsilon_0(90°, L) \sin^n \alpha_0 / \sin^n \phi \qquad \phi > \alpha_0 \geq \alpha_c$$

$$= \epsilon_0(90°, L) \qquad 90° \geq \alpha_0 \geq \phi. \qquad (5.100)$$

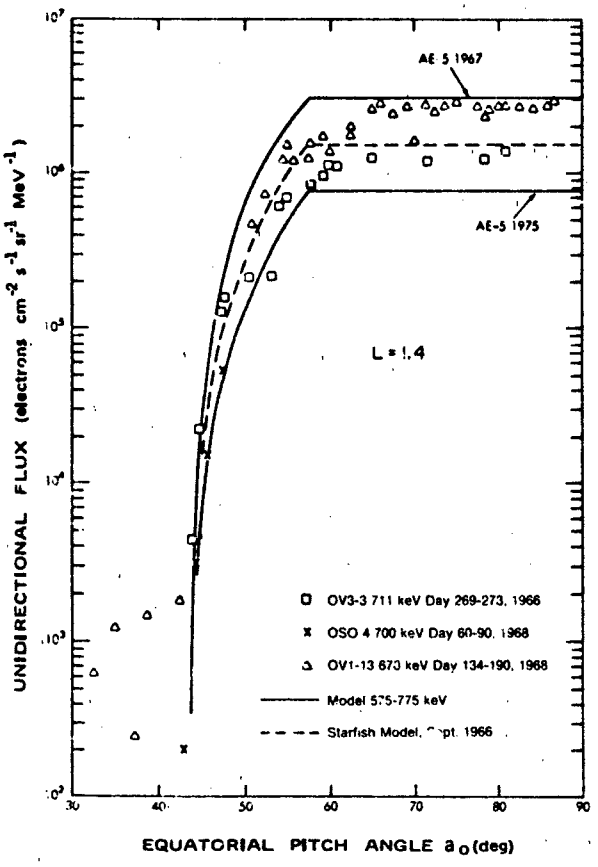


Figure 5-39. Empirical radiation belt electron fluxes: a pitch angle distribution comparison of the inner belt electron models with satellite data taken at L=1.4 [Teague and Vette, 1974].

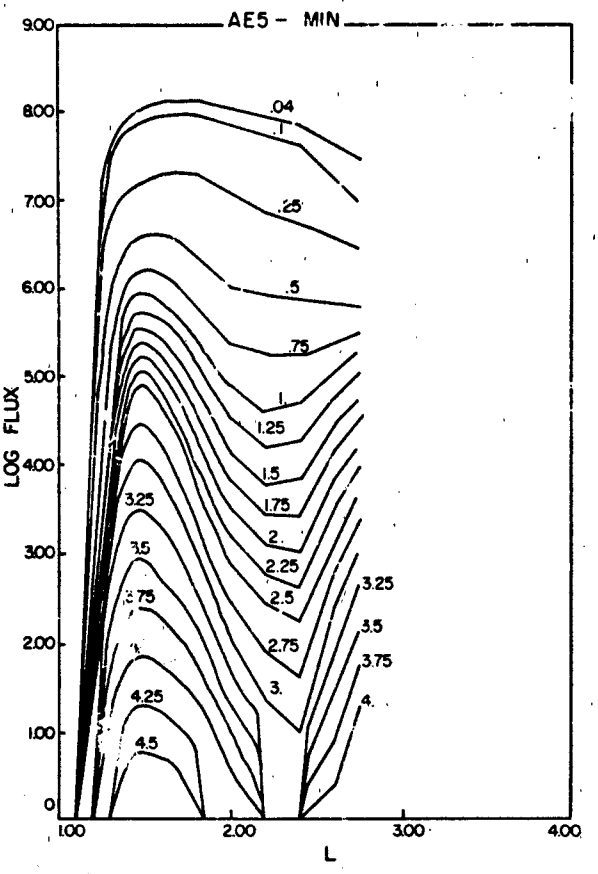


Figure 5-40. Empirical radiation belt electron observations: Equatorial flux versus L-shell for the AE5-MIN model as taken from the National Space Science Data Center (NSSDC) computer models. The flux is in units of electrons/cm²-s), and the energy range is 0.04-4 MeV.

Values for ϕ, α_c, n, m, ϵ_0 ($\alpha_0 = 90°$, L) and A($\alpha_0 = 90°$, L) for a given L-shell are given in Table 5-2.

In Equations (5.99) and (5.100) the equatorial electron pitch angle distributions are fairly flat out to an angle, ϕ, where they rapidly drop as $\sin^n\alpha$. The parameter α_{OLC} is the minimum allowed equatorial pitch angle (the loss cone angle) and corresponds to a 100 km electron mirroring altitude. These formulas should be used with caution, however, since they represent extrapolations based on an imperfect and incomplete data set. The dipole approximation gives α_{OLC} in terms of B_c from Table 5-3

$$\alpha_{OLC} = \arcsin\left[\left(\frac{B_E}{B_C L^3}\right)^{1/2}\right]. \qquad (5.101)$$

where B_C is the magnetic induction at the 100 km altitude level. Figures 5-40 and 5-41 show the equatorial electron flux profiles as given by AE-5 (1975 projected) and AE-6.

Outer Belt Electrons: The outer belt trapped electron fluxes are located between ~3.5 and ~11 R_E. Rapid changes in the magnetic field and background plasma can modify the electron flux levels within minutes. Because of the apparent coupling between magnetospheric substorm process and the outer belt trapped electron fluxes, time-averaged models have been developed. There are discernible changes in the average flux over the time period studied so that the model is given for two epochs. The NASA models are called AE-4 epoch 1964 (solar minimum) and AE-4 epoch 1967 (near solar maximum) [Singley and Vette, 1972]. The data base was acquired between 1959 and 1968 from 23 instruments on 11 satellites.

Because of the lack of azimuthal symmetry of the geomagnetic field in the outer radiation zone, studies of the radiation belt electron structure beyond L ~ 5 requires the conventional B-L coordinate system (calculated from the earth's internal magnetic field) to be augmented with the additional coordinate local time, LT.

Table 5-2 Flux model parameters for quiet-day inner belt electrons. See text for details [Teague and Vette, 1972].

Quiet-Day Model Parameters (Inner Belt Electrons)							
Reference Pitch Angle = 90 Degrees							
EPOCH = 10 67							
L (R_E)	$J(\alpha_o = 90°)$ (cm^{-2} sec^{-1} sr^{-1})	$A(\alpha_o = 90°)$ (cm^{-2} sec^{-1} sr^{-1} keV^{-1})	$\epsilon_o(\alpha = 90°)$ (keV)	M	N	PHI (deg)	B_c (G)
1.30	1.20E07	1.71E03	83.7	2.80	0.670	67.1	0232
1.35	1.70E07	2.39E03	84.3	2.20	0.660	61.4	0.234
1.40	2.23E07	3.08E03	85.1	1.70	0.650	57.5	0.238
1.45	2.81E07	3.81E03	85.7	1.20	0.640	59.0	0.241
1.50	3.41E07	4.56E03	86.5	0.93	0.630	65.0	0.245
1.55	3.99E07	5.21E03	87.5	0.92	0.620	66.0	0.249
1.60	4.49E07	5.74E03	88.4	0.91	0.610	67.0	0.253
1.65	4.79E07	6.08E03	88.8	0.90	0.600	66.5	0.257
1.70	5.10E07	6.42E03	89.1	0.89	0.590	66.0	0.262
1.75	5.46E07	6.81E03	89.5	0.88	0.580	68.0	0.265
1.80	5.77E07	7.16E03	89.8	0.87	0.570	70.0	0.268
1.85	6.00E07	7.57E03	89.0	0.86	0.545	76.0	0.271
1.90	6.11E07	7.93E03	87.8	0.85	0.520	86.0	0.274
1.95	5.84E07	7.80E03	86.5	0.83	0.500	90.0	0.277
2.00	5.38E07	7.50E03	84.7	0.80	0.480	90.0	0.280
2.10	4.69E07	7.15E03	81.0	0.79	0.470	90.0	0.286
2.20	4.15E07	7.00E03	77.0	0.78	0.460	90.0	0.292
2.30	3.61E07	6.50E03	74.5	0.77	0.450	90.0	0.298
2.40	3.11E07	6.00E03	72.0	0.76	0.440	90.0	0.304

The empirical LT dependence of the outer radiation belt electron fluxes has been determined to be

$$\log J(\alpha_o) \sim C(E, L) \cos\left[\pi \frac{LT - 11}{12}\right] \quad (5.102)$$

with LT in hours and $C(E, L \leq 5) = 0$ [Vette, et al., 1976]. This is only valid in a time average sense, and it was found that at a given local time the electron flux intensity levels varied by at least factors of 10 to 50 over the data acquisition period stated.

Given the equatorial flux ($\alpha_o = \pi/2$ or $B = B_o$) the off equatorial outer belt integral electron flux ($\alpha_o \neq \pi/2$) can be estimated by [Singley and Vette, 1972]

$$J[> E, B, L] = J[> E, B = B_o, L]\, G[B, L] \quad (5.103)$$

and

$$G[B, L] = (B - B_o)^m \left(\frac{B_c - B}{B_c - B_o}\right)^{m+1/2} \quad ; B < B_c \quad (5.104)$$

$$G[B, L] = 0 \quad ; B \geq B_c .$$

The parameters m, B_c and B_o are all empirical functions of L and are given in Table 5-3. The parameter B is the magnetic field value at the desired location off the geomagnetic equator, and B_c is the value of the magnetic field at 100 km altitude on the same field line.

Figures 5-42 and 5-43 show the AE-4 equatorial omnidirectional model electron fluxes from .04 to 4.50 MeV. Epoch 1964 represents solar minimum and epoch 1967 solar maximum. Using the above expressions with Table 5-3 and Figures 5-42 and 5-43, flux estimates can be made at non-equatorial latitudes.

Measurements from the OV1-19 satellite have indicated that the AE-4 model fluxes may be significantly too low, particularly at higher energies beyond 1 MeV [Vampola, 1977]. These newer data were averaged over periods which included two magnetic storms in 1969. The OV1-19 instrumentation measured radiation belt electrons in the 53 keV to 5.1 MeV energy range in 24 differential energy bands, which significantly improved the high energy data coverage over that which was available to construct the AE-4 models. The OV1-19 data has been incorporated into a new NASA model called AE-7 HI. Figure 5-44 shows the equatorial electron flux versus L-shell profile as predicted by the AE-7 HI model. Note,

Table 5-3. Flux model parameters for outer belt electrons. See text for details [Singley and Vette, 1972].

AE-4 B/B Model Parameters (Outer Belt Electrons)[a]			
L (R_E)	m	B_o^* (G)	B_c (G)
3.00	1.12	0.01154	0.580
3.10	0.87	0.01046	0.582
3.20	0.71	0.009511	0.585
3.40	0.66	0.007929	0.588
3.60	0.63	0.006680	0.593
4.00	0.60	0.004870	0.596
4.50	0.60	0.003420	0.599
5.00	0.60	0.002493	0.600
5.50	0.60	0.001873	0.601
6.00	0.60	0.001443	0.601
6.50	0.60	0.001134	0.602
7.00	0.60	0.000909	0.602
7.50	0.60	0.000739	0.603
8.00	0.60	0.000609	0.603
8.50	0.60	0.000507	0.6035
9.00	0.57	0.000428	0.6035
9.50	0.52	0.000363	0.604
10.00	0.44	0.000312	0.604
10.50	0.35	0.000269	0.604
11.00	0.24	0.000234	0.604

[a] $B_o^* = \frac{0.311654}{I^3}$

however, that no direct equatorial measurements have been included in the empirical model for electrons ($E > 300$ keV) above 8000 km and below $L = 5$. Much of the data base was acquired from satellites orbiting at a significant inclination to the magnetic equator, making equatorial flux representations based on these data uncertain.

Figure 5-45 shows the differences between the AE-4 and the AE-7 models at energies above 1 MeV. The AE-7 LO model is based on data taken on the AZUR satellite and is shown in Figure 5-46. Note that $E \leq 1$ MeV electrons are most likely to penetrate spacecraft shielding and contribute to the accumulated radiation dosage and damage. It is partially for this reason that the empirical model uncertainties are of interest to spacecraft designers.

Magnetic storms may cause large energetic electron flux enhancements that last for several weeks in the radiation belts. Figure 5-47 shows an example of an electron flux enhancement as observed on OGO-5 [West et al. 1981] during 1968. This example shows that the electron flux at 1.53 MeV increased by more than four orders of magnitude during one particular magnetic storm. Figure 5-48 shows the mean exponential decay time of these electron flux enhancements in days for a wide

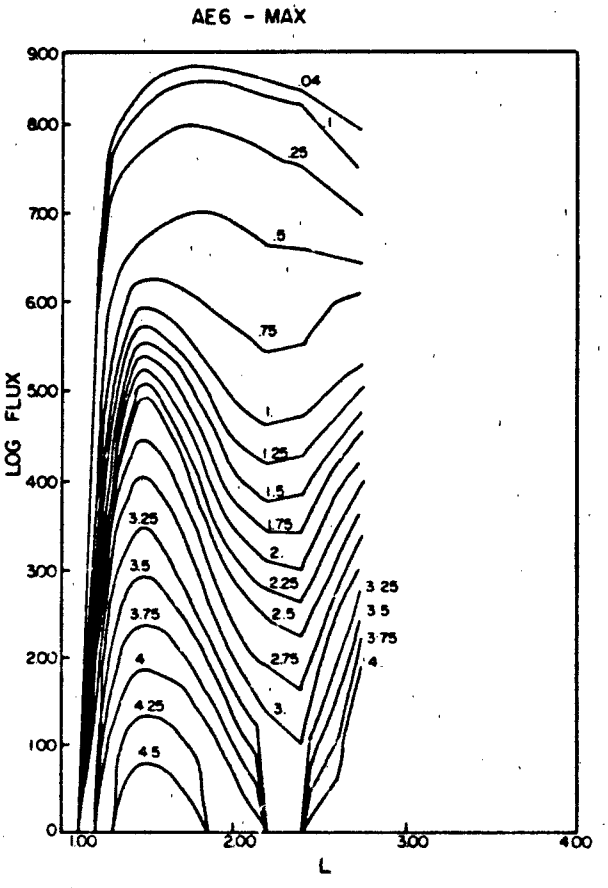


Figure 5-41 Empirical radiation belt electron observations: Equatorial flux versus L profiles for the AE-6 model as taken from the (NSSDC) computer models. The flux is in units of electrons (cm^2-s), and the energy range is 0.04-4 MeV.

range of L-shells and energies. Also plotted are the predicted electron lifetimes [Lyons et al., 1972] (see Section 5.5.1.2 on theoretical electron models) which in this comparison appear to agree with the experimental data.

5.6.1.4 Shell Splitting Effects. The outer belt particle pitch angle distributions are particularly interesting due to a phenomenon called shell-splitting. Here we shall qualitatively describe the physical process; for analytic considerations, see Roederer [1970]. Shell-splitting arises from the lack of local time (azimuthal) symmetry of the earth's magnetic field, particularly at greater distances from the earth. It is usually considered important for $L > 4$, but it should also exist at lower L-shells where the earth's magnetic multipoles become significant. Shell-splitting arises because particles drifting in longitude preserving the first two adiabatic invariants modify their pitch angle and radial location according to the asymmetric magnetic field. While the concept of dipole L-shell is useful to describe principal features of trapped particles, actual non-dipolar geometry with azimuthal asymmetry requires more generalized coordinates. One such is

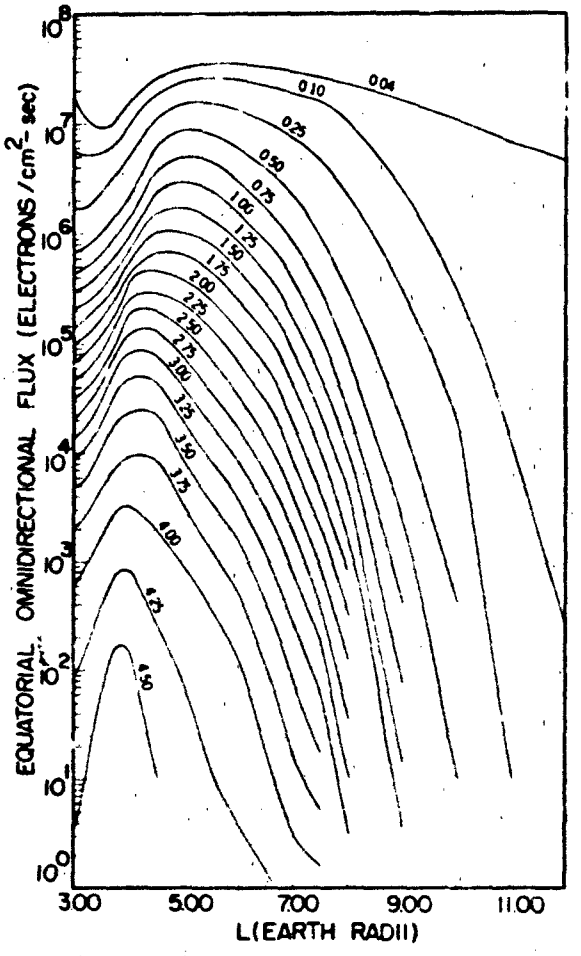


Figure 5-42. Empirical radiation belt electron fluxes. AE-4 radial profile of equatorial omnidirectional flux for various energy thresholds, *epoch 1964* [Singley and Vette, 1972].

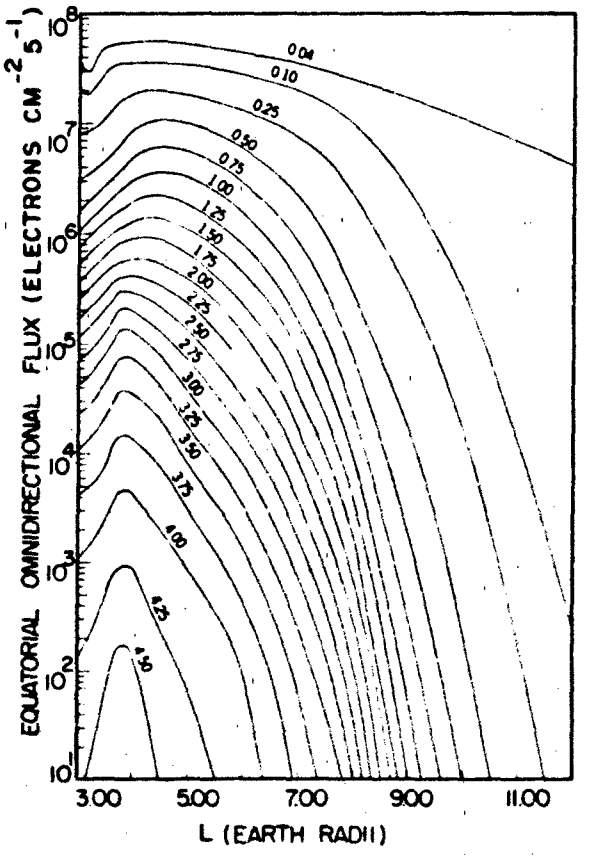


Figure 5-43. Empirical radiation belt electron fluxes. AE-4 radial profile of equatorial omnidirectional flux for different energy thresholds, *epoch 1967* [Singley and Vette, 1972].

the (strictly non-invariant) McIlwain parameter L_m [McIlwain, 1961; Stone, 1963]. L_m is defined as the equivalent dipole L-shell of a test particle having the same magnetic mirror field B_m, second adiabatic invariant, and energy as a corresponding particle in the actual, non-dipolar geometry (Figure 5-1). Shell splitting can also result from asymmetric electric fields. For mathematical details see Schulz and Lanzerotti [1974]. Figure 5-49 shows particles on the same L-shell at local noon in the noon-midnight meridian plane. When radiation belt particles drift around the earth to the midnight sector they move to a lower L-shell and smaller equatorial pitch angles preserving their first adiabatic invarient values. Those particles starting closer to the equator at noon drift to lower L-shells at midnight. Conversely, Figure 5-50 shows the position of particles at local noon having initially been on the same L-shell at local midnight. Those starting closer to the equator at midnight move closer to the magnetopause on the dayside. If they encounter the magnetopause they may become lost and there is a preferential depletion near $\alpha_0 = 90°$. This gives rise to the so-called outer zone *butterfly* distribution which is a pitch angle distribution with a minimum around an equatorial pitch angle $\alpha_0 = \pi/2$. Figure 5-51 shows a survey of the energetic electron pitch angle distributions in the near equatorial magnetosphere as measured by West [1979]. The butterfly distributions are clearly predominant in the afternoon sector after the eastward (counterclockwise) drifting electrons have interacted with the magnetopause.

Shell splitting also causes a coupling between pitch angle and radial diffusion. Any type of pitch-angle diffusion may be accompanied by radial diffusion if the B-field is azimuthally asymmetric. The direction of the radial displacement depends on the longitude at which pitch angle diffusion took place. Particles near the equator that move to lower pitch angles on the dayside will be radially displaced further from the earth on the nightside. Conversely, displacement to lower pitch angles on the nightside leads to an inward particle flux on the dayside. It is estimated that particles spend 2.3 to 3.4 of their

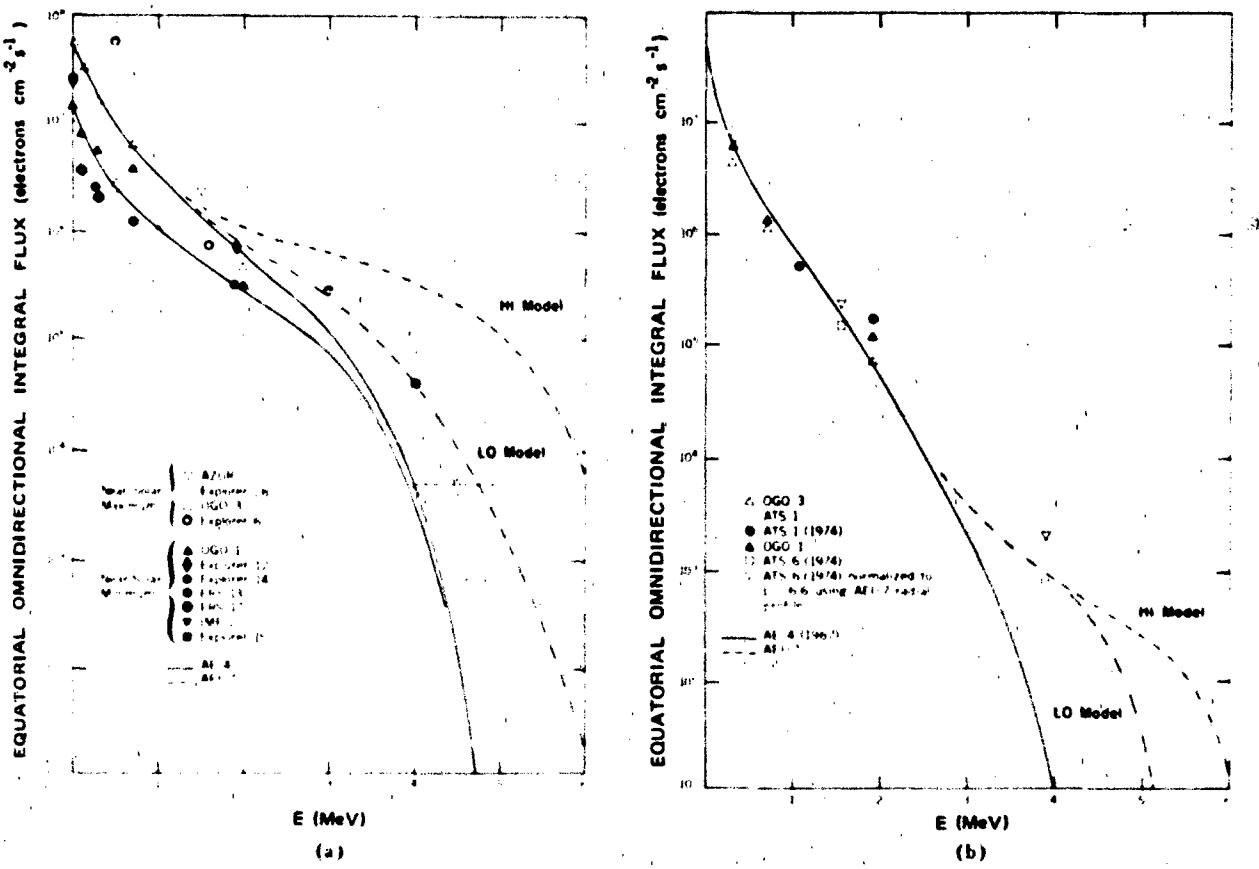


Figure 5-44. Empirical radiation belt electron fluxes.
(a) Comparison of AE1-7 Model Spectra with a number of data sets at L=4. The HI model curve is mainly based on the 0V1-19 observations from Vampola.
(b) Comparison of AE1-7 Model Spectra with a number of data sets at L=6.6 [Vette et al., 1978].

drift period on the effective dayside so that pitch angle diffusion could lead to a net energy conserving outflow of particles [Roederer, 1970]. First and second adiabatic invariant conserving inward radial diffusion as described in the theoretical modeling section would increase particle energy with inward radial motion. After undergoing many cycles of outward ϵ-conserving diffusion and inward μ-conserving diffusion, a significant local energization of trapped particles could result [Schulz and Lanzerotti, 1974 and Theodoridis, 1968], but this needs to be investigated further.

5.6.2 Geosynchronous Altitude Region ($L \approx 6.6$)

Geosynchronous altitude is 3.6×10^4 km which corresponds to an L-shell value of about 6.6. A satellite at this altitude in the plane of the earth's equator will remain fixed over the same geographical location. This feature is highly useful for communication and surveillance satellites. The natural geosynchronous charged particle environment impacts the life-time and reliability of satellites through radiation effects and spacecraft charging [Chapter 7]. Many studies of the geosynchronous environment have been made [for example, Paulikas and Blake, 1979; Young, 1979; Garrett, 1979; Baker et al., 1981; Mullen and Gussenhoven, 1983].

Near local midnight the magnetic field lines at geosynchronous altitude often depart strongly from any resemblance of dipolar shape during magnetically active periods. This effect is associated with changes in the pitch angle distribution of the particle fluxes from being peaked perpendicular to the magnetic field line to a more isotropic distribution. This and other flux changes have been used as diagnostic devices to study underlying magnetospheric processes [Higbie et al., 1978; Belian et al., 1978; Baker et al., 1978; Baker et al., 1980; Belian et al., 1981].

In this section the long term temporal behavior of energetic (> 1 MeV) electrons and the plasma environment is emphasized. Energetic electrons penetrate spacecraft shielding and may cause radiation degradation of microelectronic components. The plasma environment, of which the ions are an important component, modifies

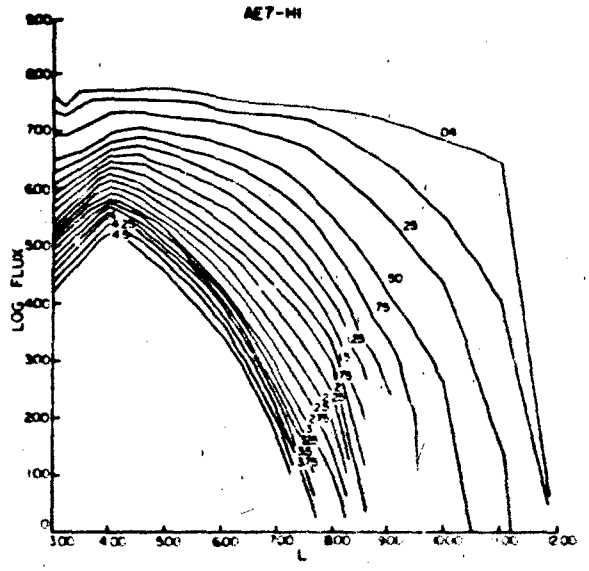


Figure 5-45 The NSSDC AE7-HI interim outer belt model for equatorial electron fluxes as a function of L. The listed energy is in MeV and the flux in units of electron $(cm^2\text{-}s)$. The discontinuous portions of these curves highlight the model's areas of least accuracy.

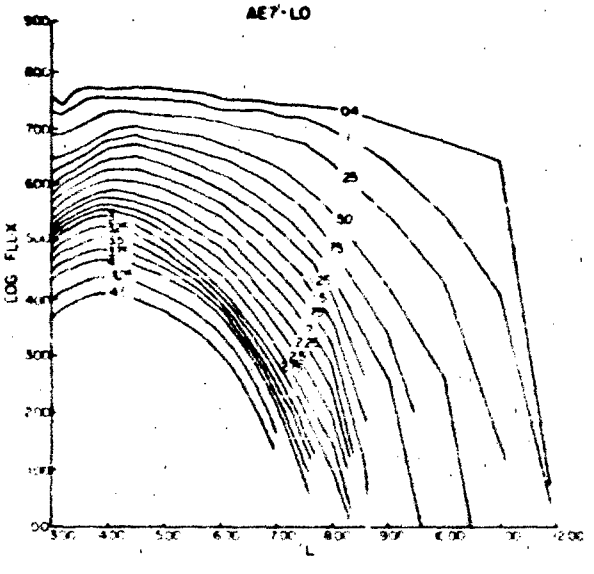


Figure 5-46 The NSSDC AE7-LO interim outer belt model for equatorial electron fluxes as a function of L. The listed energy is in MeV and the flux in units of electrons $(cm^2\text{-}s)$. The discontinuous portions of these curves highlight the model's areas of least accuracy.

the voltage to which a satellite will charge. Results from the SCATHA (Spacecraft Charging at High Altitude) satellite show that the ion composition at geosynchronous is a function of magnetic activity and local time [Mullen and Gussenhoven, 1983].

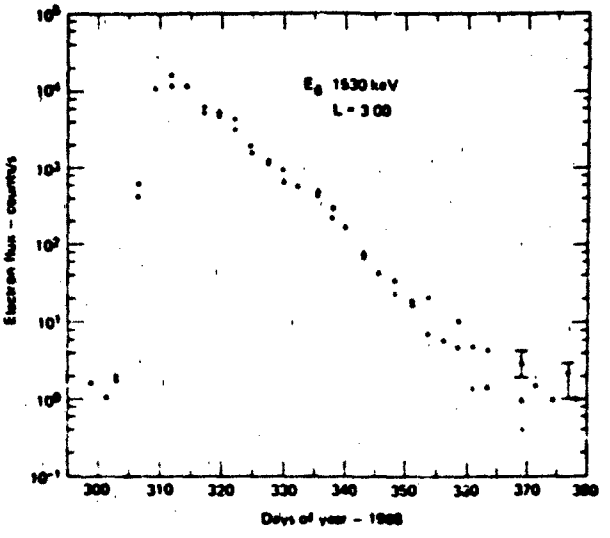


Figure 5-47 Radiation belt electron observations. An intense injection of energetic (1.53 MeV) electrons at L = 3 as reported by West et al. [1981] during October 1968. Note the four orders of magnitude increase and the subsequent exponential decay.

Energetic trapped electron flux intensities ($\epsilon \geq 1.0$ MeV) at geosynchronous altitude have been shown to be positively correlated with the average solar wind speed [Paulikas and Blake, 1979]. The 3.9 MeV integral electron flux [$J(> 3.9$ MeV)], for example, has been observed to vary by about a factor of 5 from a solar wind speed of 400 km/sec to one of 800 km/sec. Lower energy electron fluxes (140–600 keV), on the other hand show little such correlation. Sufficiently long time averages (~1 year) empirically remove the solar wind speed effects and reflect the overall average stability of the electron fluxes over longer time scales. For details see Paulikas and Blake [1979].

Energy spectra of energetic electrons in the geosynchronous altitude region measured with the SCATHA spacecraft are shown in Figure 5-52 [Mullen and Gussenhoven, 1983]. These data represent 75-day averages taken between February 1979 and February 1980, and on the average may be represented by a power law spectral dependence. Integration of the fitted curves gives integral flux levels that are consistent with the AE-4 and AE-7 models (Figures 5-42, 43, 45 and 46). This implies that the long term temporal averages of the electron fluxes at geosynchronous altitude did not materially change during the 1970's. The scatter of the individual SCATHA flux data measurements about the mean time-averaged flux is substantial. At times the observed electron fluxes differed from the mean values by well over an order of magnitude. The flux models, therefore, should be used with caution.

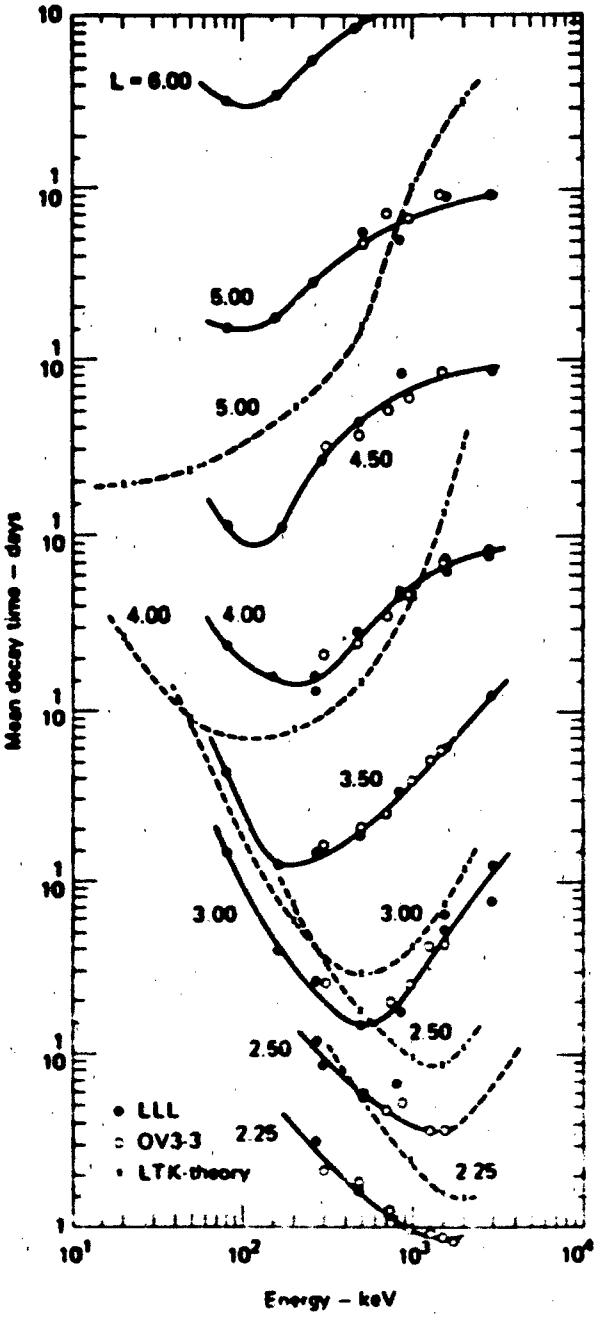


Figure 5-48. Radiation belt electron observations. Pitch angle diffusion lifetimes at constant L were in part derived from the data in the previous figure, and the predicted electron precipitation lifetimes are those of Lyons et al. [1972] [West et al., 1981].

It is useful to estimate the percentage of time that the electron flux will exceed a certain value. Figure 5-53 shows the cumulative probability distribution for the SCATHA 1.4 – 2.6 MeV energy channel in four L intervals. For example, from this figure one would

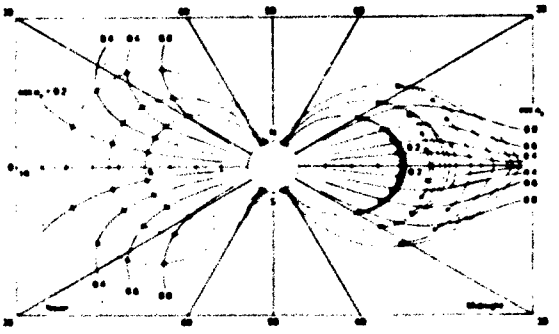


Figure 5-49. Theoretically computed shell splitting effects for particles starting on common field lines in the noon meridian. Dots represent the particles' mirror points. The curved lines give the position of mirror points for constant equatorial pitch angle α_o [Roederer, 1970].

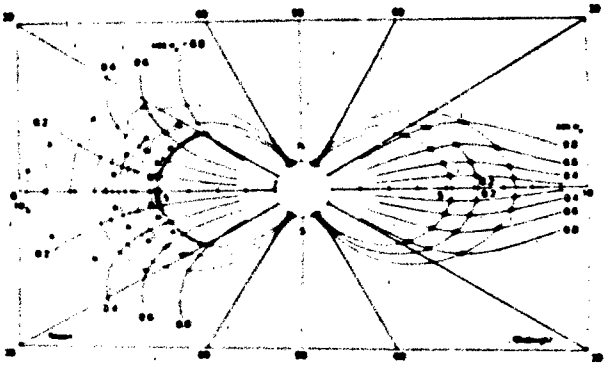


Figure 5-50. Theoretically computed shell splitting effects for particles starting on common field lines in the midnight meridian [Roederer, 1970].

expect to find 1.4 2.6 MeV electrons with a flux intensity greater than 10 electron cm^{-2} sec^{-1} sr^{-1} keV) 50% of the time in the 5.5 6.0 L-shell region. On the other hand, from the same bar graph fluxes above 100 electron cm^{-2} sec^{-1} sr^{-1} keV are expected less than 10% of the time.

Ions are a dynamic component of the radiation environment at geosynchronous altitudes. The ion composition at low energies varies with magnetic activity as illustrated in Figure 5-54 [Mullen and Gussenhoven, 1983]. During magnetically active periods (high Kp) the O^+ component becomes enhanced relative to protons. An explanation for O^+ enhancement [Kaye et al., 1981; Fennell et al., 1981] is that the storm enhanced O^+ ions are accelerated up along the magnetic field lines from the auroral ionosphere, while the protons probably also originate from the magnetotail plasma sheet that moves closer to the earth during magnetically active periods. Oxygen ion enhancements also increase the particle

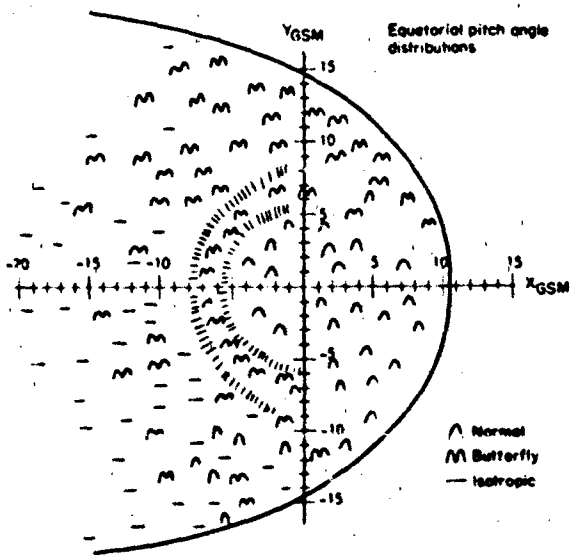


Figure 5-51 Survey of energetic electron pitch angle distributions observed in the near equatorial magnetosphere [West et al., 1979]

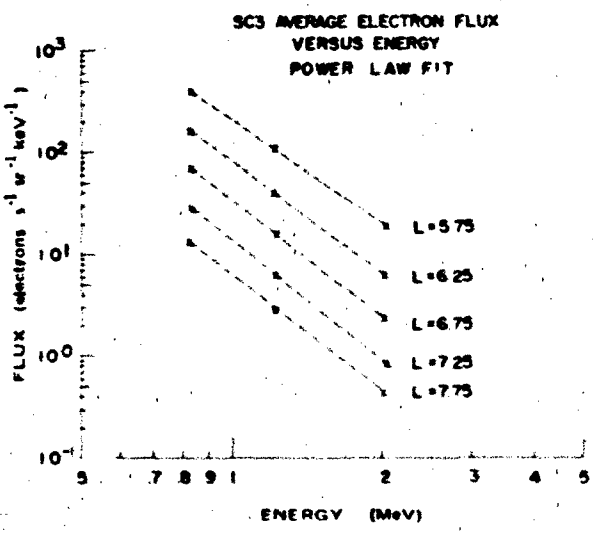


Figure 5-52 Time-averaged energetic electron energy spectra measured near geosynchronous altitude from February 1979 to February 1980. These curves represent a 75 day average and approximate a power law curve. Individual data show, however, that deviations two orders of magnitude from these means are not uncommon [Mullen and Gussenhoven, 1982].

energy density relative to that of the magnetic field. Figure 5-55 [Mullen and Gussenhoven, 1982] shows the geosynchronous altitude ratio β of the particle energy density to the magnetic field energy density as a function of local time. The individual curves represent various levels of K_p where higher K_p is a measure for higher magnetic activity [Jacobs, 1970]. A β value greater than unity

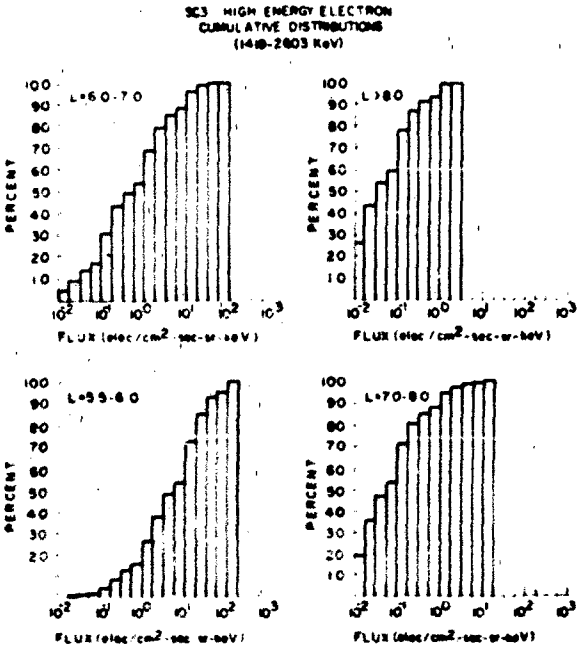


Figure 5-53 Cumulative probability that the high energy electron flux (at energies 1419-2603 keV) is less than the shown levels. This figure gives an estimate of the "spread" in the individual measurements used to obtain Figure 5-51 [Mullen and Gussenhoven, 1982].

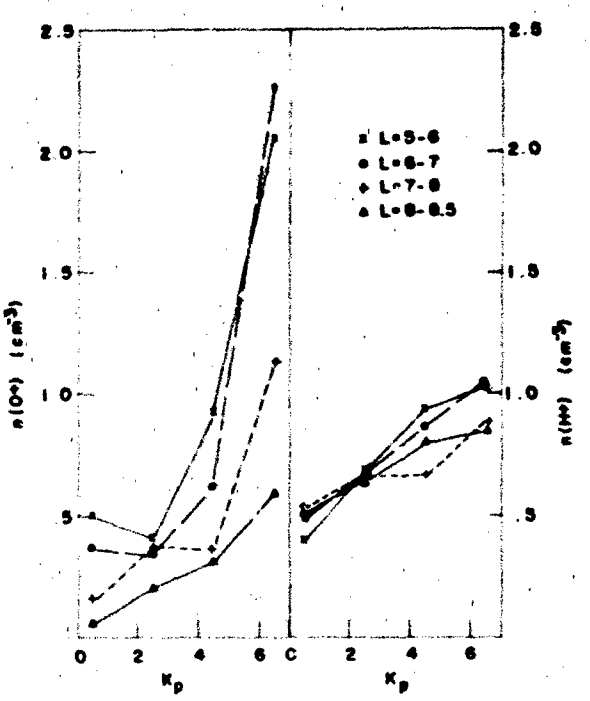


Figure 5-54 Average Oxygen (left) and hydrogen (right) number densities as determined from the Lockheed ion composition measurements on the SCATHA spacecraft (for energies 1-32 keV) versus K_p for various L-shell intervals [Mullen and Gussenhoven, 1982].

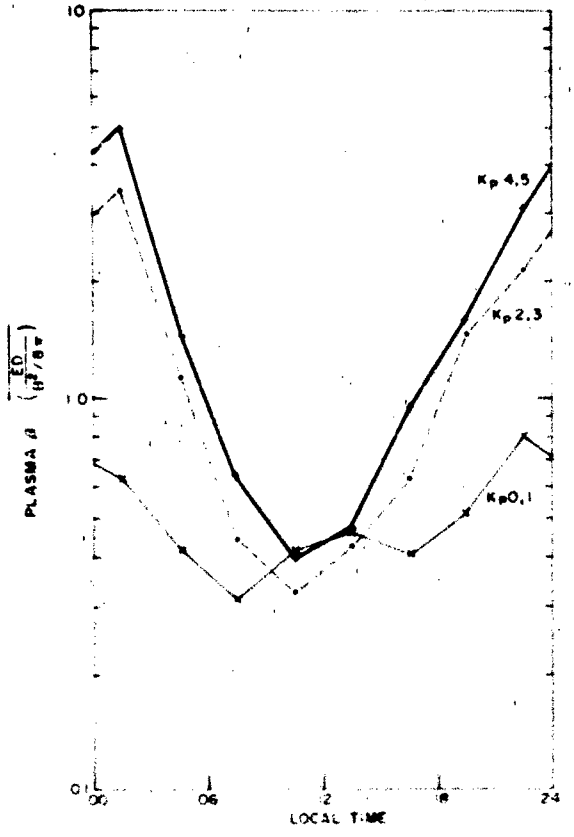


Figure 5-55 Time-average of the ratio of the particle energy density to the magnetic field energy density ($\beta \cdot 8\pi$) as a function of local time. Higher K_p implies higher magnetic activity. These data represent over 90 days averages at of geosynchronous altitude (L = 6.6) obtained during February 1979 to February 1980 [Mullen and Gussenhoven, 1982]

implies that the particles are not strongly confined by the magnetic field. A β value significantly less than unity suggests particle confinement in this region. These results indicate that the plasma processes that are operative during active periods at geosynchronous altitude are substantially different than those operative during quiet times. Garrett [1979] gives more details of the geosynchronous plasma environment.

5.7 THE RING CURRENT

The gradient-curvature drift of radiation belt particles causes differential motion that is mass and charge dependent: electrons drift eastward and positive ions drift towards the west. This constitutes an electrical current around the earth in the westward direction, called the extraterrestrial ring current. During geomagnetic disturbances, such as magnetic storms, the population of trapped particles at ~1 - 800 keV energies is substantially enhanced on L-shells between L = 3 and L = 6. As a consequence, the ring current is intensified and magnetic disturbances at the earth result.

5.7.1 Electrical Current Relations

The ring current itself produces a magnetic field that is superimposed on the earth's magnetic field. Enhancement in the ring current constitutes the cause of the magnetic field *depression* observed at mid-latitudes on the surface of the earth during the main and recovery phases of geomagnetic storms as well as magnetic field *enhancement* beyond L ~ 6.7, as illustrated in Figure 5-3. For this reason we shall consider the magnetic effects of the trapped particles.

As one can see from Equations (5.36) and (5.37), the particle gradient-curvature drift velocity is proportional to the particle energy. With the equatorial pitch angle, α_o, one may write

$$\mathbf{V}_d = \frac{\epsilon}{qB^3}\,(1 + \cos^2\alpha_o)(\mathbf{B} \times \nabla |\mathbf{B}|)\,, \qquad (5.105)$$

where electric field and gravity effects have been disregarded. Here $\mathbf{B}$ is the magnetic field due to the main (earth's internal) dipole moment and $\mathbf{M} = M\,\hat{\boldsymbol{\theta}}$ where $\hat{\boldsymbol{\theta}}$ is a unit vector in the magnetic northward direction. At the magnetic equator

$$\mathbf{B} = \frac{M}{r^3}\,\hat{\boldsymbol{\theta}} = -\frac{B_E}{L^3}\,\hat{\boldsymbol{\theta}}\,. \qquad (5.106)$$

For simplicity in the presentation, we shall consider equatorially mirroring particles only, that is, $\alpha_o = \pi/2$; for the mathematical treatment of the general case of an arbitrary pitch angle α_o, see Dessler and Parker [1959]. We then get

$$\mathbf{V}_d = c\,\frac{3\epsilon}{qM}\,r^2\,\hat{\boldsymbol{\phi}} = c\,\frac{3\epsilon}{qB_E}\,L^2\,\hat{\boldsymbol{\phi}} \qquad (5.107)$$

where $\hat{\boldsymbol{\phi}}$ is a unit vector in the eastward azimuthal direction around the earth. From Maxwell's equations (the Biot-Savart law; [Jackson, 1975]) the magnetic field generated by the drift motion of each particle is

$$\mathbf{B}_d = i \int_0^{2\pi r} \frac{d\ell}{r^2}\,\hat{\boldsymbol{\theta}} \qquad (5.108)$$

where i is the magnitude of the single-particle drift "current":

$$i = \frac{q|\mathbf{V}_d|}{2\pi r}\,, \qquad (5.109)$$

and thus

$$\mathbf{B}_d = -\frac{3\epsilon}{M}\,\hat{\boldsymbol{\theta}}\,. \qquad (5.110)$$

The minus sign indicates that the particle azimuthal drift generated field opposes the main (internal) dipole field earthward of the ring current particle population.

There is also a magnetic effect of the particle's spiral motion around the field lines. Each gyro-loop may be considered a small dipole moment $\mu = \frac{\epsilon}{B}$ (since for $\alpha_0 = \pi/2$, $\epsilon = \epsilon_\perp$), and the associated magnetic field is

$$B_g = \frac{\mu}{r^3} = \frac{\epsilon}{Br^3} = \frac{\epsilon}{B_I R_E^3} = \frac{\epsilon}{M} , \qquad (5.111)$$

which is in the direction of the internal dipole. The total perturbation at the origin due to a single equatorially mirroring particle is then

$$\Delta B_\epsilon = B_d + B_g = \frac{2\epsilon}{M} . \qquad (5.112)$$

At the equatorial surface of the earth the unperturbed (internal) dipole field is just $B_I = \frac{M}{R_E^3}$ so that

$$\frac{\Delta B_\epsilon}{B_I} = \frac{2\epsilon}{B_E^2 R_E^3} = \frac{2\epsilon R_E^3}{M^2} , \qquad (5.113)$$

and noting that the total energy in the earth's unperturbed dipole field above the earth's surface may be written as

$$U_m = \frac{1}{3} B_E^2 R_E^3 = \frac{1}{3} \frac{M^2}{R_E^3} , \qquad (5.114)$$

We may express the relative ring current single particle perturbation as

$$\frac{\Delta B_\epsilon}{B_I} = \frac{2\epsilon}{3U_m} . \qquad (5.115)$$

It turns out that this expression is valid for trapped particles in the radiation belts regardless of the equatorial pitch angle α_0 [Dessler and Parker, 1959]. In deriving Equation (5.115) it was assumed that the total energy in the ring current is less than the magnetic field energy U_M. When that is not the case Equation (5.115) is no longer strictly valid and may be in error by up to a factor of two.

By summing up the effects of all the individual particle motions in the geomagnetic field, one arrives at the total magnetic field perturbation

$$\Delta B = \sum_i \int_0^{\pi/2} \sin\alpha_0 d\alpha_0 \int_1^{L_{max}} dL \int_{\epsilon_{min}}^{\epsilon_{max}} d\epsilon \; \Delta B_\epsilon f_i(\alpha_0, L, \epsilon) \qquad (5.116)$$

where $f_i(\alpha_0, L, \epsilon)$ is the distribution function for particle species i, expressed as function of equatorial pitch angle, L-shell and energy.

Parker [1957] developed a hydromagnetic formalism that provides an alternative to this extensive integration. One may define the macroscopic particle pressures in the direction parallel and perpendicular to the local magnetic field direction

$$P_\parallel = \sum_i \iint F_i(X, v, \alpha_0) m v^2 \cos^2\alpha_0 dv d\alpha_0 \qquad (5.117)$$

$$P_\perp = \frac{1}{2} \sum_i \iint F_i(X, v, \alpha_0) m v^2 \sin^2\alpha_0 dv d\alpha_0 \qquad (5.118)$$

where F_i is the particle distribution function for a particle species i expressed in position, (X), speed, and pitch angle coordinates. The summation is extended over all particle species. The magnetic field pressure is

$$P_m = \frac{B^2}{8\pi} . \qquad (5.119)$$

With these pressure expressions the total gradient-curvature drift current can be written as

$$I_d = \frac{c}{8\pi P_m} B \times \left\{ \frac{P_\perp}{2P_m} \nabla P_m + \frac{P_\perp}{P_\parallel} (B \cdot \nabla) \frac{B}{8\pi} \right\} , \qquad (5.120)$$

where c is the velocity of light [Williams, 1982]. The corresponding gyration current of the particle distribution as a whole is

$$I_g = \frac{c}{8\pi P} B \times \left\{ \nabla P_\perp - \frac{P_\perp}{2P_m} \nabla P_m - \frac{P_\perp}{P_m} (B \cdot \nabla) \frac{B}{8\pi} \right\} . \qquad (5.121)$$

The two terms within the brackets of Equation (5.120) stem from the magnetic field gradient and field curvature respectively, and the three terms within the brackets of Equation (5.121) represent currents driven by the particle pressure gradient, the magnetic field gradient, and the magnetic field line curvature. The total current of all particles then reduces to

$$I = I_d + I_g = \frac{c}{8\pi P_m} B \times \left\{ \nabla P_\perp + \frac{(P_\parallel - P_\perp)}{P_m} (B \cdot \nabla) \frac{B}{8\pi} \right\} \qquad (5.122)$$

As a rule of thumb, it requires a total of 4×10^{22} ergs of particle kinetic energy to produce a surface magnetic field depression of ~100nT (1 nT = 1γ = 10 ^{5}G).

5.7.2 Composition and Sources

The population of trapped particles is made up primarily of electrons, protons, helium ions, carbon ions, and oxygen ions. The composition is found to vary substantially with location (such as L-shell) and with geomagnetic conditions. The heavier ions, such as He^+ and O^+, may be dominant during disturbed conditions on L-shells in the range L = 3 to L = 5, while prolonged quiet periods tend to favor H^+ (protons) above tens of keV energies. The latter is also a reasonable expectation since at typical ring current energies (~ 70 keV) the charge exchange lifetime of H^+ is longer. At lower energies, below a few keV, the H^+ lifetimes are shorter than that of He^+ and O^+ lifetimes, and the opposite may be true. Figure 5-56 shows a relative comparison of ring current ion flux observations during four different time periods. These results pertain to $\epsilon \leq 20$ keV energies while the ionic composition at higher energies remains to be investigated observationally.

Ionospheric or atmospheric ions probably form a significant fraction of the ring current population. This is inferred from the observations reported by Shelley et al. [1974] and Sharp et al. [1976]. Low-orbiting polar satellites detected the precipitation of oxygen ions (L = 6.8) during distrubed times and also detected field-aligned upward moving accelerated ions from the auroral ionosphere. Sufficient pitch angle scattering at higher altitudes could cause these upward moving ions to become stably trapped and form part of the ring current.

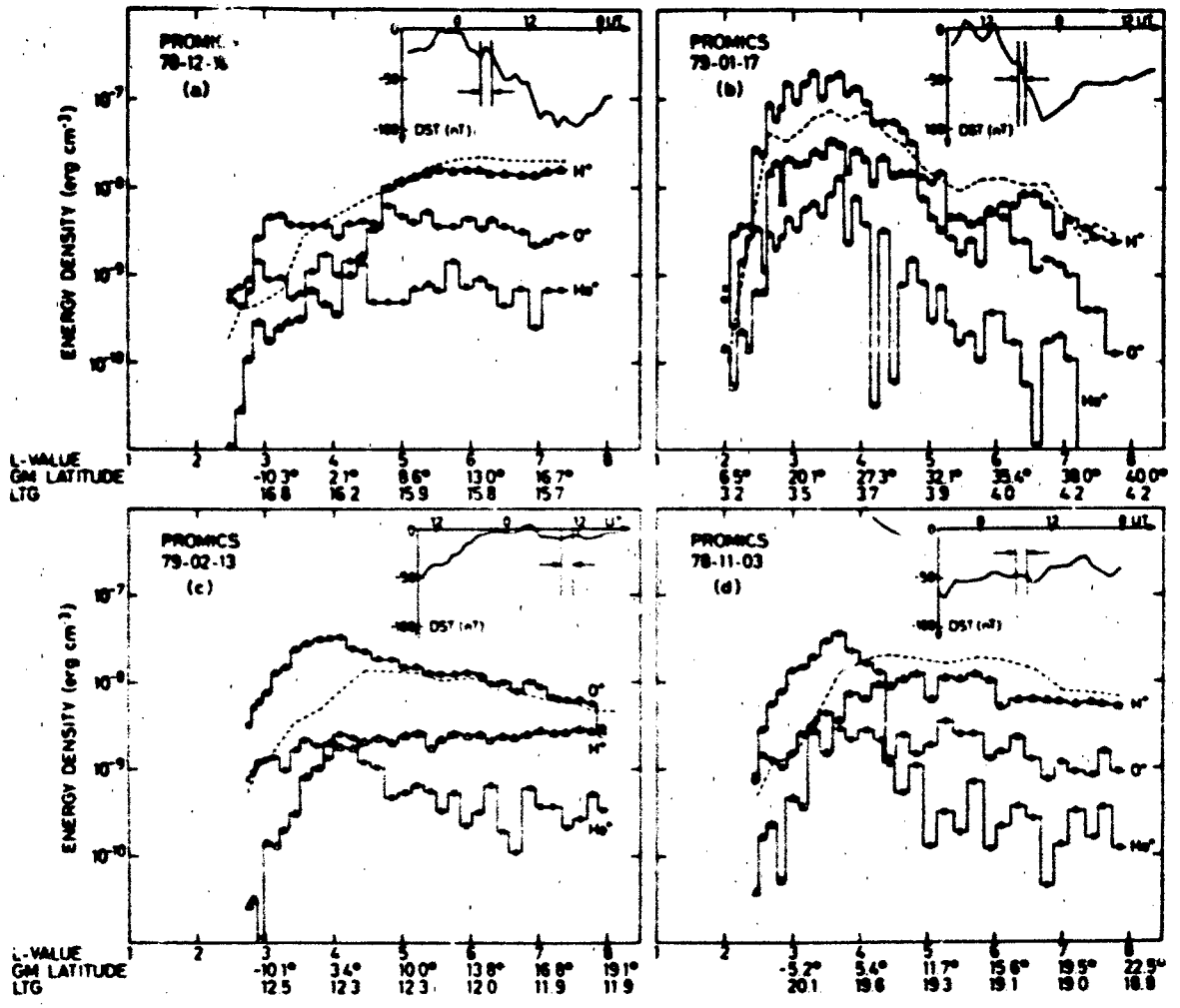


Figure 5-56 Observed energy densities of H^+, He^+ and O^+ ions in the radiation belts at ring current energies (0.2-17 keV/e ion) versus orbital parameters. The data were derived from the PROMICS experiment on the PROGNOZ-7 spacecraft. The dashed curves indicate apparent energy densities calculated from a total ion (E-q) spectrometer (at 0.1-45 keV) under the assumption that only protons were measured. The results show the importance of the heavier ions at different times and locations. The four panels represent data for four different periods, and the D_{st}-index history is also shown as a guide to the ring current activity [Lundin et al., 1980].

Based on total ion (no mass resolution) observations, it has become clear that the greatest contribution to the ring current comes from ~20–200 keV ions where the mass composition is yet unknown. This is illustrated in Figure 5-57 [Williams, 1981].

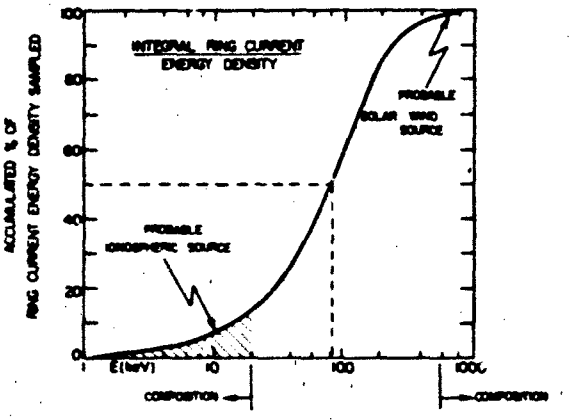


Figure 5-57. Integral representation of the ring current energy density as a function of ion energy. The contribution to the ring current energy density is greater where this curve is steepest. This is also where currently the ionic composition is unknown [Williams, 1981].

A practical measure for the overall strength of the extraterrestrial ring current is the D_{st}-index which measures the middle latitude spatially averaged decrease in the horizontal component H of the earth's surface magnetic field. $D_{st} = \langle \Delta H \rangle$. Under this definition the quiet time ring current corresponds to $D_{st} = 0$. Hourly values of the D_{st} index are published by NASA National Space Science Data Center, Goddard Space Flight Center, Maryland. Magnetic storms generally have D_{st} depressions on the order of 100 to 200 nT (very large storms may exceed $|D_{st}| = 300$ nT), and the D_{st} index may also fluctuate substantially for other geomagnetic conditions for which $|D_{st}|$ generally remains less than 50 nT. Figure 5-58 shows an example of the D_{st} index plotted for June-December 1972, and the occurrence of four magnetic storm periods in June–August, September and October/November is evident. The D_{st} index is therefore very useful to identify magnetic storms from surface magnetogram records.

5.7.3 Adiabatic Effects Produced by the Ring Current

Much of the time the magnitude of the electrical current set up by the azimuthally drifting radiation belt particles changes slowly in comparison with the ion drift period. Therefore, the third adiabatic invariant, which is proportional to the enclosed magnetic flux threading the drift path, is most often conserved following injection or acceleration. Sŏraas and Davis [1968] have shown that significant adiabatic effects will take place even for moderate values of D_{st}. To separate the adiabatic and non-adiabatic features one can transform to a $D_{st} = 0$ reference. If $j_1(\epsilon_1, L_1)$ is the equatorially mirroring flux for $\epsilon = \epsilon_1$ ($L = L_1$, $D_{st} = 0$) and $j_2(\epsilon_2, L_2)$ is the corresponding flux for $D_{st} \neq 0$ then one has

$$j_2(\epsilon_2, L_2) = (\epsilon_2, \epsilon_1) \quad j_1(\epsilon_1, L_1) \; . \qquad (5.123)$$

This follows from Liouville's theorem which states that the phase space density is constant along dynamical particle trajectories ($f = \text{const.} = j/p^2$). The equatorially mirroring energy ϵ_2 is mapped from the unperturbed ϵ_1 energy by conservation of the first adiabatic invariant:

$$\epsilon_2 = \epsilon_1 \quad (B_2, B_1) \quad , \qquad (5.124)$$

where B_1 is the value of the quiet time magnetic field induction.

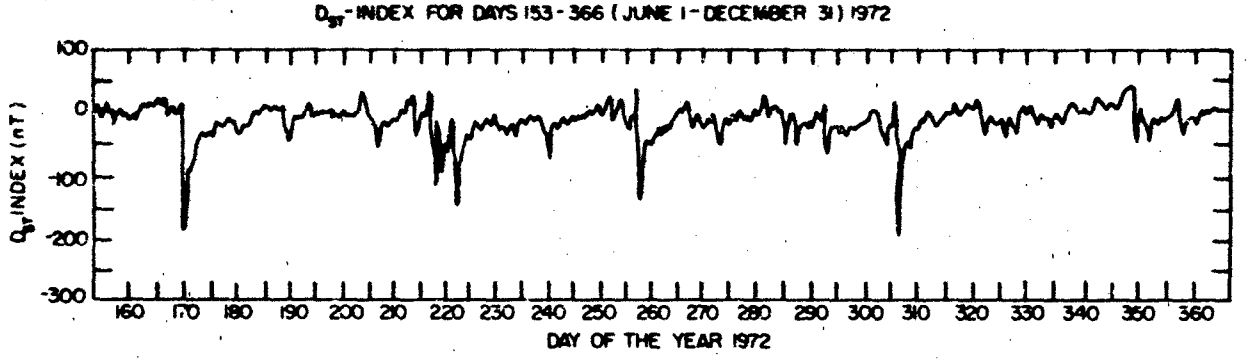


Figure 5-58. An example of the time-variation of the horizontal magnetic field component (D_{st}) at the equator. The large rapid drops in D_{st} correspond to a build up of the ring current during magnetic storms followed by subsequent decay.

For a dipole field where $B_E = 0.312$ G

$$B_1 = \frac{B_E}{L_1^3} \tag{5.125}$$

and

$$B_2 = \frac{B_1}{L_2^3} + \Delta B(r) . \tag{5.126}$$

The magnetic flux enclosed by the drift path (the third adiabatic invariant) is given by

$$\Phi_1 = 2\pi \int_0^{L_1} \frac{B_1}{L^3} L\,dL = 2\pi \int_{L_1}^{\infty} \frac{B_E}{L^3} L\,dL = \frac{2\pi B_E}{L_1} \tag{5.127}$$

and

$$\Phi_2 = 2\pi B_1 L_2^2 + 2\pi \int_0^{L_2} \Delta B(r)\, r\,dr , \tag{5.128}$$

where $\Delta B(r)$ is the magnetic field change induced by the ring current enhancement [Söraas and Davis, 1968]. Equation (5.128) uses the fact that the integral from 0 to L includes the return magnetic flux through the earth so that the net dipole magnetic flux through the entire equatorial plane is zero. This fact allows the 0 to L_1 integration interval to be replaced by one for L_1 to ∞. The ions will now equatorially mirror at L_2 where by equating Φ_1 and Φ_2

$$B_1 L_2^2 = B_1 L_1^2 + \int_0^{L_2} \Delta B(r)\, r\,dr \tag{5.129}$$

and

$$\Delta B(r) = 0.7\, D_{st}\, f(r) \tag{5.130}$$

The function $f(r)$ is shown in Figure 5-59. For a given D_{st}, L_2 can be found, and hence, B_2. Knowing B_2, ϵ_2 can be determined. In Figure 5-60 we show an illustrative example as presented by Söraas and Davis [1968].

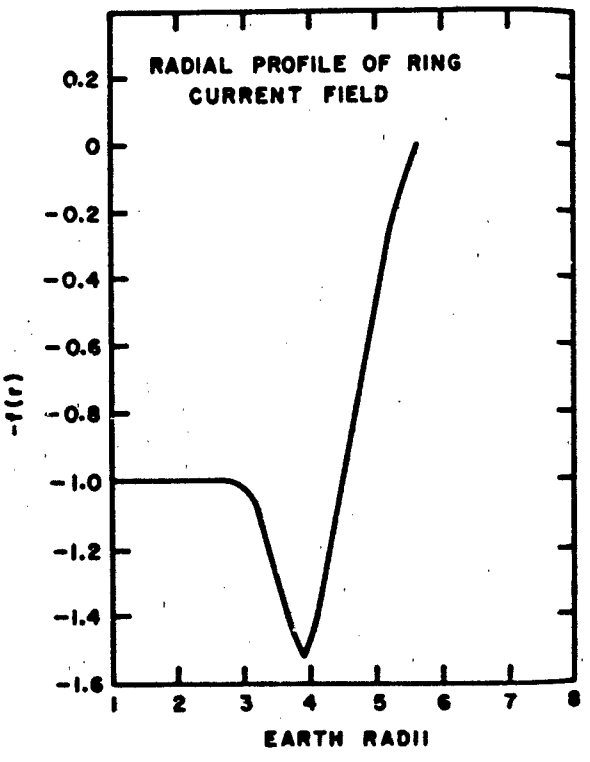


Figure 5-59. The radial dependence of the ring current magnetic field used in the calculation of adiabatic effects on trapped protons by Söraas and Davis [1968]. Notice that the decrease is greatest near $L=4$.

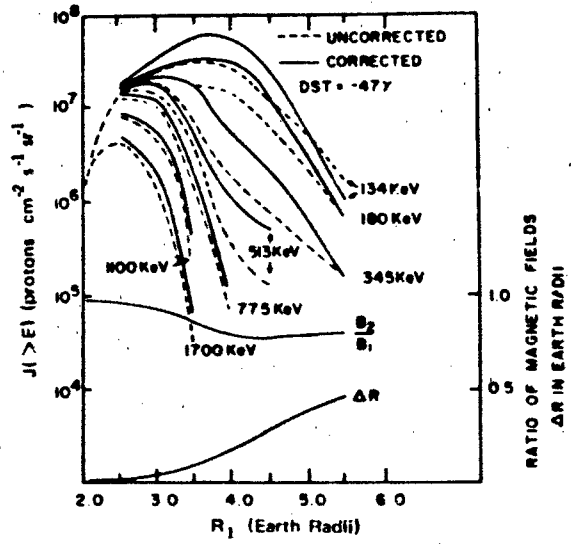


Figure 5-60. The radial proton integral energy intensity profile as measured on day 109 of 1965 when $D_{st} = -46$nT, together with the transformed profile corresponding to $D_{st} = 0$ assuming the three adiabatic invariants of motion conserved. The radial dependence of the ratio between the magnetic field after and before the build up of the ring current and the radial movement ΔR of the particles are shown in the lower part of the figure [Söraas and Davis, 1968].

5.8 RADIATION EFFECTS ON SPACE SYSTEMS

Among the known effects of particle radiation on space systems are spacecraft charging phenomena and effects of penetrating radiation on materials. Specific effects include detector malfunction and degradation.

optical system degradation, memory system alteration, and control system malfunction or failure. For manned space operations, biological effects are a major concern.

A crude measure for damage done by penetrating energetic radiation is *radiation dosage* which is measured in *rads*. This unit is defined as an energy deposition of 100 ergs (6.25×10^7 MeV) in one gram of a material substance. This definition does not distinguish between different kinds of incident radiation or different effects on the material. Radiation dosage is thus only an overall measure, and it is often necessary to examine specific interaction cross-sections when studying radiation effects.

Energy is deposited through chemical (molecular bond changes, bound electron excitation, and ionization) and nuclear (element transmutation, nuclear excitation, and induced radioactivity) interactions. The macroscopic effects are evident in device failure after a critical level of radiation exposure is reached. Most often this critical level depends directly on the nature and energy characteristics of the incident radiation.

A major concern is the on orbit lifetime of microelectronic devices that are designed to a specific level of radiation "hardness" (such as 10^4-10^5 rad). There is in many cases a trade off between orbit choice and system lifetime that must be determined.

A lowest order approximation to the expected radiation exposure effects can be estimated by combining the energy deposition rate versus incident energy curves of Janni and Radke [1979] with the expected radiation belt flux intensity deduced from previous observations, as in Section 5.6 or from theoretical modeling. A simple (but very crude) approach to estimating the radiation dosage follows.

It is assumed that shielding is equal in all directions so that a spherical shield approximation can be used. The shielding is also assumed to be aluminum or close to it in density. The incident omnidirectional particle fluxes are normalized (or scaled) to unity at a selected energy so that dosage need only be calculated as a function of spectral shape. The dose rate for a given energy spectrum is found by multiplying the resulting dosage by the model (or measured) omnidirectional flux at the selected energy. Total dosage is determined by integrating over the expected exposure time of the satellite.

Figure 5-61 shows a number of exponential spectral radiation curves normalized to unity at 1 MeV for energetic electrons up to 6 MeV. By picking the one curve that most nearly approximates the actual expected radiation energy spectrum, one has a one-parameter spectral representation, the spectral e-folding energy E_0. One then proceeds to calculate the radiation exposure for the normalized spectrum behind a certain thickness of shielding by using the curves in Figure 5-62. This process has to be averaged over the expected radiation conditions for the expected spacecraft orbit during the period of the desired orbital operations.

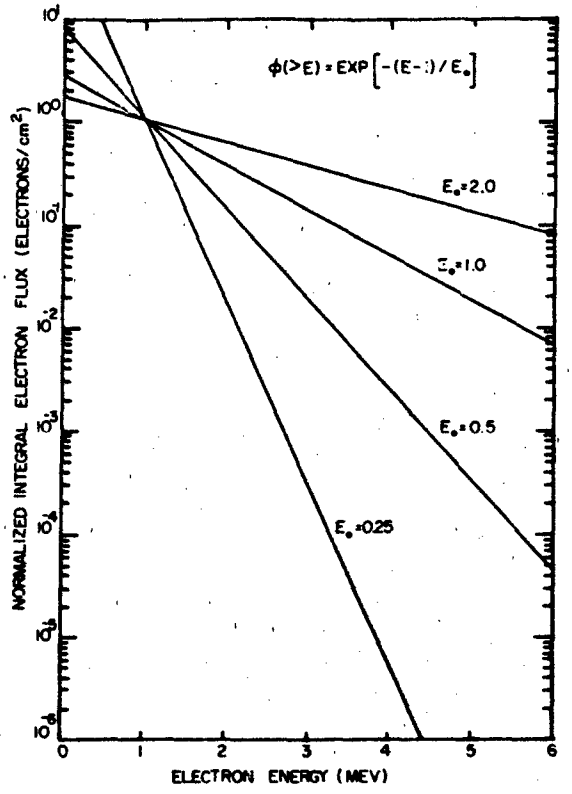


Figure 5-61. Model electron spectra for engineering applications: Normalized integral omnidirectional electron flux as a function of electron energy. By matching a model energy spectra to the curves the appropriate E_0 can be quickly found. The flux at 1 MeV is normalized to 1 electron cm² [Janni and Radke, 1979].

A similar technique also applies to ions. For protons, Figure 5-63 shows a similar set of exponential spectral radiation curves extending to 300 MeV, and Figure 5-64 gives the radiation exposure dosage as function of the aluminum shielding thickness. Notice that the shielding is generally less effective in reducing the radiation dosage due to the very energetic ions in the radiation belts.

5.8.1 Detector Malfunctions

Single particle upsets occur when a single particle creates enough free electrons to simulate a device logic state change. Single particle effects are particularly severe in small (<10 μm) sensitive regions of microelectronic devices. Figure 5-65 illustrates the incidence of a cosmic ray (very energetic heavy ion) in a single memory cell commonly used for onboard information storage. Notice that the volume where the ionization takes place is at least comparable to the sensitive cell region itself. It is presently an area of controversy whether reduction in cell size will always increase the soft error or single event upset rate; it is conceivable that with very small memory cells the ionization volume could encompass many cells.

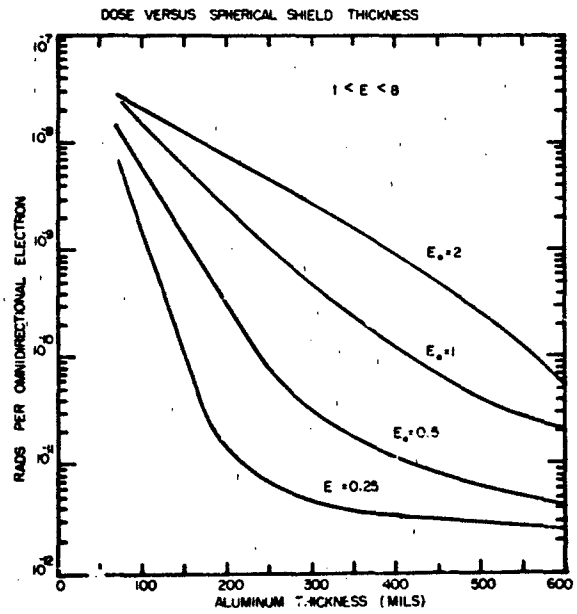


Figure 5-62. Radiation dosage experienced behind various shielding thicknesses of aluminum absorber for several incident electron energy spectra normalized to 1 electron cm^2 at 1 MeV. Multiplication by the 1 MeV flux gives the expected radiation dose rate [Janni and Radke, 1979].

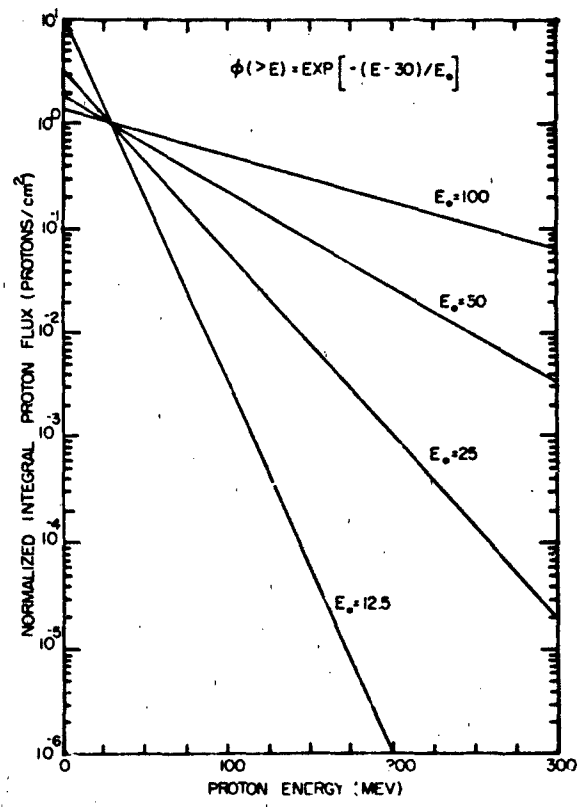


Figure 5-63. Model proton spectra for engineering applications: Normalized integral proton flux curves for estimating E_0 where the flux at 30 MeV is normalized to 1 proton cm^2 [Janni and Radke, 1979].

On the average, the effect of single-particle incidence on detector materials, such as aluminum and silicon, is the generation of one electron-hole pair per 3.6 eV of energy deposition. Thus, 1 rad of incident radiation in 1 gram of material creates 1.74×10^{13} electron-hole pairs. Even a moderately energetic radiation belt particle (for example 1 MeV) will create a large number of free charge carriers in the detector material and may lead to false signals. Prolonged exposure to energetic particles degrades the detector performance by the accumulation of material microstructural damage. For example, solid state detectors of the Al-Si-Au variety are found to have a factor of 10 increase in useful lifetime when the aluminum side is facing the radiation exposure (compared to the gold side). Very energetic and very heavy cosmic ray ions have a particularly devastating effect on detector systems; for further details see Adams and Partridge [1982] and McNulty [1981].

Figure 5-66 shows an example of a nuclear interaction occurring near a sensitive region of a radiation particle detector or other solid state device. An incident proton, for example, will stimulate a ^{28}Si nucleus to emit an alpha particle which has a short range. The recoiling ^{28}Si nucleus stops in even a much shorter distance. The combined effect can deposit tens of MeV in a small $(10 \times 20 \times 20)$ μm volume element creating a "soft" (data) error. Accelerator data [McNulty et al., 1980] indicate that the proton induced ($E > 40$ MeV) soft error rate is less than 10^{-6} soft errors/(protons-cm^2). An upper

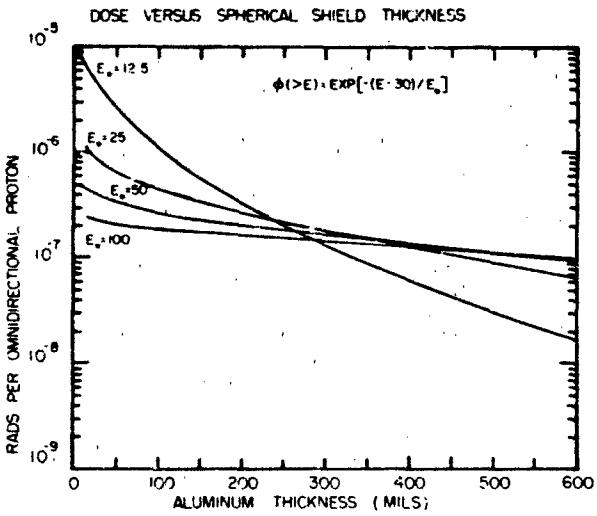


Figure 5-64. Radiation dosage experienced behind various shielding thicknesses of aluminum absorber for several incident proton spectra normalized to 1 proton cm^2 at 30 MeV. Multiplication by the omnidirectional flux at 30 MeV will give the radiation dose rate [Janni and Radke, 1979].

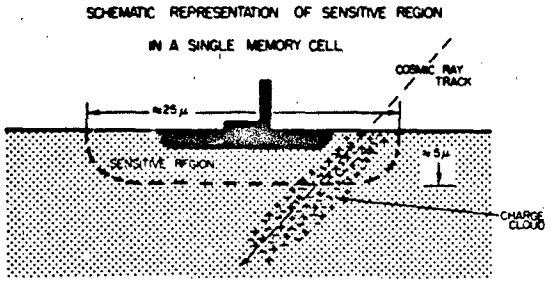


Figure 5-65. Illustration of radiation effect: Electron-hole generation near a sensitive region as a result of local ionization produced by a traversing cosmic ray or energetic particle [B. Blake, personal communication, 1982].

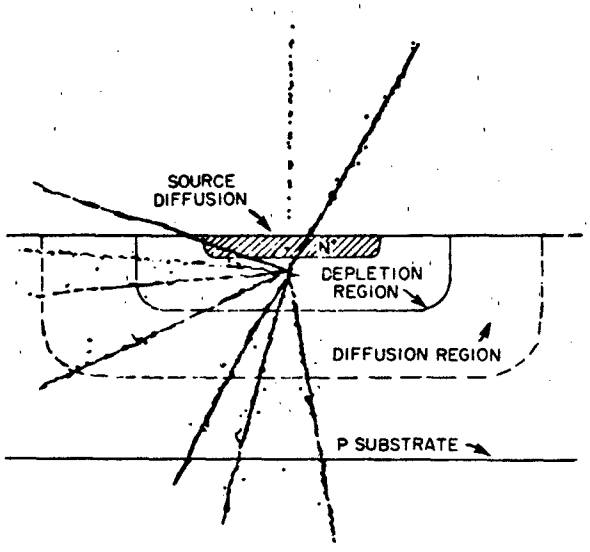


Figure 5-66. Illustration of the production of numerous secondary particles from the nuclear interaction of the primary particle in a sensitive volume of a solid state detector device [McNulty et al., 1980].

limit to the expected error rate (error sec), therefore, can be found by using the proton flux models for $E > 40$ MeV times 10^{-6}. If trapped heavy ions are sufficiently abundant they could dominate the soft error rate [Adams et al., 1981]. This is one of several purely practical motivations for measuring the energetic trapped particle composition to an adequate degree of accuracy.

5.8.2 Memory Alteration

Certain microcircuitry used in current spacecraft instrumentation has proven very susceptible to the effects of energetic heavy ions in the radiation belts and in the cosmic radiation. Memory chips and microprocessors are frequently found to have their logical states and information content severely altered by the localized energy deposition process. Similar effects can result from alpha-particle emission from nuclear interactions and from natural and induced radioactivity in the devices themselves.

If the effects are infrequent in occurrence, engineering design emphasizing redundance of the critical components could circumvent the problem. However, when the effects are frequent and or persistent this approach may not be feasible.

Figure 5-67 [E. Petersen, personal communication, 1982] shows the energy deposited in a 10 μm — thick sensitive region by different ions over a range of incident energies. The vertical scale on the right denotes the number of electrons produced. As a circuit becomes smaller and more complex, less deposited energy (charge) is needed to trigger errors.

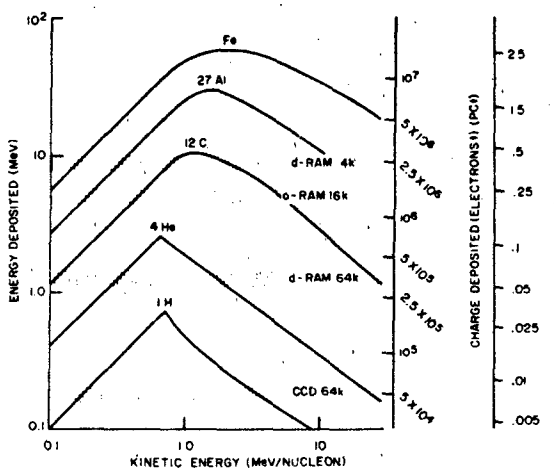


Figure 5-67. Energy deposited in 10 μm of silicon by different ions. The scale on the right shows the number of free electrons released. The low energy part of the curve occurs when the ion penetrating range is less than 10 μm so that it deposits all its energy [E. Petersen, personal communication, 1982].

5.8.3 Control System Failure

Radiation induced errors in electronic circuitry can be particularly damaging when they occur in critical circuitry such as control systems or in decision making logic. While other non-critical circuits may continue to function with false information, control systems can latch-up, that is, be switched into an undesired mode from which there may be no reset option, or the spacecraft may be damaged. Certain circuitry switching may cause burnout of electrical systems or even worse effects, particularly when propulsion, attitude, or weapons systems may be involved. For these reasons it is imperative that proper safeguards and redundancy design be considered in the early stages of spacecraft engineering.

5.8.4 Biological Effects

There is extensive literature on space biology [see for example Bacq and Alexander, 1961]. Here we shall only point out that the quiet time radiation belts at some locations present a lethal radiation dosage to a man in a space suit or even within a vehicle. But even outside the main trapped radiation zone, there are intermittent high fluxes of solar energetic particles. For example, it is believed that the energetic particle fluxes associated with the August 1972 solar flare magnetic storm event would have been extremely harmful to humans almost anywhere in space. The method presented above can also be used to estimate human radiation exposure behind different shielding designs. A definitely lethal dosage is about 500 rads [Desrosier and Rosenstock, 1960].

5.9 MAN'S IMPACT ON THE RADIATION BELTS

The activity of mankind can, to a significant degree, influence the earth's radiation environment. Examples are nuclear detonations (fission and fusion), accelerator particle beams (neutral and charged), release of chemical substances, injection of metallic powders, and electromagnetic wave energy production. The effects of some of these modification sources have not yet been studied, but for others a substantial body of data is available.

5.9.1 Nuclear Detonations

A vast number of free neutrons and other particles and a great pulse of electromagnetic energy are released in nuclear explosions. In the nuclear fission process, the fission fragments also carry significant kinetic energy. The product of nuclear fusion is generally a stable particle (He) which may be ionized. As a rule of thumb $\sim 10^{26}$ fast neutrons (each of which decay into a proton-electron-neutrino triplet) are released per megaton nuclear explosive yield.

The size of the nuclear fireball depends not only on the explosive yield, but also on the medium in which the detonation occurs. In field-free empty space the fireball will expand without limits, but in the presence of material substances or a magnetic field the fireball is effectively restrained. In a dense gas (such as below ~100 km in the earth's atmosphere), collisions between the explosion products and the atmospheric constituents dissipate much of the detonation energy as heat. About half of this energy is radiated away and the thermalized remainder is typically at 6000 K to 8000 K [Zinn et al., 1966]. At an altitude of 60 km in the earth's atmosphere, a 1 megaton fissional detonation will have a fireball radius of ~4 km, and for the same nuclear explosive yield this radius will be smaller close to the ground. The fireball itself may accelerate to velocities of several km sec due to buoyancy and shock processes. High altitude and space detonations (more than 100 km above the earth) have the fireball size limited by the magnetic field. This occurs because the explosion generates electrically charged fragments that are susceptible to the magnetic force, $q\mathbf{V} \times \mathbf{B}$, where q is the particle charge and V its velocity. A nuclear detonation of 1 megaton can have fireball expansion to ~1000 km across the magnetic field when the B-field has a value $B \approx 0.5$ G. The expansion is not magnetically limited along the field lines.

Depending on the location of the nuclear detonations, a certain fraction of the neutrons will decay within the magnetic field trapping region, and the decay products will thus constitute artificially created trapped radiation. For nuclear fission, the fission fragments also emit particles (such as electrons and α-particles) before reaching a nucleonic configuration as a stable isotope. This process further contributes to the trapped radiation, and the characteristic electron energy is 1 to 8 MeV from this source.

Studies of nuclear detonation effects have shown that even small high altitude explosions (in the kiloton range) affect the radiation belts considerably. See for example reviews by Hess [1968] and Walt [1977]. Table 5-4 gives an overview of the known radiation belt effects of the Teak, Orange, Argus 1, Argus 2, Argus 3, and Starfish nuclear detonations carried out at high altitudes by the United States, and the USSR-1, USSR-2 and USSR-3 high altitude nuclear detonations by the Soviet Union.

Energetic charged particles exiting the upper atmosphere along the geomagnetic lines of force are generally within the atmospheric bounce loss cone. In the absence of significant pitch angle scattering, such particles will follow the field lines and precipitate into the conjugate hemisphere. Empirically, a significant fraction of the nuclear detonation particles become trapped in the radiation belts. This implies that significant pitch angle scattering must take place from the angular source cone region ($\alpha_0 < \alpha_{0LC}$) to stably trapped particle orbits ($\alpha_0 > \alpha_{0LC}$). This pitch angle scattering must take place on the time scale of a single half-bounce period $\tau_b/2$ (which is of the order of seconds).

Following such an artificial injection of particles into the radiation belts, the normal radiation belt radial and pitch angle diffusion mechanisms will operate. The initial narrow injected radial distribution will broaden, and the charge exchange (for ions) and Coulomb energy degradation mechanisms will modify the characteristics of the injected distributions. Depending on the location, the artificial radiation belts may last for days or years [Walt and Newkirk, 1966; Stassinopoulos and Verzariu, 1971].

The effects of accelerator beams are likely to be similar to those of the nuclear detonation particles, but the yield (in terms of number of particles) is likely to be much smaller. On the other hand, since the beam particles may be generated over a wide range of energies

Table 5-4. Listing of high altitude nuclear detonations between 1958 and 1962 [Walt, 1977].

Event	Altitude (km)	Time (UT)	Date	Latitude	Longitude	Approximate L-Value Detonation	Yield	Characteristics of Band	Approximate Decay Time
Teak	76.8	10:50:05	01AUG58	17° N	169° W	1.12	MT Range	Low Altitude	few days
Orange	42.97	10:30:08	12AUG58	17° N	169° W	1.12	MT Range	Low Altitude	1 day
Argus 1	200	2:30:00	27AUG58	38° S	12° W	1.7	1-2 KT	Narrow Band	0-20 days
Argus 2	250	3:20:00	30AUG58	50° S	8° W	2.1	1-2 KT	Narrow Band	10-20 days
Argus 3	500	22:10:00	06SEP58	50° S	10° W	2.0	1-2 KT	Narrow Band	10-20 days
Starfish	400	09:00:29	09JUL62	16.7° N	190.5° E	1.12	1.4 MT	Wide Distribution	1-2 yrs
USSR 1		03:40:46	22OCT62			1.8		Wide Distribution	30 days
USSR 2		04:41:18	28OCT62			1.8		Wide Distribution	30 days
USSR 3		09:13	01NOV62			1.75		Narrow Band	30 days

(thermal to relativistic) a more precise study of their effects is warranted.

5.9.2 Release of Chemicals

Chemical releases for research purposes have been carried out at high altitudes. In most cases barium or lithium was released to trace magnetic field lines locally and to assess the magnitude of electric fields and upper atmosphere winds. Chemical releases into outer regions of geospace are also planned. Such programs may modify the environment locally (for example, by altering plasma wave dispersion characteristics), but are not expected to impact the radiation belts seriously unless large quantities of chemicals are used.

Extensive operations with rocket propulsion or special ion engines could, however, drastically alter the different particle populations and could lead to profound changes in the radiation belt structure. See Chapter 7. To date no comprehensive environmental impact analysis has been carried out.

5.9.3 Transmission of Radio Waves

It has been suggested that electromagnetic wave energy from tropospheric thunderstorm activity and whistler-mode waves from VLF radio transmitters can perturb the energetic electron component of the earth's radiation belts. Correlative studies indicate that energetic electron precipitation not only occurs from natural sources [Spjeldvik and Lyons, 1979] but also in correlation with strong terrestrial radio transmitter operations [Vampola and Kuck, 1978; Park et al., 1981; Imhof et al., 1981; Chang and Inan, 1983]. Precisely to what extent man's electromagnetic wave generation influences the overall radiation belt structure is not known, however.

5.9.4 Effects of Space Structures

Proposed operations of large manmade metallic and electrically insulated space structures will produce local "singular" regions in the magnetosphere. Associated with space shuttles, space platforms, or space power arrays will be a hydromagnetic wake in which the wave and particle behavior will go through a sudden change. It is not known whether or not these cavity phenomena may have a significant effect on the radiation belts themselves. For some details see Garrett and Pike [1980] and references therein.

CHAPTER 5

REFERENCES

Adams, J.H., Jr. and K. Patridge, "Do Trapped Heavy Ions Cause Soft Upsets on Spacecraft?" National Research Laboratory Memo. Rept. 4846, October 12, 1982.

Adams, J.H., Jr., R. Silverberg, and C.H. Tsao, "Cosmic Ray Effects on Microelectronics, Part I: The Near-Earth Particle Environment," National Research Laboratory Memo. Rept. 4506, August 25, 1981.

Akasofu, S.-I., *Polar and Magnetospheric Substorms*; D. Reidel, Dordrecht, Holland, 1968.

Alfvén, H. and C.G. Fälthammar, *Cosmical Electrodynamics*, Clarendon Press, Oxford, UK, 1963.

Axford, W.I., "Magnetospheric Convection," *Rev. Geophys.*, **1:** 421, 1969.

Bacq, Z.M. and P. Alexander, *Fundamentals of Radiology*, Pergamon Press, New York, 1961.

Baker, D.N., P.R. Higbie, E.W. Hones, Jr., and R.D. Belian, "High Resolution Energetic Particle Measurements at 6.6 R_E 3. Low Energy Anisotropies and Short-Term Substorm Predictions," *J. Geophys. Res.*, **83:** 4863, 1978.

Baker, D.N., P.R. Higbie, and R.D. Belian, "Multispacecraft Observations of Energetic Electron Flux Pulsations at 6.6 R_E," *J. Geophys. Res.*, **85:** 6709, 1980.

Baker, D.N. et al., "The Los Alamos Geostationary Orbit Synoptic Data Set," Los Alamos National Laboratory Report LA-8843, UC-34G, Los Alamos, N.M., August 1981.

Belian, R.D., D.N. Baker, P.R. Higbie, and E.W. Hones, Jr., "High-Resolution Energetic Particle Measurements at 6.6 R_E 2. High Energy Proton Drift Echoes," *J. Geophys. Res.*, **83:** 4857, 1978.

Belian, R.D., D.N. Baker, E.W. Hones, Jr., P.R. Higbie, S.J. Bame, and J.R. Asbridge, "Timing of Energetic Proton Enhancements Relative to Magnetospheric Substorm Activity and Its Implication for Substorm Theories," *J. Geophys. Res.*, **86:** 1415, 1981.

Blake, J.B. "Experimental Test to Determine the Origin of Geomagnetically Trapped Radiation," *J. Geophys. Res.*, **78:** 5822, 1973.

Blake, J.B., and J.F. Fennell, "Heavy Ion Measurements in the Synchronous Altitude Region," *Planet. Space Sci.*, **29:** 1205, 1981.

Bostrom, C.O., D.S. Beall, and J.C. Armstrong, "Time History of the Inner Radiation Zone," *Models of the Trapped Radiation Environment, Vol. III: Longterm Variations*, NASA SP-3024, 1971.

Brinkmann, H.C. and H.A. Kramers, "Zur Theorie der Einfangung von Elektronen Durch α-Teilchen," *Proc. Akad Wetensch. Amsterdam Afd. Natuurk.* **33:** 973, 1930.

Chandrasekhar, S., "Plasma Physics" (compiled by S.K. Trehan), *Phoenix Science Series*, University of Chicago Press, Chicago, 1965.

Chang, H.C. and U.S. Inan, "Quasi-Relativistic Electron Precipitation Due to Interactions With Coherent VLF Waves in the Magnetosphere," *J. Geophys. Res.*, **88:** 318, 1983.

Chapman, S. and J. Bartels, *Geomagnetism*, Oxford University Press, London, 1940.

Claflin, E.S., "Change Exchange Cross Sections for Hydrogen and Helium Ions Incident on Atomic Hydrogen: 1 to 1000 keV," Rep Tr-0059 (6260-20)-1, Aerospace Corp., El Segundo, Calif., 1970.

Cornwall, J.M., "Diffusion Processes Influenced by Conjugate Point Wave Phenomena," *Radio Sci.*, **3:** 740, 1968.

Cornwall, J.M., "Radial Diffusion of Ionized Helium and Protons: A Probe for Magnetospheric Dynamics," *J. Geophys. Res.* **77,** 1756, 1972.

Davidson, G.T., "The Motion of Charged Particles in the Earth's Magnetic Field," in *The Trapped Radiation Handbook*, edited by J.B. Cladis, G.T. Davidson, and L.L. Newkirk, Lockheed Palo Alto Research Laboratory, DNA 2524IH, 1977.

Desrosier, N.W. and H.M. Rosenstock, *Radiation Technology in Food, Agriculture and Biology*, Avi, Westport, Conn., 1960.

Dessler, A.J. and E.N. Parker, "Hydromagnetic Theory of Geomagnetic Storms," *J. Geophys. Res.*, **64:** 2239, 1959.

Fälthammar, C.G., "Radial Diffusion by Violation of the Third Adiabatic Invariant," in *Earth's Particles and Fields*, edited by B.M. McCormac, Reinhold, New York, 1968.

Fennell, J.F., D.R. Croley, Jr., and S.M. Kaye, "Low-Energy Ion Pitch Angle Distributions in the Outer Magnetosphere: Ion Zipper Distribution," *J. Geophys. Res.*, **86:** 3375, 1981.

Fite, W.L., R.F. Stebbings, D.J. Hummer, and R.T. Brackman, "Transfer in Proton-Hydrogen Atomic Collisions," *Phys. Rev.*, **119:** 663, 1960.

Fritz, T.A. and W.N. Spjeldvik, "Simultaneous Quiet Time Observations of Energetic Radiation Belt Protons and Helium Ions: The Equatorial α/p Ratio Near 1 MeV," *J. Geophys. Res.*, **84:** 2608, 1979.

Fritz, T.A. and B. Wilken, "Substorm Generated Fluxes of Heavy ions at the Geostationary Orbit," in *Magnetospheric Particles and Fields*, edited by B.M. McCormac, D. Reidel, Dordrecht, Holland, 1976.

Garrett, H.B., "Quantitative Models of the 0 to 100 keV Mid-Magnetospheric Particle Environment," in *Quantitative Modeling of Magnetospheric Processes*, edited by W.P. Olson, AGU, Washington, D.C., 1979.

Garrett, H.B. and C.P. Pike, (eds.) *Space Systems and Their Interactions with Earth's Space Environment*, AIAA Press, New York, 1980.

Hamlin, D.A., R. Karplus, R.C. Vik, and K.M. Watson; "Mirror and Azimuthal Drift Frequencies for Geomagnetically Trapped Particles," *J. Geophys. Res.*, **66:** 1, 1961.

Harel, M., R.A Wolf, P.H. Reiff, R.W. Spiro, W.J. Burke, F.J. Rich, and M. Smiddy, "Quantitative Simulation of a Magnetospheric Substorm 1. Model logic and overview," *J. Geophys. Res.*, **86:** 2217, 1981a.

Harel, M., R.A Wolf, R.W. Spiro, P.H. Reiff, C.K. Chen, W.J. Burke, F.J. Rich, and M. Smiddy, "Quantitative Simulation of a Magnetospheric Substorm 2. Comparison With Observations," *J. Geophys. Res.*, **86:** 2242, 1981b.

Hess, W. N., *The Radiation Belt and Magnetosphere*, Blaisdell, Waltham, Mass., 1968.

Higbie, P.R., R.D. Belian, and D.N. Baker, "High-Resolution Energetic Particle Measurements at 6.6 R_E 1. Electron Micropulsations," *J. Geophys. Res.*, **83:** 4851, 1978.

Hovestadt, D., G. Gloeckler, C.Y. Fan, L.A. Fisk, F.M. Ipavich, B. Klecker, J.J. O'Gallagher, and M. Scholer, "Evidence for Solar Wind Origin of Energetic Ions in Earth's Radiation Belts," *J. Geophys. Res. Letts.*, **5:** 1055, 1978.

Hovestadt, D., B. Klecker, E. Mitchell, J.F. Fennell, G. Gloeckler, and C.Y. Fan, "Spatial Distribution of $Z \geq 2$ Ions in the Outer Radiation Belt During Quiet Conditions," *Adv. Space Res.*, **1:** 305, 1981.

Imhof, W.L., E.E. Gaines, and J.B. Reagan, "High-Resolution Spectral Features Observed in the Inner Radiation Belt Trapped Electron Population," *J. Geophys. Res.*, **86:** 2341, 1981.

Jackson, J.D., *Classical Electrodynamics, 2nd edition, Wiley, New York, 1975.*

Jacobs, J.A., *Geomagnetic Micropulsations*, Springer-Verlag, New York, 1970.

Janni, J. and G. Radke, "The Radiation Environment and Its Effects on Spacecraft in *Quantitative Modeling of Magnetospheric Processes*," edited by W.P. Olson, Geophysical Monograph 21, AGU, Washington, D.C. 1979.

Kaye, S.M., E.G. Shelley, R.D. Sharp, and R.G. Johnson, "Ion Composition of Zipper Events," *J. Geophys.*, **86:** 3383, 1981.

Kennel, C.F. and F. Engelmann, "Velocity Space Diffusion from Weak Plasma Turbulence in a Magnetic Field," *Phys. Fluids*, **9:** 2377, 1966.

Kennel, C.F. and H.E. Petschek, "Limit on Stably Trapped Particle Fluxes," *J. Geophys. Res.*, **71:** 1, 1966.

Kennel, C.F., L.J. Lanzerotti, and E.N. Parker, (eds.) *Solar System Plasma Physics*, Vols. I–III, North Holland, Amsterdam, 1979.

Kivelson, M.G., S.M. Kaye, and D.J. Southwood, "The Physics of Plasma Injection Events," *Dynamics of the Magnetosphere*, edited by S.-I. Akasofu, D. Reidel, Dordrecht, Holland, 1980.

Lanzerotti, L.J., D.C. Webb, and C.W. Arthur, "Geomagnetic Field Fluctuations at Synchronous Orbit 2. Radial Diffusion," *J. Geophys. Res.*, **83:** 3866, 1978.

Lerche, I., "Quasilinear Theory of Resonant Diffusion in a Magneto-Active Relativistic Plasma," *Phys. Fluids*, **11:** 1720, 1968.

Lo, H.H. and W.L. Fite, "Electron-Capture and Loss Cross Sections for Fast, Heavy Particles Passing Through Gases," *Atomic Data*, **1:** 305, 1970.

Lundin, R., L.R. Lyons, and N. Pissarenko, "Observations of the Ring Current Composition at L-Values less than 4," *Geophys. Res. Lett.*, **7:** 425, 1980.

Lyons, L.R., "Plasma Processes in the Earth's Radiation Belts," in *Solar System Plasma Physics*, edited by C.F. Kennel, L.J. Lanzerotti, and E.N. Parker, North Holland, Amsterdam, 1979.

Lyons, L.R. and R.M. Thorne, "The Magnetospheric Reflection of Whistlers," *Planet. Space Sci.*, **18:** 1753, 1970.

Lyons, L.R. and R.M. Thorne, "Equilibrium Structure of Radiation Belt Electrons," *J. Geophys. Res.*, **78:** 2142, 1973.

Lyons, L.R., R.M. Thorne, and C.F. Kennel, "Electron Pitch-Angle Diffusion Driven by Oblique Whistler Mode Turbulence," *J. Plasma Phys.*, **6:** 589, 1971.

Lyons, L.R., R.M. Thorne, and C.F. Kennel, "Pitch Angle Diffusion of Radiation Belt Electrons Within the Plasmasphere," *J. Geophys. Res.*, **77:** 3455, 1972.

Lyons, L.R. and D.J. Williams, "A Source for the Geomagnetic Storm Main Phase Ring Current," *J. Geophys. Res., 85:* 523, 1980.

MacDonald, J.R. and F.W. Martin, "Experimental Electron-Transfer Cross Sections for Collisions of Oxygen Ions in Argon, Nitrogen and Helium at Energies of 7–40 MeV," *Phys. Rev. (Sect. A)*, **4:** 1965, 1971.

McIlwain, C.E., "Coordinates for Mapping the Distribution of Magnetically Trapped Particles," *J. Geophys. Res.* **66:** 3681, 1961.

McIlwain, C.E., "Redistribution of Trapped Protons During a Magnetic Storm," *Space-Research V*, North Holland, Amsterdam, 1965.

McNulty, P.J., R.C. Wyatt, G.E. Farrell, R.C. Filz and P.L. Rothwell, "Proton Upsets in LSI Memories in Space," in Prog. Astronaut. Aeronaut. Vol. 71, *Space Systems and Their Interactions with Earth's Space Environment*, edited by H.B. Garrett and C.P. Pike, AIAA Press, New York, 1980.

Mott, N.F. and H.S.W. Massey, *The Theory of Atomic Collisions*, Clarendon Press, Oxford, UK, 1952.

Mullen, E.G. and M.S. Gussenhoven, "SCATHA Environmental Atlas," AFGL TR-83-0002, ADA131456, 1938.

Park, C.G., C.S. Lin, and G.K. Parks, "A Ground-Satellite Study of Wave-Particle Correlations," *J. Geophys. Res.*, **86:** 37, 1981.

Parker, E.N., "Newtonian Development of the Dynamical Properties of Ionized Gases of Low Density," *Phys. Rev.* **107:** 924, 1957.

Parsignault, D.R., E. Holeman, and R.C. Filz, "Solar Cycle Modulation of the 55-MeV Proton Fluxes at Low Altitudes," *J. Geophys. Res.*, **86:** 11439, 1981.

Paulikas, G.A. and J.B. Blake, "Effects of the Solar Wind on Magnetospheric Dynamics: Energetic Electrons at the Synchronous Orbit," in *Quantitative Modeling of Magnetospheric Processes*, edited by W.P. Olson, AGU, Washington D.C. 1979.

Retterer, J.M., J.R. Jasperse, and T.S. Chang, "A New Approach to Pitch Angle Scattering in the Magnetosphere," *J. Geophys. Res.* **88:** 201, 1983.

Roederer, J.G., *Dynamics of Geomagnetically Trapped Radiation*, Springer-Verlag, New York, 1970.

Rossi, B. and S. Olbert, *Introduction to the Physics of Space*, McGraw-Hill, New York, 1970.

Rothwell, P.L. and L. Katz, "Enhancement of 0.24–0.96 MeV Trapped Protons During May 25, 1967 Magnetic Storm," *J. Geophys. Res.*, **78:** 5490, 1973.

Sagdeev, R.Z. and A.A. Galeev, *Nonlinear Plasma Theory*, W.A. Benjamin, New York, 1969.

Sawyer, D.M. and J.I. Vette, "AP-8 Trapped Proton Environment for Solar Maximum and Solar Minimum," NSSDC/WDC-A-R&S 76-06, NASA-GSFC TMS-72605, December 1976.

Schulz, M., "Geomagnetically Trapped Radiation," *Space Sci. Rev.*, **17:** 481, 1975a.

Schulz, M., "Particle Saturation of the Outer Zone, A Non-Linear Model," *Astrophys. Space Phys.*, **29:** 233, 1975b.

Schulz, M. and L.J. Lanzerotti, *Particle Diffusion in the Radiation Belts*, Springer-Verlag, New York, 1974.

Schulz, M. and G.A. Paulikas, "Secular Magnetic Variation and the Inner Proton Belt," *J. Geophys. Res.*, **77:** 744, 1972.

Sharp, R.D., R.G. Johnson, and E.G. Shelly, "The Morphology of Energetic O^+ Ions During Two Magnetic Storms, Temporal Variations," *J. Geophys. Res.*, **81:** 3283, 1976.

Shelley, E.G., R.G. Johnson, and R.D. Sharp, "Morphology of Energetic O^+ in the Magnetosphere," in *Magnetospheric Physics*, edited by B.M. McCormac, D. Reidel, Dordrecht, Holland, 1974.

Singley, G.W. and J.I. Vette, "The AE-4 Model of the Outer Radiation Zone Electron Environment," NSSDC 72-06, 1972.

Solov'ev, E.S., R.N. Il'in, V.A. Oparin, I.T. Serenkov, and N.V. Federenko, "Capture and Loss of Electrons by Fast Nitrogen and Oxygen Atoms and Ions in Air, Nitrogen and Oxygen," *Sov. Phys.-Tech. Phys.*, **37:** 777, 1975.

Söraas, F. and L.R. Davis, "Temporal Variations of the 10 keV to 1700 keV Trapped Protons Observed on Satellite Explorer 26 During First Half of 1965," NASA TMX-63320, GSFC, August 1968.

Spjeldvik, W.N., "Equilibrium Structure of Equatorially Mirroring Radiation Belt Protons," *J. Geophys. Res.*, **82:** 2801, 1977.

Spjeldvik, W.N., "Expected Charge States of Energetic Ions in the Magnetosphere," *Space Sci. Rev.* **23:** 199, 1979.

Spjeldvik, W.N. and T.A. Fritz, "Theory for Charge States of Energetic Oxygen Ions in the Earth's Radiation Belts," *J. Geophys. Res.*, **83:** 1583, 1978a.

Spjeldvik, W.N. and T.A. Fritz, "Energetic Ionized Helium in the Quiet Time Radiation Belts: Theory and Comparison with Observation," *J. Geophys. Res.*, **83:** 654, 1978b.

Spjeldvik, W.N. and T.A. Fritz, "Observations of Energetic Helium Ions in the Earth's Radiation Belts During a Sequence of Geomagnetic Storms," *J. Geophys. Res.*, **86:** 2317, 1981a.

Spjeldvik, W.N. and T.A. Fritz, "Energetic Heavy Ions with Nuclear Charge $Z \geq 4$ in the Equatorial Radiation Belts of the Earth: Magnetic Storms," *J. Geophys. Res.*, **86:** 2349, 1981b.

Spjeldvik, W.N. and L.R. Lyons, "On the Predictability of Radiation Belt Electron Precipitation into the Earth's Atmosphere Following Magnetic Storms," *Proceedings of the International Symposium on Geophysical Predictions, IV*, B-59, U.S. Department of Commerce, Boulder, 1980.

Spjeldvik, W.N. and R.M. Thorne, "The Cause of Storm After Effects in the Middle Latitude D-Region," *J. Atmos. Terr. Phys.*, **37:** 777, 1975.

Stassinopoulos, E.G. and P. Verzariv, "General Formula for Decay Lifetimes of Starfish Electrons," *J. Geophys. Res.*, **76:** 1841, 1971.

Stebbings, R.F., W.L. Fite, and D.G. Hummer, "Charge Transfer Between Atomic Hydrogen and N^+ and O^+," *J. Chem. Phys.*, **33:** 1226, 1960.

Stix, T.H., *The Theory of Plasma Waves*, McGraw-Hill, New York, 1962.

Stone, E.C., "Physical Significance and Application of L, B_o and R_o to Geomagnetically Trapped Particles," *J. Geophys. Res.*, **68:** 4157, 1963.

Stormer, L., *Polar Aurora*, Clarendon Press, Oxford, 1955.

Teague, M.J., K.W. Chan, and J.I. Vette, "AE-6: A Model Environment of Trapped Electrons for Solar Maximum," NSSDC/WDC-A-R&S 76-04, May 1976.

Teague M.J., N.J. Schofield, K.W. Chan, and J.I. Vette, "A Study of Inner Zone Electron Data and Their Comparison with Trapped Radiation Models," NSSDC/GSFC, August 1979.

Teague, M.J. and E.G. Stassinopoulos, "A Model of the Starfish Flux in the Inner Radiation Zone," NASA-GSFC TMX-66211, December 1972.

Teague, M.J. and J.I. Vette, "The Inner Zone Electron Model AE-5," NSSDC/WDC-A-R&S 72-10, November 1972.

Teague, M.J. and J.I. Vette, "A Model of the Trapped Electron Population for Solar Minimum," NSSDC/WDC-A-R&S-74-03, NASA, April 1974.

Theodoridis, G.C., "Bimodal Diffusion in the Earth's Magnetosphere: 1. An Acceleration Mechanism for Trapped Particles," *Ann. Geophys.*, **24:** 944, 1968.

Tinsley, B.A., "Evidence that the Recovery Phase Ring Current Consists of Helium Ions," *J. Geophys. Res.*, **81:** 6193, 1976.

Vampola, A.L., "A New Study of the Outer Zone Electron Environment, A Hazard to CMOS," SAMSO-TR-77-127, June 1977.

Vampola, A.L. and G.A. Kuck, "Induced Precipitation of Inner Zone Electrons," *J. Geophys. Res.*, **83:** 2542, 1978.

Vette, J., I. King, W. Chan, and M.J. Teague, "Problems in Modelling the Earth's Trapped Radiation Environment," AFGL-TR-78-0130, ADA059273, 1978.

Vette, J.I., M.J. Teague, D.M. Sawyer, and K.W. Chan, "Modelling the Earth's Radiation Belts" in *Solar-Terrestrial Prediction Proceedings, Vol. 2*, edited by R.F. Donnelly, U.S. Department of Commerce, NOAA, Boulder, 1979.

Walt, M., "History of Artificial Radiation Belts" in *The Trapped Radiation Handbook*, edited by J.B. Cladis, G.T. Davidson, and L.L. Newkirk, Lockheed Palo Alto Research Laboratory, DNA 2524H, Revised January, 1977.

Walt, M. and L.L. Newkirk, "Addition to Investigation of the Decay of the Starfish Radiation Belt," *J. Geophys. Res.* **71:** 1966.

West, H.I., Jr., R.M. Buck, and G.T. Davidson, "The Dynamics of Energetic Electrons in the Earth's Outer

Radiation Belt During 1968 as Observed by the Lawrence Livermore National Laboratory's Spectrometer OGO-5," *J. Geophys. Res.*, **86:** 2111, 1981.

West, H.I., Jr., "The Signatures of the Various Regions of the Outer Magnetosphere in the Pitch Angle Distributions of Energetic Particles," in *Quantitative Modelling of Magnetospheric Processes*, edited by W.P. Olson, AGU, Washington D.C. 1979.

Williams, D.J., "An Overview of Radiation Belt Dynamics" in *Proceedings of the Air Force Geophysics Laboratory Workshop on the Earth's Radiation Belts: January 21–27, 1981*, edited by R.C. Sagalyn, W.N. Spjeldvik, and W.J. Burke, AFGL TR-81-0311, ADA113959, 1981.

Williams, D.J., "The Earth's Ring Current: Causes Generation and Decay," *Space Sci. Rev.* **34:** 223, 1983.

Young, D.T., "Ion Composition Measurements in Magnetospheric Modelling" in *Quantitative Modelling of Magnetospheric Processes*, edited by W.P. Olson, AGU Washington D.C. 1979.

Zinn, J., H. Hoerline, and A.G. Petschek, "The Motion of Bomb Debris Following the Starfish Test," in *Radiation Trapped in the Earth's Magnetic Field*, D. Reidel, Dordrecht, Holland, 1966.

Chapter 6

GALACTIC COSMIC RADIATION AND SOLAR ENERGETIC PARTICLES

D.F. Smart and M.A. Shea

6.1 NOMENCLATURE AND DEFINITIONS

Energetic charged particle radiation is usually characterized by the energy of the particle in electron volts. An electron volt is the kinetic energy a charged particle gains by being accelerated through a potential difference of one volt. The amount of energy in one electron volt is rather small; particle energies are usually given in units of keV (10^3 eV), MeV (10^6 eV), GeV (10^9 eV), and for extremely high energies, TeV (10^{12} eV). (An electron volt is equal to 1.602×10^{-12} erg.) The total energy of a nucleon E_T is the sum of the kinetic energy E plus rest-mass energy $m_o c^2$

$$E_T = E + m_o c^2 = mc^2, \tag{6.1}$$

where m_o is the rest mass, m the relativistic mass, and c the speed of light in vacuo. The rest mass energy of a proton $m_o c^2$ is equal to 938.232 MeV, the rest mass of a neutron is 939.526 MeV, and the rest mass of an electron is equal to 510.984 keV. It is customary in the literature to give values of the kinetic energy per nucleon for heavier particles.

The total energy is related to the particle momentum p by

$$E_T = [p^2c^2 + m_o^2c^4]^{1/2}. \tag{6.2}$$

In the absence of electric fields and of time-varying magnetic fields, E_T and p are constants of the particle's motion.

Observations of the intensity of corpuscular radiation are reported in various units depending upon the detection method. The symbol "J" is normally used to designate the flux, but the precise units must always be specified to avoid ambiguity. There are a number of ways to specify particle flux. The unidirectional differential intensity J(E) is the flux (number per unit time) of particles of a given energy per unit energy interval in a unit solid angle about the direction of observation, incident on a unit area perpendicular to the direction of observation; the units are usually particles $cm^{-2}s^{-1}sr^{-1}$ MeV $^{-1}$. Unidirectional integral intensity J(>E) is the intensity of particles with energy greater than a threshold energy E

$$J(>E) = \int_E^\infty J(E)dE, \tag{6.3}$$

with units of particles $cm^{-2}s^{-1}sr^{-1}$. Omnidirectional intensities are J(E) or J(>E) integrated over the 4π steradians solid angle, with units of particles $cm^{-2}s^{-1}MeV^{-1}$ or particles $cm^{-2}s^{-1}$ respectively.

A graph of the differential energy spectrum gives J(E) plotted against E; a graph of the integral energy spectrum gives J(>E) plotted against E. The differential energy spectrum J(E) and the integral energy spectrum J(>E) are commonly expressed in two ways, either as a power law in energy,

$$J(E) = J_o E^{-\gamma}, \tag{6.4}$$

$$J(>E) = J_o \frac{E^{-\gamma+1}}{\gamma - 1}, \tag{6.5}$$

where γ is the differential spectral index, or in the exponential form

$$J(E) = J_o \exp\frac{-E}{E_o} \tag{6.6}$$

$$J(>E) = J_o E_o \exp\frac{-E}{E_o} \tag{6.7}$$

where E_o is defined as the characteristic e-folding particle energy, that is, the value over which the flux will fall to 1/e of its previous value. There is a considerable hazard in extrapolating any spectral form beyond the range of the data from which it was derived, and many scientists give the range of values for which specific spectral forms are valid.

When the effects of corpuscular radiation on personnel are of primary concern, particle flux intensities or counting rates must be converted to dose or dose rates. The rad is the unit of absorbed energy; one rad is 100 erg absorbed per gram of absorbing material. A rem (roentgen equivalent man) is the dose absorbed that produces the same biological

effect as one rad of x-rays or γ rays. The relation between the rem and the rad is one rem = rad × rbe where rbe is the relative biological effectiveness. Conversion of observed particle fluxes to absorbed dose rate, however, is not straight-forward and is sometimes very difficult requiring extremely detailed information on the particle composition and energy spectrum. The conversion depends in a complex way on the energies and kinds of particles and on geometrical configurations of the absorbers and direction of the incident particles, as well as on the different absorbing properties of materials. Computer codes exist [see, for example, Seltzer, 1980] that will convert incident energetic particle fluences into radiation dose behind a specified shield thickness. See Haffner [1967] for a detailed discussion of radiation dosage and the relative sensitivity of various organs to radiation.

The magnetic rigidity P of a particle is a measure of its resistance to a magnetic force that deflects the particle from a straight-line trajectory. The rigidity, with units of momentum per unit charge, is defined as

$$P = \frac{pc}{q}, \tag{6.8}$$

where q is the charge of the particle. If pc is electron volts, then q is the number of electronic charge units and P is in volts. Convenient units are MV (10^6V) and GV (10^9V).

Conversion from rigidity to energy can be accomplished by solving Equation (6.2) for the desired quantity; however, the conversion can be done conveniently by employing the relativistic parameter γ. Conversion from rigidity to energy in terms of kinetic energy per nucleon is

$$E_A = (\gamma - 1)E_{oA} \tag{6.9}$$

where E_A is the kinetic energy per nucleon, and E_{oA} is the rest mass energy per nucleon. Conversion from kinetic energy per nucleon to rigidity is

$$P = \frac{A}{Z}[(\gamma^2 - 1)^{0.5}]E_{oA} \tag{6.10}$$

where A is the atomic number and Z is the atomic charge.

The relativistic parameter γ can be computed from either the cosmic ray kinetic energy or the cosmic ray rigidity.

$$\gamma = \frac{E_A + E_{oA}}{E_{oA}} \tag{6.11a}$$

or

$$\gamma = \left[\left(\frac{PZ}{A/E_{oA}}\right)^2 + 1\right]^{1/2}. \tag{6.11b}$$

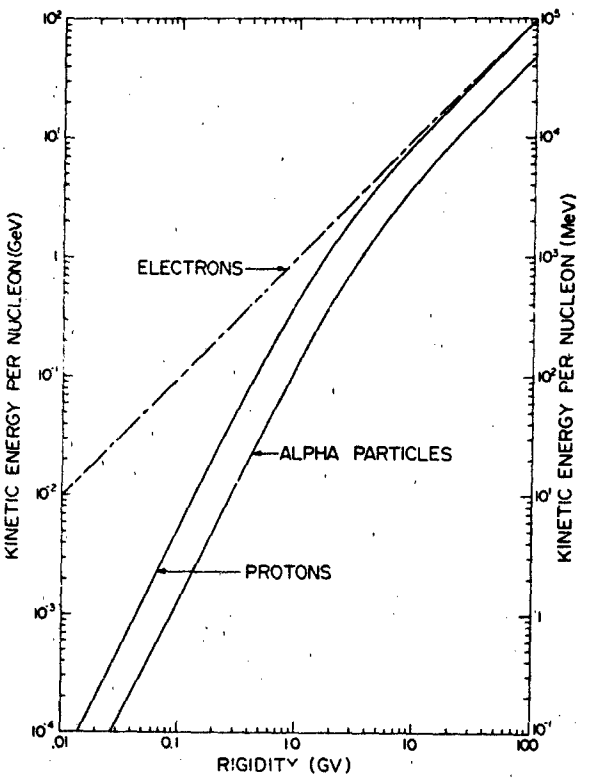


Figure 6-1. Conversion from magnetic rigidity to kinetic energy per nucleon for electrons, protons and alpha particles.

Figure 6-1 shows the relationship between the rigidity of protons, alpha particles, and electrons and their kinetic energy per nucleon. Since most heavy nuclei have an A/Z ratio of approximately two, the alpha particle curve can be used to approximate the rigidity to kinetic energy per nucleon of the heavier nuclei. To obtain the total kinetic energy of a heavy nuclei the energy scale must be multiplied by the total number of nucleons (as an example, by 4 to obtain the total kinetic energy of an alpha particle).

On many occasions the magnetic rigidity is used in place of the kinetic energy in describing the flux spectrum. In particular, in lieu of Equation (6.6), a frequently used expression is

$$J(P) = J_o \exp\frac{-P}{P_o}. \tag{6.12}$$

The earth's magnetic field acts as a momentum analyzer on cosmic rays incident upon the earth's atmosphere. Only those cosmic rays having a momentum per unit charge (that is, rigidity) exceeding that of the threshold of the observation point may be detected at any specified point on the earth's surface or in the earth's magnetosphere. The cutoff or threshold rigidity of cosmic rays is the minimum rigidity that permits

a charged particle to arrive from a specified direction at a given latitude and longitude; particles of lower rigidity are not observed at the specified location and direction because of this geomagnetic cutoff. The geomagnetic field is configured such that particles of progressively lower rigidity are detected as the distance north or south of the equatorial region increases. See Section 6.2.3 for more detailed discussions. More specialized concepts that consider the "bending" of particle trajectories through the magnetic field and allow a mapping of observational directions inside the magnetosphere to directions in interplanetary space are referred to as asymptotic directions and asymptotic cones of acceptance. A more detailed description of these concepts and their use is given by McCracken et al. [1968].

6.2 GALACTIC COSMIC RADIATION

Galactic cosmic radiation is composed of high energy nuclei believed to propagate throughout all space unoccupied by dense matter. Its origin is still a matter of scientific debate and may have both galactic and extragalactic sources. The flux of galactic cosmic radiation is believed to be essentially isotropic outside the heliosphere; propagation effects inside the heliosphere result in an anisotropy of ~ 1%.

Primary cosmic radiation is, by definition, the cosmic radiation incident on the earth's atmosphere. Cosmic rays propagating through the atmosphere undergo nuclear collisions and generate secondary cosmic rays; these secondary cosmic rays consist of all known nuclear and sub-nuclear species.

High energy cosmic ray particles contain a large amount of kinetic energy, and the deposition of this energy can permanently affect the material through which the cosmic ray nucleus passes. In the case of small, state of the art, solid state electronic devices operating in space, the passage of a cosmic ray through a circuit element can generate enough electrons in the sensitive volume to change the state of the circuit element and cause "soft errors" or permanent damage. A recent examination of the effects of cosmic radiation on microelectronics is given by Adams et al. [1981].

6.2.1 Primary Cosmic Radiation

The primary cosmic radiation observed at the earth's orbit consists of approximately 83% protons, 13% alphas, 1% nuclei of atomic number $Z > 2$, and 3% electrons. This composition extends over an energy range from a few hundred MeV to $>10^{20}$ eV. There are no known local planetary sources for the high energy ions observed in the cosmic radiation, but the electron component below about 20 MeV is dominated by Jovian electrons.

The intensity of cosmic rays observed at the earth's orbit is solar cycle dependent undergoing a solar cycle modulation as the inverse of the solar sunspot number cycle. This is discussed more fully in Section 6.2.2.1. The isotropic flux exposure to galactic cosmic radiation in space at sunspot minimum is ~ 4 protons $cm^{-2}s^{-1}$ resulting in a yearly integrated exposure of ~ 1.3×10^8 protons/cm^2. The isotropic flux exposure to galactic cosmic radiation at sunspot maximum is ~ 2 protons $cm^{-2}s^{-1}$ resulting in a yearly integrated exposure of ~ 7×10^7 protons/cm^2. The galactic cosmic radiation converted to integrated dose results in exposures ranging from 4 rads per year at sunspot maximum to 10 rads per year at sunspot minimum [West et al., 1977].

The differential energy spectra of all high energy cosmic ray nuclei above ~ 1 GeV/nucleon exhibit a spectrum proportional to a power law of the form $dJ/dE \propto E^{-\gamma}$, where E is the kinetic energy per nucleon and γ is the spectral index. Below ~ 1 GeV/nucleon, the differential spectrum of cosmic ray nuclei observed at the earth's orbit deviates from a simple power law. The differential spectrum becomes flatter with decreasing energy until a maximum in the differential intensity is reached around a few hundred MeV/nucleon. Below the maximum, the differential intensity decreases monotonically to a few tens of MeV/nucleon. The primary cosmic radiation spectrum observed for protons and alpha particles in the inner heliosphere around the earth's orbit is shown in Figure 6-2. At the vicinity of the earth, the low energy portion of the spectrum changes with time. These changes are mainly the effects of solar modulation, and are illustrated by the shaded and hatched areas of Figure 6-2.

6.2.1.1 Elemental Composition of Primary Cosmic Radiation. Nuclei heavier than helium comprise only about 1% of the total primary cosmic radiation and have a total integral intensity of about 25 particles $m^{-2}s^{-1}sr^{-1}$. Cosmic ray nuclei with $Z > 2$ are classified into various charge groups such as light (L), medium (M), light-heavy (LH), and very heavy (VH). The L-group nuclei include those with $3 < Z \leq 5$, the M-group with $6 \leq Z \leq 8$, the LH-group with $9 \leq Z \leq 14$, the H-group with $15 \leq Z \leq 19$ and the VH-group with $20 \leq Z \leq 28$. The charge group from manganese ($Z = 25$) to nickel ($Z = 28$) is commonly referred to as the iron group.

During their travel from the source regions to the vicinity of the earth, the cosmic rays interact with the interstellar medium, and some fragmentation occurs resulting in a depletion of the heavy charged primary cosmic rays and an increase of the lighter nuclei compared to the "Universal Abundance" shown in Figure 6-3. The relative abundance of the various nuclei observed at the earth for energies greater than 0.45 GeV/nucleon are shown in Table 6-1 [Lezniak and Webber, 1978].

The abundance of the elements N, Na, Al, S, Ar, Ca, Cr, and Mn is enhanced in the cosmic radiation observed in the heliosphere as a result of fragmentation in the intersellar medium. Almost all of the abundance of the elements Li, Be, B, F, Cl, K, Si, Ti, and V is due to fragmentation of heavier cosmic

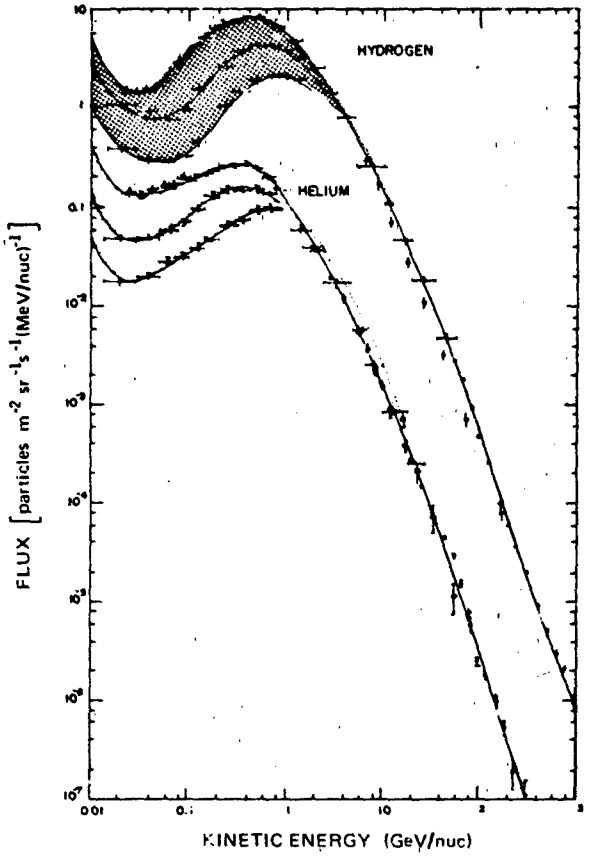


Figure 6-2. Primary cosmic ray differential energy spectrum. The upper envelope indicates the solar minimum spectrum while the lower envelope indicates the solar maximum spectrum. The shaded area indicates the range of the solar modulation over a solar cycle. The hydrogen spectrum in this figure has been multiplied by a factor of 5 so the modulated portion of the spectrum avoids merging with the top of the helium spectrum.

ray nuclei (that is, higher charge number elements) during their propagation through the interstellar medium. The ratios of the cosmic ray elemental abundance to the solar system elemental abundances are generally ordered by atomic parameters such as the first ionization potential, at least up through Z = 40 (as shown in Figure 6-4), with exceptions to this rule at H, He, C, N, and possibly Ne and Mo.

6.2.1.2 The Anomalous Cosmic Ray Component. The low energy portion of the cosmic ray spectrum is quite variable reflecting its dependance on solar modulation. In 1972, during the decline of the 20th solar sunspot cycle, anomalies in the energy spectra and composition were noted at energies ≤70 MeV/nucleon. These differences persisted through sunspot minimum and continued throughout the rising portion of the 21st solar sunspot cycle but seem to have disappeared after the maximum of the 21st solar cycle. This behavior leads to the suggestion that the anomalous

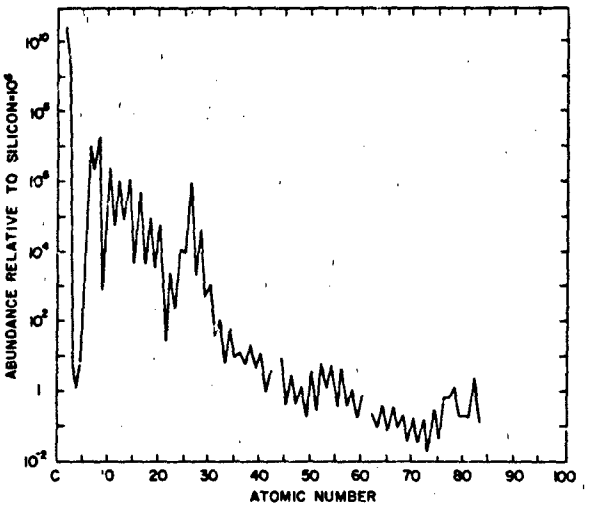


Figure 6-3. The universal abundance of the chemical elements in nature relative to silicon = 10^6. These results were obtained from studies of meteorites, our sun and other stars. [Data from Cameron, 1981.]

Table 6-1. Cosmic-ray composition as observed at 1 AU.

Charge	>450 MeV/Nuc
He	44700 ± 500
Li	192 ± 4
Be	94 ± 2.5
B	329 ± 5
C	1130 ± 12
N	278 ± 5
O	1000
F	24 ± 1.5
Ne	158 ± 3
Na	29 ± 1.5
Mg	203 ± 3
Al	36 ± 1.5
Si	141 ± 3
P	7.5 ± 0.6
S	34 ± 1.5
Cl	9.0 ± 0.6
A	14.2 ± 0.9
K	10.1 ± 0.7
Ca	26 ± 1.3
Sc	6.3 ± 0.6
Ti	14.4 ± 0.9
V	9.5 ± 0.7
Cr	15.1 ± 0.9
Mn	11.6 ± 1.0
Fe	103 ± 2.5
Ni	5.6 ± 0.6

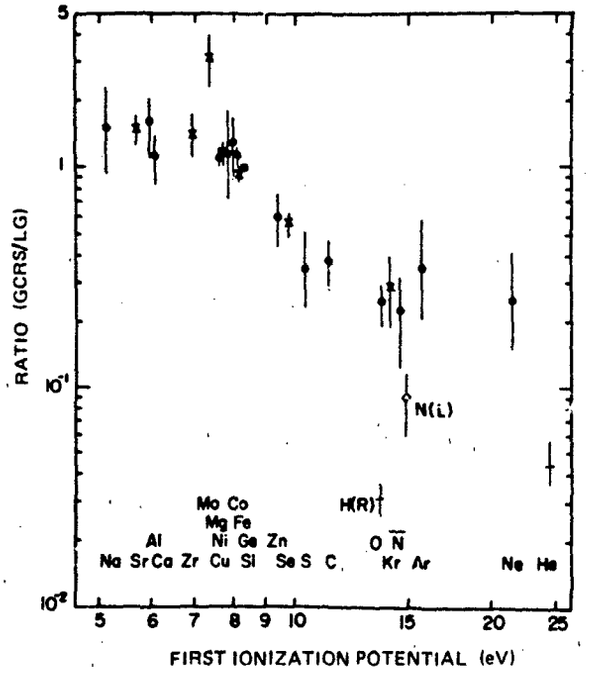


Figure 6-4. Ratio of the galactic cosmic ray source (GCRS) to the "local galactic" (LG) elemental abundances at 1 AU ordered by the first ionization potential [after Mewaldt, 1981].

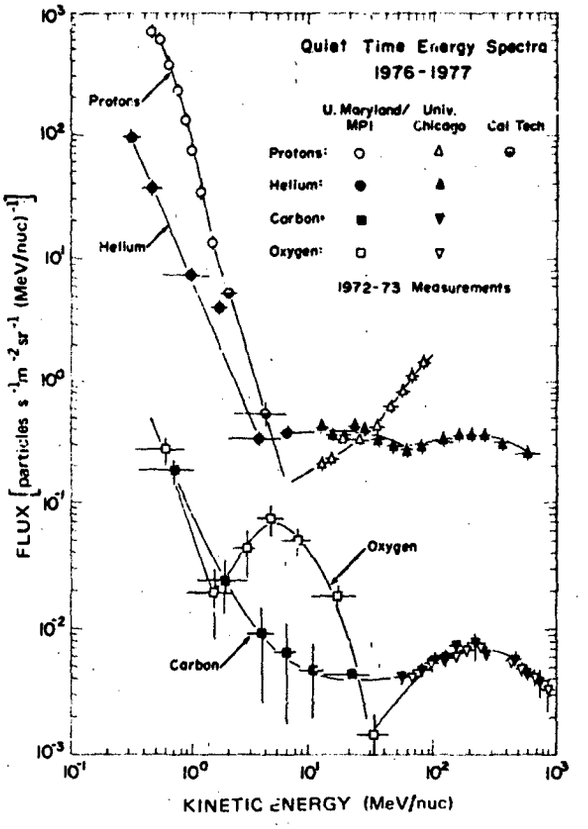


Figure 6-5. The "anomalous" cosmic ray spectrum. The differential energy spectra of hydrogen, helium, carbon, and oxygen observed in the interplanetary medium near 1 AU during the solar minimum in 1976–1977 during quiet-times. The "anomalous cosmic ray" component appears between ~2 and ~30 MeV/nucleon and is characterized by large overabundance of He and O compared to H and C, respectively [Gloeckler, 1979].

component may be solar magnetic cycle dependent and only present during alternate cycles. The anomalous component resulted in the low energy cosmic ray spectra shown in Figure 6-5. The "anomaly" can be seen by comparing this spectra with that of the previous solar minima, the upper envelope of the low energy portion of the spectra shown in Figure 6-2. In a review of this anomalous component, Gloeckler [1979] noted that the helium spectra above ~ 2 MeV/nucleon was relatively flat, and between ~ 5 and 30 MeV/nucleon He was more abundant than protons. The oxygen spectrum possessed an unusual hump between ~ 2 and 20 MeV/nucleon with no comparable feature for carbon, and at ~ 5 MeV/nucleon oxygen was some 10 to 20 times more abundant than carbon. In addition to helium and oxygen, nitrogen and neon were observed to be more abundant in the anomalous component than in the ≥100 MeV/nucleon galactic cosmic rays. The abundances relative to carbon of the major elements in the anomalous component are compared to the composition of galactic cosmic rays in Table 6-2. Assuming that 2–30 MeV/nucleon carbon is of predominately galactic origin during quiet times, the elements He, N, O, Ne, and possibly Fe were from 5 to 20 times more abundant in the anomalous component than in the normal cosmic ray composition. Within statistical uncertainties the composition of the other elements in the anomalous component region were comparable to that expected from the galactic cosmic ray flux.

Table 6-2. Abundances relative to carbon of the anomalous component and galactic cosmic rays.

Element	Anomalous Component (2-30 MeV/nucleon)[a]	Galactic Cosmic Rays (>100 MeV/nucleon)
H	40 ± 4	275
He	90 ± 10	50
C	≡ 1	≡ 1
N	3 ± 1	0.225
O	18 ± 4	1.0
Ne	1.3 ± 0.4	0.175
Mg	0.3 ± 0.2	0.23
Si	$0.2^{+0.25}_{-0.15}$	0.17
Fe	0.6 ± 0.2	0.12

(a) Approximate energy range of the measurements in MeV/nucleon.

6.2.1.3 Cosmic Ray Electrons. The interplanetary cosmic ray energetic electrons of non-solar origin measured at 1 AU has two components, a galactic cosmic ray electron component and a Jovian electron component. All electron flux in the heliosphere below about 25 MeV are of Jovian origin [Eraker and Simpson, 1981]. At the earth's orbit, for energies greater than ~ 100 MeV, the galactic cosmic ray electron flux dominates. A composite primary electron spectrum is shown in Figure 6-6. The solar cycle modulation dominates the variability in the electron spectrum from about 100 MeV to about 5 GeV.

Jupiter's magnetosphere is a source of relativistic electrons that can be observed throughout the entire heliosphere. Within several AU of the planet Jupiter, the flux is modulated with the characteristic 10-hour planetary rotation period. The MeV Jovian electron flux observed at the earth has a 13-month "seasonal" characteristic. This "seasonal" effect results from the relative position of Earth and Jupiter as the planets revolve around the sun in their respective orbits. When the interplanetary magnetic field lines passing near the earth connect to the Jovian magnetosphere (the Jovian magnetosphere being a minimum of 5 AU in length), the propagation of the Jovian electrons from Jupiter to the earth is more efficient. However, the propagation characteristics are still quite variable, depending on the status of the interplanetary medium. The most favorable connection of the interplanetary magnetic field lines between the earth and Jupiter repeats at approximate 13-month intervals with corresponding peaks in the electron flux during the "electron season" which may last for several months. Before the presence of the Jovian electrons was established, these "quiet time electron increases" were unexplained.

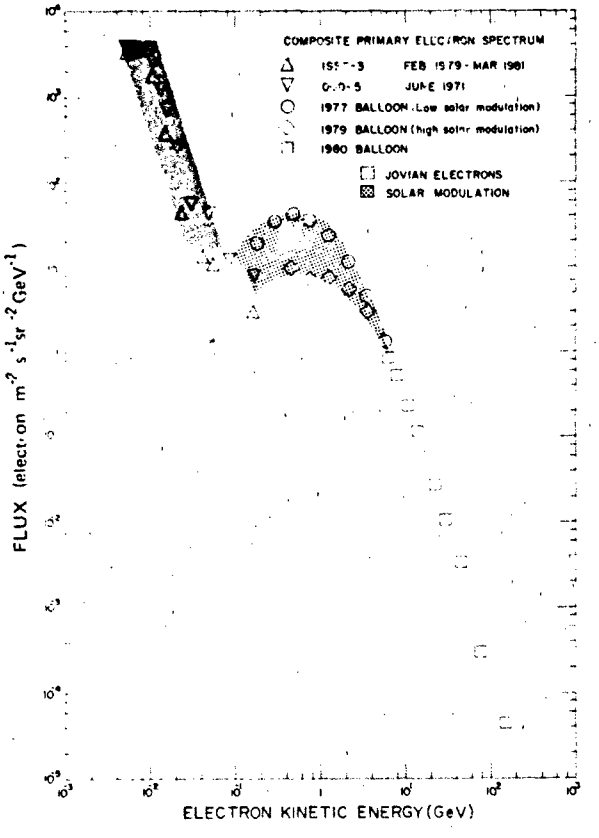


Figure 6-6. Composite primary cosmic-ray electron differential energy spectrum [courtesy F. Meyer, University of Chicago, Private Communication]

6.2.2 Cosmic Ray Flux Modulations

Cosmic ray flux modulations can be divided into three groups. The long term modulation (millions of years), solar cycle modulations (usually referred to as the 11-yr cycle), short term modulations (consisting of cyclic variations ranging from a semi-diurnal variation to a 27-day variation, and transient variations consisting of decreases and increases associated with various solar phenomena).

There is conjecture about variations in the galactic cosmic ray sources over very long time scales. The analysis of cosmic ray exposure to meteorites and lunar rocks is both limited and complicated. So far, statistics and uncertainty in the determinations only allow the conclusion that the current cosmic ray composition has not changed by more than a factor of two over the last 10^8 years. The isotopic composition of the cosmic radiation, specifically the ^{10}Be isotope presumed to be a spallation product of cosmic ray interaction in the interstellar medium, results in a probable cosmic ray age of ~ 15–17 million years.

Cosmogenic isotopes produced by the cosmic ray interactions in air, such as ^{26}Al, ^{10}Be, and ^{14}C, and then frozen in sea sediments, ices or biological materials produce somewhat better limits on possible cosmic ray variations. The ^{10}Be isotope sediments over the last two million years show less than a 30% variation in the cosmic radiation.

The carbon-14 variations are most likely due to the strength of the solar modulation effect. These variations of a few percent have the amplitude and phase expected from historical and recent sunspot data, cosmic ray modulation observations, and CO_2 transport theory [Forman and Schaffer, 1979].

6.2.2.1 Solar Cycle Modulations. The galactic cosmic radiation exhibits an intensity variation with ~ 11-yr periodicity, as shown in Figure 6-7. The cosmic ray intensity changes at the earth are inversely correlated with the sunspot number and lag changes in sunspot number by 9 to 12 months. The increase in solar activity modulates the galactic cosmic rays, through the agency of the solar wind, in such a manner that an increase in solar activity corresponds to a decrease in the cosmic ray intensity. From solar cycle minimum to solar cycle maximum, the energy density of the primary galactic cosmic rays in the vicinity of the earth decreases by about 40%. The cosmic ray intensity, as ob-

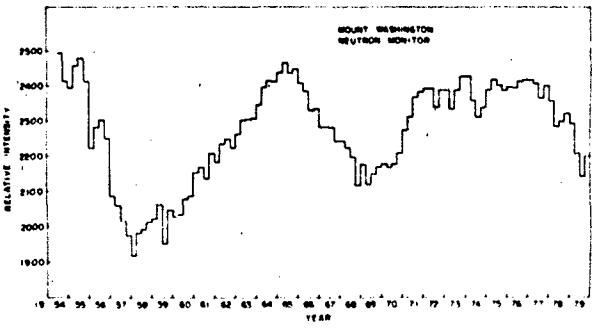


Figure 6-7. Illustration of the ~ 11-yr solar cycle modulation of galactic cosmic rays observed at the earth by the Mt. Washington neutron monitor from 1954 thru 1979.

served at the earth, does not change smoothly from maximum to minimum value. It decreases by a series of sharp drops followed by partial recovery until the minimum intensity is reached near the maximum in solar activity. The recovery to maximum cosmic ray intensities may be either a slow rise, as observed from 1958–1964, or a rapid increase as observed from 1970–1972. These two variations in increasing cosmic ray intensity are also observed in the ionization chamber records going back to 1933, giving rise to a 22-yr solar cycle variation. The historical record of sunspot observations shows longer (80–200 yr) periods modulating the amplitude of the sunspot number cycle. The variations in the carbon-14 record seem to correlate with these same periods. See, for example, Stuiver and Quay [1980].

Theory of Solar Cycle Modulation. There has been considerable theoretical work attempting to model the modulation of cosmic rays by the interplanetary medium. The physical model is based on a solar system filled with an expanding fully ionized and highly conducting plasma, the solar wind (see Chapter 3), which contains frozen-in irregular magnetic fields. Cosmic rays undergo many scatterings from these irregularities and execute a random walk in the solar wind. The cosmic ray population outside the heliosphere diffuses inward, and during this diffusive process it undergoes deceleration by the adiabatic cooling associated with the expansion of the solar wind. The parameters required to define the transport equation and its solution are the diffusion coefficient (generally a function of radius and energy), the solar wind velocity, and the interstellar energy spectrum.

The basis of current solar modulation theory is the Fokker-Planck equation for the modulated number density U(r E) per unit kinetic energy at heliocentric radius r and kinetic energy E. Gleeson and Axford [1967, 1968] have given this equation as

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2VU) - \frac{1}{3r^2}\frac{\partial}{\partial E}(r^2V)\frac{\partial}{\partial E}(\alpha EU) - \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2K\frac{\partial V}{\partial r}\right) = 0 \quad (6.13)$$

where V is the solar wind velocity and K the effective diffusion coefficient in the interplanetary magnetic field irregularities. The diffusion coefficient is generally a function of $\beta = (v/c)$, magnetic rigidity and radial distance. The factor $\alpha = (\gamma + 1)/\gamma$, where γ is the relativistic factor. Numerical techniques for solving Equation (6.13) have been developed by cosmic ray modulation theory specialists. These techniques require the specification of the interstellar density spectrum $U(E_0)$, the functional form of the diffusion coefficient, and appropriate boundary conditions. Numerical solutions of equation (6.13) yield the modulated spectra at 1 AU.

The solar cycle modulation of the cosmic ray spectra is described quantitatively by the "modulation parameter," ϕ. This parameter is defined as

$$\phi = \int_r^{R_1} \frac{V}{3K_1(r)}dr \quad (6.14)$$

where V is the solar wind velocity and $K_1(r)$ is the radial part of the diffusion coefficient. The diffusion coefficient K is usually treated as a separable function of radius and rigidity. In practice the actual value of the cosmic ray modulation parameter at a specific instant of time is difficult to determine. A reference level (such as the counting rate of a stable neutron monitor) is correlated with the historically derived modulation parameters as shown in Figure 6-8, and

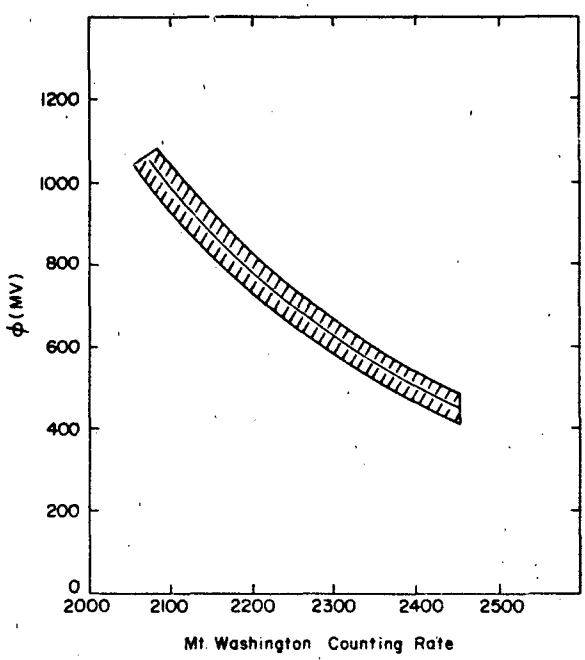


Figure 6-8. Correlation between the cosmic ray modulation parameter ϕ and the counting rate of the Mt. Washington neutron monitor.

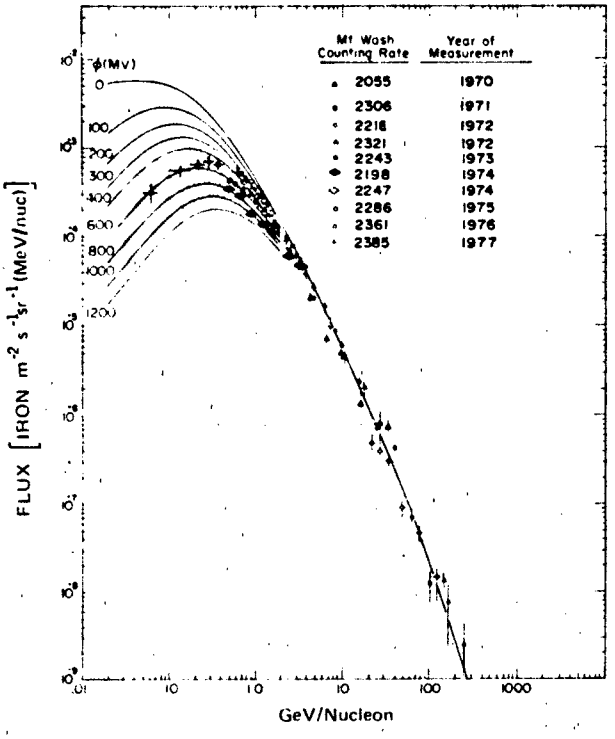


Figure 6-9. Differential energy spectra for cosmic-ray iron nuclei. The shaded area indicates the range of the solar cycle modulation.

the current neutron monitor counting rate is used as a reference modulation level.

Solutions of Equation (6.13) result in an analytic description of the cosmic ray spectrum as a function of kinetic energy and solar modulation parameter.

$$J(E,\phi) = A\frac{[E(E + 2E_{0A})(E + \phi + M)^{\gamma}]}{[(E + \phi)(E + 2E_{0A} + \phi)]^{\gamma}}, \quad (6.15)$$

where E is the kinetic energy of the cosmic ray nuclei in MeV/nucleon, E_{0A} is the rest mass energy of the cosmic ray nuclei in MeV/nucleon and ϕ is the modulation parameter. To model the cosmic ray proton spectrum Garcia-Munoz et al. [1975] present solutions of the cosmic ray spectrum for various modulation levels and elements of the cosmic ray flux. Figure 6-9 shows modulated differential iron spectra for various values of the modulation parameter ϕ. The shaded areas indicate the range of the modulation observed during the past few solar cycles.

6.2.2.2 Short Term Modulations. Short term modulations can be subdivided into two types, cyclic variations and transient variations.

Cyclic Variations. There are ~ 27-day variations in the observed cosmic ray intensity within the heliosphere, at least in the regions explored from 0.3 to 20 AU. These variations are statistical averages associated with the structure of the interplanetary medium corotating with the sun. The amplitude of these variations, as observed by ground-level neutron monitors, is quite variable, normally a few percent, ranging from about 1% to 5%.

For cosmic ray detectors located on the earth there are also diurnal and semi-diurnal variations in phase with the earth's rotation. The amplitude of these diurnal variations is variable ranging from a maximum of a few percent to a minimum of about 0.1%. Figure 6-10 shows an enhanced daily variation continuing for several days. During stable interplanetary conditions the maximum amplitude occurs at ~ 1800 h local solar time. This corresponds to an asymptotic direction roughly along the direction of the sun-earth Archimedean spiral path continuing outward from the earth into interplanetary space.

The semi-diurnal variation is usually less than 0.1% and maximizes in both directions along a line 135° west of the earth-sun line, a direction perpendicular to the mean interplanetary magnetic field direction at the orbit of the earth.

The presence of a sidereal variation reflecting a galactic anisotropy is still not firmly established, and if present would have an extremely small amplitude of a few hundredths of a percent.

Transient Variations. There can be both sudden decreases, and more rarely, increases in the cosmic radiation flux that result from solar activity. Sudden decreases in intensity, called Forbush decreases (named after the original discoverer) are associated with sudden increases in the plasma density and magnetic flux emitted from the sun. The large Forbush decrease events are generally associated with "large"

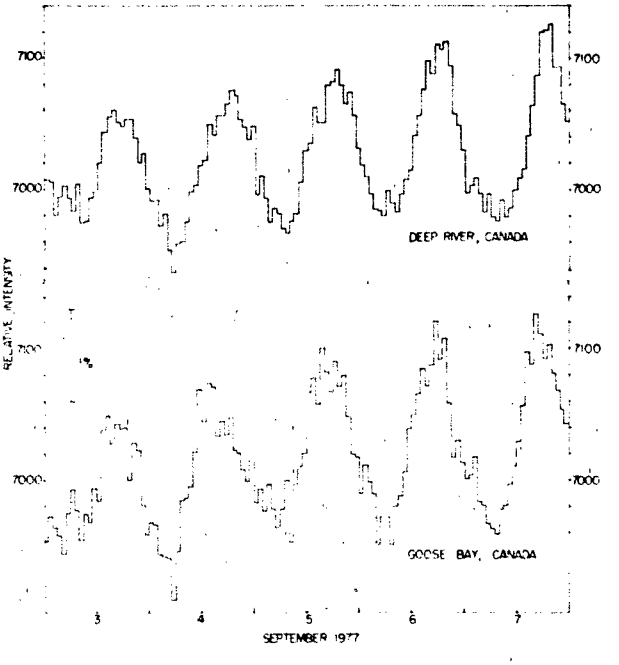


Figure 6-10. Illustration of an enhanced cosmic ray daily variation.

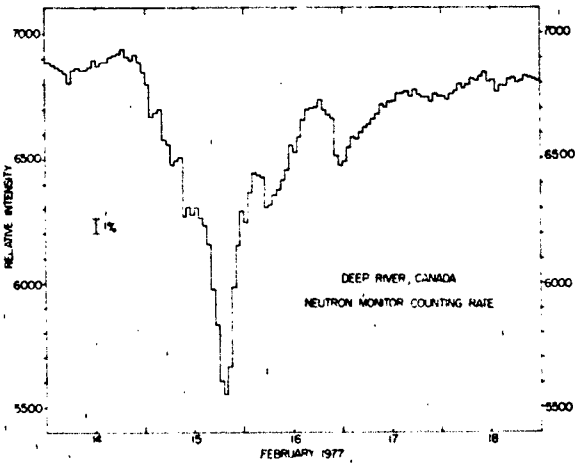


Figure 6-11. Illustration of a cosmic ray Forbush decrease.

solar flares and interplanetary shock structures. The magnitude of these decreases is quite variable ranging from a few percent to as great as 35% in the most extreme case yet recorded by neutron monitors on the earth. An example of a Forbush decrease is shown in Figure 6-11.

There are two types of "cosmic ray" increases attributable to the sun. The solar "cosmic ray" event is discussed in Section 6.3. There are other occasional cases when extremely energetic solar flare generated shocks propagating through the interplanetary medium can accelerate the ambient particle population to higher energies. An example is the 4 August 1972 cosmic ray event thought to be the result of Fermi type acceleration of particles between two propagating interplanetary shock structures [Levy et al., 1976].

6.2.3 Geomagnetic Effects

Cosmic rays being charged nuclei experience a $\mathbf{V} \times \mathbf{B}$ force that continuously alters their path as they propagate through the geomagnetic field. When analyzing cosmic ray observations acquired on the earth to deduce the intensity and anisotropy of the cosmic ray flux in interplanetary space, it is necessary to make allowance for the "magnetic bending" of the cosmic ray trajectory through the magnetosphere. Methods that correct cosmic ray observations for the actual amount of the "geomagnetic bending" each particle has undergone are very specialized and beyond the scope of this handbook. Detailed information on cosmic ray asymptotic directions of approach, asymptotic cones, and variational coefficients is given by McCracken et al. [1965, 1968].

6.2.3.1 Cosmic Ray Cutoff Rigidities. Cosmic ray cutoff rigidities are values that specify the minimum rigidity a charged particle must possess to be observed at a specified position in the geomagnetic field from a specified direction. Rigidity is a canonical coordinate used in cosmic ray studies which is element independent. When translating these to the different elements of the cosmic ray spectrum, a different curve is obtained for each of the elements contained in the cosmic ray flux. If a primary cosmic ray particle is sufficiently rigid (that is, having a rigidity higher than the cutoff rigidity), it can penetrate through the geomagnetic field to an observation point; if the rigidity of the particle is less than the cutoff rigidity, it cannot penetrate through the magnetic field to the specified location from a specified direction. When considering charged particle propagation through the geomagnetic field, it is convenient to use the rigidity (momentum per unit charge) of the particle. To determine the energy per nucleon of a specific cosmic ray element, rigidity to energy tables can be utilized. See Section 6.1 for additional discussion.

Cosmic ray cutoffs are a function of geomagnetic latitude, altitude, and the zenith and azimuthal directions of the incident particle at the detection point. They have a maximum value in the earth's equatorial region and a minimum value in the polar regions. Unfortunately, the cosmic ray cutoff is not a simple value. There is an upper cutoff above which all particles are allowed, a lower cutoff below which all particles are forbidden, and a penumbral region between these two values where the transmission of cosmic

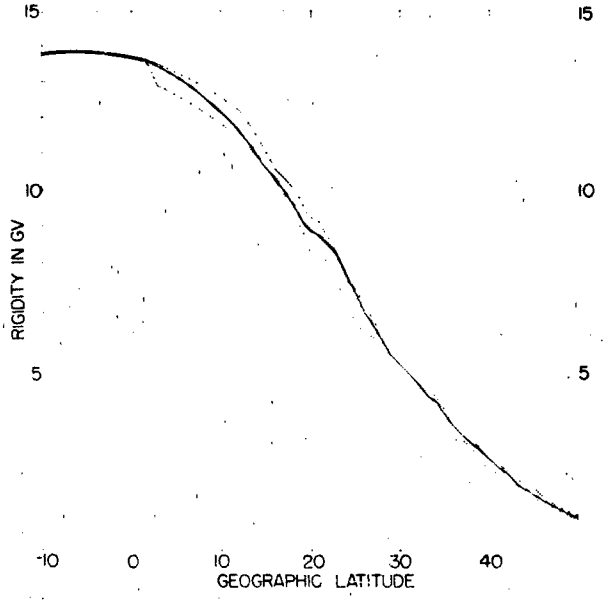


Figure 6-12. Illustration of the width (shaded area) of the cosmic ray vertical cutoff penumbra as a function of latitude along the 260°E meridian. The solid line indicates the effective geomagnetic cutoff rigidity along this meridian. The "sharpness" of the penumbral width variations are exaggerated by only considering the vertical direction; however, the magnitude of the variation is correct.

radiation is chaotic. The cosmic ray penumbra exists near the cosmic ray cutoff rigidity where there is a complex series of allowed and forbidden particle trajectories that has so far defied all attempts to systematically order it. The effective cosmic ray cutoff is a practical value corrected for the transmission through the cosmic ray penumbra. An example of the variation of the vertical cosmic ray cutoff along the 260°E longitude meridian is shown in Figure 6-12. The width of the cosmic ray penumbra shown in this figure is somewhat typical of its worldwide behavior.

The primary reason that precise values of the cosmic ray cutoffs are relatively difficult to obtain is that the equation of charged particle motion in a magnetic field does not have a solution in closed form. Relatively accurate geomagnetic cutoff rigidities can be derived by the numerical integration of cosmic ray trajectories in mathematical models of the geomagnetic field; however, the calculation of these values for a large number of locations and directions involves a formidable amount of computational time. The precision of these calculations is limited only by the accuracy of the geomagnetic field model utilized.

There are long term secular variations in the cosmic ray cutoff rigidities directly reflecting the long term secular changes in the geomagnetic field. Over an approximate 10-yr period these changes are sufficient to be experimentally observed, and for very precise analyses these secular variations should be considered.

Because of the complexity of cutoff rigidity calculations, a number of approximations are generally employed. For experimental analyses, the specific approximation utilized is usually determined by the precision of the measurement being studied; as cosmic ray experiments become more precise, more accurate cutoff rigidities are required.

For the majority of cosmic ray studies on the earth's surface or in the atmosphere, it is often sufficient to know the vertical cutoff rigidity at each observational location. World grids of vertical cutoff rigidities from which values for intermediate locations and time periods can be inter-

Table 6-3. Trajectory-derived effective vertical cutoff rigidities for Epoch 1980.

Geographic Latitude	Geographic Longitude (E)											
	0	15	30	45	60	75	90	105	120	135	150	165
70	0.27	0.34	0.39	0.44	0.48	0.51	0.51	0.55	0.58	0.60	0.61	0.57
65	0.60	0.69	0.80	0.87	0.91	0.94	0.98	1.03	1.12	1.23	1.18	1.11
60	1.16	1.36	1.43	1.59	1.62	1.68	1.70	1.80	1.96	2.05	2.12	2.06
55	2.00	2.29	2.45	2.53	2.67	2.73	2.84	2.93	3.12	3.31	3.31	3.15
50	3.32	3.59	3.83	3.94	4.06	4.20	4.34	4.45	4.69	5.00	4.97	4.69
45	4.99	5.20	5.35	5.44	5.66	5.81	6.08	6.31	6.59	6.96	6.96	6.38
40	6.95	7.44	7.59	7.73	8.07	8.54	8.99	9.23	9.57	9.99	9.82	9.05
35	9.77	9.74	10.01	10.42	10.88	11.27	11.39	11.63	11.95	12.18	11.69	10.67
30	11.49	11.83	12.10	12.51	13.09	13.82	14.19	14.31	14.23	13.97	13.46	12.75
25	13.25	13.68	14.03	14.38	14.86	15.37	15.69	15.70	15.47	15.05	14.46	13.76
20	14.17	14.61	14.99	15.39	15.91	16.43	16.73	16.68	16.36	15.85	15.21	14.54
15	14.63	15.10	15.54	15.99	16.54	17.07	17.35	17.27	16.91	16.37	15.75	15.14
10	14.70	15.19	15.67	16.17	16.75	17.29	17.57	17.50	17.15	16.63	16.06	15.56
5	14.41	14.88	15.38	15.94	16.57	17.11	17.41	17.38	17.07	16.61	16.15	15.77
0	13.80	14.22	14.73	15.34	16.00	16.56	16.87	16.90	16.67	16.31	15.97	15.74
−5	12.94	13.27	13.77	14.41	15.10	15.64	15.97	16.07	15.94	15.68	15.50	15.42
−10	11.86	12.11	12.57	13.23	13.90	14.40	14.71	14.88	14.84	14.69	14.65	14.77
−15	10.45	10.63	11.08	11.75	12.32	12.80	13.06	13.24	13.17	13.18	13.39	13.69
−20	8.87	8.89	9.26	9.74	10.24	10.45	10.55	10.69	10.75	10.66	10.62	11.87
−25	7.28	7.29	7.63	7.93	8.02	7.71	7.28	7.26	7.42	7.64	8.36	9.19
−30	6.11	5.84	5.84	5.80	5.58	5.40	5.19	5.14	5.09	5.38	5.84	6.60
−35	5.05	4.49	4.37	4.38	4.12	3.85	3.47	3.41	3.34	3.55	4.10	4.90
−40	4.03	3.62	3.38	3.26	2.89	2.58	2.21	2.04	2.04	2.22	2.53	3.29
−45	3.33	2.88	2.53	2.38	2.00	1.54	1.28	1.12	1.10	1.22	1.47	2.01
−50	2.76	2.27	1.97	1.64	1.30	0.94	0.71	0.56	0.53	0.60	0.81	1.15
−55	2.17	1.72	1.45	1.12	0.82	0.56	0.33	0.24	0.21	0.25	0.38	0.57
−60	1.69	1.29	1.03	0.76	0.69	0.30	0.15	0.09	0.06	0.08	0.14	0.26
−65	1.29	0.96	0.72	0.48	0.29	0.14	0.06	0.00	0.00	0.00	0.03	0.10
−70	0.84	0.66	0.46	0.30	0.16	0.08	0.00	0.00	0.00	0.00	0.00	0.02

polated for Epochs 1955, 1965, 1975, and 1980 have been calculated by numerical integration [Shea et al., 1968; Shea and Smart, 1975; 1983]. An iso-rigidity contour map of effective vertical cosmic ray cutoffs for Epoch 1980.0, reflecting the use of the newer and more accurate internal magnetic field model [Peddie, 1982], is shown in Figure 6-13. The values for the effective vertical cosmic ray cutoff rigidities for this same epoch are tabulated in Table 6-3.

The Störmer Approximation. The classic work of Störmer on the motion of charged particles in a dipole magnetic field resulted in the following equation for the cosmic ray cutoff rigidity P_c

$$P_c = M\frac{\cos^4\lambda}{R^2[1 + (1 - \sin\varepsilon \sin\phi \cos^3\lambda)^{1/2}]^2} \quad (6.16)$$

where λ is the geomagnetic latitude, ε is the zenith angle, and ϕ is the azimuthal angle measured from the magnetic north. When this equation is normalized to the earth's dipole

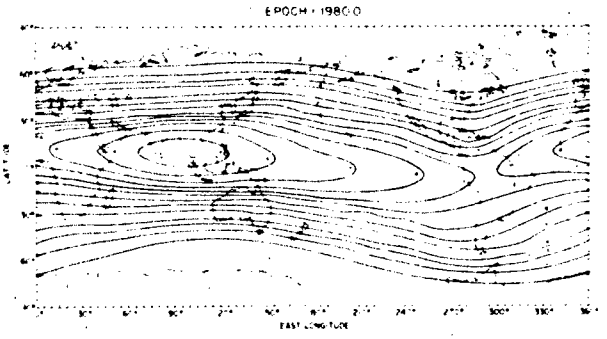


Figure 6-13. Iso-rigidity contours of vertical cosmic ray effective cutoff rigidities. The units of cutoff rigidity are in GV.

Table 6-3. (*Continued*)

Geographic Longitude (E)												
180	195	210	225	240	255	270	285	300	315	330	345	Geographic Latitude
0.47	0.36	0.23	0.13	0.04	0.00	0.00	0.00	0.00	0.06	0.11	0.19	70
0.99	0.77	0.49	0.31	0.17	0.09	0.05	0.05	0.09	0.17	0.30	0.43	65
1.76	1.38	0.98	0.65	0.40	0.23	0.16	0.16	0.23	0.39	0.64	0.90	60
2.85	2.28	1.75	1.23	0.81	0.49	0.36	0.36	0.50	0.80	1.21	1.70	55
4.25	3.46	2.81	2.05	1.42	0.98	0.75	0.73	0.95	1.43	2.13	2.85	50
5.60	4.85	4.15	3.16	2.37	1.75	1.35	1.27	1.62	2.47	3.35	4.35	45
7.96	6.47	5.49	4.60	3.63	2.77	2.14	2.05	2.57	3.79	5.08	6.15	40
9.52	8.99	7.67	6.11	5.17	4.21	3.24	2.99	3.83	5.26	7.36	9.11	35
11.69	10.48	9.65	8.78	6.99	5.54	4.38	4.02	5.14	7.54	10.02	10.94	30
13.04	12.38	11.74	10.88	9.63	7.74	6.02	5.51	7.02	9.98	11.78	12.68	25
13.90	13.33	12.84	12.23	11.32	9.35	7.67	6.61	8.49	11.67	12.88	13.63	20
14.59	14.09	13.67	13.19	12.46	11.05	9.27	8.64	11.01	12.60	13.50	14.12	15
15.10	14.68	14.30	13.88	13.35	12.36	11.27	11.06	12.22	13.11	13.78	14.25	10
15.43	15.07	14.73	14.36	13.92	13.22	12.45	12.29	12.73	13.33	13.76	14.04	5
15.53	15.25	14.93	14.60	14.21	13.71	13.14	12.80	12.97	13.32	13.48	13.55	0
15.35	15.17	14.92	14.62	14.28	13.86	13.36	12.99	12.98	13.09	13.00	12.85	−5
14.87	14.82	14.66	14.43	14.15	13.81	13.37	12.97	12.81	12.69	12.36	11.92	−10
14.01	14.17	14.15	14.03	13.84	13.57	13.19	12.77	12.47	12.15	11.53	10.79	−15
12.72	13.15	13.36	13.42	13.35	13.17	12.85	12.42	11.99	11.47	10.49	9.47	−20
10.19	11.24	11.98	12.56	12.67	12.63	12.38	11.94	11.40	10.52	9.35	8.04	−25
7.98	9.45	9.40	10.75	11.79	11.92	11.78	11.34	10.56	9.57	8.15	6.88	−30
5.65	6.62	8.01	8.73	9.62	11.05	11.09	10.55	9.73	8.50	6.88	5.95	−35
4.15	4.84	5.60	6.76	8.18	9.73	10.08	9.63	8.80	7.35	6.18	5.00	−40
2.69	3.30	4.28	4.99	6.01	7.87	8.89	8.52	7.74	6.80	5.34	4.15	−45
1.65	2.24	2.94	3.79	4.58	5.61	7.05	7.41	6.77	5.47	4.27	3.42	−50
0.95	1.36	1.94	2.64	3.35	4.29	4.90	5.18	4.90	4.25	3.48	2.77	−55
0.51	0.76	1.20	1.77	2.27	2.98	3.75	4.01	3.82	3.39	2.72	2.14	−60
0.22	0.42	0.69	1.05	1.55	1.96	2.46	2.72	2.63	2.40	2.02	1.61	−65
0.10	0.22	0.40	0.60	0.90	1.20	1.51	1.67	1.75	1.59	1.33	1.11	−70

moment M and the distance from the dipole center R is expressed in earth radii, the constant terms evaluate to 59.6. For locations on the earth the vertical cutoff rigidity is very useful, and since the zenith angle is zero the dipole vertical cutoff is given by

$$P_c = \frac{14.9 \cos^4\lambda}{R^2}. \qquad (6.17)$$

Variations of the Störmer cutoff equation can be employed when normalized to the earth's actual magnetic field. An improvement in accuracy can be obtained by correcting for the displacement of the earth's effective magnetic center from the geocenter. A number of magnetic coordinate systems can be employed; however, the McIlwain B-L coordinate system [McIlwain, 1961] is very useful since in a dipole field $\cos^2\lambda = R/L$. The worldwide grid of trajectory-derived effective vertical cosmic ray cutoff rigidities, normalized to the L parameter, results in

$$P_c = 16.0\, L^{-2}. \qquad (6.18)$$

This form is useful where the geomagnetic field retains its basic dipolar structure; it can be used throughout the magnetosphere wherever a useful L coordinate can be calculated. A practical limit for the region of applicability seems to be within the domain of the $L = 4$ shell. At higher L-shells increasing errors will be encountered due to the external current systems of the magnetosphere.

Recent work indicates that the Störmer equation can be generalized to estimate the cutoff at other directions and altitudes if a local normalization value is available. If a vertical (more properly, dipole radial) cutoff is known, the cutoff at other directions may be computed using a normalized form of Equation (6.16) and an appropriate magnetic coordinate system. If a cutoff at one altitude is known, the cutoff at another altitude can be computed. A convenient way is to employ the L^{-2} dependence of Equation (6.18).

Cosmic Ray Cutoffs at High Latitudes. The earth's magnetosphere has a dynamic topology reflected in the daily variation of the cosmic ray cutoff at high magnetic latitudes. In regions where the magnetic field lines are "open" (that is, traceable to the distant magnetospheric tail), the cosmic ray cutoff is essentially zero. In regions of the magnetosphere where the magnetic field line topology changes from open to closed as the earth rotates (magnetic latitudes from $\sim 60°$ to $\sim 75°$), there is a significant daily variation in the value of the cosmic ray cutoff rigidities. Current magnetospheric models are inadequate to precisely model these changes and are limited to providing upper limits of the cutoff rigidity values. The best available average high latitude cutoff values have been obtained by polar orbiting satellites [Fanselow and Stone, 1972] and are shown in Figure 6-14.

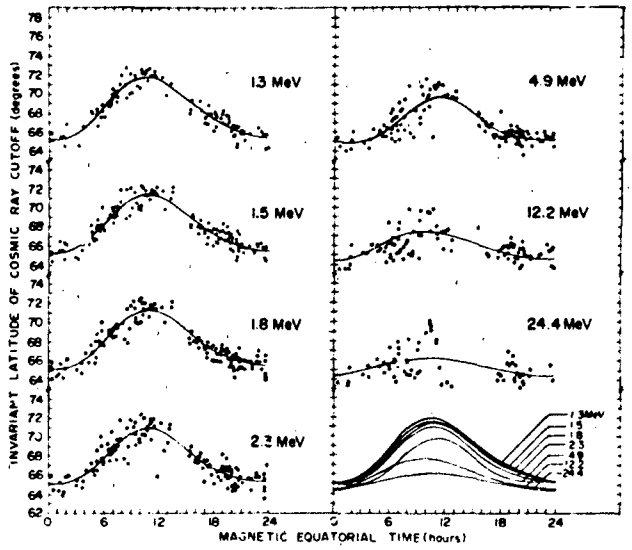


Figure 6-14. Daily variation of the high latitude geomagnetic cutoffs.

6.2.4 Charged-Particle Flux Within the Atmosphere

The intensity and composition of the cosmic rays observed within the atmosphere depend on the quantity of absorbing material traversed before observation in addition to the cutoff rigidity of the observation point. Atmospheric conditions, especially barometric pressure, have an appreciable effect on the measured intensity; hence, cosmic-ray intensities are usually reported in terms of atmospheric depth (mass of air per unit area above the observation point) or of air pressure at the observation point rather than the altitude of the observation. The ionization rate measured within the atmosphere depends upon the amount of matter above the point of observation and on its distribution in height. At a given altitude, the value of the pressure in millibars is about 2% less than the atmospheric depth in grams per square centimeter. Figure 6-15 is a plot of atmospheric depth as a function of altitude.

Primary cosmic rays incident upon the top of the earth's atmosphere interact with air nuclei producing high-energy secondary cosmic rays. These secondary particles, in turn, interact with other nuclei and produce additional secondaries. Figure 6-16 illustrates a nuclear cascade process initiated by a primary cosmic ray. The production of secondary components becomes significant at about 55 km (4-mb pressure with the local intensity reaching a maximum (the Pfotzer maximum) at approximately 20 km (56 mb). The intensity of secondaries then decreases from the Pfotzer maximum to the surface of the earth as the particles lose energy by additional collisions until the majority either decay or are absorbed. An excellent compilation assembling cosmic ray

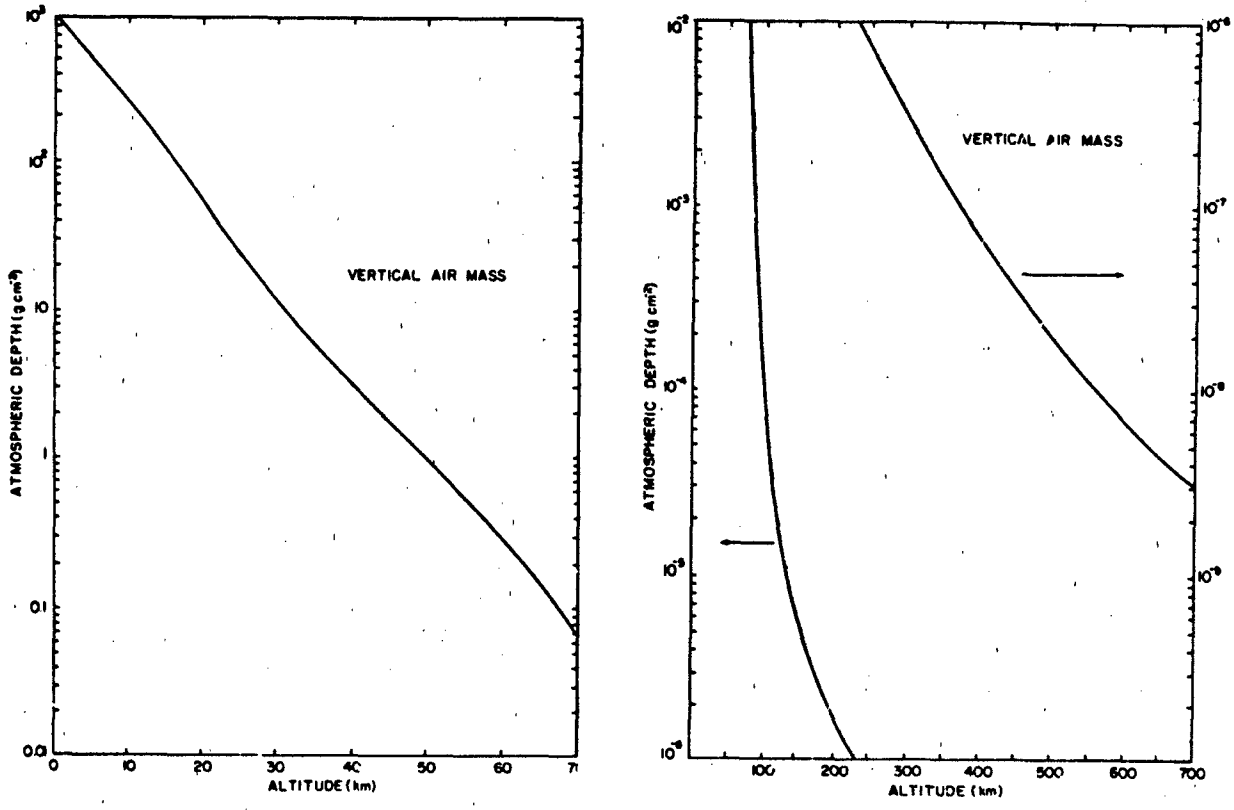


Figure 6-15. The mass of air per unit area in a vertical column extending upward from a height H above sea level [Ely, 1962].

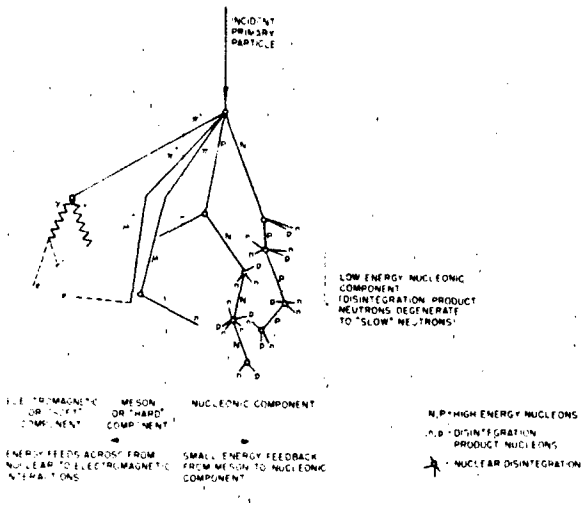


Figure 6-16. Schematic diagram of a cosmic ray shower. N and P are high energy neutrons and protons while the lower case n and p are used to denote disintegration product neutrons and protons. Pions, mesons, electrons, positrons, and gamma rays are indicated by conventional symbols.

observations on the earth has been prepared by Allkofer and Grieder [1984].

In theory, the intensity of any specified secondary component can be derived from a knowledge of the primary spectrum and the specific yield functions. Computer codes exist that will convert a specified primary cosmic ray spectrum into probable secondary components at any location and specified atmospheric depth. See O'Brien [1970], O'Brien [1979] and references therein. An example of the computed and experimentally measured ionization as a function of altitude is shown in Figure 6-17.

6.2.4.1 The Secondary Cosmic Ray Ionizing Component. The cosmic ray secondaries most commonly measured on the earth's surface are π and μ mesons, neutrons, protons, electrons, and γ-ray photons. The secondary cosmic rays are often classified into three major components: the hard component comprised of primarily relativistic muons, the nucleonic component comprised of locally produced protons and neutrons and the soft component, comprised of electrons and γ rays. The flux and composition of the secondary cosmic rays are the result of random interaction probabilities; the instantaneous flux of ionizing secondaries can have a considerble variation from the average flux. For

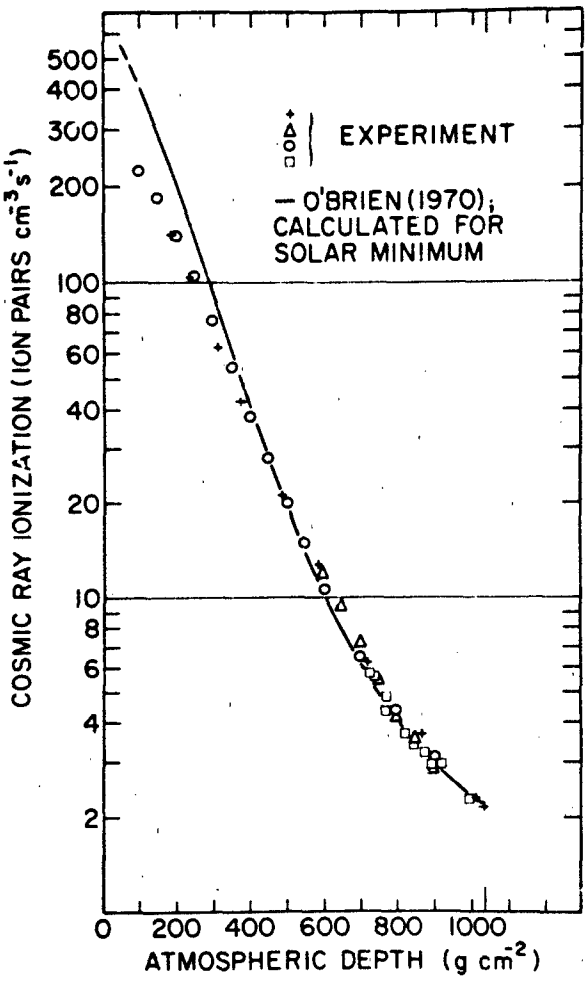


Figure 6-17. Computed and experimental altitude variation of cosmic ray ionization. The data are normalized to a sea level value of 2.15 ion pairs per second per cubic centimeter at sea level (Kykler and Liboff, 1978).

example the intensity in the core of the cosmic ray shower may be orders of magnitude above the "average" for a few microseconds.

Typical ionization rates in the polar atmosphere as a function of atmospheric depth and solar-cycle variation are shown in Figure 6-18. The variation of ionization with latitude obtained by Neher [1967] during the 1965 solar minimum, plotted as a function of cosmic ray cutoff rigidity, is illustrated in Figure 6-19. The most common high energy ionizing component of the secondary cosmic radiation is the cosmic ray μ meson. These muons must be relativistic to traverse the atmosphere (half life $= 2 \times 10^{-6}$ s), are only weakly ionizing (the energy loss of a μ meson in air is 2.2 MeV cm^2 g^{-1}) and at sea level have a pole to equator latitude dependence of the order of ~ 10%. The typical μ meson spectrum is shown in Figure 6-20. The μ meson intensity as a function of altitude is shown in Figure 6-21.

The high energy secondary cosmic radiation can initiate nuclear interactions in whatever matter they penetrate. The nuclear fragments of these interactions resemble "stars" when recorded in photographic emulsions and examined under high magnification. "Star counts" can be used to estimate the probable rate of interactions in semi-conductors; these interactions can cause errors in solid state logic circuits. A compact tabulation of observed star counts is given in Table 6-4.

6.2.4.2 The Secondary Cosmic Ray Neutron Component. The secondary cosmic ray neutron component is generated within the atmosphere by the nuclear interaction of high energy cosmic ray nuclei with atmospheric atoms. There are no neutrons in the primary cosmic ray flux. At high energies the "knock-on" process dominates the neutron production, while at lower energies the "neutron evaporation" process dominates the neutron production. These neutrons interact with other atmospheric atoms and produce radioactive isotopes. The secondary neutron flux in the atmosphere is responsible for the generation of cosmogenic isotopes such as ^{14}C (half life of 5730 years), ^{26}Al (half life

Table 6-4. Cosmic ray induced nuclear interactions (per cubic centimeter per day).

Type of Star (Number of Prongs)	2 GV Cutoff Rigidity						10 GV Cutoff Rigidity	
	Atmospheric Depth (g/cm^2)						Atmospheric Depth (g/cm^2)	
	14.9	47.4	63.7	81.3	121	677	14.9	47.4
>2	2390	2030	2150	2040	1610	22.0	575	425
>9	618	352	372	235	290	1.2	238	132
>16	223	122	128	27	116	0.29	105	61

[Data sources: Rossi [1952, 1964] and references therein.]

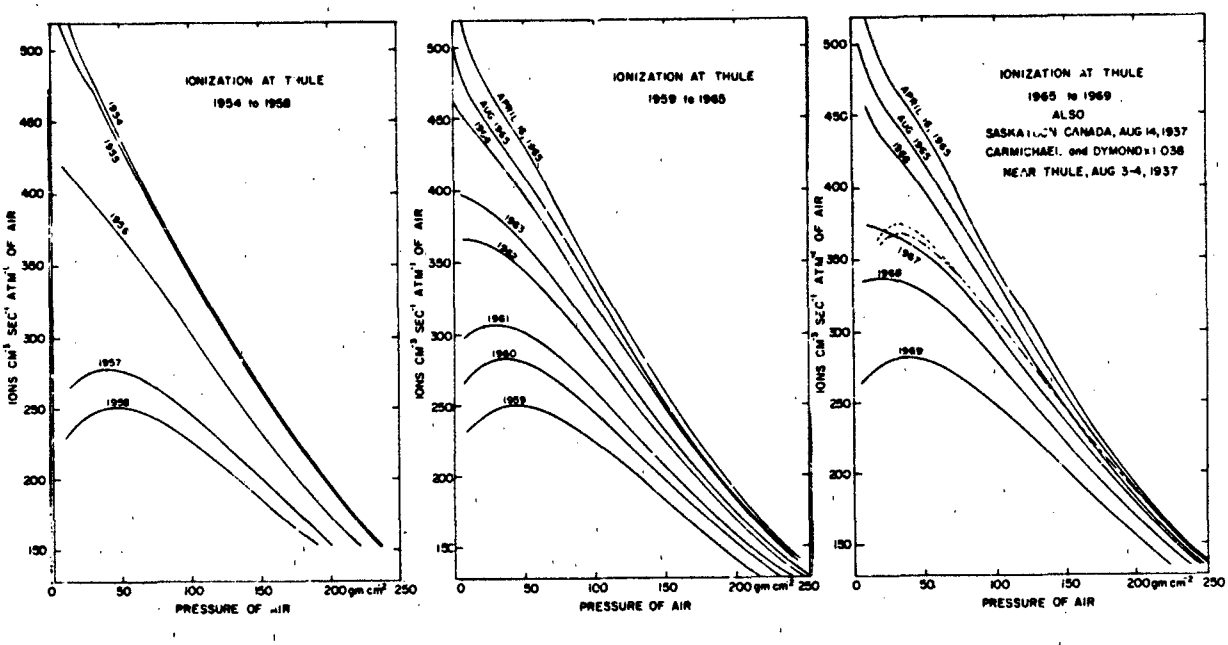


Figure 6-18. Yearly averaged cosmic ray ionization rate per atmosphere of air as a function of atmospheric depth [Neher, 1971].

of 7.4 × 10^5 years), and ^{10}Be (half life of 1.5 × 10^6 years). The carbon-14 isotope enters the biosystem and can be used to date dead biological materials. Other long lived isotopes may collect in the polar ice or sea sediments. The differential energy spectrum of neutrons at various depths in the atmosphere is shown in Figure 6-22 [Armstrong et al., 1973]. These results are typical of midlatitudes. The fast neutron flux in the atmosphere at various depths and cutoff rigidities for both solar minimum and solar maximum conditions is illustrated in Figure 6-23 [Light et al., 1973].

6.2.5 Cosmic Ray Albedo

Some small fraction of the secondary cosmic radiation will escape the atmosphere and contribute to the total radiation flux exposure to earth orbiting satellites. The electrons and protons will spiral along the earth's magnetic field lines to the opposite hemisphere. Measurements of the albedo proton flux at the "top" of the atmosphere for high and midlatitudes are given in Figure 6-24 while the albedo electron flux at the "top" of the atmosphere is given in Figure

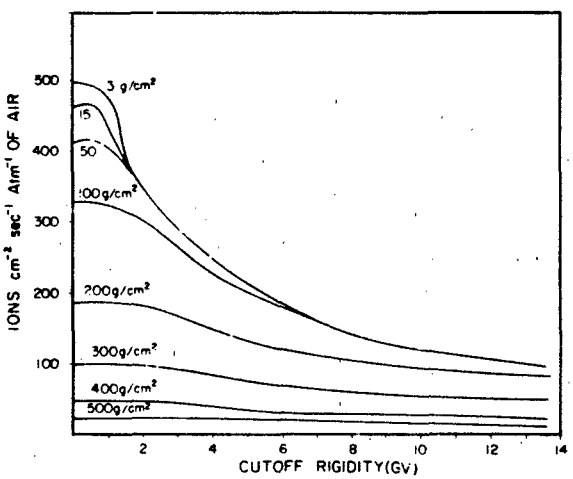


Figure 6-19. Cosmic ray ionization rate as a function of altitude and geomagnetic cutoff rigidity.

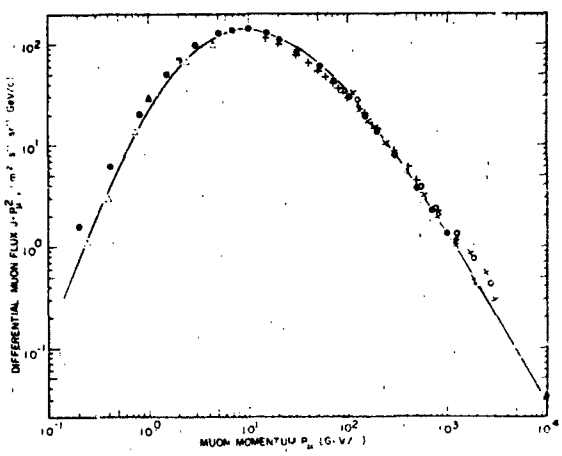


Figure 6-20. Differential vertical flux of muons at sea level. Measured values are indicated by the various symbols while the solid line is a mathematical calculation. [Murakami et al., 1979].

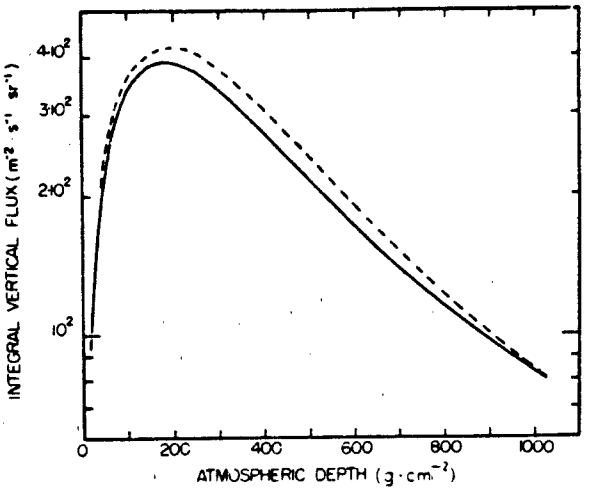


Figure 6-21. Altitude variation of the muon intensity >0.3 GeV. The solid line represents older values (circa 1948) [Murakami et al., 1979].

6-25. The electron albedo includes approximately equal numbers of electrons and positrons. There appears to be a large variability in the available electron albedo measurements that is not yet resolved. Typical numbers in the energy range 0.1 to 2.5 GeV are $e^+ = 77, e^- = 67\ m^{-2}s^{-1}sr^{-1}$. The results of satellite measurements in the equatorial region at 230 km are given in Figure 6-26. There is an approximate factor of 2 increase in these numbers in the high latitude region.

About 10% of the secondary cosmic ray neutron flux in the upper atmosphere will escape into space. The cosmic ray albedo neutron flux escaping from the atmosphere is not bound by the earth's magnetic field and will later decay (neutron half life ~ 11.7 minutes) into an electron and proton, which may be trapped in the earth's magnetic field, and an anti-neutrino. The albedo neutron flux is an important source of the high energy proton flux trapped in the earth's magnetosphere (see Chapter 5). A model "global averaged" albedo neutron spectrum is given in Figure 6-27. Near the earth there is an approximate 1:7 ratio of the albedo neutron flux from the equatorial atmosphere as compared with the flux from the polar atmosphere.

6.3 ENERGETIC SOLAR PARTICLES (SOLAR COSMIC RAYS)

Part of the energy released in solar flares may be nuclei accelerated to high energies and released into space. These solar particle events are commonly referred to by a number of descriptive names such as solar cosmic ray events (SCR), solar proton events, solar electron events, polar cap absorption events (PCA), and ground level events (GLE). Each of these names results from and is descriptive of a measurement technique. Cosmic ray sensors on the earth recorded the first observed energetic solar flare particle outbursts, and hence the name solar cosmic ray events was derived. Later, when balloon borne cosmic ray detectors observed particle events not detected at the earth's surface, the name ground level event was introduced to distinguish between those events detected at ground level and those observed only near the top of the atmosphere. This name is still used to identify solar particle events containing relativistic protons. Beginning in the 1950s the earth's polar ionosphere was shown to respond to solar flare protons with energies from about 1 to 50 MeV, and the term PCA events became synonymous with solar proton events. Spacecraft measurements are now many orders of magnitude more sensitive than earth-based measurements and allow measurements of the composition of solar particle events. The term solar electron event was introduced after 1965 to distinguish the species of particles present. However, the older definitions are still used to provide continuity of historical data bases.

The most commonly measured components of solar particle events are protons and electrons. The first positive measurements of solar neutrons occurred in 1980 [Chupp et al., 1982]. Solar neutron fluxes at the earth will be extremely rare as the neutron half life of ~ 11.7 min implies all but the most energetic solar neutrons will have decayed before they can reach the earth's orbit.

6.3.1 The Solar Particle Source

The source of the energetic solar particles is generally the solar flare (see Chapter 1). There is a time correlation between the observation of an increase of the particle flux in interplanetary space and the occurrence of a solar flare. At low energies, there is a correspondence between the level of solar activity (such as the number of flares and sub-flares) and the low energy solar particle flux in interplanetary space. Some flares are more copious energetic particle producers than others, the more copious solar particle generators generally being solar flares rich with x-ray and high frequency radio emission (particularly in the centimeter and millimeter wave lengths).

Many of the particles accelerated to high energies in a solar flare will escape from the solar corona into the interplanetary medium. As the energetic charged particles move into the interplanetary medium they will be guided along the existing spiral magnetic field pattern. As a result, both the intensity and the spectrum observed at the earth depend somewhat on the relative positions of the earth and the flare on the sun. For example, a solar wind velocity of 400 km/s produces a spiral field that connects the earth to the solar longitude ~ 55° west of the central meridian of the sun as viewed from the earth. The actual degree of "well connectedness" between the earth and the sun depends on in-

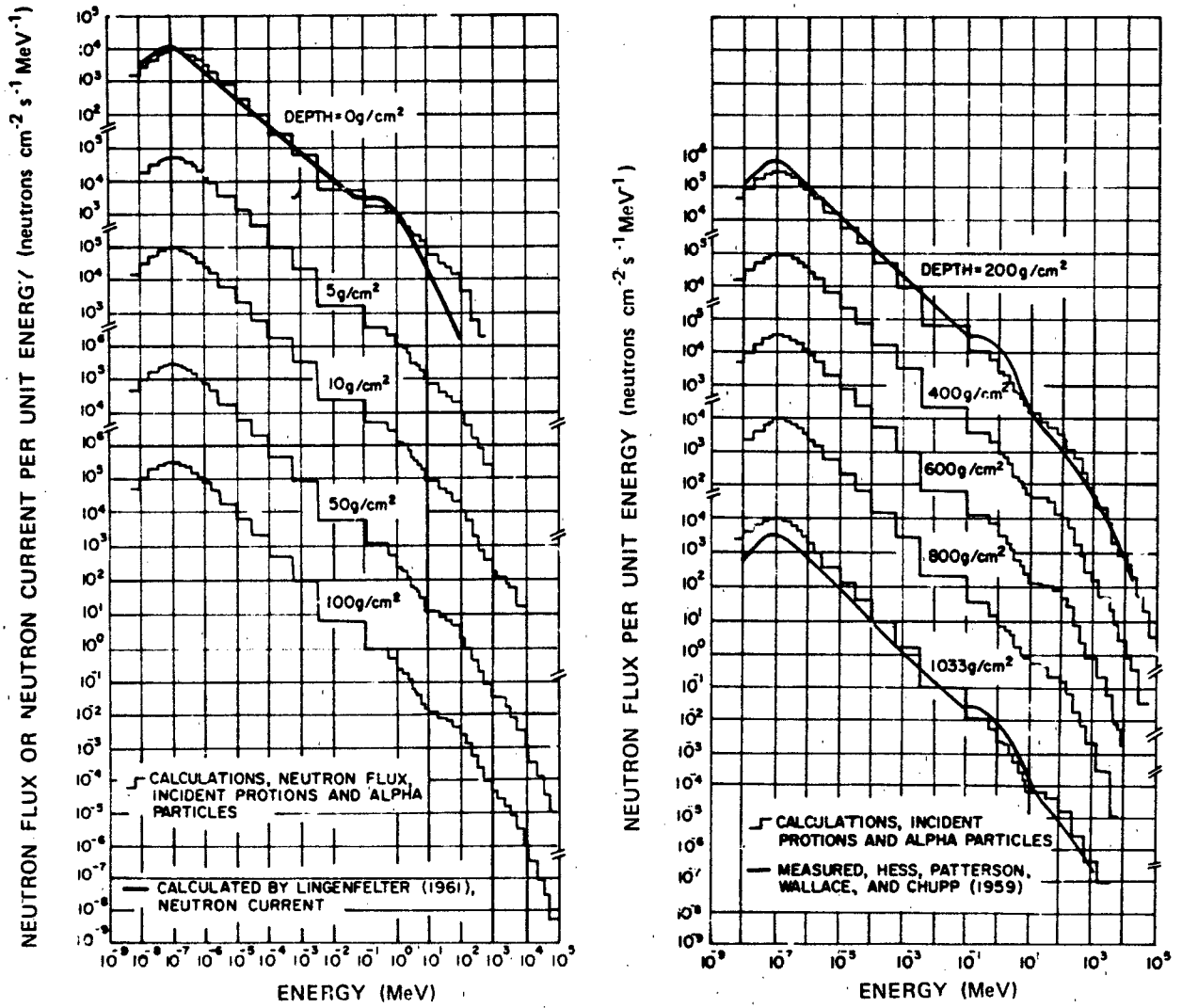


Figure 6-22. Neutron spectra at various depths from the top of the atmosphere. The data are adjusted to solar minimum conditions at the midlatitude geomagnetic latitude of ~42°, (equivalent to a vertical cutoff rigidity of 4.5 GV).

terplanetary conditions at the time of the flare, and these conditions are highly variable and unpredictable.

6.3.2 The Size and Frequency of Solar Particle Events at the Earth

Major solar particle events occur at random, with a frequency that varies from approximately one every two months to one every two years. Major solar particle events consist primarily of sizeable fluxes of energetic electrons and energetic protons. These solar flare initiated solar particle events may also contain a small and apparently variable flux of heavier elements. The frequency of solar particle events is correlated with the solar activity cycle; however, there is a large variance in the distribution. Significant solar particle events have occurred during solar minimum. There can be relatively long periods between significant solar particle events during the solar maximum. The elemental composition of the accelerated ions is not radically different from the elemental composition of the solar atmosphere, although composition differences as a function of energy exist as discussed in Section 6.3.3. The "size" of a solar particle event depends on the energy of measurement, and consequently the "size" is dependent on the energy spectra. The particle events near the earth that result from these solar flares may last from a few hours up to a maximum of approximately ten days. The *Catalog of Solar Proton Events*

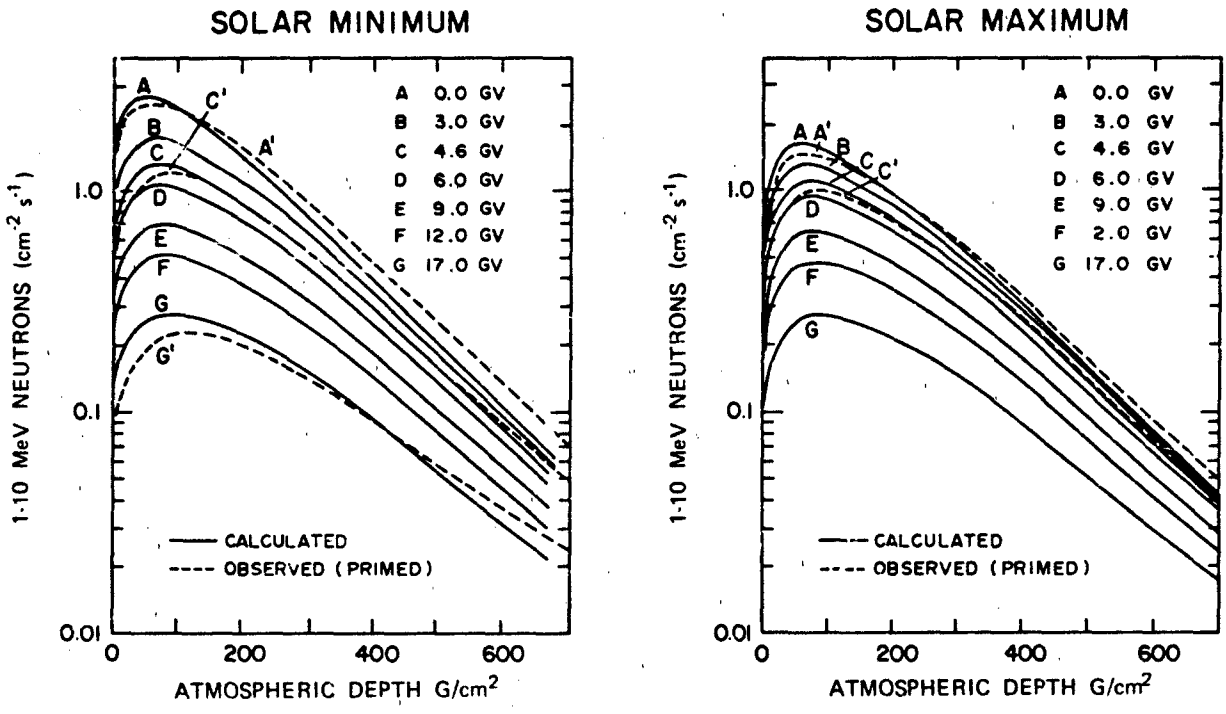


Figure 6-23. Fast neutron flux versus atmospheric depth for various cutoff rigidities for solar minimum conditions (left panel) and solar maximum conditions (right panel).

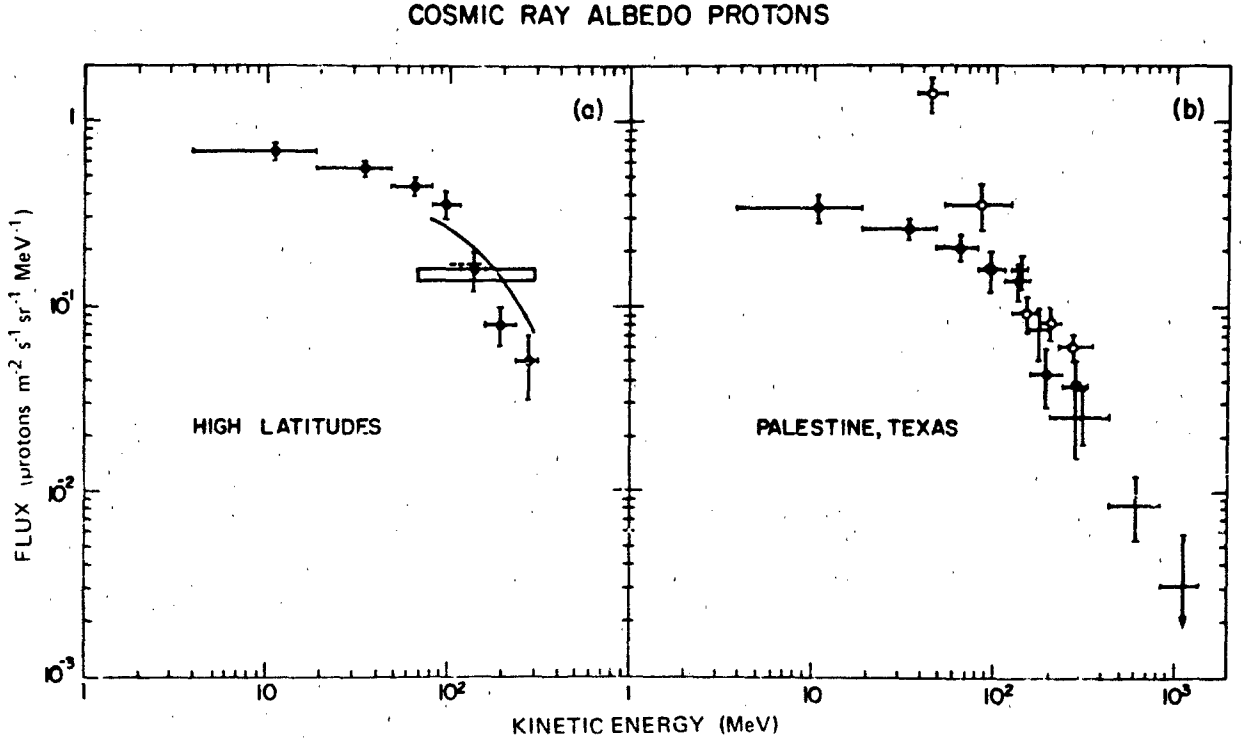


Figure 6-24. Measured differential energy spectrum of the albedo proton flux at the "top" of the atmosphere, at two different locations [Wenzel et al., 1975].

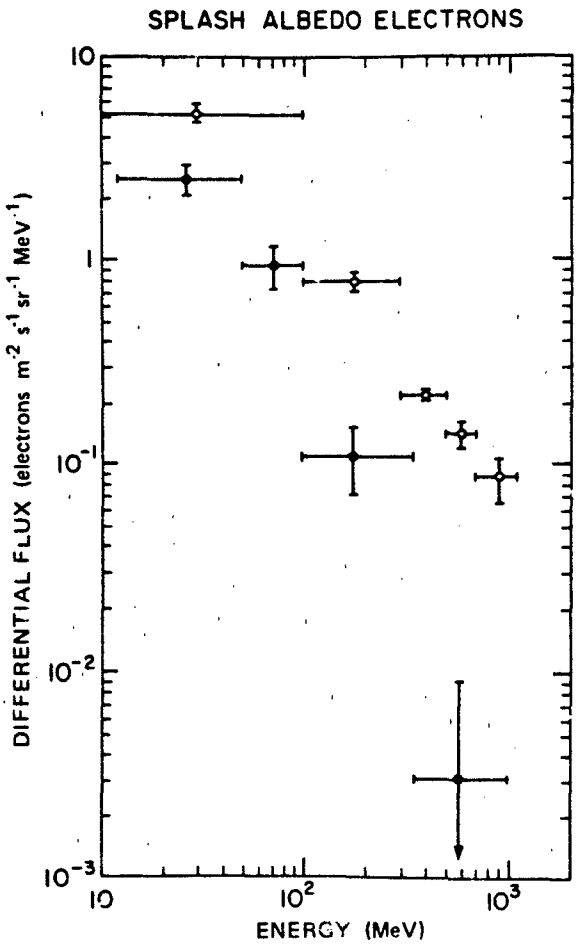


Figure 6-25. Measured differential kinetic energy spectrum of the albedo electron flux at the "top" of the atmosphere. The solid circles are measurements at Palestine, Texas (vertical cutoff rigidity ~4.5 GV), and the open circles are measurements at high latitudes (Ft. Churchill, Canada). Adapted from Israel [1969].

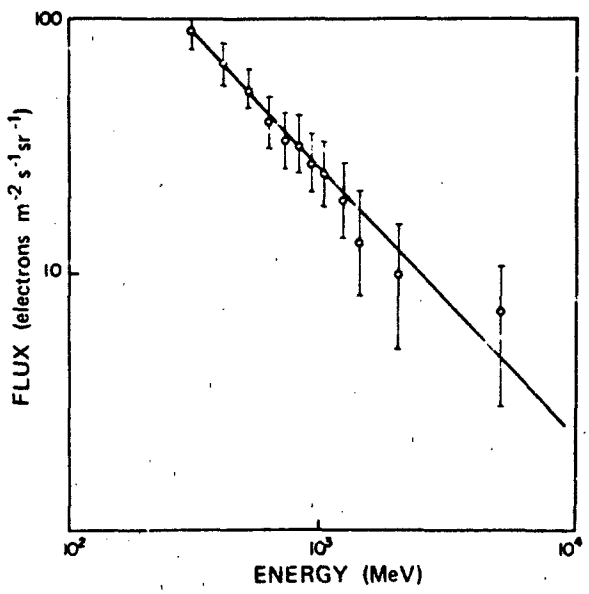


Figure 6-26. Spacecraft measurements of the low altitude albedo electron flux [Kuznetsova, 1978].

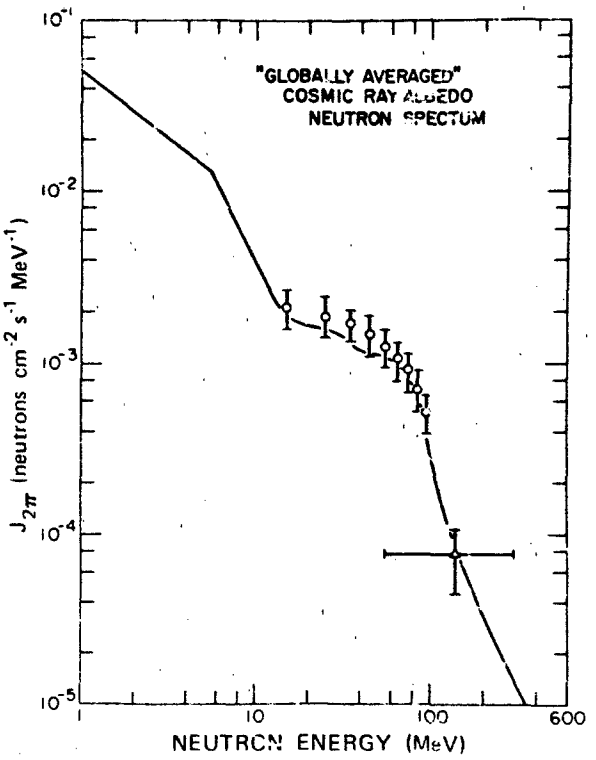


Figure 6-27. A model "global averaged" albedo neutron spectrum [Claflin and White, 1974].

1955–1969 [Dodson et al., 1975] and *Catalog of Solar Proton Events 1970–1979* [Akiniyan et al., 1983] contain considerable detail in describing solar particle events and associated solar circumstances.

6.3.2.1 Solar Proton Events. A list of the major solar proton events observed at 1 AU at the orbit of the earth is given in Table 6-5. The solar proton event frequency distribution has a period of ~ 11 years (the solar sunspot cycle), and there is usually one anomalously large event in each solar cycle. Apart from these anomalous events, the remaining events seem to be distributed as though the log (to base 10) of their sizes has a normal statistical distribution. This is called a log-normal distribution. King [1974] has found that the distribution of the solar proton fluences F is a log-normal distribution above any energy threshold. The

Table 6-5. Major solar proton events observed at the earth.

Solar Particle Event Onset Date	PECS	Maximum Neutron Monitor Increase (Percent)	Maximum 30 MHz Riometer Absorption (db)	Peak Flux (cm^{-2} s^{-1}) (onmidirectional) >30 MeV	Integrated Intensity (cm^{-2}) >30 MeV
1956 Feb 23	X 3 4	4554	13	6 200*	6.5×10^{8}*
1956 Aug 31	X 3 0	—	4.9	150*	2.5×10^{7}*
1957 Jan 20	X 2 0	—	4.1	2 500*	3×10^{8}*
1958 Mar 23	X 2 0	—	3.2	1 200*	[2×10^{8}]†*
1958 Mar 25	X 3 0	—	10	1 500*(sc)	
1958 Jul 7	X 4 0	—	23.7	1 500*	3×10^{8}*
				2 000*(sc)	
1958 Aug 16	X 4 0	—	>15	200*	2×10^{7}*
1958 Aug 22	X 3 0	—	10.6	500*	5×10^{7}*
1958 Aug 26	X 4 0	—	16.6	1 100*	5.3×10^{7}*
1958 Sep 22	X 3 0	—	10.6	50*	6×10^{6}*
1959 May 10	X 4 0	—	22	6 000*	7×10^{8}*
				8 000*(sc)	
1959 Jul 10	X 4 0	—	20	4 000*	8.8×10^{8}*
1959 Jul 14	X 4 0	—	23.7	11 000*	1.1×10^{9}*
1959 Jul 16	X 4 2	10	21.2	16 000*	8.1×10^{8}*
				18 000(sc)	
1960 Apr 1	1 2 0	—	3	50*	2.7×10^{6}*
1960 Apr 5	1 2 0	—	3.1	40*	2×10^{6}*
1960 Apr 28	(2)2 0	—	3	300*	2.5×10^{7}*
1960 May 4	(1)2 4	290	5	200*	7×10^{6}*
1960 May 13	(1)2 0	—	4.5	60*	4×10^{6}*
1960 Sep 3	2 2 1	4	2.7	240*	4×10^{7}*

Date	PECS	Maximum Neutron Monitor Increase (Percent)	Maximum 30 MHz Riometer Absorption (db)	Peak Flux (cm^{-2} s^{-1}) directional >30 MeV	Peak Flux (cm^{-2} s^{-1}) onmidirectional >30 MeV	Integrated Intensity (cm^{-2}) >30 MeV
1960 Nov 12	4 4 4	135	21.2		12 000*	1.4×10^{9}*
1960 Nov 15	4 4 3	88	>20		6 000*	5.2×10^{8}*
1960 Nov 20	(3)3 2	8	5		1 000*	6×10^{7}*
1961 Jul 11	X 1 0	—	1		20*	2×10^{6}*
1961 Jul 12	(2)4 2	—	17		120*	1.0×10^{8}*
1961 Jul 18	3 3 3	24	11		2 500*	2.1×10^{8}*
1961 Jul 20	(1)2 2	7	5		300*	9×10^{6}*
1961 Sep 28	2 2 0	—	3.3	11.0	150*	2.4×10^{6}*
1963 Sep 20	1 2 0	—	4	4.0	30	1×10^{6}
1965 Feb 5	1 1 0	—	1.3	50.0	250	1×10^{6}
1966 Mar 24	1 2 0	—	1.6	15.0		~1.0×10^{5}
1966 Jul 7	1 2 1	3	2.1	30.0		4.4×10^{5}
1966 Aug 28	1 2 0	—	4	15.0		1.0×10^{5}
1967 May 25	3 3 0	—	11	32.0		1.7×10^{6}
1967 May 28	2 2 0	—	4.1	27.0		1.3×10^{6}
1967 Dec 3	1 2 0	—	1.8	10.5		4.6×10^{5}
1968 Jun 9	2 3 0	—	6.5	12.4		8.9×10^{5}
1968 Sep 28	1 1 0	—	1.2	0.1		[6.9×10^{5}]†
1968 Sep 29	1 2 1	1	1.7	19.0		

Date	PECS	Maximum Neutron Monitor Increase (Percent)	Maximum 30 MHz Riometer Absorption (db)	Peak Flux (cm^{-2} s^{-1} sr^{-1}) (directional) >30 MeV	Integrated Intensity (cm^{-2}) >30 MeV
1968 Oct 4–6	1 2 0	—	1.5	6.3	2.6×10^5
1968 Oct 31	2 3 0	—	5.5	10.0	
					[1.2×10^6]†
1968 Nov 1	2 3 0	—	4	11.7	
1968 Nov 18	2 3 3	14	12.5	404.0	1.7×10^7
1968 Dec 3	1 2 0	—	(slight)	1.7	
					[3.2×10^6]†
1968 Dec 5	2 3 0	—	4.7	31.0	
1969 Feb 25	1 2 3	16	2.1	41.5	2.1×10^6
1969 Mar 30	1 1 2	9	1.4	13.0	1.3×10^6
1969 Apr 11	3 3 0	—	>16	123.0	1.6×10^7
1969 Nov 2	3 3 0	—	13	737.0	2.1×10^7
1970 Jan 31	1 2 0	—	3	6.2	2.7×10^5
1970 Mar 6	1 3 0	—	5	0.9	1.0×10^5
1970 Mar 29	1 2 0	—	1.8	20.2	1.7×10^6
1970 July 24	1 2 0	—	4.5	0.8	5.8×10^4
1970 Aug 14	2 2 0	—	3.0	2.7	3.9×10^5
1970 Nov 5	1 2 0	—	3.5	1.7	2.8×10^5
1971 Jan 24	3 3 3	26	14.5	408.0	2.7×10^7
1971 Apr 6	1 2 0	—	3.8	5.0	2.0×10^5
1971 Sep 1	2 3 3	16	5.2	162.0	1.3×10^7
1972 May 28	1 2 0	—	2.6	2.7	5.3×10^5
1972 Aug 4	4 4 2	20	>22	21 000.0	6.2×10^8
1972 Aug 7	3 4 2	8	3.1	384.0	3.0×10^7

PECS. The notation PECS stands for the three digit proton event classification system as defined by Smart and Shea [1971] as follows:

First Digit	Second Digit	Third Digit
E > 10 MeV Satellite Measured Proton Intensity Digit $cm^{-2}s^{-1}sr^{-1}$	Daylight Polar 30 MHz Riometer Absorption	Sea Level Neutron Monitor Increase
−3 From 10^{-3} to 10^{-2}	—	—
−2 From 10^{-2} to 10^{-1}	—	—
−1 From 10^{-1} to 10^{0}	—	—
0 From 10^{0} to 10^{1}	No measurable increase	No measurable increase
1 From 10^{1} to 10^{2}	Less than 1.5 dB	Less than 3%
2 From 10^{2} to 10^{3}	From 1.5 dB to 4.6 dB	From 3% to 10%
3 From 10^{3} to 10^{4}	From 4.6 dB to 15 dB	From 10% to 100%
4 Greater than 10^{4}	Greater than 15 dB	Greater than 100%

The first digit represents the measurement of the E > 10 MeV proton flux by a satellite within the earth-moon system, the second digit represents the 30 MHz absorption measured by a sunlit polar riometer, and the third digit represents the response of a high latitude sea level neutron monitor. An "X" in the first digit indicates that not enough satellite data are available to make a determination of the maximum flux of protons greater than 10 MeV. A digit within parenthesis is provisional, based on partial data.

*Proton flux data derived from Polar Cap absorption measurements.

(sc) Peak flux associated with geomagnetic storm.

[ ]† Composite integrated proton flux of 2 separate events.

Table 6-6. The parameters of the log-normal distributions for ordinary solar flares. Parameters are shown for the integral omnidirectional fluence for the entire solar event in protons cm^{-2}, and the peak omnidirectional flux in protons $cm^{-2} s^{-1}$.

	E > 10 MeV	E > 30 MeV	E > 60 MeV	E > 100 MeV
Log Omnidirectional Integral Fluence (mean ± σ)	8.27 ± 0.59	7.28 ± 0.75	6.63 ± 0.95	5.77 ± 1.24
Log Omnidirectional Peak Flux (mean ± σ)	3.27 ± 0.64	2.37 ± 0.82	1.88 ± 0.78	—

Data from King [1974].

means and standard deviations of the distributions for four energy thresholds are shown in Table 6-6.

6.3 2.2 Solar Electron Events. The data base for solar electron events begins with their identification in 1965. Since it is difficult to measure electrons in a high flux of penetrating protons, it was not until 1965 that the detector technology used on spacecraft was capable of uniquely distinguishing between electrons and protons. Solar electrons are present in all major solar particle events and were probably present, but not identified, in major events prior to 1965.

Solar electron events are characteristically separated into two groups: low energy solar electrons (energy ranges in the 10s of KeV) and high energy electrons (energies in the MeV range). (Since the rest mass energy of an electron is 0.511 MeV, electrons with kinetic energy in the MeV range are relativistic, and their velocity is nearly the speed of light).

Low energy electrons are emitted by many solar flares (even small ones), and there is a very good correlation between Type III radio bursts and the measurement of low energy solar electrons for "favorably" located solar flares.

Table 6-7. Average abundances relative to oxygen of energetic particle populations in the interplanetary medium.

	Solar Flare Particle Events			Solar Material	
Element	Normal (1–20 MeV)	Iron-Rich (1–5 MeV)	Corotating Particle Streams (1–10 MeV)	Corona	Photosphere
H	4600	300	2800	1780	1445
He	70	40	170	150	91*
Li	~0.0005				$\sim 1.4 \cdot 10^{-8}$
Be	<0.0007				$\sim 2 \cdot 10^{-8}$
B	<0.0007				$< 1.8 \cdot 10^{-8}$
C	0.54	0.23	1.05	1.0	0.6
N	0.13			0.2	0.126
O	≡1	≡1	≡1	≡1	≡1
F	<0.002				
Ne	0.16	0.39	0.2	0.063	0.054
Na	0.016			0.005	0.0028
Mg	0.18	0.54	0.13	0.079	0.050
Al	0.016			0.005	0.0048
Si	0.13	0.44	0.087	0.079	0.065
P	<0.002			0.0005	0.0005
S	0.026	0.33	0.06	0.0251	0.023
Ar	~0.004			0.01	0.0015
Ca	0.012			0.0045	0.0032
Fe	0.15	1.3	0.095	0.093	0.05

Data from Cheynet [1979].
*Note: The photosphere He abundance is not accurately measured and there is a large variance between the results of different measurement methods

The frequency of the low energy solar electron events is highly correlated with solar flare frequency throughout the solar sunspot cycle. The location of the solar flare is important since low energy solar electron event detection at the earth is generally limited to heliolongitudes within ± 90° of a position about 50° West heliographic longitude.

High energy solar electron events have a distinctly different time profile from the low energy solar electron events. The high energy solar electron event profile is similar to the time-intensity profile of the 90 to 100 MeV protons. The high energy solar electrons characteristically arrive at 1 AU after the observation of the low energy solar electrons emitted from the same flare. See Simnett [1974] for a detailed study of this phenomena.

There is a special class of solar electron events called "scatter free" events. These events are very impulsive, with short interplanetary propagation times, and short rise and decay times. The data indicate that the solar electrons in these events undergo very little scattering in the interplanetary medium between the sun and the earth. The scatter free events have an upper energy limit of about 1 MeV, and detection at the earth is restricted to a relatively narrow solar flare source region about ± 30° from the Archimedean spiral path from the sun to the earth. See Section 6.3.4 for a discussion of the most favorable propagation path between the sun and the earth.

6.3.3 Elemental Composition of Solar Particle Events

Average values of the elemental abundances of ≥ 1 MeV/nucleon solar particles in relatively large solar flare events (peak intensities exceeding ~ 100 protons $cm^{-2}s^{-1}sr^{-1}MeV^{-1}$) with a "normal" composition are given in Table 6-7 and shown in Figure 6-28 as solid bars, open rectangles (element groups) and solid triangles (upper limits). The abundances relative to oxygen of the individual elements or element groups are derived from a number of individual flare particle events. There is signifiant variability in the composition of normal flare particle events, not only from one event to the next but sometimes during a given event. The vertical extent of the solid bars and open rectangles indicates the degree of variability in the elemental composition. For the large events, variations in the abundances from flare to flare, and with time during single flares, give no indication that physical mechanisms preferentially accelerate any particular species over the energy range of observations. Systematic changes in the elemental composition have been observed during individual solar particle events and in successive flares from the same active region. In addition to the velocity dispersion there is a charge-to-mass dependent dispersion in the time-to-maximum for elements having equal velocities.

Despite these variabilities the basic composition of solar flare particles is similar to the composition of the solar

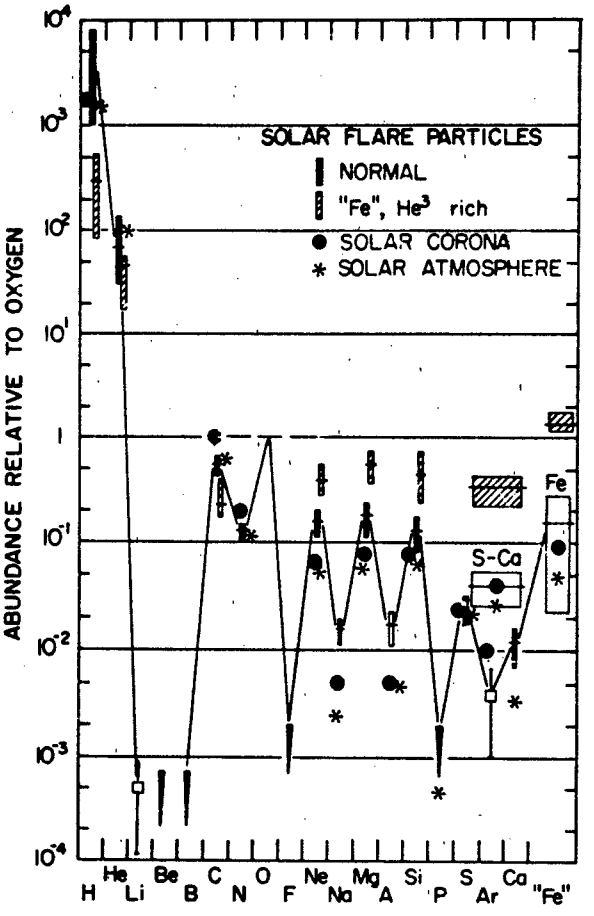


Figure 6-28. Abundances of elements and element groups normalized to oxygen in the solar corona (filled circles), solar atmosphere or photosphere (stars) and in solar flare particle events. The abundances for relatively large (peak intensities > ~ 100 protons $cm^{-2}s^{-1}sr^{-1}MeV^{-1}$) flare particle events with a normal composition are indicated by solid vertical bars, solid triangles (upper limits), open squares (single measurement with error bars) and open rectangles (element groups). The composition for "Fe"-rich (also ^{3}He-rich) solar flare particle events is indicated by cross hatched bars and rectangles and is based on abundance measurements in the 1 to 4.6 MeV/nucleon range. The vertical extent of the bars for H, He, C, N, Ne, Mg, Si, S-Ca, and Fe in both the normal and Fe-rich events indicates the degree of variability in the composition between different solar flare particle events [Gloeckler, 1979].

corona and photosphere whose abundances are also given in Table 6-7. There is overall agreement in the abundances of He, C, N, and S through Fe; the elements Li, Be, B, F, and P are rare, and the odd Z (Z is the atomic number) elements N, F, Na, Al, and P are far less abundant than their even Z neighbors. The energetic solar particle composition seems to be consistently enriched in the elements Ne to Si, and the solar flare C/O ratio, although consistent with the solar photospheric abundance, is lower than the currently acceptable coronal abundance values. At this time

little is known about elements heavier than Fe except that their abundances are very low.

The elemental composition of the solar particle flux is basically similar for small, medium and large events; there is, however, a tendency for the He/O ratio to be somewhat larger (~ 70 to 90) for the small and medium events than the corresponding ratio for the large solar particle events [Gloeckler, 1979].

Our knowledge of the isotropic composition of solar flare particles is at this time rather limited and generally confined to the lighter elements. Typical values of the H and He isotropic ratios in the energy range of ~ 1 to 10 MeV/nucleon averaged over a number of solar flare particle events are $^2H/^1H = 7(+10, -6) \cdot 10^{-6}$, $^3H/^1H \sim 3 \cdot 10^{-6}$ and $^3He/^4He \sim (9 \pm 4) \cdot 10^{-4}$. The isotropic abundances of solar flare Ne in the energy range from 11 to 60 MeV/nucleon have recently been reported to be $^{20}Ne/^{22}Ne = 7.6 \pm 2$. This value is distinctly lower than the comparable and well established solar wind ratio of 13.1 ± 0.6. The low abundances of the carbon, oxygen, and neon isotopes ($^{13}C/C \sim 10^{-2}$, $^{17}O/O < 3 \cdot 10^{-3}$, $^{18}O/O \sim 2 \cdot 10^{-3}$, $^{21}Ne/Ne < 10^{-2}$) indicate that element production by nuclear spallation reactions in the solar atmosphere is not significant.

6.3.3.1 Charged States of Solar Particles. Direct and indirect determinations of the ionization states of solar particles indicate incomplete stripping of the heavy ions over an extended energy range up to at least ~ 5-10 MeV/nucleon. At high energies (≥10 MeV/nucleon), the ionic component of solar particle events seems to be completely stripped of electrons. Direct measurements have shown that, in general, the charge states of C, O and Fe below 1 Mev/charge in energetic solar particle events are (a) consistent with the ionization states of these elements in the solar wind (charge states of 6, 6 and 12 for C, O and Fe respectively), (b) do not vary substantially from event to event, and (c) remain constant over the energy range of the measurements of ~ 50 to 1000 keV/charge. A number of indirect determinations of charge states of Fe at higher energies (~ 1 to 10 MeV/nucleon) also indicate partial ionization.

6.3.3.2 Corotating Energetic Particle Streams. Corotating energetic particle events are modest increases in the intensity of protons and α-particles in the range of < 1 to ~ 20 MeV/nucleon, that persist for three to six days and often reappear a number of times at ~ 27-day intervals. These corotating or 27-day recurrent increases are not associated with solar flares and do not exhibit the velocity dispersion often observed at the onset of solar flare particle events but are correlated with high-speed solar wind streams and interplanetary magnetic field structures corotating with the sun. Measurements of radial gradients and particle anisotropies in corotating events combined with the temporal-spatial correlations observed at ~ 4 to 5 AU between the ~ 1 MeV/nucleon corotating particles and the magnetic field and plasma turbulence in the corotating interaction region (CIR) provide evidence for the continuous acceleration of particles in the interplanetary medium between ~ 2 and 5 AU. The acceleration mechanisms proposed most frequently include both statistical processes resulting from magnetic and plasma turbulences that are generated when high and low speed solar wind streams collide in the corotating interaction region, and acceleration at the forward and reverse shocks which bound the corotating interaction region.

Of the possible sources for the corotating energetic particles, the most likely candidates are either the high-energy tail of the solar wind or a possible but so far undetected, low energy(< 200 keV/nucleon) residual particle population, presumably of solar origin.

6.3.4 Prediction of Solar Proton Events

The capability to predict a phenomenon such as a solar proton event illustrates our understanding of the processes controlling such a phenomenon. Solar protons are accelerated in solar active regions during solar flare events. The x-ray, radio and optical emissions during the solar flare event are the indicators that particle acceleration is occurring. The "U-shaped" spectral signature in the radio emission peak power spectrum [Castelli et al., 1967; Castelli and Guidice, 1976] can be interpreted as the indicator that solar protons are being released from the solar active region (see Chapter 11). After the energetic particles are released from the accelerating region, they will travel through the solar corona and some will propagate outwards along the interplanetary magnetic field lines to the position of the earth. A computerized code to predict solar proton intensities at the earth after the occurrence of a specific solar flare has been developed by Smart and Shea [1979].

6.3.4.1 Propagation of Solar Particles. The propagation of solar protons from the flare site to the earth can be separated into two distinct phases. The first phase is diffusion from the flare site through the solar corona to the "foot" of the Archimedean sprial path formed by the interplanetary magnetic field line between the sun and the earth. The second phase is the propagation in the interplanetary medium from the sun to the earth along the interplanetary magnetic field lines. Figure 6-29 provides an illustration of the propagation concepts.

Propagation in the Solar Corona. The angular distance the solar particles travel in the solar corona from the presumed source (that is, the solar flare site) to the foot of the Archimedean spiral path from the sun to the earth is designated by the symbol θ in Figure 6-29. In a heliographic coordinate system the solar flare will have coordinates defined as solar latitude λ_f and heliocentric longitude ϕ_f, where ϕ_f is positive west of central meridian, and ϕ_f is negative east of central meridian. If the "foot" of the Ar-

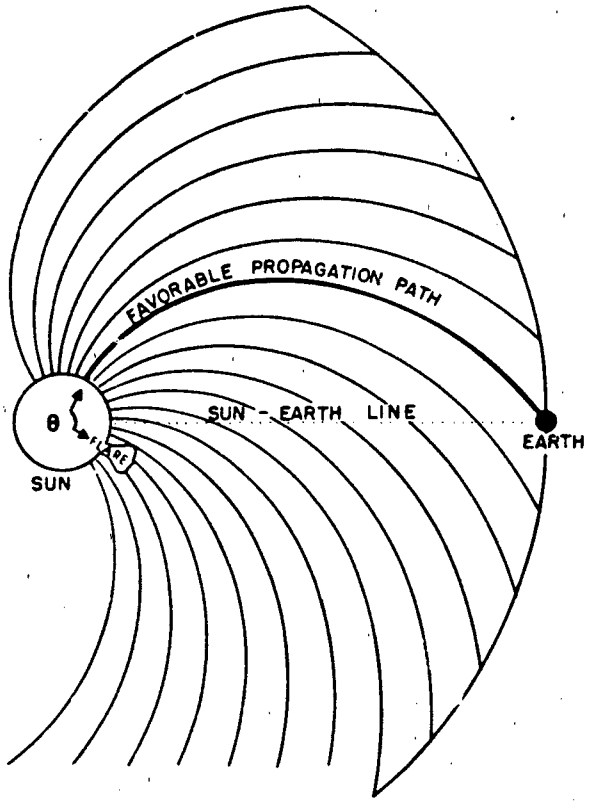


Figure 6-29. Illustration of the concept of solar particle propagation from the sun to the earth. The coronal propagation distance θ is illustrated by the heavy arc on the sun. Interplanetary propagation proceeds along the Archimedean spiral path from the sun to the earth.

chimedean spiral is at heliocentric longitude ϕ_A, then the heliocentric longitudinal distance from the flare site to the "foot" of the Archimedean spiral will be $|\phi_A - \phi_f|$. It is assumed that coronal propagation is a function of θ, and from diffusion theory, it is expected to be proportional to θ^2 [Wibberenz, 1974].

As a result of diffusion in the solar corona and the inherent assumption that some stochastic processes are operating, there will be a solar particle gradient existing in the solar corona such that the proton intensity decreases as a function of distance from the flare site. The observational evidence shows that the gradient may vary from case to case. An observer at one astronomical unit who is connected via the interplanetary magnetic field to the heliographic position of the flaring region would observe the maximum possible particle intensity. An observer whose interplanetary magnetic field connection is at a distance of θ from the flaring location would observe a reduced intensity. Observational data suggest an average gradient may be about one order of magnitude per radian of heliocentric distance.

Propagation in the Interplanetary Medium. After the energetic solar particles propagate through the solar corona

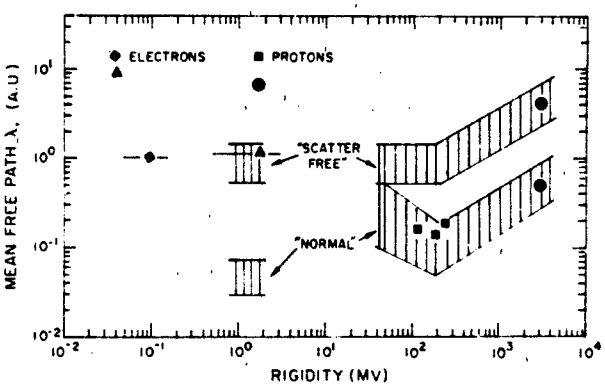


Figure 6-30. Interplanetary mean free path $\lambda_{\parallel}$ as a function of charged particle rigidity. These data have been derived from an analysis of many solar particle events. The shaded areas indicate the range of "scatter free" (upper region) and "normal" (lower region) propagation in the interplanetary medium.

and are released into the interplanetary medium, they propagate along the interplanetary magnetic field lines. (See Chapter 3 for a detailed description of the solar wind.)

The spiral structure of the interplanetary magnetic field shown in Figure 6-29 is highly idealized. On a microscale there are many irregularities in the interplanetary medium. The energetic solar particles propagating through the interplanetary medium will undergo random scattering and the same general theory described in Section 6.2.2.1 applies. The mean free path between scattering centers is quite variable, depending on both the particle energy and the amount of turbulance in the interplanetary medium. A summary of the general range of mean free paths determined from a number of solar particle events is illustated in Figure 6-30.

During the sun to earth phase of their propagation, the MeV solar particles appear to have an average mean free path length of the order of 0.1 to 0.3 AU; consequently, over the 1 AU radial distance, diffusion perpendicular to the interplanetary magnetic field line is negligible. The energetic particles travel essentially along the interplanetary magnetic field lines with a velocity that is a function of the particle energy. The minimum propagation delay will be for particles that essentially travel along the interplanetary magnetic field lines with miminal scattering, so for "scatter free" onsets the propagation time from the sun to the earth will be the distance traveled along the Archimedean spiral path divided by the particle velocity.

Some scattering certainly occurs in the interplanetary medium, and to determine the time of maximum intensity, diffusion theory is applicable. Almost all theories involving solar particle transport in the interplanetary medium show that the time of maximum is proportional to the square of the distance traveled. The distance travelled can be obtained by integrating along the Archimedean spiral path. Normalized to the earth's orbit at 1 AU, in a heliographic coordinate system, the Archimedean spiral path length, D, is

$$D = 1/2\left[\sqrt{\phi_A^2 + 1} + \frac{\ln\phi_A + \sqrt{\phi_A^2 + 1}}{\phi_A}\right] \quad (6.19)$$

where ϕ_A is the heliographic longitude of the "foot point" where the Archimedean spiral path to the earth connects to the sun. If the solar wind plasma carries "frozen in" magnetic fields, then the heliographic longitude of the "foot point" of the interplanetary magnetic field line that extends from the sun to the earth is given by $\phi_A = (\omega_s r)/V_{sw}$, where ω_s is the solar synodic rotation rate (13.3 degrees/day), r is 1 AU, and V_{sw} is the velocity of the solar wind in km/s at the earth. If the Archimedean spiral equation in the polar form is normalized to 1 AU, then $\phi_A = 404/V_{sw}$ where 404 is the nominal solar wind velocity when the solar rotation is equal to one radian during the time it takes the plasma element to travel 1 AU.

The velocity of a proton is given by

$$\beta = \left[1 - \frac{1}{(E/m_pc^2 + 1)^2}\right]^{1/2} \quad (6.20)$$

where E is the kinetic energy of the particle in MeV, β is the velocity of the particle compared with the velocity of light, and the rest energy of the proton m_pc^2 has the value of 938.232 MeV. (Note: $\beta = V/c$.)

The minimum time in hours required for a proton to propagate along the Archimedean spiral path from the sun to the earth with zero pitch angle and no scattering is $0.133\ D/\beta$.

6.3.4.2 Particle Increase Onset Time. The delay time from the solar flare until the onset of the particle increase at the earth is the combination of the coronal propagation time and the time for the particles to propagate from the sun to the earth along the interplanetary magnetic field lines.

The distribution of onset times expected for 30 MeV protons for nominal solar wind speeds is shown in Figure 6-31. In this figure the heavy solid line has the functional form of $4\theta^2$; the data points shown on the figure are taken from Barouch et al. [1971] and indicate typical variations that may be expected. The minimum in the figure corresponds to a flare at the heliographic longitude of the "foot point" of the nominal Archimedean spiral path between the sun and the earth (57.2° West of central meridian). The initial onset will be for the fastest (highest energy) protons. The onset time is given by

$$T_i \text{ (in hours)} = \frac{0.133\ D}{\beta} + 4\theta^2, \quad (6.21)$$

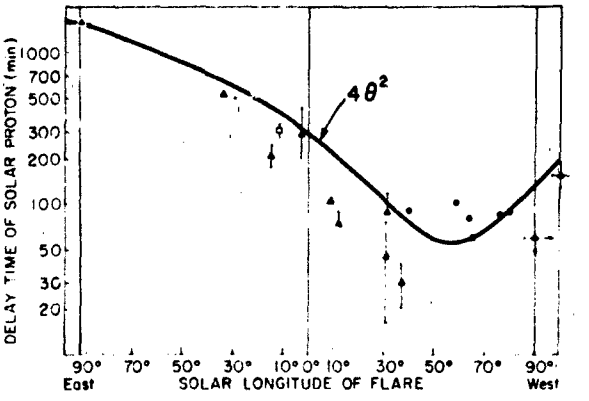


Figure 6-31. Distribution of observed onset times of 30 MeV protons as a function of solar longitude. The data points are the measurements of Barouch et al. [1971]

where D is the distance along the length of the Archimedean spiral from the sun to the earth and β is the proton velocity. This calculation can be repeated for each proton energy (velocity) desired to obtain the velocity dispersion of the predicted onset times as a function of energy.

6.3.4.3 Time of Maximum Particle Intensity. The distribution of the time of the maximum of 20 to 80 MeV protons observed at the earth for a number of proton events [Van Hollebeke et al., 1975] as a function of solar longitude is illustrated in Figure 6-32 where the heavy solid line has the functional form of $8\theta^2$. The minimum in the curve corresponds to a flare at the "foot point" of the Archimedean spiral path between the earth and the sun which would be about 57.2° for a nominal solar wind of 404 km/s.

The time of maximum can be computed by assuming the particles have an average bulk velocity that is about half

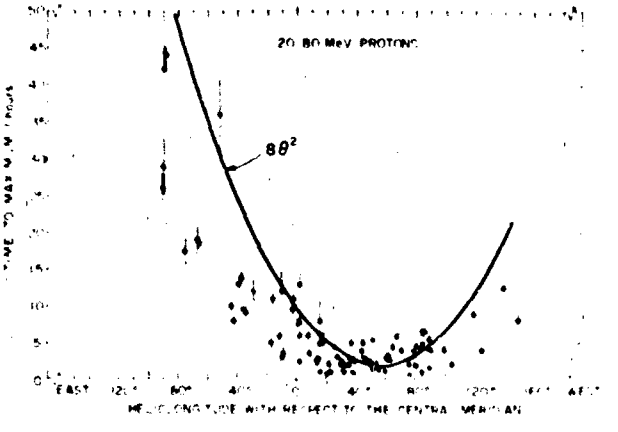


Figure 6-32. Distribution of observed times of maximum of 20 to 80 MeV protons as a function of solar longitude. The data points are the measurements of Van Hollebeke et al. [1975]

of the particle speed that would be computed from the particle energy. Most theoretical models set the time of maximum proportional to the square of the distance traveled along the magnetic field line from the sun. The time of maximum intensity is computed by

$$T_m \text{ (in hours)} = 2.0\frac{0.133\ D^2}{\beta} + 8\theta^2. \qquad (6.22)$$

6.3.4.4 Magnitude of Maximum Intensity. The maximum intensity is predicted by converting an electromagnetic emission parameter (usually radio flux or x-ray flux) to a proton flux. Many different algorithms can be used but currently the best predictors are the relationships between integrated radio flux and peak proton flux. The maximum particle intensity is expected to be along the Archimedean spiral path from the flare. The maximum particle intensity expected at the earth is derived by attenuating the flux through the coronal gradient over the heliocentric distance in the corona between the flare position and the solar equatorial longitude of the "foot point" of the Archimedean spiral path from the sun to the earth.

6.3.4.5 Proton Event Decay. The decaying portion of a solar particle event is normally exponential in character with a nominal time constant for the flux to decay by a factor of 1/e although there is a slight energy dependence, and there can be a large variability between events.

The decaying portion of the event can be modeled after the principles of diffusive particle transport in the interplanetary medium. One such method is collimated convection [Roelof, 1973]. This results in a 1/e decay constant of

$$T_d \text{ (in hours)} = \frac{3D}{4V_{sw}(|\gamma| + 1)\,3600}. \qquad (6.23)$$

where T_d is the 1/e decay constant, D is the distance along the Archimedean spiral path, V_{sw} is the solar wind velocity, and γ is the differential energy spectral exponent.

Effect of Coronal Gradient on Apparent Decay Rate. The assumption that there is a particle gradient in the corona combined with the solar rotation, results in a decay rate at the earth that is dependent on the solar flare location. After the particle flux maximum, the flux along the interplanetary magnetic field lines can be considered as a corotating structure being convected out into space. If the particle source is a flare on the eastern hemisphere or near central meridian of the sun, the particle coronal longitudinal flux gradient rotates toward the "foot point" of the Archimedean spiral from the sun to the earth; thus the apparent decay rate will be slower than the radial convection rate out from the sun. If the particle source rotates away from the "foot point," as is the case for flares that occur on the extreme western hemisphere of the sun, then the solar rotation and the coronal particle longitudinal gradient will combine to increase the flux decay rate at the earth.

6.3.4.6 Differential Energy Spectra. The slope of the differential energy spectrum of solar proton events observed at the earth is summarized by Van Hollebeke et al. [1975] and shown in Figure 6-33. Assuming that the hardest spectra (the minimum in the solid curve of Figure 6-33) should be along the interplanetary magnetic field lines leading away from the flare site, the average spectral slope is observed to vary as

$$\gamma_t = -2.7\left(1.0 + \frac{\theta}{2}\right). \qquad (6.24)$$

6.3.4.7 Prediction of Polar Cap Absorption. It is possible to convert predicted proton flux to prediction of the profile of the riometer absorption in the polar ionosphere. It has been shown that riometer absorption is proportional to the square root of the intergal proton flux above some specified energy. A convenient relationship is $J_{(>10)} = 10A^2$, where $J_{(>10)}$ is the flux of protons with energy greater then 10 MeV and A is the absorption in decibels for a 30 MHz riometer. The relationships developed by Sellers et al. [1977] are simple equations that are relatively independent of the slope of the solar proton energy spectrum. These authors selected the integral proton flux greater than 5.2 MeV and 2.2 MeV to be a good predictor of the riometer absorption in the sunlit polar ionosphere and the night time polar ionosphere respectively. These equations are

$$A \text{ (day)} = 0.6[J_{(>5.2)}]^{1/2} \qquad (6.25)$$

$$A \text{ (night)} = 0.25[J_{(>2.2)}]^{1/2}. \qquad (6.26)$$

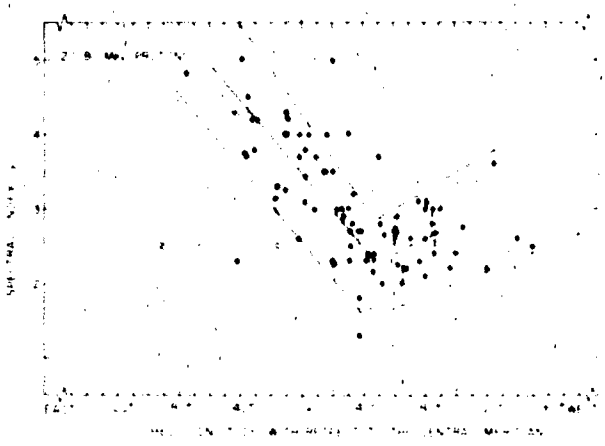


Figure 6-33. Variation of the absolute value of spectral index γ in the 20–80 MeV range as a function of the heliolongitude. The dashed contour lines enclose 92% of the "normal" events. The solid line is a least squares fit that can be represented by $2.7[1 + \theta/2]$.

CHAPTER 6

REFERENCES

Adams, J.H. Jr., R. Silverberg, and C.H. Tsao, "Cosmic Ray Effects on Microelectronics, Part I: The Near Earth Particle Environment," NRL Rept. 4056, Naval Research Laboratory, Washington, D.C. 1981.

Akiniyan, S.T., G.A. Bazilevskaya, V.N. Ishkov, et al., *Catalog of Solar Proton Events 1970—1979*, Institut Zemnogo Magnetizma, Ionosfery I Rasprostraneniya Radiovoln (Institute of Earth Magnetism, Ionosphere and Radiowave Propagation), Academy of Science, USSR, 1983.

Allkover, O.C. and P.K.F. Grieder, *Cosmic Rays on Earth, Physics Data 25-1*, Fachinformationszentrum Energie, Physik, Mathematik, Gmbh, Karlsruhe, Germany, 1984.

Armstrong, T.W., K.C. Chandler, and J. Barish, "Calculations of Neutron Flux Spectra Induced in the Earth's Atmosphere by Galactic Cosmic Rays," *J. Geophys. Res.*, **78:** 2715, 1973.

Barouch, E., M. Gros and P. Masse, "The Solar Longitude Dependence of Proton Event Decay," *Sol. Phys.*, **19:** 483, 1971.

Cameron, A.G.W., "Elementary and nuclidic abundances in the solar system," in *Essays in Nuclear Astrophysics*, edited by C.A. Barnes, D.D. Clayton, and D.N. Schramm, Cambridge University Press, Cambridge, England, 1981.

Castelli, J.P., and D.A. Guidice, "Impact of Current Solar Radio Patrol Observations," *Vistas in Astron.*, **19:** 355, 1976.

Castelli, J.P., J. Aarons, and G.A. Michael, "Flux Density Measurements of Radio Bursts of Proton-Producing and Nonproton Flares," *J. Geophys. Res.*, **72:** 5491, 1967.

Chupp, E.L., D.J. Forrest, J.M. Ryan, J. Heslin, C. Reppin, K. Pinkau, G. Kanbach, E. Rieger, and G.H. Share, "A Direct Observation of Solar Neutrons Following the 0188 UT Flare on 1980 June 21," *Astrophys. J. Lett.*, **263:** L95–L99, 1982.

Claflin, E.S. and R.S. White, "A Study of Equatorial Inner Belt Protons from 2 to 200 MeV," *J. Geophys. Res.*, **79:** 959, 1974.

Dodson, H.W., E.R. Hedeman, R.W. Kreplin, M.J. Martres, V.N. Obridko, M.A. Shea, D.F. Smart, and H. Tanaka, *Catalog of Solar Particle Events 1955–1969*, Vol. 49 of the Astrophysics and Space Science Library, edited by Z. Svestka and P. Simon, D. Reidel, Dordrecht, Holland, 1975.

Ely, J.T.A., "Atmospheric Depth and Effective Solid Angle for Radiation Penetrating the Atmosphere," *Geophysics Research Papers No. 74* (AFCRL 62-260), Air Force Cambridge Research Laboratories, 1962.

Eraker, J.H. and J.A. Simpson, "Origins of the Low Energy Relativistic Interplanetary Electrons," Proceedings of the 17th International Cosmic Ray Conference Paris 1981 *Conference Papers*, **3:** 279, Centre d'Etudes Nucleaires, Saclay, France, 1981.

Fanselow, J.L. and E.C. Stone, "Geomagnetic Cutoffs for Cosmic-Ray Protons for Seven Energy Intervals between 1.2 and 39 MeV," *J. Geophys. Res.*, **77:** 3999, 1972.

Forman, M.A. and O.A. Schaeffer, "Cosmic Ray Intensity Over Long Time Scales," *Rev. Geophys. and Sp. Phys.*, **17:** 552, 1979.

Garcia-Munoz, M., G.M. Mason, and J.A. Simpson, "The Anomalous He component in the Cosmic Ray Spectra of ≤ 50 MeV per Z Nucleon During 1972–1974," *Astrophys. J.*, **202:** 265, 1975.

Gleeson, L.J. and W.I. Axford, "Cosmic Rays in the Interplanetary Medium," *Astrophys. J.*, **149:** L115, 1967.

Gleeson, L.J. and W.I. Axford, "Solar Modulation of Galactic Cosmic Rays," *Astrophys. J.*, **154:** 1011, 1968.

Gloeckler, G., "Composition of Energetic Particle Populations in Interplanetary Space," *Rev. Geophys. Space Phys.*, **17:** 569, 1979.

Haffner, J.W., "Radiation and Shielding in Space," *Nuclear Science and Technology 4*, Academic Press, New York, 1967.

Hess, W.N., H.W. Patterson, R. Wallace, and E.L. Chupp, "Cosmic-ray Neutron Energy Spectrum," *Phys. Rev.*, **116:** 445, 1959.

Israel, M.H., "Cosmic Ray Electrons Between 12 MeV and 1 Gev in 1967," *J. Geophys. Res.*, **74:** 4701, 1969.

King, J.H., "Solar Proton Fluences for 1977–1983 Space Missions," *J. Spacecraft and Rockets*, **11:** 401, 1974.

Kurnosova, L.V., "Charged Particle Flux at 200 to 300 km above the Earth" in *Cosmic Rays in the Stratosphere and Near Space*, edited by N.G. Basov, Lebedev, Trudy 88, Proceedings of the P.N. Lebedev Institute, **88:** 137, 1976. Engl transl Consultants Bureau, New York, 1978.

Kykler, G.C. and A.R. Liboff, "Absolute Cosmic Ray Ionization Measurement in a 900-Liter Chamber," *J. Geophys. Res.*, **82:** 5539, 1978.

Levy, E.H., S.P. Duggal, and M.A. Pomerantz, "Adiabatic Fermi Acceleration of Energetic Particles Between Converging Interplanetary Shock Waves," *J. Geophys. Res.*, **81:** 51, 1976.

Lezniak, J.A. and W.R. Webber, "The Charge Composition and Energy Spectra of Cosmic-Ray Nuclei from 3000 MeV per Nucleon to 50 GeV Per Nucleon," *Astrophys. J.*, **223:** 1978.

Light, E.S., M. Merker, H.J. Verschell, R.B. Mendell, and S.A. Korff, "Time Dependent World Wide Distribution of Atmospheric Neutrons and of Their Products, 2. Calculations," *J. Geophys. Res.*, **78:** 2741, 1973.

Lingenfelter, R.E., "The Cosmic Ray Neutron Leakage Flux," *J. Geophys. Res.*, **68:** 5633, 1963.

McCracken, K.G., U.R. Rao, B.C. Fowler, M.A. Shea, and D.F. Smart, "Cosmic Ray Tables," *IQSY Instruction Manual No. 10*, IQSY Committee, London, 1965.

McCracken, K.G., U.R. Rao, B.C. Fowler, M.A. Shea, and D.F. Smart, "Cosmic Rays (Asymptotic Directions, etc)" in Chapter 14 of *Annals of the IQSY, Vol. 1*, (Geophysical Measurements: Techniques, Observational Schedules and Treatment of Data) edited by C.M. Minnis, 198, MIT Press, Cambridge, Mass., 1968

McDonald, F.R., ed., *Solar Proton Manual*, National Aeronautics and Space Administration Technical Report, NASA TR-R-169, 1963.

McIlwain, C.E., "Coordinates for Mapping the Distribution of Magnetically Trapped Particles," *J. Geophys. Res.*, **66:** 3681, 1961.

Mewaldt, R.A., "The Elemental and Isotopic Composition of Galactic Cosmic Rays," Proceedings of the 17th International Cosmic Ray Conference, *Conference Papers*, **13:** 49, Centre d'Etudes Nucleaires, Saclay, France, 1981.

Murakami, K., K. Nagashima, S. Sagisaka, Y. Mishima, and A. Inoue, "Response Functions for Cosmic-Ray Muons at Various Depths Underground," *Nuovo Cimento, 2C:* 635, 1979.

Neher, H.V., "Cosmic Ray Particles that Changed from 1954 to 1958 and 1965," *J. Geophys. Res.*, **72:** 1527, 1967.

Neher, H.V., "Cosmic Rays at High Latitudes and Altitudes Covering Four Solar Maxima," *J. Geophys. Res.*, **76:** 1637, 1971.

O'Brian, K., "Calculated Cosmic Ray Ionization in the Lower Atmosphere," *J. Geophys. Res.*, **75:** 4357, 1970.

O'Brian, K., "Secular Variations in the Production of Cosmogenetic Isotopes in the Earth's Atmosphere," *J. Geophys. Res.*, **84:**423, 1979.

Peddie, N.W., "International Geomagnetic Reference Field: The Third Generation," *J. Geomag. Geoelectr.*, **34:** 309, 1982.

Roelof, E.C., "New Aspects of Interplanetary Propagation Revealed by 0.3 MeV Solar Proton Events in 1967," in *Solar-Terrestrial Relations*, 411, University of Calgary, Canada, 1973.

Rossi, B., *High Energy Particles*, Prentice-Hall, New York, 1952.

Rossi, B., *Cosmic Rays*, McGraw Hill, New York, 1964.

Sellers, B., F.A. Hanser, M.A. Stroscio, and G.K. Yates, "The Night and Day Relationships betweeen Polar Cap Riometer Absorption and Solar Protons," *Radio Sci.*, **12:** 779, 1977.

Seltzer, S. "SHIELDOSE: A computer code for Space-shielding Radiation Dose Calculations," NBS Technical Note 1116, National Bureau of Standards, U.S. Department of Commerce, Washington, D.C., May 1980.

Shea, M.A. and D.F. Smart "A Five by Fifteen Degree World Grid of Calculated Cosmic-Ray Vertical Cutoff Rigidities for 1965 and 1975," 14th International Cosmic Ray Conference, *Conference Papers*, **4:** 1298–1303, 1975.

Shea, M.A., and D.F. Smart, "A World Grid of Calculated Cosmic Ray Vertical Cutoff Rigidities for 1980.0," 18th International Cosmic Ray Conference, *Conference Papers*, **3:** 415, 1983.

Shea, M.A., D.F. Smart, and John R. McCall, "A Five Degree by Fifteen Degree World Grid of Trajectory-Determined Vertical Cutoff Rigidities," *Can. J. Phys.*, **46:** S1098, 1968.

Simnett, G.M., "Relativistic Electron Events in Interplanetary Space," *Space Sci. Revs.*, **16:** 257, 1974.

Smart, D.F. and M.A. Shea, "Solar Proton Event Classification System," *Sol. Phys.*, **16:** 484, 1971.

Smart, D.F. and M.A. Shea, "PPS76 - A Computerized "Event Mode" Solar Proton Forecasting Technique," *Solar-Terrestrial Predictions Proceedings, Vol. 1:* Prediction Group Reports, edited by R.F. Donnelly, Environment Research Laboratories, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, Boulder, CO, 406, 1979.

Stassinopoulos, E.G., and J.H. King, "An Empirical Model of Energetic Solar Proton Fluxes with Application to Earth Orbiting Spacecraft," NASA X-601-72-487, 1972.

Stuiver, M. and P.D. Quay, "Changes in Atmospheric Carbon-14 Attributed to a Variable Sun," *Science*, **207:** 11, 1980.

Van Hollebeke, M.A.I., L.S. Ma Sung, and F.B. McDonald, "The Variation of Solar Proton Energy Spectra and Size Distribution with Heliolongitude," *Sol. Phys.*, **41:** 189, 1975.

Wenzel, K.-P., E.C. Stone, and R.E. Vogt, "Splash Albedo Protons Between 4 and 315 MeV at High and Low Geomagnetic Latitudes," *J. Geophys. Res.*, **80:** 358, 1975.

West, G.S. Jr., S.J. Wright, and H.C. Euler, (eds.), "Space and Planetary Environment Criteria Guidelines for use in Space Vehicle Developments," NASA-TM-78119, 1977.

Wibberenz, G., "Interplanetary Magnetic Fields and the Propagation of Cosmic Rays," *J. Geophys.*, **40:** 667, 1974.

Chapter 7

THE CHARGING OF SPACECRAFT SURFACES

H.B. Garrett

The buildup of static charge on satellite surfaces is an important issue in the utilization of satellite systems. The analysis of this spacecraft environmental interaction has required important advances in basic charging theory and the development of complex codes to evaluate the plasma sheaths that surround satellites. The results of these theories and calculations have wide application in space physics in the design of systems and in the interpretation of low energy plasma measurements.

In this chapter, those aspects of charge buildup on satellite surfaces relevant to the space physics community are summarized. The types of charging processes, models of charge buildup, satellite sheath theories, and charging observations are described with emphasis on basic concepts.

As many books and monographs on specific aspects of the charging of bodies in space have appeared in the last two decades, it is difficult to cover all areas in detail in a chapter of this nature. Rather, the chapter is limited to the charging of spacecraft surfaces in the near-earth magnetosphere. Rocket measurements are only briefly treated. The reader is referred to books by Singer [1965], Grard [1973a], Rosen [1976], Pike and Lovell [1977], and Finke and Pike [1979] that contain papers on the charging of natural bodies such as the moon [Manka, 1973; Freeman et al., 1973], dust particles [Feuerbacher et al., 1973] and other planetary bodies [Shawhan et al., 1973]. A brief historical review of spacecraft charging is followed by a discussion of observations. Following a description of the major charging mechanisms, specific spacecraft charging models are examined with emphasis on basic concepts. A substorm worst case environment is included. The chapter concludes with a discussion of spacecraft charge mitigation techniques. A more detailed version of this chapter can be found in Garrett [1981].

7.1 SPACECRAFT CHARGING HISTORICAL PERSPECTIVE

The historical roots of spacecraft charging analysis can be traced to the early electrostatic probe work of Langmuir [Langmuir, 1924; Mott-Smith and Langmuir, 1926]. Not only is the Langmuir probe still an important space plasma instrument, but as will be discussed, much of Langmuir's analysis is applicable to current spacecraft charging problems. This is true as in a very real sense the space vehicle itself can be considered as a "floating probe." This theme, of the vehicle as a probe, will form the basis of much of this chapter.

Probe theory has developed into an important subfield of plasma physics in its own right, see review by Chen [1965]. It has only been with the advent of rockets and ultimately satellites that the charging of objects in space has become a major separate area of concern. The first period of charging studies, in fact, was concerned with the potential of interstellar dust grains. One of the earliest studies, that of Jung [1937], found that photoemission and electron accumulation were probably the dominant processes in interstellar space. This subject was extended and put on a firm physical footing by Spitzer [1941; 1948], Spitzer and Savedoff [1950], and others [Cernuschi, 1947; Opik, 1956]. Depending on the assumed "sticking" probability of the electrons, the photoemission yield, and the ambient environment, the estimate satellite to space potentials for these early studies ranged between -3 and $+10$ V.

With the advent of rocket-borne sensors in the early 1950s, spacecraft charging emerged as a discipline. Perhaps the first example of a spacecraft charging effect is in a paper by Johnson and Meadows [1955]. They theorized that shifts in the ion peaks measured by their RF mass spectrometer above 124 km could be explained by a negative rocket potential of -20 V. The first treatment of the charging of a macroscopic object was published the next year by Lehnert [1956] who estimated satellite to space potentials of -0.7 to -1.0 V when ion ram effects were included.

Not only did 1957 see the launch of Sputnik, but it ushered in a second phase in spacecraft charging studies. Gringauz and Zelikman [1958] discussed the distribution of charge (or sheath) around a space vehicle in the ionosphere and the influence of photoemission and satellite velocity. Jastrow and Pearse [1957], while neglecting photoemission but including ram effects, computed the drag on a satellite caused by charged particles in the ionosphere. Their study for the ionosphere assumed $T_E \gg T_I$ so that the estimated potentials were between -10 V (night) and -60 V (day), values which are too high as $T_E \gtrsim T_I$ is the actual case [Brundin, 1963]. Another basic assumption of their study, as in the related study of Chang and Smith [1960], is that

the ion density is little changed in the immediate vicinity of the vehicle.

Zonov [1959] and Beard and Johnson [1960] analyzed the effects of electric fields induced by the movement of a satellite across the earth's magnetic field; these fields can be quite important for large structures. Beard and Johnson [1961] also discussed the effects of emitting charged particles from a vehicle and the limitations on vehicle potential in the ionosphere, another issue which is still of concern (see Section 7.4.1). In the same period, the first satellite potential measurements were made by Sputnik 3. Krassovsky [1959] found a potential of -6.4 V and $T_e = 15\,000$ K at 795 km.

The first review of spacecraft charging appeared in 1961 [Chopra, 1961]. Despite difficulties with this review (Chopra predicted high positive potentials since the photoelectron flux he assumed was too high), it can be used as a convenient marker for the end of the second phase of the study of spacecraft charging. By 1961 most of the elements of current spacecraft charging theory were in place. Preliminary observations by rockets and satellites had confirmed that charging existed and, in agreement with some theories, was on the order of a volt (typically negative) in the ionosphere. Photoemission and the ambient electron flux were recognized as dominant sources and $\mathbf{v} \times \mathbf{B}$ effects had also been considered. On the negative side, secondary emission and backscatter had not really been adequately considered [Whipple, 1965], charged particle drag (which was ultimately shown to be of less importance than errors in the neutral drag coefficient (see reviews by Brundin [1963] and deLeeuw [1967])) was still of primary concern, and self-consistent solutions of the particle trajectories and fields had not been carried out. Only monoenergetic or Maxwellian distributions were being considered.

The new phase of studies in spacecraft charging actually started somewhat earlier than 1961 with Bernstein and Rabinowitz's probe study. Bernstein and Rabinowitz [1959] introduced a means of calculating particle trajectories in the vicinity of a probe for the case where collisions could be ignored. The importance of their study (see Section 7.4.4) was that their method allowed a self-consistent solution of the time-independent Vlasov equation and Poisson's equation. Their work was later adapted to spacecraft by a number of investigators. Davis and Harris [1961], by a more simplified method, calculated the shielding of a rapidly moving sphere in the ionosphere by estimating ion trajectories in a fixed electron sheath. Their results concerning the distribution of ions in the wake, despite their neglecting the ion thermal velocity, are consistent with Explorer 8 measurements reported in the same year [Bourdeau et al., 1961]. These latter results are probably the first accurate spacecraft potential measurements between 425 km and 2300 km. The ion current, which varied strongly as a function of angle relative to the satellite velocity vector, was found to agree with Whipple's [1959] theory (see Equation (7.23)). A satellite potential of -0.15 V was observed along with $\mathbf{v} \times \mathbf{B}$ effects ($\sim$0.14 V). During this period more believable potential measurements were also reported for rocket probes as exemplified by Sagalyn et al. [1963] who found potentials of -0.4 V (150 km) to -1.7 V (450 km).

In a 1961 paper, Kurt and Moroz [1961] predicted potentials of -3.2 to 4 V outside the radiation belts. They also predicted potentials as high as -20 kV in the radiation belts. Although only crude estimates, their predictions anticipated the -20 kV potentials observed in eclipse on the geosynchronous ATS 6 satellite in the 1970s. These and other results were compared in an excellent review of the effects of charged particles on a satellite by Brundin [1963]. Another good review from this period, in which the validity of various ionospheric measurements were discussed, is that of Bourdeau [1963]. The most complete works of this period, however, are in the first book concerned with spacecraft charging [Singer, 1965] and the thesis of Whipple [1965]. Whipple's thesis brings together most of the preceding results in an analysis of the roles that secondary emission, backscatter, photoemission, and magnetic field effects have in spacecraft charging. As such, it completes the third period of spacecraft charge analysis—a period marked by a realization of the importance of spacecraft charging for plasma measurements and of the importance of self-consistent calculations. Quantitative measurements also became available for the first time. The period 1965 to the present has been primarily one of refinement and extension of these 1961–1965 results to higher altitude regimes and more complex geometric situations. It represents the "fourth period" of spacecraft charging and is the concern of this chapter.

7.2 SPACECRAFT CHARGING OBSERVATIONS

7.2.1 Rocket Measurements

Rocket measurements of the ionospheric plasma have been routine since the early 1950s. The first observations of spacecraft potentials were probably the RF spectrometer measurements on a rocket by Johnson and Meadows [1955]. They made use of the differences in the energy shifts of different ionized species to estimate a vehicle potential of -20 V above 120 km. Sagalyn et al. [1963], employing 2 spherical electrostatic analyzers mounted on a Thor rocket, found much lower potentials of -0.4 V (150 km) to -1.7 V (450 km). Their measurements are in good agreement with subsequent satellite measurements in the same region. Narcisi et al. [1968] also found rocket potentials of about -0.5 V in the D and E regions although their results may have been affected by the potential distribution near the rocket [Parker and Whipple, 1970]. As discussed by Parker and Whipple [1970], however, there are for these and similar satellite measurements difficulties in interpreting the results

since the detailed particle trajectories must be considered in determining the actual instrument responses. Further, as probably happened to Johnson and Meadows [1955], the electric fields near the rockets may be perturbed by exposed potential surfaces. See, for example, the results of the recent "tethered payload" experiment of Williamson et al. [1980], in which a potential of +10 V on the main payload induced a potential of −5 V on the secondary payload 40 m away.

Olsen [1980] (see also Winckler [1980]) has described potential measurements from a number of rocket beam experiments. These ranged from the Aerobee flight [Hess, 1969] through the Echo series [Hendrickson, 1972; Winckler, 1976] down to the recent ARAKS [Cambou et al., 1978], Precede, and Exceede [O'Neil et al., 1978a, and b]. Although extensive literature exists on many of these flights, the majority is concerned with the beam aspects. Potentials, when observed, were typically a few tens of volts positive or negative relative to the ambient plasma (Jacobsen and Maynard [1980], however, have reported potentials of hundreds of volts on the POLAR 5 rocket experiment). As an example, voltages of +4 V (108 km) to +28 V (122 km) were observed by Precede (electron beam voltage: ~2.5 kV; current: ~0. 8A O'Neil et al., [1978a]).

An exception to the rocket beam flights just mentioned was the "Spacecraft Charging Sounding Rocket Payload" [Cohen et al., 1979; Mizera et al., 1979]. This flight tested prototypes of the positive and negative charge ejection systems, the transient pulse monitor, and the rocket surface potential monitor subsequeatly flown on the P78-2 SCATHA satellite. Additionally, a thermal emissive probe, a bipolar-intersegment voltmeter, and a retarding potential analyzer were also flown. The unneutralized beams repeatedly varied the rocket ground potential between −600 V and +100 V (Figure 7-1, Cohen et al. [1979]). Potentials as high as +1100 V were observed on conducting surfaces and +400 V on insulators [Mizera et al., 1979] relative to spacecraft ground. The variations in the vehicle to space potential correlated well with the ambient plasma density.

7.2.2 Satellite Measurements

Satellite measurements at low earth orbit have been made primarily by retarding potential analyzers and similar current collection probes. The earliest satellite observations of spacecraft charging were by the ion trap experiment on Sputnik 3 which measured potentials in the −2 V to −7 V range [Krassovsky, 1959], however, Whipple [1959] estimated −3.9 V. As has been discussed, the first well documented measurements of satellite potential at low altitudes were by the Explorer 8 where potentials of −0.15 V were observed between 425 and 2300 km [Bourdeau et al., 1961]. Such low negative values are typical of this region: Reddy et al. [1967], −0.5 V at 640 km on TIROS 7; Samir [1973], −0.71 V to −0.91 V between 600 and 900 km (Explorer 31); Goldan et al. [1973], −0.7 V between 400 and 650

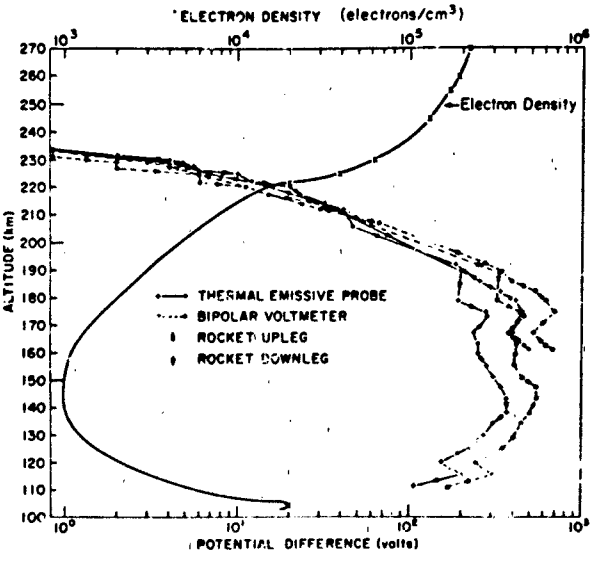


Figure 7-1. Vehicle-to-ambient potential difference and electron density (modeled) as a function of altitude [Cohen et al., 1979].

km (OGO 4); Sagalyn and Burke [1977], −1.5 V to 4 V in the plasma trough and −0.5 to −1 V in the polar cap at 2500 km (INJUN 5); and Samir et al. [1979 a, b], −0.1 to −1.3 between 275 to 600 km (AE-C). The highest values observed at low altitudes were by Hanson et al. [1964; 1970] who estimated potentials of −6 V at 240 km to −16 V at 540 km and Knudsen and Sharp [1967] who recorded −15 V at 516 km. Even higher values of −40 V were observed by Sagalyn and Burke [1977] at 2500 km on INJUN 5. These latter values are probably valid as they were observed in the auroral zone during impulsive precipitation events and at night. Although most of these results were for eclipse conditions, there was no unambiguous effect due to photoelectrons at these altitudes (this is not true for higher altitudes).

Higher voltage variations, particularly in sunlight, are seen in the plasmasphere proper at altitudes above 2500 km. On OGO 5, Norman and Freeman [1973] found potentials of −7 to −10 V at 1.1 R_I. Between 2 to 6 R_I, as the satellite crossed the plasmapause, the voltage varied from −5 to +5 V. At 8 R_I the potential reached +20 V (note that in eclipse the potential fell below −3.5 V). Ahmed and Sagalyn [1972], employing spherical electrostatic analyzers on OGO 1, calculated potentials of −3 to −6 V beyond the plasmapause and −11 to −8 V in the plasmasphere. On the same satellite, Taylor et al. [1965], employing an RF spectrometer, estimated −15 V at low altitudes (1500–2700 km) to −0 V at 30 000 km. As discussed, however, this potential variation may have resulted from the interaction of the exposed positive electrodes on the spacecraft solar cells with the environment. Whipple et al. [1974] reported potentials between 0 and −5.4 V in the plasmasphere on OGO 3 in sunlight. Montgomery et al.

[1973] observed potentials of +100 V in the high latitude magnetotail at 18.5 R_E on Vela 6. (As there were apparently no ion measurements at the time of these estimates, there may be some uncertainty in the method used as a result of possible differential charging [see Whipple, 1976b, or Grard et al., 1977].) During eclipse in the same region, they estimated the potential to be +15 V.

As the satellite potential in eclipse is proportional to the electron temperature, it is not surprising that the most spectacular potential variations have been for geosynchronous satellites in the plasmasheet ($T_E \sim 10$ keV). The best documented and most extensive set of such observations come from the University of California at San Diego (UCSD) particle experiments on the geosynchronous satellites ATS 5 and ATS 6. The large potentials observed by these satellites were first reported and explained by DeForest [1972; 1973], and provided a major impetus to the discipline of spacecraft charging. A typical example of a −10 000 V eclipse charging event in spectrogram format for day 59, 1976, for ATS 6 is presented in Figure 7-2a. In this type

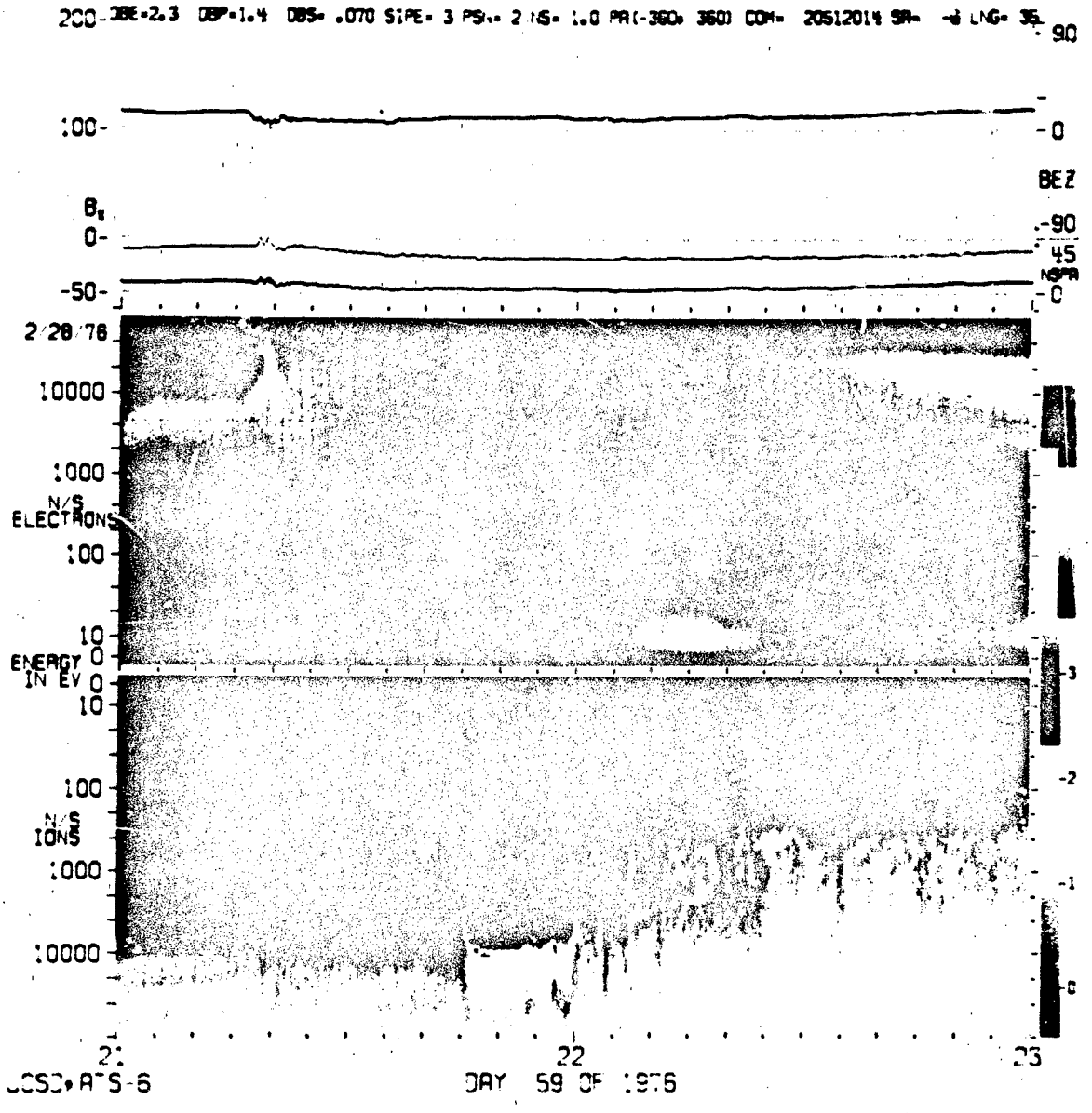


Figure 7-2a. Spectrogram of the UCSD particle detectors on ATS 6 for day 59, 1976 showing a 10-kV charging event between 2140 and 2200 UT.

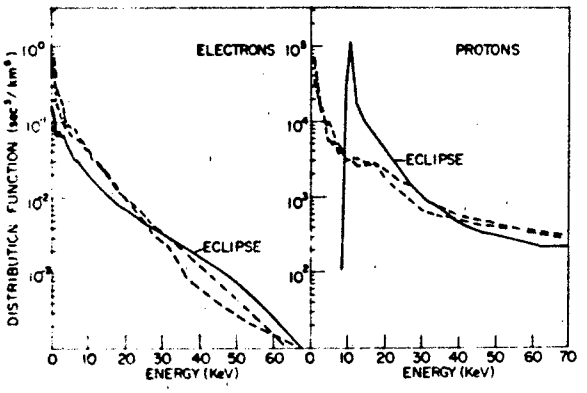


Figure 7-2b. Electron and ion distribution functions versus energy for day 59, 1976. (Dashed lines represent the spectra before and after eclipse; solid lines represent eclipse spectra.)

of plot [DeForest and McIlwain, 1971] the Y axis is particle energy, the X axis is time, and the Z axis (shading) is particle count rates. The figure is interpreted as follows. When the satellite entered eclipse (~2140 UT), the photoelectron flux went to zero shifting the already (100 V) negative potential even more negative. This negative satellite potential V_s accelerated the positive ions as they approached the satellite, adding energy qV_s to each particle. Zero energy positive ions thus appear as a bright band between 2140 and 2200 UT at an energy (10 keV) equal to q times the potential V_s. Electrons, in contrast, were decelerated giving the dropout in the electron spectra between 2140 and 2200 UT. The spectra for before and after eclipse are presented in Figure 7-2b demonstrating this 10 keV shift. Potentials as high as -2000 V [Reasoner et al., 1976] in sunlight and $-20\,000$ V in eclipse have been observed on ATS 6 [S.E. DeForest, 1978, private communication].

Another effect of potential variations, differential charging, is also visible in the ATS 5 and ATS 6 spectrograms. In Figure 7-3, a spectrogram from day 334, 1969, for ATS 5 is presented for the detector looking parallel to the satellite spin axis. Between 0500 and 1100 UT a feathered pattern is visible in the ions and a dropout in the electrons below 750 eV is observed. These patterns are not visible in the detectors looking perpendicular to the satellite spin axis. The explanation [DeForest, 1972; 1973] is that a satellite surface near the detectors has become differentially charged relative to the detectors resulting in a preferential focusing of the ion fluxes and a deficiency in the electron fluxes to the parallel detectors.

The frequency of occurrence of the ATS 5 and ATS 6 charging events in eclipse are presented in Figure 7-4 [Garrett et al., 1978]. The daylight charging events on ATS 6 have been found by Reasoner et al. [1976] (see also Johnson et al. [1978]) to be anticorrelated with encounters with the plasmapause. The level of charging has, however, been found to correlate with Kp (Figure 7-5). To be discussed later are the ATS 5 and ATS 6 beam experiments [Goldstein and DeForest, 1976; Olsen, 1980; and Purvis and Bartlett, 1980]. These studies presented evidence that while electron emission reduced satellite charging, neutral plasma emission was necesssary to achieve zero satellite to space potentials.

Limited observations are availiable in the solar wind and in the vicinity of the other planets. The Vela satellites typically experienced potentials of +3 to +5 V in the solar wind [Montgomery et al., 1973]. Whipple and Parker [1969a] computed potentials of +2.2 V and +10 V for OGO 1 and IMP 2, respectively, in the solar wind. The Voyager spacecraft observed potentials of +1 V to +10 V [Scudder et al., 1981]. Although preliminary, similar potentials were apparently measured by the Voyager in Jupiter's outer magnetosphere while slightly negative potentials occurred inside the denser, cooler regions of Io's plasma torus [Scudder et al., 1981; see also comments in Grard et al., 1977].

The most complete spacecraft charging measurements are those being made currently by the P78-2 SCATHA satellite (Figure 7-6). P78-2 SCATHA was launched in January of 1979 into a near-synchronous orbit (5×7 R_E). The satellite is specifically designed to study spacecraft charging as is evidenced by the extensive list of scientific and engineering instruments (Table 7-1). P78-2 SCATHA has confirmed current ideas concerning the charging process. At the same time, information on the sheaths surrounding a satellite has been obtained. An abbreviated list of observations follows:

1. Arcs were observed under different potential conditions (eclipse, sunlight, beam operations, etc. [Koons, 1980]).
2. The satellite surface potential monitor (SSPM) has determined the response of a number of materials to both natural and artificial charging events [Mizera, 1980]. Sample potentials of over -1000 V have been observed relative to spacecraft ground (see Figure 7-26b).
3. Natural charging events of -1000 V or greater have been observed with one event in excess of -8000 V.

These observations are the first to include simultaneous data on the plasmas, magnetic and electric fields, surface potentials of dielectrics and other surface materials, arcing, and surface contaminants.

The SCATHA data set has been of particular value in defining a "worst case" charging environment for the geosynchronous orbit. In Table 7-2 are listed two "worst case" examples from SCATHA as adapted from Mullen et al. [1981] and Mullen and Gussenhoven [1982]. For comparison with earlier estimates, a "worst case" example from ATS 6 has also been included [Deutsch, 1981]. The plasma moments and single maxwellian temperatures [described in Garrett, 1979] are averaged over all angles while for SCATHA the 2-maxwellian values are for components parallel and perpendicular to the magnetic field. The SCATHA example on Day 114 (24 April 1979, 0650 UT, 23.1 MLT)

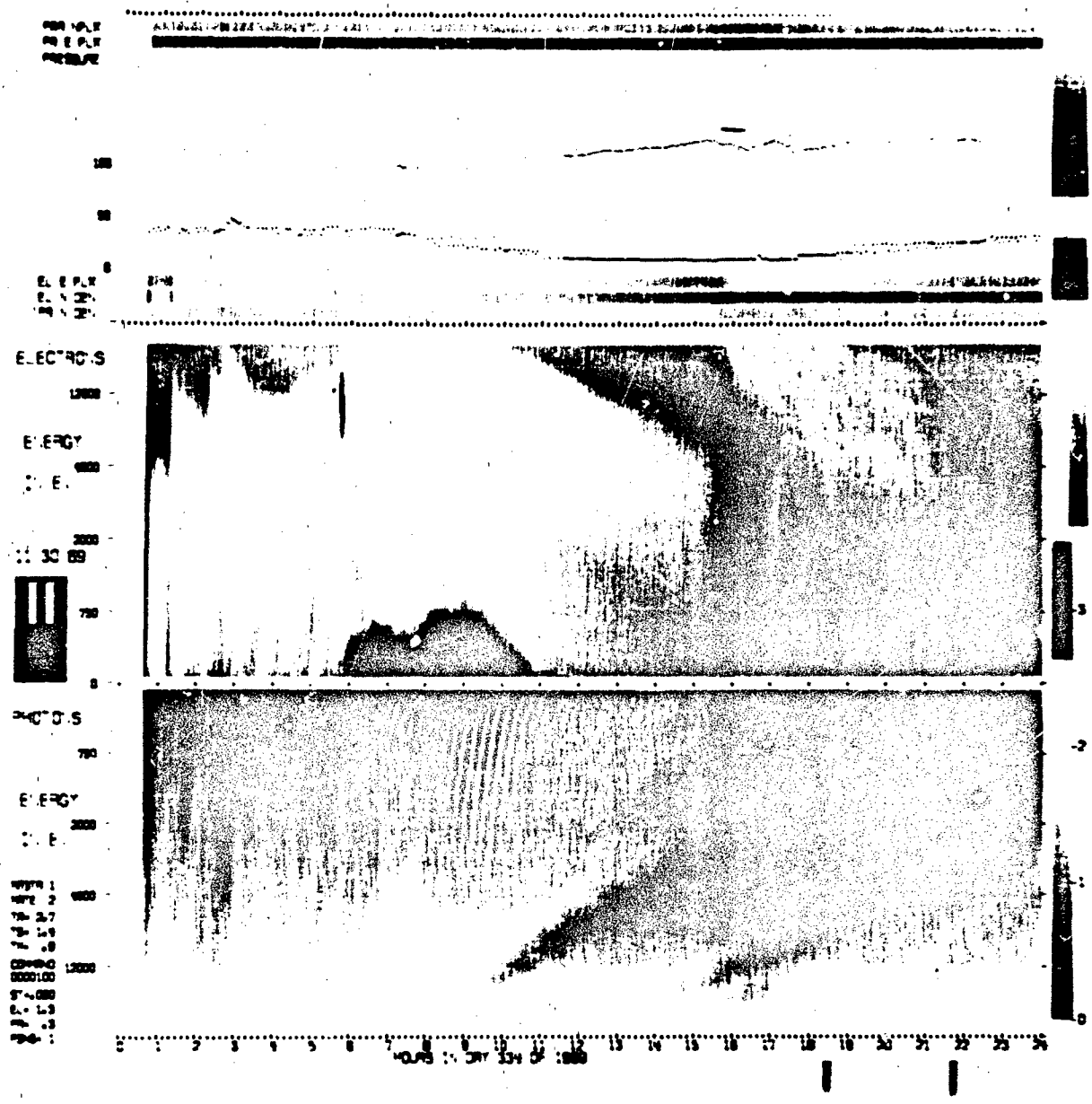


Figure 7-3 Day 334, 1969, spectrogram from ATS 5 [DeForest, 1973]. The feathered pattern in the ions between 0700 and 1100 UT and the corresponding loss of low energy electrons is real and the result of differential charging.

was particularly well documented. For this example, it was found that the vehicle frame potential closely followed the electron current between 33 keV and 335 keV. It is this current that apparently caused the observed high negative vehicle potentials (-340 V in sunlight; estimated to be -16 kV in eclipse). Maximum surface material charging levels for this event were: -3.8 kV on quartz fabric, -6.4 kV on silvered Teflon, and -1.5 kV on aluminized Kapton. The average ratio of oxygen ions to hydrogen ions was 0.4 during the event.

7.2.3 The Effects of Spacecraft Charging

Aside from direct measurements of spacecraft charging, indirect indicators that reflect the effects of spacecraft charg-

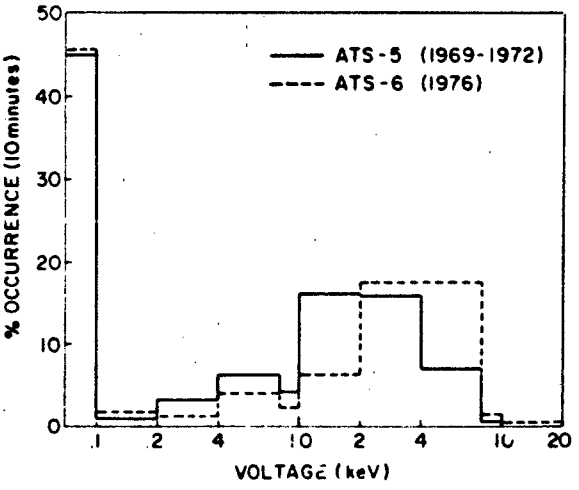


Figure 7-4 Occurrence frequency of ATS 5 (1969-1972) and ATS 6 (1976) eclipse potentials (10-min intervals)

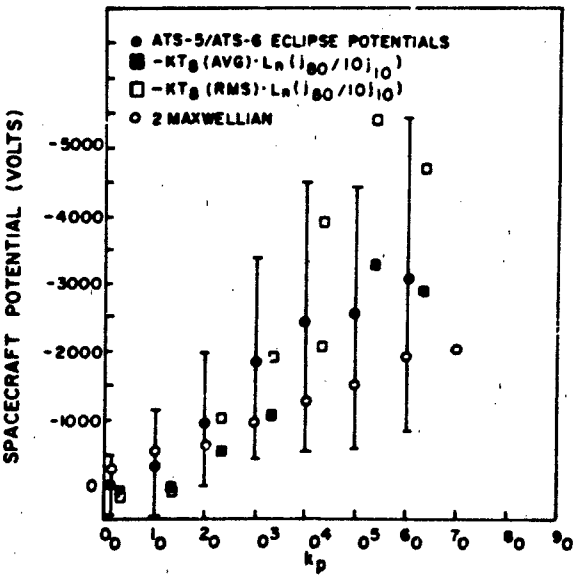


Figure 7-5 Statistical occurrence frequency of observed variations of ATS 5 and ATS 6 eclipse potentials as a function of Kp (solid dots) and various theoretical predictions [Garrett et al., 1979]

ing also exist. Typical is the plot of satellite operational anomalies as a function of local time at geosynchronous orbit presented in Figure 7-7 [McPherson and Schober, 1976]. The cause of this clustering near local midnight is believed to be spacecraft charging—intense fluxes of energetic electrons associated with injection events are encountered near local midnight which lead to charge buildup and arcing and hence to control circuit upsets and operational anomalies.

Given that differential charging can take place, whether through potential differences on adjoining surfaces or through charge deposition in dielectrics, arcing can occur. Arcing, defined as the rapid (~nanosecond) rearrangement of charge by punchthrough (breakdown from dielectric to substrate), by flashover (propagating subsurface discharge), blowoff (arc to surface), between surfaces, or between surfaces and space, is not well understood. A typical arc discharge pulse [Balmain et al., 1977] is plotted in Figure 7-8. Balmain [1980] finds that surface discharges display characteristics that scale with variations in specimen area according to well defined power laws: peak current scales as the 0.50 power of the area, released charge as 1.00, energy dissipation as 1.50, and pulse duration as 0.53. Although still preliminary, various attempts are also underway to theoretically model these arcing phenomena [Muelenberg, 1976; Beers et al., 1979].

Another result for surface arcs [Stevens, 1980; Nanevicz and Adamo, 1980] in the laboratory is that the breakdown potential on a negative surface varies from −100 V at low earth orbit to −10 000 V at geosynchronous orbit implying that arcing should not be a common occurrence. In Figure 7-9 [Shaw et al., 1976] the arcing rate on a geosynchronous satellite shows a steady increase with the daily geomagnetic index ap. As arcing is common even at low levels of geomagnetic activity, a discrepancy exists between laboratory

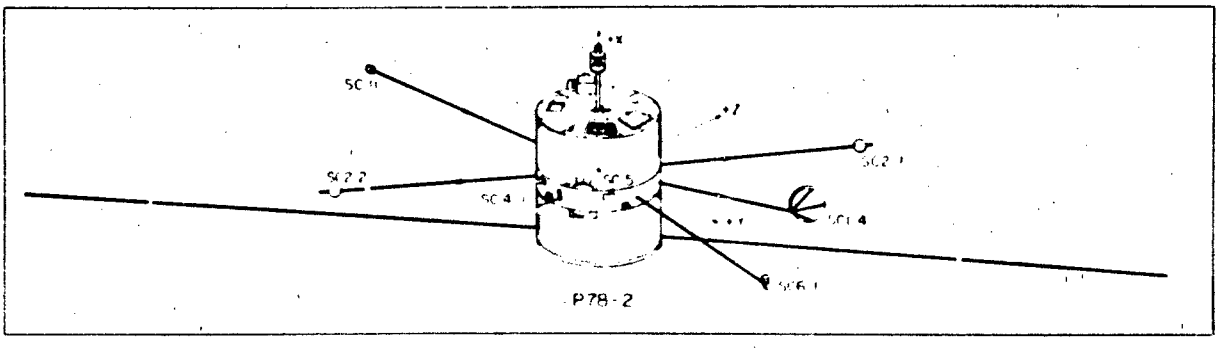


Figure 7-6 The P78-2 SCATHA satellite. The dimensions are approximately 1.3 m wide by 1.5 m high

CHAPTER 7

Table 7-1 Principal investigators/sponsors for P78-2 SCATHA

Experiment Number	Title	Principal Investigator/ Sponsor	Address
SC1	Engineering Experiments	Dr. H.C. Koons/ USAF/AFSC/SD	The Aerospace Corporation P.O. Box 92957 Los Angeles, CA 90009
SC2	Spacecraft Sheath Electric Fields	Dr. J.F. Fennell/ USAF/AFSC/SD	The Aerospace Corporation P.O. Box 92957 Los Angeles, CA 90009
SC3	High Energy Particle Spectrometer	Dr. J.B. Reagan/ Office of Naval Research	Lockheed Palo Alto Rsch Lab 3251 Hanover Street Palo Alto, CA 94304
SC4	Satellite Electron and Positive Ion Beam System	H.A. Cohen/ USAF/AFSC/AFGL	AFGL/PHG Hanscom AFB, MA 01731
SC5	Rapid Scan Particle	Capt. D. Hardy/ USAF/AFSC/AFGL	AFGL/PHG Hanscom AFB, MA 01731
SC6	Thermal Plasma Analyzer	P.C. Sagalyn/ USAF/AFSC/AFGL	AFGL/PH Hanscom AFB, MA 01731
SC7	Light Ion Mass Spectrometer	Dr. D.L. Reasoner/ Office of Naval Research	NASA Marshall Space Flight Center, Code BS-23 Huntsville, AL 35815
SC8	Energetic Ion Composition Experiment	Dr. R.G. Johnson/ Office of Naval Research	Lockheed Palo Alto Rsch Lab 3251 Hanover Street Palo Alto, CA 94304
SC9	UCSD Charged Particle Experiment	Dr. E.C. Whipple/ Office of Naval Research/USAF/AFSC/ SD	University of California B019 Dept. of Physics La Jolla, CA 92093
SC10	Electric Field Detector	Dr. T.L. Aggson/ Office of Naval Research	NASA Goddard Space Flight Center, Code 625 Greenbelt, MD 20771
SC11	Magnetic Field Monitor	Dr. B.G. Ledley/ Office of Naval Research	NASA Goddard Space Flight Center, Code 625 Greenbelt, MD 20771
ML12	Spacecraft Contamination	Dr. D.F. Hall/ USAF/AFSC/AFML	The Aerospace Corporation P.O. Box 92957 Los Angeles, CA 90009
TPM	Transient Pulse Monitor	Dr. J.E. Nanevicz/ USAF/AFSC/SD	SRI Menlo Park, CA

Table 7-2 "Worst" case geosynchronous environments. The moments, TAVG, and TRMS are averaged over all angles. The SCATHA 2-Maxwellian parameters are for fluxes parallel and perpendicular to the magnetic field. ATS 6 2-Maxwellian parameters are averaged over all directions.

SOURCE		DEUTSCH [1981]		MULLEN ET AL. [1981]		MULLEN AND GUSSENHOVEN [1982]	
DATE		DAY 178, 1974 ATS 6		DAY 114, 1979 SCATHA		— SCATHA	
		ELECTRONS	IONS	ELECTRONS	IONS	ELECTRONS	IONS
(ND) (cm^{-3})		0.112E + 01	0.245E + 00	0.900E + 00	0.230E + 01	0.300E + 01	0.300E + 01
(J) (nA cm^{-2})		0.410E + 00	0.252E − 01	0.187E + 00	0.795E − 02	0.501E + 00	0.159E − 01
(ED) (eV cm^{-3})		0.293E + 05	0.104E + 05	0.960E + 04	0.190E + 05	0.240E + 05	0.370E + 05
(EF) (eV cm^{-2} s^{-1} sr^{-1})		0.264E + 14	0.298E + 12	0.668E + 13	0.430E + 13	0.151E + 14	0.748E + 12
N1 (cm^{-3})	parallel	—	0.882E − 02	0.200E + 00	0.160E + 01	0.100E + 01	0.110E + 01
	perpendicular	—	—	0.200E + 00	0.110E + 01	0.800E + 00	0.900E + 00
T1 (eV)	parallel	—	0.111E + 03	0.400E + 03	0.300E + 03	0.600E + 03	0.400E + 03
	perpendicular	—	—	0.400E + 03	0.300E + 03	0.600E + 03	0.300E + 03
N2 (cm^{-3})	parallel	0.122E + 01	0.236E + 00	0.600E + 00	0.600E + 00	0.140E + 01	0.170E + 01
	perpendicular	—	—	0.230E + 01	0.130E + 01	0.190E + 01	0.160E + 01
T2 (eV)	parallel	0.160E + 05	0.295E + 05	0.240E + 05	0.260E + 05	0.251E + 05	0.247E + 05
	perpendicular	—	—	0.248E + 05	0.282E + 05	0.261E + 05	0.256E + 05
TAVG (eV)		0.160E + 05	0.284E + 05	0.770E + 04	0.550E + 04	0.533E + 04	0.822E + 04
TRMS (eV)		0.161E + 05	0.295E + 05	0.900E + 04	0.140E + 05	0.733E + 04	0.118E + 05

and *in situ* measurements which underscores the need for further analysis.

The effects of arcing are somewhat better understood than the process itself. The current pulse (Figure 7-8) generates an electrical pulse in spacecraft systems either by direct current injection or by induced currents due to the associated electromagnetic wave (see Proceedings of the IEEE Conference on Nuclear and Space Radiation Effects, 1979). Balmain [1980] and Nanevicz and Adamo [1980] have analyzed the effects of arc discharges on material surfaces. Balmain [1980] gives numerous examples of holes and channels of micron size in dielectric surfaces. Nanevicz and Adamo [1980] find additional, large scale physical damage to solar cells such as fracturing of the cover glass.

Of more immediate concern to the space physics community than arcs, however, are the effects of spacecraft charging on plasma measurements. There are numerous ways that charging can complicate the interpretation of low energy plasma measurements. These can be loosely defined as shifting of the spectra in energy, preferential focusing or exclu-

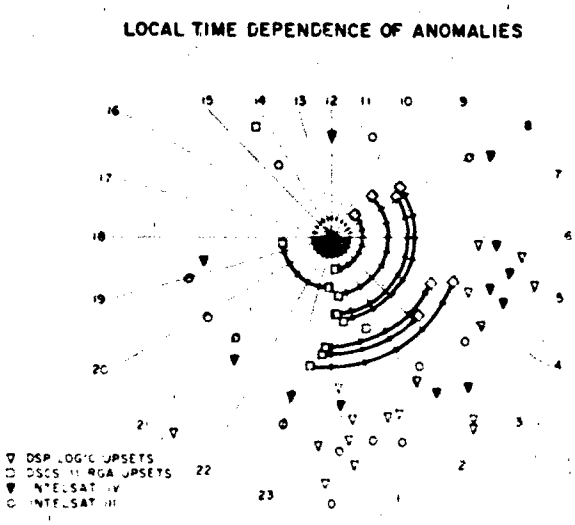


Figure 7-7. Local time plot of satellite operations anomalies; radial distance has no meaning [McPherson and Schober, 1977].

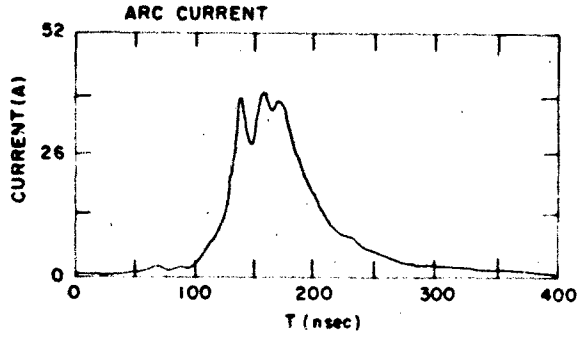


Figure 7-8. Oscilloscope tracing of discharge currents into a conductor supporting a mylar specimen [Balmain et al., 1977].

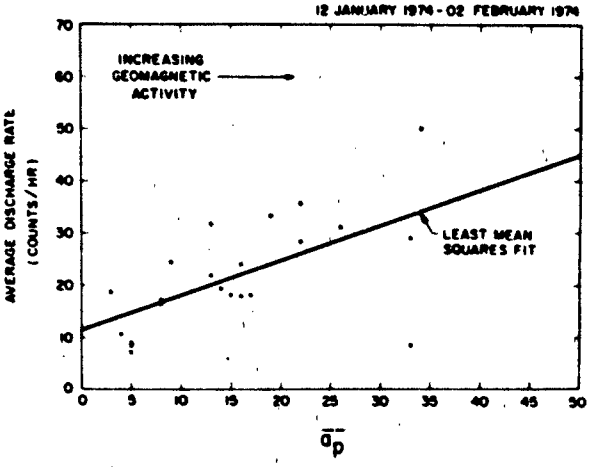


Figure 7-9 Number of arcs per hour as a function of daily ap for a geosynchronous satellite [Shaw et al., 1976].

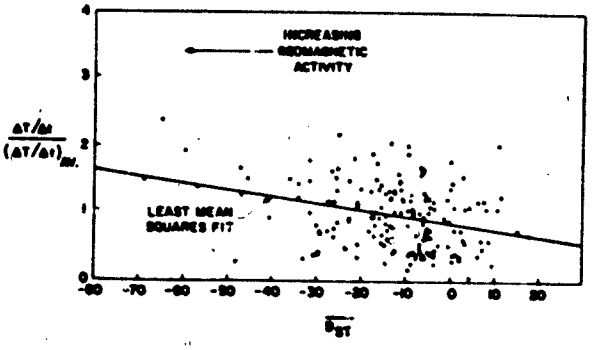


Figure 7-10 Heating rate per day divided by the average heating rate for the entire time period as a function of the daily Dst index [Nanevicz and Adamo, 1980].

sion of particles of a particular energy or direction, and contamination of measurements by secondaries, backscattered electrons, and photoelectrons. Each of these effects is briefly described below.

As described earlier in reference to the ATS 6 observations (Figure 7-2), a potential difference between the vehicle and the ambient plasma can raise or lower the energy of the incoming particles. In the case of differential flux measurements such as the ATS 6 electrostatic analyzers, the shift in energy is easily detected in the attracted species (Figure 7-2). For devices that measure the total current (such as a Langmuir probe) more subtle techniques must be employed which involve intimate knowledge of the current-voltage characteristic.

The most difficult effects to correct are due to particle focusing or exclusion. As demonstrated in Figure 7-3, differential charging of surfaces near a detector can result in focusing or defocusing at specific energies. Just as the geometry of the vehicle can shadow the field of view of a detector, the potential gradients near the detector can distort the particle trajectories. Methods for estimating particle fluxes in the vicinity of such charged surfaces are discussed in Section 7.4.

Secondary electrons, backscattered electrons, and photoelectrons are emitted by spacecraft surfaces. If the vehicle is positively charged, these low energy particles can be reattracted to the vehicle. They constitute an extra and undesirable current source that can confuse ambient electron measurements. Even in the case of negative potentials, differential potentials or space charge potential minima can result in significant secondary and backscattered electrons and photoelectron currents into the detector. This current is easily misinterpreted as a positive potential effect if only electron measurements are available.

Contaminant ions, due to thrusters (ionic or chemical) or outgassing of satellite materials, can be trapped within the satellite sheath and preferentially deposited on negatively charged spacecraft surfaces. Cauffman [1973] (see also Jemiola [1978; 1980]) has estimated that as much as 50 Å of material can be deposited on charged optical surfaces in as little as 100 days. In Figure 7-10 [Nanevicz and Adamo, 1980] the heating rate of sensors on a geosynchronous satellite apparently rose with geomagnetic activity. This is believed to be due to increased contaminant deposition during periods of geomagnetic activity and, therefore, increased charging (Figure 7-5). Such deposition may also alter secondary emission and photoelectron properties.

Another effect related to charging observed in the laboratory is parasitic power loss. This is anticipated to be important at low orbital altitudes due to interactions between the ambient plasma and exposed high voltage surfaces such as solar cells. McCoy et al. [1980] and Stevens [1980] have estimated this effect both theoretically and by laboratory experiments in large vacuum chambers to result in a 10% power loss for voltages in excess of 5000 V at low earth orbit. They also find that arcing begins at low earth orbit at -100 V and is a significant problem for potentials of -1000 V at geosynchronous orbit. Apparently, the power loss and damage due to such arcing is much more serious on high voltage arrays than parasitic power loss.

7.3 CURRENT MECHANISMS

The spacecraft charging effects reported in the previous section result from current balance, that is, all the various currents of charged particles to and from the satellite surface must balance. Such current balance is valid for a uniform, conducting satellite surface or for a non-conducting surface in the limit of a point. Unless otherwise stated, the former case (for a uniform conducting satellite) will be assumed (if

a non-uniform satellite surface is considered, then the current balance for each isolated surface and the resistive, capacitive, and inductive currents between surfaces must also be considered). The calculation of the satellite surface potential consists of 2 steps. First, the currents to the satellite surface are determined as functions of ambient conditions, satellite geometry, and potential. Second, a satellite potential is found so that current balance is achieved. The calculation of the currents to the satellite surface is very dependent on the fields in the vicinity of the space vehicle. These not only depend on the geometry of the vehicle but also on the sheath or the cloud of charged particles trapped near the satellite. The formulas for calculating the current sources, given the incident particle distribution at each satellite surface, will be described in this section.

7.3.1 Time Scales

In determining the validity of the assumption of current balance, an important issue to consider is the time scales for which it is applicable. Basic electrostatic considerations give order of magnitude estimates of these time scales. As an example, assume the satellite is a conducting sphere of radius r and has a capacitance of $C_s (\propto r)$. The time scale $\tau_{s/s}$ for the charging of a conducting sphere relative to space in the earth's magnetosphere is then [Katz et al., 1977]

$$\tau_{s/s} \approx \frac{C_s V}{4\pi r^2 J} \approx 2 \times 10^{-4}\,s \qquad (7.1)$$

where $r \approx 1$ m, V is the satellite potential relative to space (~ 1 kV), and J is the ambient flux (~ 0.5nA cm^{-2}).

Unfortunately satellites normally are covered with thermal blankets that consist of thin dielectrics deposited over conducting substrates. The capacitance of a given dielectric area A of thickness s can be estimated by $C_D (\propto A/4\pi s)$ so that the time scale τ_D is [Katz et al., 1977]

$$\tau_D \approx \frac{C_D V}{AJ} \approx 1.6s \qquad (7.2)$$

where $s \approx 0.1$ cm, $V \approx 1$ kV, and $J \approx 0.5$nA cm^{-2}.

Other important time scales are the charging time of large, isolated surfaces relative to each other (from seconds to perhaps hours, depending on surface details), the plasma frequency (10^4-10^7 Hz), and the gyrofrequency (10^3-10^6 Hz for electrons and 10-10^3 for H^+); lower values are representative of geosynchronous orbit where $n \approx 1$ cm^{-3} and $B \approx 100$nT, and the higher values for the ionosphere, where $n \approx 10^6$ cm^{-3} and $B \approx 3 \times 10^4$nT. Typically environmental changes take place in minutes or longer, although time scales on the order of the plasma frequency and gyrofrequency are observed. Arcing durations are 10^{-6}-10^{-5} s. These time scales imply at least three different ranges for the validity of current balance. Except for the highest frequency ambient electron variations, for plasma frequency variations at low altitudes, and arcing, satellite to space current balance can be realized. Thin dielectrics can respond to typical environmental variations (note this includes ambient variations due to satellite spin modulations, which are usually a few seconds to a minute). Current balance of large surfaces relative to each other may, however, not be achieved (this behavior is illustrated in Figure 7-25b). Thus, current balance is expected to be valid in a number of interesting cases although care must be exercised in the vicinity of the plasma frequency and, in the other extreme, of large isolated dielectric surfaces.

7.3.2 Incident Particle Fluxes

The major natural source of potentials of 10 kV or higher on satellite surfaces is the ambient space plasma. Although space plasma is seldom representable in terms of a single temperature and density, the Maxwell-Boltzmann distribution function is a useful starting point for describing the ambient plasma conditions that generate these large potentials. Given the distribution function f for an isotropic Maxwell-Boltzmann plasma

$$f(v_i) = n_i \left(\frac{m_i}{2\pi kT_i}\right)^{3/2} \exp\left(\frac{-m_i v_i^2}{2kT_i}\right) \qquad (7.3)$$

where n_i = number density of species i, m_i = mass of species i, T_i = temperature of species i, v_i = velocity of species i, k = Boltzmann constant, and f = distribution function.

The current flux to a surface in the absense of an electric field is

$$J_{no} = q_i \int\int\int v_i \cdot \mathbf{n}\, f\, d^3V, \qquad (7.4a)$$

Assuming isotropy,

$$J_{no} = \left(\frac{q_i n_i}{2}\right)\left(\frac{2kT_i}{\pi m_i}\right)^{1/2} \qquad (7.4b)$$

where J_{no} is the current density per unit area for 0 potential, $\mathbf{n}$ is the unit normal to surface, q_i is the charge on species i, and d^3V is the volume element in velocity space.

As will be shown in Section 7.4, when the effects of the spacecraft potential, sheath or plasma anisotropies, and deviations of the ambient plasma from a Maxwellian distribution are considered, the simple distribution function of Equation (7.3) is no longer valid and the integration of Equation (7.4) becomes difficult. Even so, Equation (7.4)

or a modification of it, is accurate for many practical purposes. Approximate values of T and n for various space plasmas are tabulated in Table 7-3.

7.3.3 Photoelectron Currents

The photoelectron current from a surface is a function of satellite material, solar flux, solar incidence angle, and satellite potential (see review by Lucas, [1973]). In Figure 7-11 adapted from Grard [1973b] are plots of the 4 functions necessary to describe the photoelectron current. In Figure 7-11 the solar flux S is plotted as a function of energy E (or wavelength). The details of the spectrum change with solar activity and can vary greatly if the sunlight reaching the spacecraft is attenuated by the atmosphere [Garrett and DeForest, 1979]. Also shown in Figure 7-11 are the electron yield per photon for normal incidence, W (E), and the total photoelectron yield H(E) (= W(E)S(E)) as functions of energy for aluminum oxide [Grard, 1973b]. The total current density for zero or negative satellite potential and normal incidence, as derived by Grard, is

$$J_{PHO} = \int_0^\infty W(E)\, S(E)\, dE = \int_0^\infty H(E)\, dE. \quad (7.5)$$

J_{PHO} for a variety of materials is tabulated in Table 7-4 [Grard, 1973b].

If the satellite is positively charged, the ambient photoelectron current is attracted to the surface. As this return current is a function of potential and geometry the energy spectrum of the electrons for a given incident monochromatic photon must be known to calculate it accurately. Grard [1973b] has carried out these calculations for several materials and different probe geometries (see also Whipple, [1965]).

Table 7-3. Estimated plasma parameters for various environments. Most values are rough estimates. See appendix.

Region	Altitude	N_e (cm^{-3})	Ions	Characteristic Energy (eV)		λ_D (m)		V (km/s)	J_{PH} (nA-cm^{-2})	Potential (V)**					
										Sunlight			Eclipse		
				I*	E	I*	E			1-D	RAM	3-D	1-D	RAM	3-D
Venus															
	200km	10^5	O^+, O_2^+	0.05	0.3	0.005	0.01	8	8	-1.2	-1.0	-0.83	-1.8	-1.2	-0.88
	1500km	10^5	O^+	0.2	1	0.33	0.74	8	8	-6.0	-6.0	-2.4	-5.6	-4.4	-2.9
Earth															
	150km	10^5D*	O^+, O_2^+, NO^+	0.1	0.2	0.007	0.01	8	2	-1.1	-0.7	-0.55	--	--	--
		10^5N	NO^+	0.05	0.1	0.05	0.07	8	2	--	--	--	-0.58	-0.33	-0.37
	1000km	10^4D	O^+	0.3	0.4	0.04	0.05	8	2	-1.3	-1.2	-1.2	--	--	--
		10^4N	H^+	0.2	0.2	0.03	0.03	8	2	--	--	--	-0.75	-0.73	-0.52
	$3.5R_E$	10^3	H^+	1	1	0.23	0.23	3.7	2	-1.6	-1.6	-1.4	-3.8	-5.2	-2.5
Geosynchronous	$5.62R_E$	2	H^+	5000	2500	370	260	3	2	-2.0	-1.9	-2.0	-8500	-23000	-6500
High Latitude		0.1	H^+	200	200	330	330	800	2	-15	-15	-15	-750	-490	-500
Jupiter															
Cold Torus	$3.5R_J$	50	S^+, O^+, O^{++}	0.5	0.5	0.74	0.74	44	0.08	-7.5	-5.9	-7.2	-2.3	-1.2	-1.6
	$5.5R_J$	1000		2	1	0.33	0.23	69	0.08	-3.8	-2.2	-3.1	-4.2	-2.3	-3.3
Hot Torus	$6.0R_J$	1000	S^+, O^+	40	10	1.5	0.74	75	0.08	-37	-34	-33	-39	-34	-33
	$8.0R_J$	100		80	20	6.6	3.3	100	0.08	-65	-60	-60	-78	-70	-65
Plasma Sheet	$8.0R_J$	12	H^+, S^+	50	50	15	15	150	0.08	-110	-110	-94	-190	-170	-130
	$15R_J$														
Outer Magnetosphere	--	0.01	H^+	1000	1000	230	230	250	0.08	-9.6	-9.5	-9.5	-3600	-4400	-2500
Solar Wind															
	0.3AU	50	H^+	40	65	6.6	8.5	500	20	-4.6	-4.9	-4.4	-260	-150	-160
	1.0AU	2	H^+	10	50	17	37	450	2	-7.8	-8.0	-7.3	-230	-120	-110
	5.2AU	0.2	H^+	1	10	17	53	400	0.08	-7.4	-8.0	-6.0	-50	-18	-21

*D = Day, N = Night

**See Appendix for description of computations and captions.

Underlined values are "preferred" estimates.

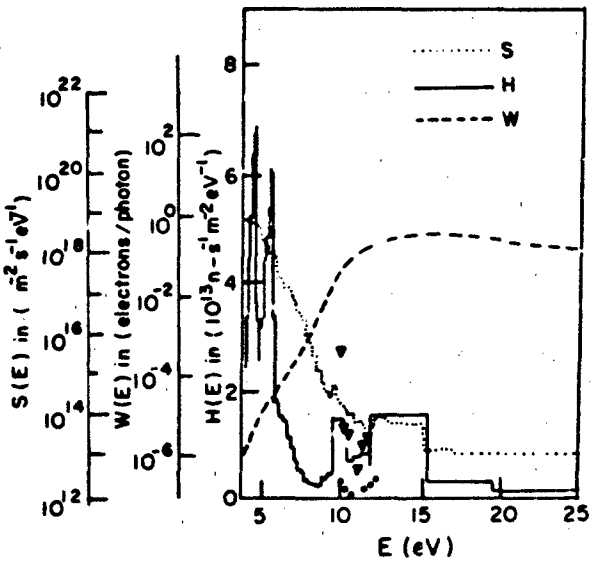


Figure 7-11. Composite plot of W(E), the electron yield per photon; S(E), the solar flux; and their product, H(E), the total photoelectron yield, as function of energy E for aluminum oxide [Grard, 1973].

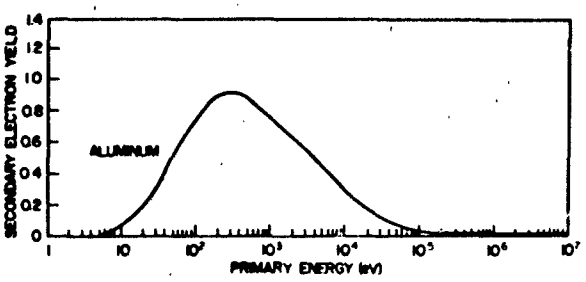


Figure 7-12a. Secondary electron yeild δ_i due to incident electrons of energy E impacting on aluminum [Whipple, 1965]

7.3.4 Backscattered and Secondary Electrons

The impact of ambient electrons and ions on a spacecraft surface generates backscattered and secondary electrons—backscattered and secondary ion fluxes being insignificant. These fluxes, though often neglected in charging calculations can exceed the incident fluxes under some circumstances. Although a clear distinction between secondary and backscattered electrons is not always possible, backscattered electrons are those ambient electrons reflected back from the surface with some loss of energy [Sternglass, 1954] (see also reviews by Dekker [1958], Hachenberg and Brauer [1959], Gibbons [1966], and Lucas [1973].) Secondary electrons are emitted as a result of energy deposition by incident electrons or ions [Sternglass, 1957; Willis and Skinner, 1973, Chung and Everhart, 1974; Baragiola et al., 1979]. Each of the three has a characteristic emission spectrum.

The equation for the current density due to secondary emission, assuming an isotropic flux and ignoring other angular variations is (see, however, Whipple, [1965]):

$$J_{s_i} = \frac{2\pi q_i}{m_i^2} \int_0^\infty dE' \int_0^\infty g_i(E', E)\, \delta_i(E) E\, f_i(E)\, dE \quad (7.6)$$

where J_{s_i} = secondary electron flux due to incident species i (usually assumed to be electrons, e^-, or protons, H^+), g_i = emission spectrum of secondary electrons due to incident species i of energy E, δ_i = secondary electron yield due to incident species i of energy E, E' = secondary electron energy, and f_i = distribution function of incident particles at surface.

Typical curves for δ_i for e^- and H^+ impacting on aluminum from Whipple [1965] are plotted in Figure 7-12a and 7-12b (see also Sternglass [1957], Willis and Skinner [1973], Chung and Everhart [1974], and Baragiola et al. [1979]). The function g_i is assumed to be independent of incident energy and incident particle species. The normal-

Table 7-4. Photoelectron emission characteristics of various spacecraft materials

Material	Work Function (eV)	Photoelectron Saturation Flux ($10^{12} n_e$/s-m^2)	Saturation Current Density (μ A/m^2)
Aluminum Oxide	3.9	260	42
Indium Oxide	4.8	190	30
Gold	4.8	180	29
Stainless Steel	4.4	120	20
Aquadog	4.6	110	18
LiF on Au	4.4	90	15
Vitreous Carbon	4.8	80	13
Graphite	4.7	25	4

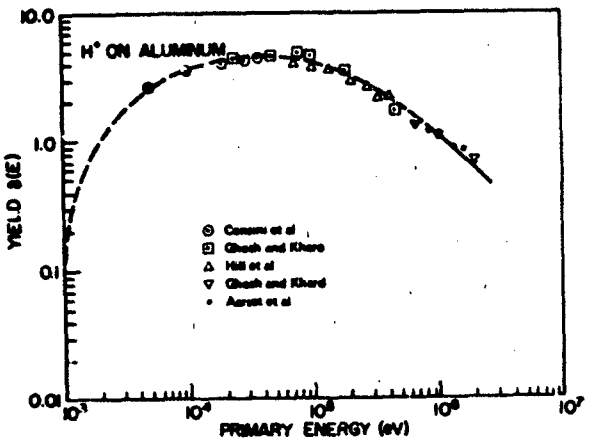


Figure 7-12b. Secondary electron yield δ_i due to incident ions of energy E impacting on aluminum [Whipple, 1965].

ized curve [Whipple, 1965] for aluminum is plotted in Figure 7-12c.

For backscattered electrons, the current density is given by DeForest [1972]

$$J_{BSE} = \frac{2\pi q_E}{m_E^2} \int_0^{\infty} dE' \int_E^{\infty} B(E',E)\, E\, f_E(E)\, dE, \quad (7.7)$$

where

$$B(E', E) = G\left(\frac{E'}{E}\right)\frac{1}{E}$$

and G is the percentage of electrons scattered at fraction E'/E of the incident energy E.

Sternglass [1954] has published experimental measurements of backscatter parameters for different materials. An estimate of G for Al from his data is plotted in Figure 7-13. For negative potentials, the backscatter flux is roughly 20% of the incident flux.

For both secondary and backscattered currents, the actual values are dependent on angle of incidence which has been ignored in the preceding discussion. Further, the secondary and backscatter properties of actual satellite surfaces which are invariably oxidized or contaminated are not well known. Currently the lack of knowledge in this area is one of the major deficiencies in spacecraft charging theory.

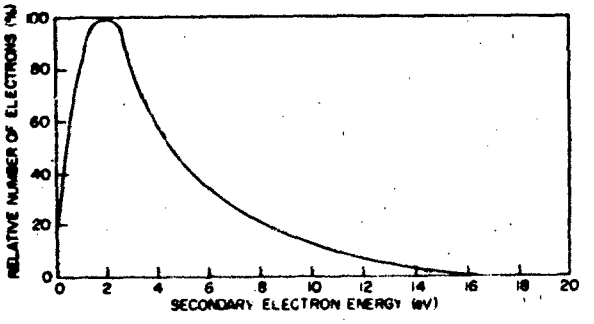


Figure 7-12c. Emission spectrum of secondary electrons due to incident electrons or ions impacting on aluminum [Whipple, 1965].

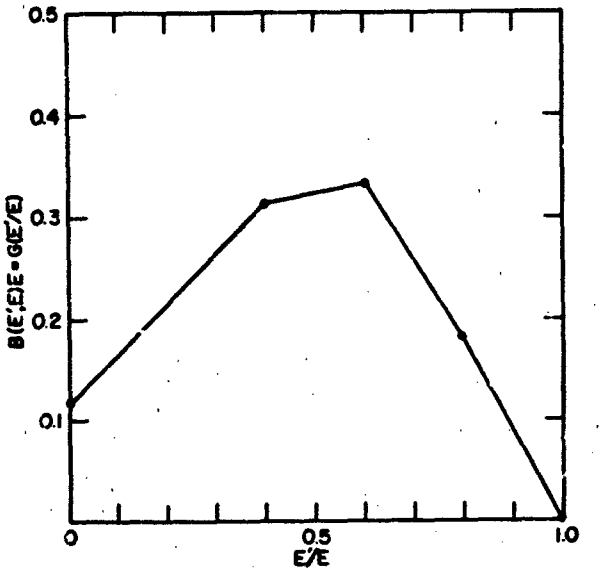


Figure 7-13. Graph of G (E'/E) as a function of (E'/E) where G is approximately the percentage of electrons scattered at a given energy E' as a result of an incident particle of energy E [Sternglass, 1954].

7.3.5 Magnetic Field-Induced Current Distortions

A problem in low earth orbit often encountered by electric field experiments employ long (~10 m or longer) antennas or booms is the induced electric field due to the satellite crossing the earth's magnetic field lines. Briefly, a satellite moving relative to a plasma (assumed to have a zero electric field with velocity v_s will see an electric field E in its rest frame given by

$$\mathbf{E} = \frac{\mathbf{v}_s \times \mathbf{B}}{C} = 10^{-8}(\mathbf{v}_s \times \mathbf{B})\ \text{V/cm} \quad (7.8)$$

where E is in V/cm, v_s is in cm/s, B is in G, and C is the speed of light.

An earth-orbiting satellite will see a maximum induced $\mathbf{v} \times \mathbf{B}$ field at low altitudes on the order of ~0.3 V/m. The local fields and current flows in the vicinity of the vehicle will be distorted by this effect.

Besides the $\mathbf{v} \times \mathbf{B}$ current, the magnetic field also induces anisotropies in the particle fluxes. Ambient fluxes, secondaries, beam fluxes, and charged particle wakes are

all controlled to a greater or lesser extent by the magnetic field. Whipple [1965] and Parker and Murphy [1967] have analyzed some of the effects of these magnetic field-induced anisotropies on spacecraft charging (see also reviews by Brundin [1963] and Guverich et al. [1970]) and find that the electron flux can be reduced by as much as a factor of 2, but as a rule these are ignored in spacecraft charging calculations. McCoy et al. [1980] has suggested that this constraint of electrons to field lines may become of real concern for very large structures which are much larger than a gyroradius.

7.3.6 Motion-Induced Effects

In low earth orbit, the velocity of a satellite is 7.5 km/s. Within the plasmasphere, where particle energies are 1 eV or less, the ion thermal velocity is 10 km/s or less. This implies that a plasma wake would be formed around the vehicle. Chang et al. [1979] have estimated that motion-induced effects on the satellite potential due to such a wake are important for ion ram to thermal velocity ratios as low as 0.1. These motion-induced effects would be present even for neutral particles and result in there being large asymmetries in the flow between the leading surfaces of the satellite and the rear for even small objects such as booms. The motion of the satellite can also induce space charge variations in the vicinity of the vehicle [Parker, 1977]. The depletion of electrons and ions in the satellite wake can distort the flow of the particles to the vehicle. A typical observation is presented in Figure 7-14 in terms of the normalized electron fluxes in the wake of the Explorer 31 satellite [Samir and Wrenn, 1969]. Although simple models of this phenomena will be discussed below, reviews by Brundin [1963], Kasha [1969], Liu [1969], and Al'Pert [1976] and papers by Gurevich et al. [1970] and Kunemann [1978] should be consulted for details.

7.3.7 Charge Deposition by Energetic Particles

The deposition of charge in dielectrics by high energy particles is a well known phenomena in nuclear physics (see, for example, Gross and Nablo [1967]; Evdokimov and Tubalov [1974]; and Frederickson [1979]) and has been proposed as a source of satellite charging [Meulenberg, 1976]. As an example [Frederickson, 1980], electrons between 10 keV and 100 keV lose energy at a rate of $R = 10^6$ to 5×10 eVg^{-1}cm^2 depending on energy and the material involved ($R \cdot \rho$, where ρ is the density of the material, gives the energy loss rate per cm). Electrons of 1 MeV will typically penetrate several millimeters into a solid. Fields of the order of M V/mm are necessary to retard such incident fluxes. Fields of this magnitude are more than adequate to cause electrical breakdown in a dielectric. Figure 7-15 is a theoretical plot of the potential in a PCV

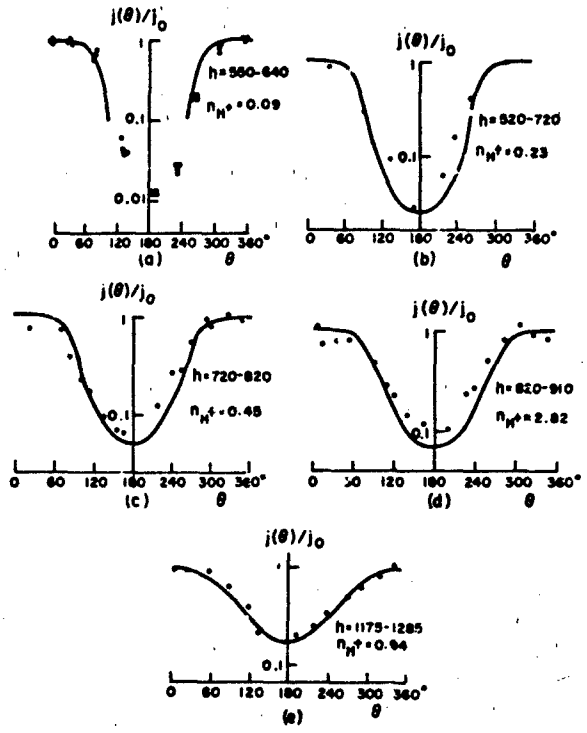


Figure 7-14. Normalized electron current versus angular position of the plasma probe on Explorer 31 [Samir and Wrenn, 1969]; 180° would correspond to the center of the wake (that is, opposite the direction of movement). Also shown are theoretical fits. The altitude h and ambient density n are indicated [Gurevich et al., 1970].

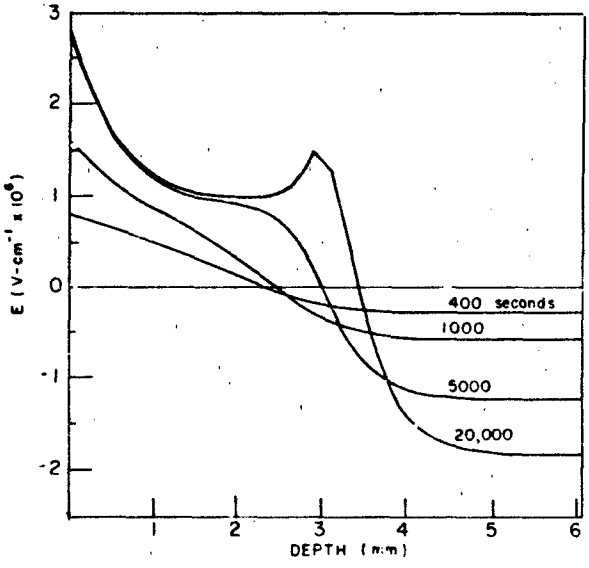


Figure 7-15. The electric field in 6 mm PVC sample as a function of exposure time to a 4.66×10^{-10} A/cm^2, 1 MeV electron beam [Frederickson, 1980].

sample exposed to 1-MeV electrons as a function of time [Frederickson, 1980].

As satellites in the earth's radiation belts or the hostile environment of Jupiter can experience high dosages of energetic particles, this effect becomes important if long mission lifetimes are desired. Recent evidence [Treadway et al., 1979] indicates that the effects of charging and arcing may be substantially altered by the buildup of charge in dielectrics. As yet, this process has not been included in spacecraft charge modeling. Its inclusion, however, will be critical to a complete understanding of the charging/arcing process.

7.3.8 Artificial Charging Mechanisms

Artificial mechanisms that effect spacecraft charging are numerous and include electron and ion beams and exposed, high potential surfaces such as the junctions between solar cells [Stevens, 1980]. Recently, beam sources have been actively exploited both as probes of the satellite sheath and as a means of controlling the spacecraft potential [Goldstein and DeForest, 1976; Olsen, 1980; Purvis and Bartlett, 1980; Cohen et al., 1979]. Voltages in the KV range and currents between mA and A are typical of these systems [Winkler, 1980]. The Excede 2 test [O'Neil et al., 1978b], in fact, ejected 3 kV electrons at 10 A. An example of a theoretical beam calculation will be given in Section 7.4 as an illustration of the complexities involved in analyzing such experiments.

Parker [1979, 1980], Stevens [1980], McCoy et al. [1980], and Rieff et al. [1980] have carried out calculations of the currents to exposed potential surfaces for large, high voltage structures. They find that a major effect is to induce large voltage gradients in the satellite sheath that must be considered in the proper design and use of such systems. The potentials can lead to multipacting and preferential deposition of ion contaminants. Another possible difficulty associated with very high positive potentials is the "pinhole" effect. As discussed in Kennerud [1974], the insulation on such surfaces can become punctured. Even microscopic holes can result in exceptionally high focusing of electrons—in some cases the pinhole can completely defeat the insulation.

Finally, a number of experiments have been conducted in vacuum chambers. Currently, the most extensive programs are those at NASA Johnson [McCoy et al., 1980] and NASA Lewis [Stevens, 1980; Purvis et al., 1977; and their colleagues]. The NASA Johnson chamber tests have involved testing of high potential surfaces and of rocket beam sources prior to flight [Konradi, private communication, 1980] in an environment resembling the ionosphere. As discussed in Bernstein et al. [1978; 1979], these latter tests have concentrated on the analysis of the so-called "beam-plasma discharge" (BPD) in electron beam experiments. They have found that the electron beam at a critical current transitions from one well defined by single particle dynamics to a much larger beam spread marked by an increase in the local plasma density. The critical current I_C when this occurs is proportional to the beam energy, V_B, such that $I_C \propto V_B^{3/2}$, an equation resembling Equation (7.14). The NASA Lewis studies are part of a coordinated effort to validate the NASCAP code (see Section 7.4.6) and have included testing of the P78-2 SCATHA experiments [Stevens et al., 1980a].

7.4 SPACECRAFT CHARGING THEORY

The basic equation expressing current balance for a given surface in an equilibrium situation is, in terms of the current:

$$I_E(V) - [I_I(V) + I_{SE}(V) + I_{SI}(V) + I_{BSE}(V) + I_{PH}(V) + I_B(V)] = I_T \quad (7.9)$$

where V = satellite potential, I_E = incident electron current on satellite surface, I_I = incident ion current on satellite surface, I_{SE} = secondary electron current due to I_E, I_{SI} = secondary electron current due to I_I, I_{BSE} = backscattered electrons due to I_E, I_{PH} = photoelectron current, I_B = active current sources such as charged particle beams or ion thrusters, and I_T = total current to satellite (at equilibrium, $I_T = 0$).

In this section methods of solving Equation (7.9) for V so that $I_T = 0$ will be described. The basic problem is the solution of Equation (7.9) subject to the constraints of Poisson's equation:

$$\nabla^2 V = 4\pi q(n_s + n_E - n_I) \quad (7.10)$$

and the time-independent collisionless Boltzmann (or Vlasov) equation:

$$\mathbf{v} \cdot \nabla f_i - \frac{q_i}{m_i} \nabla V(r) \cdot \nabla_v f_i = 0, \quad (7.11)$$

where n_E = local electron density, n_I = local ion density, n_s = surface-emitted electron density, ∇, and ∇_v = gradient operators with respect to position and velocity space respectively.

7.4.1 Analytic Probe Theory General Considerations

The most important concept in probe theory is that of the Debye length λ_D, the distance over which a probe or satellite significantly perturbs the ambient medium. (Note that this is only one definition of the satellite sheath. The sheath thickness is not only dependent on satellite potential and charge, but through the so-called "presheath", influences the plasma up to the order of the satellite radius

[Parker, 1980]. More precisely, electrons and ions form a cloud around any charged surface that is a function of the particles' energies and densities and the potential on the surface. As an illustration, the solution to Poisson's equation for spherical symmetry is

$$V(r) = q\,\frac{e^{-r/\lambda_D}}{r} \qquad (7.12)$$

where $\lambda_D = (KT/4\,\pi\,q^2 n_o)^{1/2}$ and n_o = ambient charge density. This is the classical Debye length for the assumption of "linearized space charge." The "linearized space charge" approximation can be extended to spherical and other multidimensional geometries; an analytical fit for the spherical case is discussed by Parker [1980]. Values for λ_D are listed in Table 7-3.

The terminology "thick" sheath and "thin" sheath derives from the assumption that the region over which the satellite affects the ambient plasma or is screened from the ambient plasma is either larger or smaller than the characteristic dimensions of the spacecraft. Usually this reduces to determining whether λ_D is long (thick sheath) or short (thin sheath) compared to the spacecraft radius. From Table 7-3, for a 1 m radius spacecraft with no exposed potential surfaces in low earth orbit, the term thin sheath is appropriate. For geosynchronous orbit, unless the satellite is 10 m or larger, the thick sheath approximation is appropriate. If active surfaces (that is, exposed potentials driven by the satellite systems) or a substantial photoelectron population are present, these limiting criteria will change, but even so for most satellite studies they are very useful (Opik [1965] has carefully considered this issue of the appropriate screening criteria for the general case of an arbitrary central force).

As a first example, consider a large structure such that the characteristic scale of the sheath is significantly smaller than r_s, the radius of the surface (the thin sheath assumption). Assuming at the surface ($X = r_s$) the potential is V_o and that the surface is nearly planar relative to the sheath dimensions, at distance $Y (= X - r_s)$ from the surface (only the attracted species is considered)

Poisson's Equation
$$\frac{d^2V(Y)}{dY^2} \simeq -4\,\pi\,q\,n(Y) \qquad (7.13a)$$

Current continuity
$$J = q\,n(Y)\,v(Y) = \text{constant} \qquad (7.13b)$$

The particles are assumed to have $E = 0$ in the ambient space:

Energy conservation
$$\frac{1}{2}\,mv(Y)^2 + q\,V(Y) = 0. \qquad (7.13c)$$

In order to solve these equations, it is assumed that V and dV(Y)/dY are 0 at some distance $Y = S$ which determines the sheath thickness (called the "space-charge limited" assumption). The solution becomes

$$J = \frac{1}{9\,\pi}\left(\frac{2q}{m}\right)^{1/2}\frac{V_o^{3/2}}{S^2} \qquad (7.14)$$

This is the "space-charge limited diode model" solution or Child-Langmuir model for a plane. If J is replaced by $J_o = K^* q n_o (kT/m)^{1/2}$ (where $(2\,\pi)^{-1/2} < K^* < 1$, the lower limit corresponds to Equation (7.4b); the upper limit is for a monoenergetic flow of energy $E \equiv 1/2\ kT$) then the sheath thickness can be estimated:

$$S = \frac{1}{3}\left(\frac{2}{qKT}\right)^{1/4}(V_o)^{3/4}\left(\frac{1}{\pi\,n_o}\right)^{1/2}\left(\frac{1}{K^*}\right)^{1/2}. \qquad (7.15)$$

This sheath thickness determines the region over which charge is collected and is important in determining the maximum current that can flow to a probe for a given V.

Although seldom utilized in spacecraft charging calculations, a form of Equation (7.14) was employed by Beard and Johnson [1961] to obtain the potential to which a vehicle can be charged by an electron emitter in the ionosphere. This calculation placed an upper limit on the current that could be drawn at a given potential ignoring magnetic field shielding effects. When included [Parker and Murphy, 1967], the maximum current was reduced by as much as a factor of 10. Linson [1969], including turbulence, found values between these limits, however. These studies imply that a 1 m sphere emitting a 0.5A electron beam at shuttle altitudes would have a potential of 10^4–10^6V, depending on the magnetic field [Liemohn, 1977], thereby seriously inhibiting a 10–100 keV beam. Recently, Parker [1980] (see also Kennerud [1974], McCoy et al. [1980], and Parks and Katz [1981]), in order to estimate the sheaths around large space structures, have developed an analytic expansion for the thin sheath approximation to a sphere. He has investigated the applicability of this estimate of the sheath to extremely high applied potentials in the ionosphere by comparing it to the results of a self-consistent numerical calculation (see Section 7.4.2). Although his rigorous computations (for a Debye length to satellite radius ratio of 1 to 100) deviate at low potential values ($qV/kT \leq 10$) from the space charge limited model, the results approach each other for very high potentials ($qV/kT \sim 400\,000$).

The preceding theory concerning the thin sheath assumes that sheath effects dominate the current flow to the satellite. This places emphasis on Poisson's equation and the space charge around the satellite. In the opposite extreme, the sheath and space charge are ignored so that, to first order, Laplace's equation holds. In practical terms this translates into the assumption that $\lambda_D \gg r_s$. For spherical symmetry, conservation of energy and angular momentum imply that

for an attracted particle approaching the satellite from infinity

$$\frac{1}{2} m v_o^2 = \frac{1}{2} m v (r_s)^2 + q V(r_s) \qquad (7.16a)$$

$$m R_f v_o = m r_s v (r_s) \qquad (7.16b)$$

where v_o is the velocity in ambient medium, R is the impact parameter, and R_f is the R for a grazing trajectory for a vehicle of radius r_s.

Solving for the impact parameter R_f (only particles having $R < R_f$ will reach r_s):

$$R_f^2 = r_s^2 \left(1 - \frac{2qV(r_s)}{mv_o^2}\right) \qquad (7.17)$$

$(R_f - r_s)$ is equivalent to the sheath thickness S defined for a thin sheath as it also is the size of the region from which particles can be drawn.

The total current density striking the satellite surface for a monoenergetic beam is

$$J(V) = \frac{I}{4\pi r_s^2} = J_o \left(1 - \frac{2qV(r_s)}{mv_o^2}\right) \qquad (7.18)$$

where I is the total current to spherical satellite equal to the ambient current that would pass through an area $4\pi R_f^2$ and J_o is the ambient current density outside the sheath equal to $I/4\pi R_f^2$. This is the so-called "thick sheath, orbit-limited" current relation.

The thick sheath results are readily extended to more complex distributions. For most spacecraft charging problems, Boltzmann's equation [Equation (7.11)] in the absence of collisions reduces to Liouville's theorem. For a complex distribution function, F, Liouville's theorem states that:

$$F(v) = F'(v') \qquad (7.19)$$

where F' and v' are the distribution function and velocity at the surface of the spacecraft. F and v are the ambient distribution and velocity at the end of the particle trajectory connected to the satellite surface where F' and v' are measured. Prokopenko and Laframboise [1977, 1980] have derived (based on Equation (7.19)), in more general terms the current density to a sphere, infinite cylinder, and infinite plane (that is, three-, two-, and one-dimensions) for a Maxwell-Boltzmann distribution for the orbit-limited solution. Their results (see original derivation in Mott-Smith and Langmuir [1926]) for the attracted species are

$$J_i = J_{io} \cdot \begin{cases} (1 + Q_{is}) & \text{Sphere} \\ [2(Q_{is}/\pi)^{1/2} + e^{Q_{is}} \operatorname{erfc}(Q_{is}^{1/2})] & \text{Cylinder} \\ (1) & \text{Plane} \end{cases} \qquad (7.20)$$

and for the repelled species $J_i = J_{io} e^{Q_{is}}$ where

$$Q_{is} = \pm \frac{qV}{kT_i} \quad (+ \text{ for electrons, } - \text{ for ions}) \qquad (7.21)$$

The current-voltage characteristics implied by Equations (7.20) and (7.21) for the three geometries are plotted in Figure 7-16. Several important conclusions can be drawn from Equations (7.20) and (7.21) and Figure 7-16. First Prokopenko and Laframboise's results for a sphere are identical to Equation (7.18) (with $^1/_2 mv_o^2$ replaced by kT_i) for

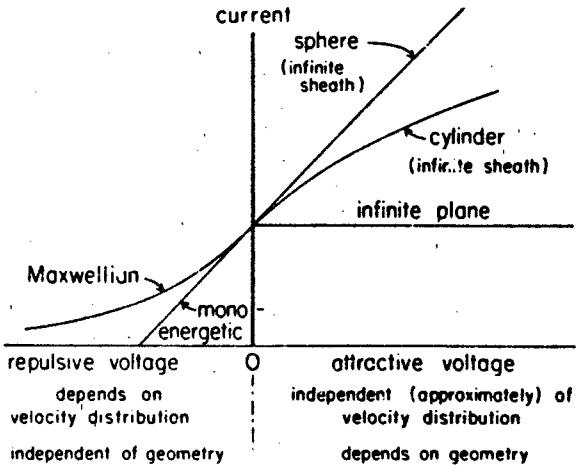


Figure 7-16. Qualitative behavior of Equations (7.20) and (7.21) for a Langmuir probe [Parker and Whipple, 1967]. (Reprinted with permission of Academic Press © 1967)

a thick sheath. It should also be readily apparent that the planar solution is conceptually equivalent to a thin sheath. Thus, Equation (7.20) additionally gives a qualitative picture of how the probe characteristics change as the ratio of the Debye length to satellite radius is varied from small values (thin sheath or planar) to large values (thick sheath or spherical). This was done explicitly by Whipple et al. [1974] and Whipple [1977] for the ratio of the Debye length to satellite radius and is closely related to the parametrization method of Cauffman and Maynard [1974].

7.4.2. Analytic Probe Theory Thick Sheath Models

At this point a complete analytic theory for the case of a thick sheath, spherical probe has been developed. Substituting into Equation (7.10) and assuming that the secondary and backscatter terms can be parametrized, for an ambient Maxwellian plasma

$$A_E \cdot J_{EO} \cdot [1 - SE(V,T_E,n_E) - BSE(V,T_E,n_E)] \cdot \exp(qV/kT_E) - A_I \cdot J_{IO} \cdot [1 + SI(V,T_I,n_I)] \cdot \left(1 - \frac{qV}{kT_I}\right) - A_{PH} \cdot J_{PHO} \cdot f(X_m) = I_T = 0 \qquad V < 0 \tag{7.22}$$

where

J_{EO} = ambient electron current density [Equation (7.4)]

J_{IO} = ambient ion current density [Equation (7.4)]

A_E = electron collection area ($4\pi r_s^2$ for a sphere)

A_I = ion collection area ($4\pi r_s^2$ for a sphere)

A_{PH} = photoelectron emission area (πr_s^2 for a sphere)

SE, SI, BSE = parametrization functions for secondary emission due to electrons and ions and backscatter

J_{PHO} = saturation photoelectron flux (Table 7-4)

$f(X_m)$ = percent of attenuated solar flux as a function of altitude X_m of center of sun above the surface of the earth as seen by satellite.

Equation (7.22) is appropriate for a small (<10 m), uniformly conducting satellite at geosynchronous orbit in the absence of magnetic field effects. To solve the equation, V is varied until $I_T = 0$. A number of examples of this procedure, assuming SI = SE = BSE = 0, are tabulated in Table 7-3 for various plasma regions.

If it is assumed that SI, SE, and BSE are constants (that is, ~3., ~0.4, and ~0.2 for Al), then Equation (7.2) predicts [Garrett and Rubin, 1978] that in eclipse, the potential between the satellite and space is

$$V \approx -T_E \tag{7.23}$$

The satellite to space potentials observed by UCSD electrostatic detectors for twenty-one ATS 5 and four ATS 6 eclipses are plotted in Figure 7-17 versus the electron temperature. As the geosynchronous plasma is not necessarily Maxwellian, two different "temperatures," T_0 (AVG) (= 2/3 · energy density/number density) and T_0 (RMS) (= 1/2 · energy flux/number flux), are presented (these would be equal if the plasma was actually Maxwellian). The agreement is good considering that the range of potentials is between −300 V and −10 000 V. The existence of a

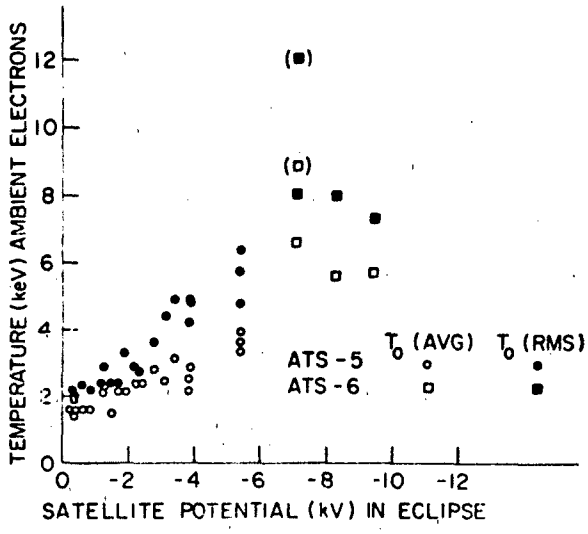


Figure 7-17. Observed temperature of ambient electrons versus satellite potential in eclipse. T_0(AVG) is two thirds of energy density/number density; T_0(RMS) is one half of energy flux/number flux.

threshold temperature below which charging does not occur is real and is due to the fact that at an intermediate energy (usually a few hundred eV), the secondary yield is greater than 1 [Rubin et al., 1978]. The electron temperature must be several times greater than this threshold energy before charge buildup occurs.

The agreement can be significantly improved for Figure 7-17 if the actual ambient spectra are utilized in the integration of Equations (7.4), (7.6), and (7.7) instead of a Maxwellian in computing the currents. This method of using the actual particle spectra to estimate the currents has been extensively employed by DeForest [1972], Knott [1972], Garrett and DeForest [1979], and Prokopenko and Laframboise [1977, 1980] in the calculation of satellite to space potentials for geosynchronous spacecraft. The potentials of Figure 7-17 are recalculated in this manner and plotted in Figure 7-18a [Garrett and DeForest, 1979]. Results are presented both for particle spectra immediately before entry or immediately after exit from eclipse ("sunlit") and for eclipse ("eclipsed"); differences are attributed to the digitization of the spectra. As the exact secondary response of the satellite surface was not known, SE, SI, and BSE were assumed to follow the Al curves of Figures 7-12 and 7-13. Their absolute amplitudes were then varied until the observations were fit in a least squares sense [Garrett and DeForest, 1979].

Given the validity of the calibration method, the effects of a varying photoelectron flux on the satellite to space potential can be studied during eclipse passage. If the satellite position is known, the photon flux reaching the satellite can be calculated from first principles [Garrett and Forbes

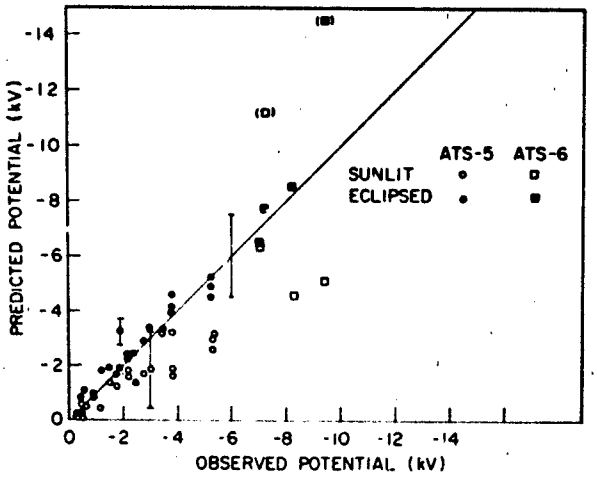


Figure 7-18a. Predicted and observed potentials in eclipse for ATS 5 and ATS 6. Solid symbols are for calculations using the spectra in eclipse. Open symbols are for calculations using the spectra measured in sunlight

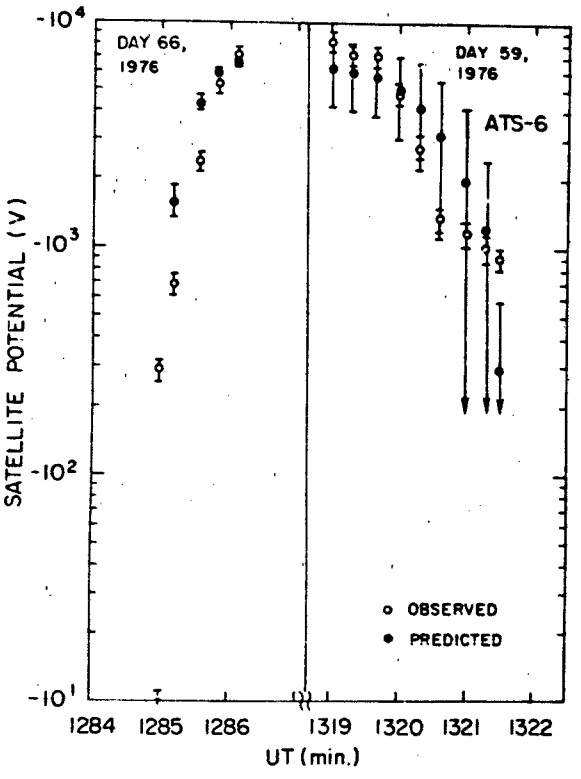


Figure 7-18b. Observed and predicted potentials for the entry into eclipse of ATS 6 on day 66, 1966, and for eclipse exit on day 59, 1976.

1981]. If the satellite surface materials were known, then the photoelectron current, $J_{PHO} \cdot f(X_m)$, could also be calculated from first principles. Although the exact surface response is in actuality not known, adequate approximations can be derived [Garrett and DeForest, 1979; Garrett and Forbes, 1981]. J_{PHO} in Equation (7.22) is then varied to fit observations. Estimates of the varying potential on ATS 6 during eclipse entry and exit by this technique are compared with actual observations in Figure 7-18b [Garrett and DeForest, 1979]. This eclipse model has proven valuable in estimating photoelectron flux and potential variations for ATS 5, ATS 6, Injun 5, and P78-2 SCATHA.

The results of Prokopenko and Laframboise [1977, 1980], using the spectra suggested by Knott [1972], are particularly important because they established the existence of muiltiple roots for Equation (7.22) (see also Sanders and Inouye [1979]). Multiple roots imply that a satellite can undergo rapid voltage variations in response to small environmental perturbations and that adjacent surfaces can come to radically different potentials for the same conditions. To obtain their result, they solved an equation equivalent to Equation (7.22) (that is, local current balance) for the spherical, the cylindrical, and planar assumptions of Equations (7.20) and (7.21) and for eclipse conditions. They found that the potential is markedly increased for a planar probe relative to a spherical probe.

7.4.3 Analytical Probe Theory Thin Sheath and Related Models

Analytic probe theory can also be utilized to estimate satellite to space potentials in the 250–700 km range. As discussed in Brundin [1963], in the absence of magnetic forces, photoelectrons, and secondaries, Equation (7.9) reduces to

$$A_E \cdot J_{EO} \cdot C_E \cdot e^{qV/kT_E} - A_I \cdot J_{IR} \cdot C_I = 0 \ (V < 0) \quad (7.24)$$

where

J_{EO} = ambient electron current density (Equation (7.4)),

J_{IR} = ion ram current density (ignoring ion thermal velocity)

$= q_I n_I v_s$,

v_s = satellite velocity,

A_E = electron collection area ($4 \pi r_s^2$ for a sphere),

A_I = ion collection area (πr_s^2 for a sphere),

r_s = satellite radius

C_E = electron shielding factor (= 1, no electron wake assumed; ≈ 1/2, wake on rear half).

C_i = ion shielding factor {(= 1, complete shielding of ambient ions or thin sheath;

$= \left(1 - \frac{2qV}{m_i v_\infty^2}\right)$, no shielding or thick sheath, (Equation (7.13)).

Samir [1973] and Samir et al. [1979b], assuming no electron wake and no ion shielding (note that this is a thick sheath assumption which is usually inappropriate in this region), have compared the predictions of Equation (7.24) with observations. In spite of the simplicity of Equation (7.24) and the assumption of a thick sheath, they predict the satellite to space eclipse potential (typically ~ −0.75V) for a large range of ambient conditions to a factor of 2.5 or better.

A simple approximation for estimating the current to an isolated point on a planar satellite surface in the ionosphere as a function of the surface normal relative to the velocity vector is that of Whipple [1959] and Bourdeau et al. [1961] (see also Tsien [1946] and Chang et al. [1979]):

$$I_i = \alpha q n_i A_i \left[v_\infty \cos\theta \, (1/2 + 1/2 \, \mathrm{erf}(x)) + \frac{a e^{-x^2}}{2\sqrt{\pi}} \right] \quad (7.25)$$

where

$$x = \frac{v_\infty \cos\theta}{a} - \left(\frac{qV}{kT_i}\right)^{1/2}.$$

θ = angle between sensor normal and velocity vector,
A_i = collection area,
α = grid transparency function,
a = most probable ion thermal velocity.

This equation, the so-called "planar approximation", is good for short Debye lengths but becomes inaccurate for long Debye lengths [Parker and Whipple, 1970] and more complicated orbital trajectory calculations must be carried out. Even so, Bourdeau et al. [1961] found it to be a good approximation to the ion current relative to the velocity vector measured by the ion planar probe on Explorer 8 (425–2300 km altitude). Their results for V = 0 are plotted in Figure 7-19 (note that there is some question as to the best value for a, see Bourdeau et al. [1961]).

A simple analytic theory closely related to thin sheath probe methodology and capable, within limits, of explaining the satellite wake structure in the ionosphere has also been developed. It (see reviews in Gurevich et al. [1970] and Al'pert [1976]) is based on the thin sheath assumption,

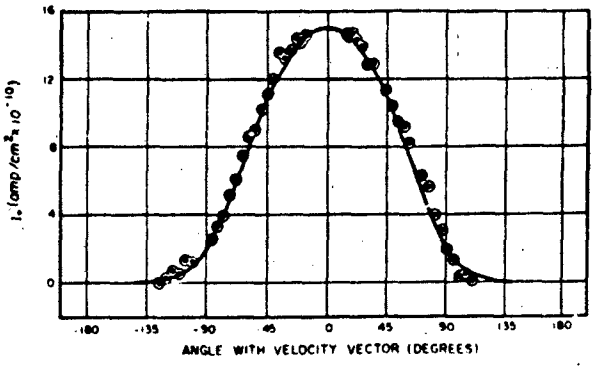


Figure 7-19. Variations in the positive ion current density with angle relative to the satellite velocity vector. The data and figure are from Bourdeau et al. [1961]. The fitted line is given by Equation (7.29) with a assumed to be 3.6 km/s; the satellite potential is assumed to be 0.

$r_\infty \gg \lambda_D$, the Debye length in the ionosphere being much smaller than the satellite's characteristic dimensions. This allows the neglect of the effects of the satellite potential except very close to the satellite surface. The left hand side of Equation (7.10) can then be ignored, so that it becomes:

$$n_E - n_i \simeq 0 \quad (7.10^*)$$

This is the so-called "quasi-neutrality" assumption [for example, Gurevich et al., 1966, 1968, 1973; Grebowski and Fischer, 1975; Gurevich and Dimant, 1975; Gurevich and Pitaveskii, 1975]. Equation (7.11) for the ions can be reduced by making use of the "hypersonic" character of the motion of a body in the ionosphere (that is, M, the square root of the ratio of the ion kinetic energy to electron thermal energy, is much greater than 1). First, this implies that the gradient of the potential *perpendicular* to the flow direction is much greater than the gradient *along* the flow direction so that this latter term can be ignored. Second, the thermal velocity of the ions in the direction *along* the flow direction can be neglected relative to the satellite velocity v_∞ in the ionosphere. Based on these assumptions, Equation (7.11) for ions can be rewritten as

$$v_\infty \frac{\partial f_i}{\partial z} + \mathbf{v}_\perp \cdot \frac{\partial f_i}{\partial \mathbf{r}_\perp} - \frac{q}{m_i} \frac{\partial V}{\partial \mathbf{r}_\perp} \cdot \frac{\partial f_i}{\partial \mathbf{v}_\perp} = 0 \quad (7.11^*)$$

where $r_\perp$ and $v_\perp$ correspond to components normal to the direction of motion and z is in the v_∞ direction.

For the electrons, a Boltzmann distribution ($n_E = n_{E0} \cdot \exp(qV/KT_E)$) is usually assumed. The resulting system of equations for the ions does not contain the ion thermal velocity along the direction of motion and can be put into a dimensionless, self-similar form [Al'pert et al., 1965; Gurevich et al., 1970] that resembles classical hypersonic aerodynamic equations. Depending on the char-

acteristics of the assumed plasma conditions, the wake variations for a number of simple geometries can be analytically solved. These range from the extreme assumption of a true neutral flow as reviewed in Gurevich et al. [1970] and Al'pert [1976]—charged particle variations in the wake mirroring the neutral variations—to plasma flows around infinite half-planes, wedges, plates, cylinders, and discs. The predictions for one such analytic solution [Gurevich et al., 1970] are compared with Explorer 31 observations in Figure 7-14 from Samir and Wrenn [1969] for various distances r/r, from the satellite. The theory is not considered reliable at angles greater than ~120° as the electron density can differ greatly from the ion density in this region of maximum rarefaction [Gurevich et al., 1970].

Although severe constraints (primarily $r_s \gg \lambda_D$ and neglect of the ion thermal velocity) have been placed on the realm of applicability of this "hypersonic," quasi-neutral theory, it does allow the analytic study of the effects of geometry, magnetic field, and, of more importance, ionic composition on the wake. As discussed in Gurevich et al. [1970, 1973] and Samir et al. [1980], variations in ion composition play a critical role in the details of the expansion of the ion population into the rarefied wake region behind the satellite. Although these results are useful, in most practical situations a finite Debye length is a necessary assumption. This greatly complicates any theoretical computation and requires the advanced probe theory of Section 7.4.4.

An application of analytic probe theory based on local current balance is in the so-called circuit models. The single probe theory introduced so far does not explicitly consider the problem that satellites consist of a variety of surfaces, including dielectrics, and that each surface can charge to a different potential if isolated from the others. In order to explicitly model this differential charging effect, the coupling currents between surfaces must be included in Equation (7.9). Circuit models [Robinson and Holman, 1977; Inouye, 1976; Massaro et al., 1977], as this class of models is termed, consist of many "probes," each representing a particular point or surface on the satellite (a dielectric surface, for example, would be approximated by one or more individual points). Besides the ambient, secondary, backscattered, and photoelectron currents considered in the single probe model, the coupling currents to each point, J_{RLC}, are included in J to estimate the currents between surfaces. Time variations are explicitly handled by including inductive and capacitive elements which have finite charging times (local current balance with space is assumed at each instant in time) making these models applicable to a wider range of problems than the single probe models. Inouye [1976] and Massaro et al. [1977] have utilized such models to assess the effects of geomagnetic storm variation, varying solar angle, isotropic fluxes, etc., on individual satellite surfaces as a function of time. Their results, although subject to the difficulties associated with simple probe analysis, indicate that dielectric surfaces can charge in tens of minutes but may take hours to discharge. Further, charge up can be a long term (days) process and be dependent on the time history.

7.4.4 General Probe Theory

Whereas analytic probe theory is applicable to a number of practical problems, it does not allow for the complex geometric and space charge effects of the satellite sheath on particle trajectories. As a consequence, it is severely limited in its quantitative accuracy (the parametrization method of Cauffman [1973]; Cauffman and Maynard [1974]; is one qualitative means of studying these effects using analytic probe theory). In order to include the effects of the sheath on particle trajectories, it is necessary to seek simultaneous, self-consistent solutions of Equations (7.10) and (7.11). Typically, this is not analytically possible and an iterative procedure must be employed. First, V (**r**) is assumed so that f_i (**r**,**v**) can be computed subject to Equation (7.10). The number densities, n_i, are then computed from

$$n_i(\mathbf{r}) = \iiint f_i(\mathbf{r},\mathbf{v})\, d^3V. \qquad (7.26)$$

Given the n_i (**r**), V (**r**) is computed according to Equation (7.10). The process is iterated until a consistent set of value of n_i(**r**), V(**r**), and f_i (**r**,**v**) at grid points surrounding the surface are found. Then the J_i's at the surface r_s are found from the generalized form of Equation (7.4)

$$J_i(\mathbf{r}_s) = q_i \iiint \mathbf{v}\cdot\mathbf{n}\, f_i(\mathbf{r}_s,\mathbf{v})\, d^3V. \qquad (7.27)$$

The theories to be discussed assume the time-independent form of the Vlasov equation. Equation (7.10) is then just a restatement of Liouville's theorem, namely, that f_i (**r**,**v**) is constant along a particle trajectory in a potential V(**r**). To determine f_i ($\mathbf{r}_p$,**v**) at a point $\mathbf{r}_p$ for a particle of a given energy, all that is required is to find the intersection of that particle's trajectory with a surface where f_i is known. The trajectory may either be traced from the point backwards to the surface (the inside-out procedure) or from a surface to the point (outside-in procedure). According to Parker [1976a], the inside-out trajectory tracing method is preferable in that the points at which the density is calculated can be picked at random and is suitable for both electrons and ions. It has the disadvantage that the trajectory information computed for a given point is lost in moving to the next, increasing the computer time. The outside-in method can be readily adapted to time-dependent simulations but introduces difficulties in choosing trajectories so that the density at an arbitrary point can be determined. As a result

many time-consuming trajectory calculations are required. The computation can be greatly reduced, however, by using the "flux tube" method. As Parker [1976a] notes, this adaptation of the outside-in method, in which all the particle trajectories between two reference trajectories are assumed to be similar, ignores the possibility of orbit crossings or reversals and is only suitable for axisymmetric bodies and cold ion beams [Davis and Harris, 1961]. Unless otherwise stated Parker's inside-out method is assumed in the following. The integration of Equation (7.26) then reduces to determining the limits on the trajectories that intersect r_p as these will be the limits on the integral.

The precise method of determining the trajectories that intersect r_p differentiates the various probe calculations [Parker, 1978b, 1980]. Common to all, however, is the so-called orbit classification scheme. Consider the following arguments from classical mechanics (see for example, Goldstein [1965], Bernstein and Rabinowitz [1959], Whipple [1976a], Parker [1977, 1980]). The equation for energy conservation in a spherically symmetric potential V(r) is

$$E = \left[q_i V(r) + \frac{L^2}{2mr^2} \right] + \frac{1}{2} m v_r^2 \qquad (7.28)$$

where E is total energy, L is $m v_\phi r$ = angular momentum, v_ϕ is total tangential velocity, and v_r is radial velocity.

The two terms in brackets are called the "equivalent (or effective) potential" U(r,L) [Goldstein, 1965]. For illustration assume the potential V(r) is attractive and of the form r^{-1}. For various values of L we find that U has the shapes plotted in Figure 7-20a. Looking at the contour L_p, four types of orbits can be defined relative to point $r = r_s$. Adopting the terminology of Parker [1973, 1980], they are (Figure 7-20b) as follows:

Type 1: The particle has sufficient kinetic energy and small enough angular momentum to reach r_s from infinity or to reach infinity from point r_s. These orbits would contribute only once to a density integral at r_s.

Type 2: The particle starts at infinity but never reaches r_s, being repelled at some minimum distance. If the minimum distance is inside the region being integrated over, it contributes twice as it goes both in and out [Parker, 1980] and zero if it is outside the region of interest.

Type 3: The particle starts at r_s, but is reflected at some distance back to r_s. These particles contribute twice to the integral if r_s lies outside the turning point.

Type 4: Type 4 orbits are trapped orbits that circulate around the satellite. These are normally ignored as the orbits can only be populated by collisions that are assumed zero in most satellite studies. As Parker [1980] notes, this assumption has never been justified rigorously.

As a specific example of these calculations, we will review the spherically symmetric models of Parker [1973, 1975, 1976b, 1980]. These models have been successfully compared to the "Equivalent Potential Formulation" [Parker, 1975] and the particle-pushing code of Rothwell et al. [1976] (Figure 7-23a). The typical proceedure follows. Assuming spherical symmetry, Equations (7.26) and (7.27) can be reduced to two-dimensional integrals. It is also more

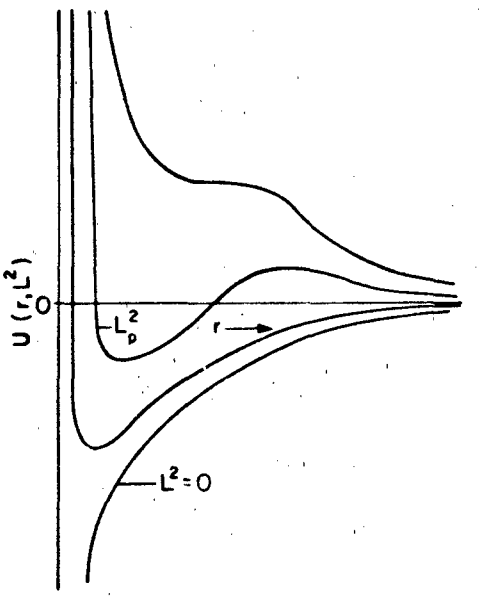


Figure 7-20a. The "equivalent potential" U (r,L) as a function of r for various values of L.

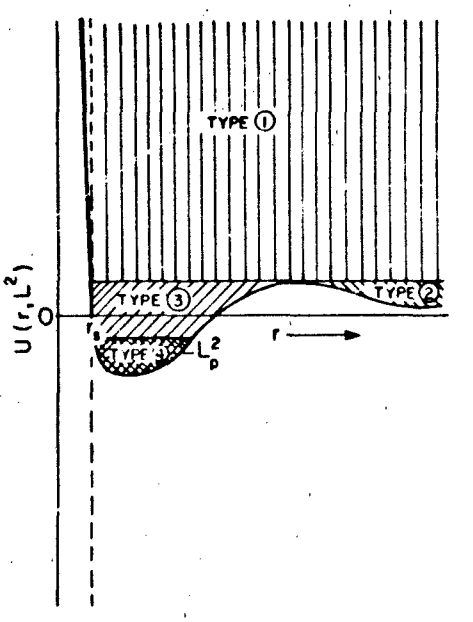


Figure 7-20b. Same as a, only for the L_p contour illustrating the four classes of orbits.

convenient to work in terms of E and L^2 so that the equations become

$$n_i(r) = \frac{\pi}{m^{1/2} r^2} \int\int \frac{f_i(E,L^2)\, dE\, dL^2}{(2m\,|E - q\,V(r)| - L^2/r^2)^{1/2}} \qquad (7.29a)$$

$$J_i(r) = \frac{q\,\pi}{m^{1/2} r^2} \int\int f_i(E,L^2)\, dE\, dL^2. \qquad (7.29b)$$

The integrations are now over the allowable ranges of E and L^2.

There are two common ways of defining the allowable range of integration in the (E,L^2) space. The first scheme makes use of the fact that the energy E must be greater than U(r) if the trajectories are to exist:

$$E > \left[q_i\,V(r) + \frac{L^2}{2mr^2} \right] = U(r). \qquad (7.30)$$

The maxima are found in U(r). This is the "Effective Potential Formulation" [Parker, 1980] and has been utilized by Bernstein and Rabinowitz [1959], Laframboise [1966], and Chang and Bienkowski [1970].

The other approach is to define a function:

$$g \equiv r^2\,|E - q_i V(r)| > \frac{L^2}{2m} \qquad (7.31)$$

where g is called the turning point function [Parker, 1980]. To classify the orbits, the minima in g are found. This technique is termed the "Turning-Point Formulation" and has been utilized by Bohm et al. [1949], Allen et al. [1957], Medicus [1961], and Parker [1973, 1975, 1976b]. Although the two methods are equivalent, Parker [1975] indicates that the Turning-Point Formulation is simpler and more efficient.

After obtaining equations of the form of Equations (7.29a) and (7.29b), solutions are found at a given point in space assuming V(r) is known. That is, Equations (7.29a) and (7.29b) are broken up into integrals corresponding to the different orbit types (note that type 4 orbits are normally ignored). $f_i(E,L^2)$ is assumed known at the probe surface (0 for no emission or the appropriate values for secondary or photoelectron emission) and, typically, assumed to be a Maxwellian in the ambient medium.

The next step is to determine the integral bounds on E and L^2. A typical example in terms of the Turning-Point Formulation is presented in Figure 7-21 [Parker, 1980] and should be compared with Figure 7-20. For a given value of r and E the limits on L^2 are determined. Once the integral ranges of E and L^2 are known for r, $n_i(r)$ and $J_i(r)$ are

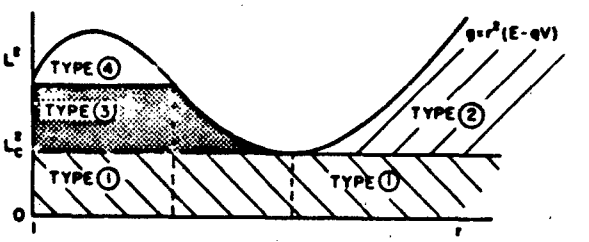


Figure 7-21. Turning-point formulation of Parker [1980]. Example illustrating domains of four types of orbits in (R, LK^2) space. Four in this plot is in units of r_c.

computed. V(r) is found by numerical integration from Poisson's equation given the $n_i(r)$ at grid points around the surface. The process is then iterated until a consistent solution is found. Recently, Parker (1979, 1980) employed the spherically symmetric model to calculate the sheath of a body of radius 100 Debye lengths and for a voltage of 400 000 kT/q, the most extreme combination of size and voltage solved rigorously to date.

More general geometric situations require actual numerical trajectory tracing and are discussed in Parker [1973; 1977; 1978a, b; 1979], Parker and Whipple [1967, 1970], Whipple [1977], and Whipple and Parker [1969a, b]. While the above procedure for spherical symmetry is qualitatively representative of the calculations represented by these papers, numerical particle tracing is required if more realistic geometries such as for the truncated cylinder or "pillbox" illustrated in Figure 7-22 are to be studied. The final potential contours, in this case for a directed plasma flow showing the effects of a wake [Parker, 1978a], arrived at by the iteration process are plotted in Figure 7-22.

The preceding theoretical studies of Parker, Whipple, and others have been particularly useful in studying the effects of differential charging and space-charge potential minima. Differential charging, as distinct from space-charge potential minima, has been demonstrated by these and similar efforts to result from wake effects [Parker, 1967a, 1978a, b], from asymmetric photoelectron emission (see, for example, Grard et al. [1973]; Fahleson [1973]; Whipple [1976b]; Prokopenko and Laframboise [1977b, 1980]; Besse and Rubin, [1980]), and from exposed potentials (Reiff et al. [1980]; Stevens [1980]). The effects of the space-charge potential minimum produced by emitted-electron space charge has been investigated by Soop [1972, 1973], Schroder [1973], Parker [1976b], Whipple [1976a], and Rothwell et al. [1977] [see also Guernsey and Fu, 1970; Grard and Tunaley, 1971; and Grard et al., 1973]. Whipple [1976b] and Parker [1976a] found that such barriers, which are typically a few volts, are inadequate to account for the ATS 6 observations of trapped photoelectrons and secondaries, and that a differential charging barrier must be invoked. Such charging barriers can significantly affect observations and their existence must be considered in designing satellite instrumentation.

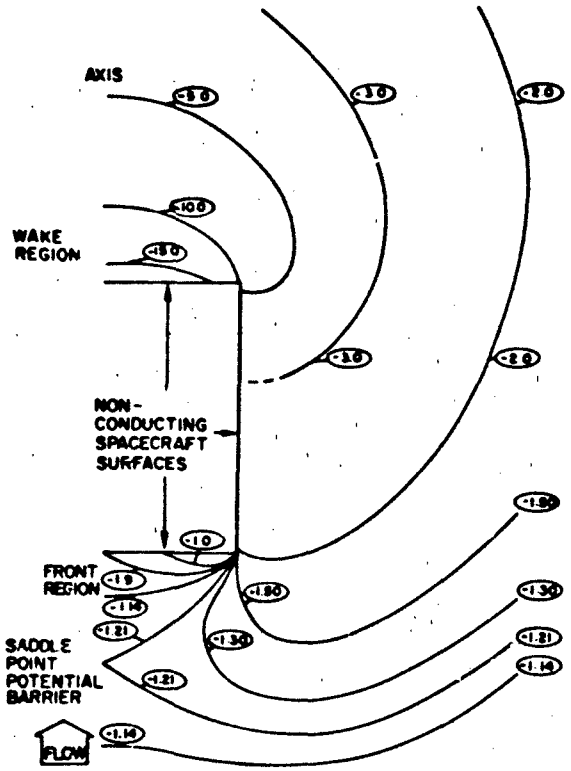


Figure 7-22. Differential charging of nonconducting spacecraft by directed plasma flow (equipotential contours are in units of kT/q) [Parker, 1978a].

7.4.5 Numerical Simulation Techniques

Although general probe theory can be applied to a number of interesting and important cases, it has not been extended much beyond spherical or cylindrical geometries nor does it take into consideration time variations in the sheath. Numerical techniques have been developed that, though retaining many of the basic concepts of probe theory, allow explicit inclusion of time dependency or geometry. These models are capable of handling time variations on the order of the plasma frequency or complex shapes such as the shuttle or the P78-2 SCATHA satellite.

The most straightforward numerical techniques conceptually, though perhaps the most demanding computationally, are the so-called "particle pushing" codes. As originally presented by Albers [1973] and Rothwell et al. [1976, 1977] for a spherical geometry and by Soop [1972, 1973] and Mazzella et al. [1979] for cylindrical geometry, many individual or groups of similar particles are followed simultaneously by computer as they move through the satellite sheath. At each time step the potential on a given particle or set of particles due to all the other particles is computed by a fast Poisson equation solver. The particle group is incremented in time using this net potential. The next group is then moved based on the other particles. The process is iterated in time with the computer keeping track of surface interactions (backscattered and secondary electrons, photoelectrons, and the outer boundary) and interactions between the satellite fields and the particles. Although plasma simulation codes of this type have been extensively employed in plasma physics, they are just beginning to be used for the plasma probe problem.

Although still limited by computer capacity to relatively simple spherical and cylindrical geometries, these models have been invaluable in studying the detailed effects of space charge and the time-history of the plasma sheath. Results for a spherical model of this type from Rothwell et al. [1976, 1977] are presented in Figures 7-23a and b. Of considerable interest are comparisons in Figure 7-23a between the code and a steady state solution for the case of strong electron emission [the "PARKSSG" code of Parker, 1976b]. The probe in this case was biased at +2 V and the electrons emitted at an energy of 1eV, the other parameters are as indicated (secondary emission has been ignored). The results are meant to resemble isotropic photoemission and show very good agreement between the two very different types of codes. Figure 7-23b illustrates the important ability that such codes have in simulating the time-dependent behavior of satellite charging following the turn-on of photoemission as a function of different ambient conditions. The rapid (2–10μs) rise time of the probe in response to

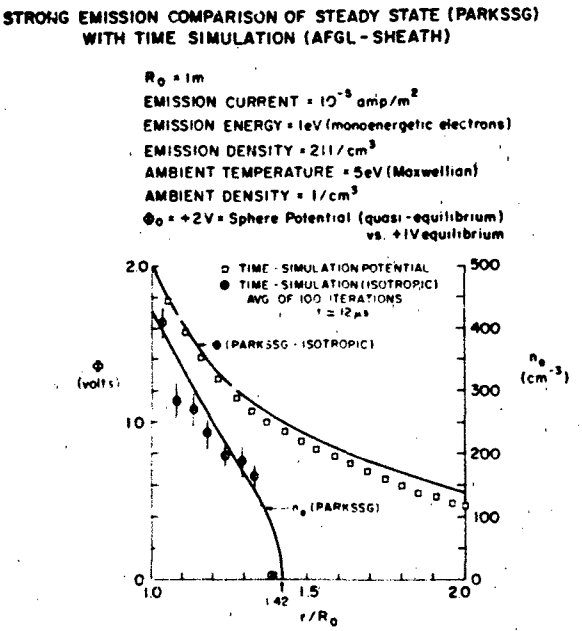


Figure 7-23a. Comparison of steady state sheath potential and density for 2 models: the steady state PARKSSG code and the AFGL-SHEATH code [Rothwell et al., 1977].

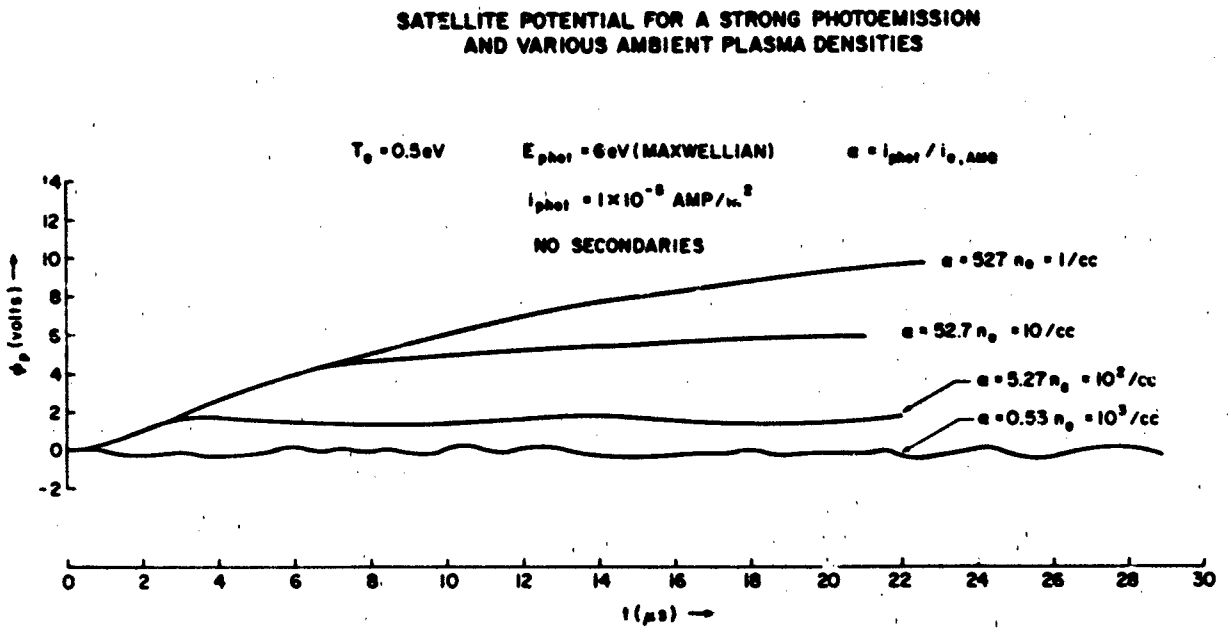


Figure 7-23b. Spacecraft voltage transients following the "turn-on" of the plasma at t = 0. The results are for a sphere (no secondary emission). α is the ratio of photoelectron to ambient electron temperature, and E_{phot} is the mean energy of the photoelectrons assumed to be Maxwellian [Rothwell et al., 1977].

the turn-on of photoemission could be significant in causing satellite transients as only ~2 V are necessary to trigger many circuits.

The cylindrical, particle pushing code AFSIM (*Air Force Satellite Interactions Model*) [Mazzella et al., 1979] has been employed to follow the effects of an electron beam emitted by a satellite. A 300 eV beam at 200 μA was emitted into a vacuum from a satellite initially at 0 potential. The computations showed that the satellite potential rose to +300 V at 36 μs, at which point a portion of the beam was trapped in the sheath and began to orbit the satellite. A space-charge potential formed four spacecraft radii away with most of the beam being reflected back to the satellite. Figure 7-24 shows the configuration 65 μs after turn on. The circle represents the satellite and the beam is visible as a collimated source while the comet-like structure to the right is the cloud of previously trapped beam electrons orbiting the satellite.

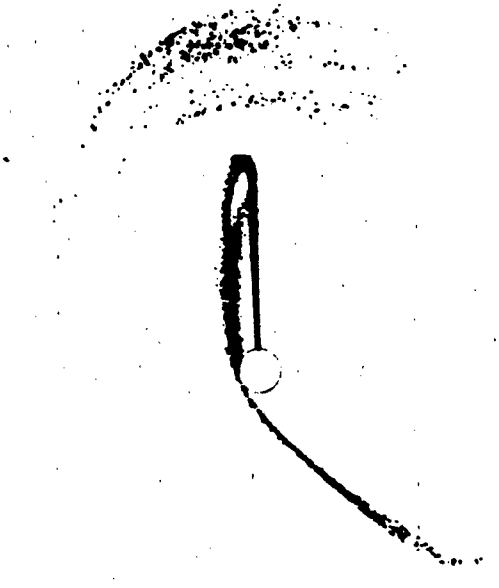


Figure 7-24. Sheath simulation using the AFSIM code [Mazzella et al., 1970] at 65 μs after turn-on of a 30 eV, 200 μA electron beam.

7.4.6 NASCAP

Particle pushing codes are very useful in studying the detailed behavior of particles in time and in the sheaths surrounding a probe. They are still, however, limited to relatively simple geometries because of computer limitations and are subject to numerical instabilities. Recently an alternative model has become available that explicitly treats the effects of geometry. This is the NASCAP (NASA Charging Analyzer Program) computer code [Katz el al., 1977,

1979; Schnuelle et al., 1979; Roche and Purvis, 1979; Rubin et al., 1980] for the thick sheath limit (the model is currently being extended to thin sheaths). This code combines a solution of the Poisson equation and a probe charging model along with a complex graphics package to compute the detailed time behavior of charge deposition on spacecraft surfaces. An approximate circuit model of the satellite is used to estimate voltage changes during time steps for an implicit potential solver. The propagation of particle beams through the satellite sheath is computed by orbit tracing in the sheath field. The code has several options available ranging from a less-detailed code capable of calculating the differential potential on a simple laboratory material sample to a detailed code capable of modeling individual booms and sensor surfaces on a complex satellite such as the P78-2 SCATHA satellite [Rubin et al., 1980]. The primary intent of the model is to compute the effects of satellite geometry on the satellite photosheath, sheath field, and surface material potentials.

The first step in the NASCAP model is to insert the satellite geometry and material content. Using simple 3-dimensional building blocks, the code has provisions for modeling surfaces ranging from cubes and planes to complete satellites with their booms and solar panels (Figure 7-25a). In the second stage the currents to the satellite surfaces are computed by treating each surface element as a current-collecting probe. The inside-out technique may then be used to model the charge distribution in the photosheath. The program is stepped forward in time, local current balance being assumed at each time step. Time-dependency and variations in differential potential can thus be modeled as demonstrated in Figure 7-25b. The analysis is carried out on successively larger, coarser nested grids [Katz et al., 1977], allowing the trajectories of particles emitted from the satellite to be traced in the space surrounding the satellite (Figure 7-25c).

The NASCAP program combines the best elements of

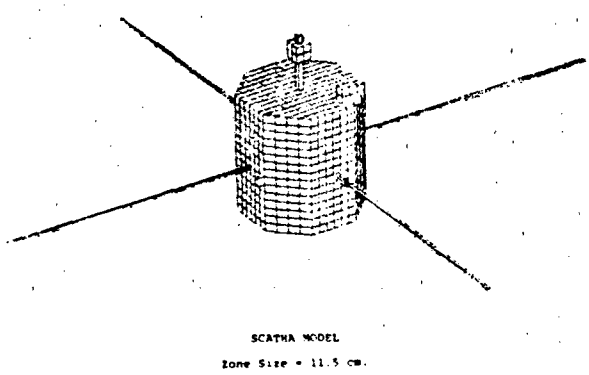


Figure 7-25a. NASCAP computer simulation of the P78-2 SCATHA satellite [Schnuelle et al., 1979]. Compare with Figure 7-6.

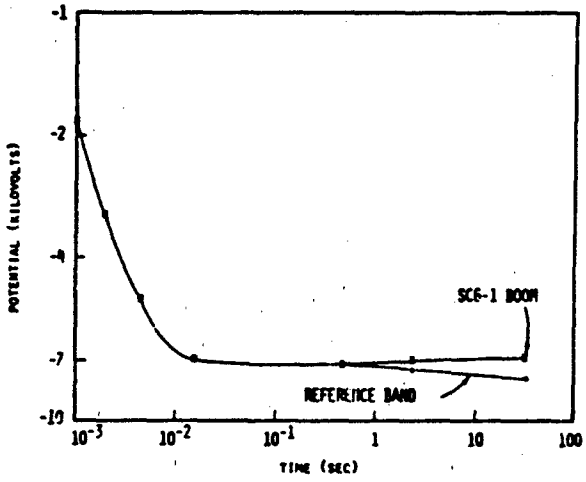


Figure 7-25b. NASCAP computer simulation time-dependence of potential for two surfaces on the P78-2 SCATHA satellite [Schnuelle et al., 1979].

the circuit element model, the inside-out technique, and Langmuir probe theory in one model along with the added advantage of detailed graphical results. Unfortunately, with this gain in capability, the code has become large and requires hours of computing time for the more detailed models. Two steps have been taken to alleviate these problems. First, efficient versions have been developed for specific com-

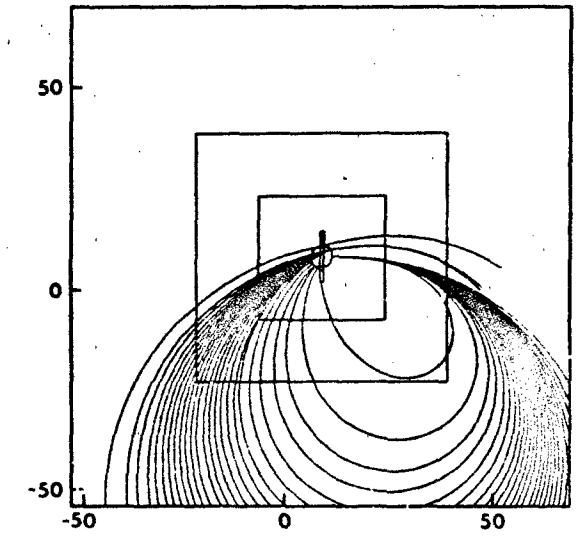


Figure 7-25c. Particle trajectories from an electron emitter as computed by the NASCAP code in the presence of a magnetic field [Rubin et al., 1980]

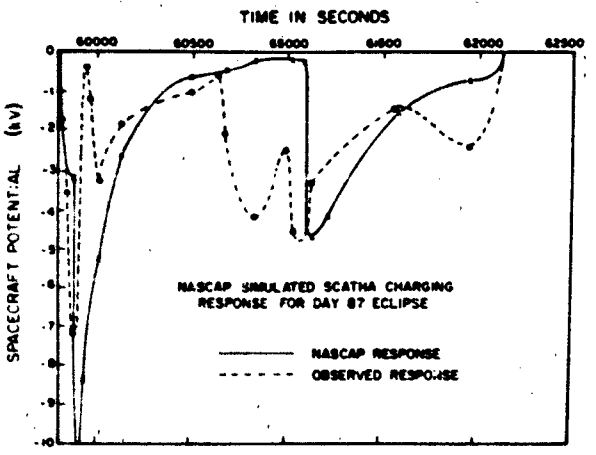


Figure 7-26a. Observed satellite to space potential (P78-2 SCATHA) between 1630 and 1730 UT on day 87, 1979. Also shown are NASCAP predictions for the same period [Schnuelle et al., 1981].

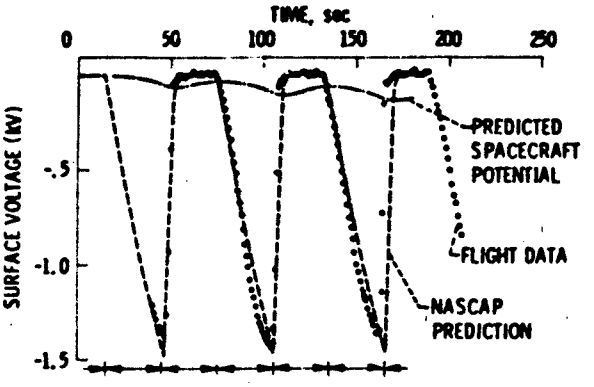


Figure 7-26b. Observed potential difference measured by the P78-2 SCATHA satellite surface potential monitor experiment between a Kapton sample and the satellite ground. Also shown are NASCAP predictions from same period [Stevens et al., 1980a].

putational tasks such as the SCATHA version discussed in Rubin et al. [1980]. Secondly, in conjunction with a number of laboratory and *in situ* satellite experiments, an attempt is being made to verify all aspects of the code and determine what are the critical input parameters [Purvis et al., 1977; Roche and Purvis, 1979; Stevens et al., 1980a]. The model has been applied to many satellite charging problems such as large cavities, exposed potentials, and arbitrary geometries not considered previously [Purvis, 1980; Stevens, 1980; Stevens and Purvis, 1980].

Preliminary results of comparisons between the P78-2 SCATHA measurements and NASCAP have recently become available. Some of these results are plotted in Figures 7-26a and 7-26b. In Figure 7-26a [Schnuelle et al., 1981], the observed satellite to space potential between 1630 and 1730 UT on day 87, 1979 is plotted. The data, consisting of potential and plasma measurements during an eclipse, were input into the NASCAP "one-grid" model (as the name implies, this computation makes use of only the inner most NASCAP computational grid—Figure 7-25c) at ~1 minute (1 spin period) intervals. Although the NASCAP simulation misses the two minor jumps in potential, it reproduces the two major ones and is in excellent quantitative agreement with the data considering the uncertainties in material properties. NASCAP also responds more slowly to environmental changes than the actual data. As Schnuelle et al. [1981] note, this is due to the 1 minute time steps imposed on NASCAP by the data whereas the real environment is changing continuously.

Figure 7-26b [Stevens et al., 1980c], comparing the NASCAP predictions with the surface potential of a Kapton sample (SC1-2—see Table 7-1) on P78-2-SCATHA, is interesting for two reasons. First, it illustrates kV changes in the potential between the Kapton and the satellite ground as the spin frequency of the satellite. Second, it illustrates the power of NASCAP in predicting such rapid variations (the slight time-lag between the data and theory is due to the actual satellite spin period being slightly faster than the assumed 1 rpm). This agreement is, in part, much better than the prediction of the satellite to space potential as the Kapton sample properties were obtained prior to flight in ground test simulations—an important consideration in future studies.

7.5 PREVENTION OF SPACECRAFT CHARGING

Although varying the satellite to space potential allows the measurement of very low energy plasma, charge buildup on satellite surfaces is not in general a desired phenomenon. In order to eliminate or at least limit the worst effects of spacecraft charging, several techniques have been developed. Although the obvious solution is to develop systems that can withstand the worst effects, this is not always a feasible or desirable method. Alternatives to this "brute force" method will be described in this section.

The simplest method for preventing spacecraft charging effects is to employ sound design techniques—use conducting materials where possible and proper grounding techniques. These techniques are detailed in a design guideline handbook recently completed at NASA Lewis [Purvis et al., 1984]. Although a large satellite to space potential can occur, differential charging, the major spacecraft charging problem, is significantly reduced by these procedures. Several different methods have been developed to assure an adequately conducting surface. As an example, non-conducting surfaces on the GEOS series of geosynchronous

satellites were coated with indium oxide. As solar cells are the primary non-conducting surfaces on GEOS and as indium oxide is sufficiently transparent to sunlight so that it does not degrade their operation, this technique has been quite successful [G.L. Wrenn, private communication, 1980] at keeping satellite differential potentials near zero and, because of the secondary emission properties of indium oxide, the satellite to space potential between zero and −1000 V even in eclipse. Such coating techniques, however, can be expensive and difficult if large surfaces are involved. Furthermore, they do not reduce the hazards associated with charge deposition in dielectrics and in the case of "pinholes" (Section 7.3.8), may be ineffective.

Another technique that may be applicable to large surfaces involves the use of electron and ion emitters. Grard [1977] and Gonfalone et al. [1979] discussed the application of such systems to actual satellite systems. The latter paper described the successful application of a low current (mA) electron emitter on the ISEE-1 satellite. The ISEE-1 is in a highly elliptical (300 km to 23 R_E) orbit so that it spends a long time in the solar wind. The ISEE-1 results indicated that the electron cloud emitted by the satellite gun successfully clamped the potential of the satellite at a few volts positive to the ambient solar wind plasma. Purvis and Bartlett [1980] reported results from ion and electron emitters on the geosynchronous ATS 5 and ATS 6 satellites. These results indicated that whereas electron emission alone reduced the satellite to space potential to ~0, it did not significantly reduce the charge on dielectrics. Use of an ion emitter and neutralizer together not only clamped the satellite to space potential at 0, but also, through the cloud of ions, neutralized the negative charge on the dielectrics. Similar success was demonstrated by the beam experiments on the P78-2 SCATHA satellite [H.A. Cohen, private communication, 1980]. There may be some difficulties, however, with these techniques as reducing the surface charge may enhance dielectric breakdown [A.R. Frederickson, private communication, 1980] between the deposited charge and the surface. Also there is the possibility of contamination of the satellite environment by the beam ions.

Careful selection of satellite materials can reduce spacecraft charging. Although thermal control surfaces which are necessary on many satellites generally consist of dielectric materials, careful selection of the materials according to their secondary emission properties and bulk conductivity can reduce charge buildup. Rubin et al. [1978] have demonstrated that for materials with a secondary emission greater than 1, the plasma temperature must be several times the energy at which this occurs if a satellite is to charge up. Again, however, the increased secondary electron population could contaminate low energy ($E \lesssim 10$ eV) plasma observation.

Several other techniques have been proposed (for example, see Beattie and Goldstein [1977] for methods of protecting the Jupiter probe) and recent results indicate that dielectric materials may be altered by the arcing process in a manner which greatly reduces future arcing [A.R. Frederickson, private communication, 1980]. It is currently thought, however, that the techniques described above are adequate in reducing charging. Basically, spacecraft charge prevention is a matter of good design technique—ground well, avoid cavities in which charge can be deposited, and avoid exposed potentials.

7.6 CONCLUSIONS

Before concluding this chapter, a brief summary of the major accomplishments of this fourth period of charging analysis is in order. Probably the major step forward has been the growing realization by the space physics community of the role of spacecraft charging. Before the geosynchronous observations of 10(kV)and higher potentials, spacecraft charging was considered to be a nuisance. Since that time, however, spacecraft charging analysis has become an important adjunct to plasma experiments and to satellite design. On the negative side, however, there is still apparent confusion on the part of some experimentalists as to how to correct low energy measurements and an unwillingness on the part of satellite designers to spend the necessary time in properly designing their satellites. Both of these problems have proven hard to solve.

In the area of plasma measurements, the spacecraft charging theory necessary for their interpretation can be said on the basis of this review to be quite sophisticated. The simple probe theory of Langmuir and his successors has been shown to give adequate order of magnitude estimates of the gross effects of spacecraft charging. The introduction of various sophistications such as the satellite velocity or the satellite sheath [Cauffman and Maynard, 1974, for example] have made this theory applicable to many practical cases. The development of finite element models has allowed fairly sophisticated engineering studies. The original methods of Bernstein and Rabinowitz have grown into the intricate, advanced trajectory codes of Parker, Whipple, and others. In combination with various simplifications these techniques have yielded straightforward methods of correcting Langmuir probe data. Even secondary emission, photoelectron emission, and velocity effects can be modeled.

With the advent of electrostatic analyzers and their ability to provide both detailed spectral information and mass discrimination, experimental information on charging effects has increased enormously. As a result, active studies of the sheath population and fields in the vicinity of a spacecraft under a variety of ambient conditions have become feasible. These have been carried out in detail on the GEOS and P78-2 SCATHA satellites for the geosynchronous orbit. Similar experiments are planned for the early shuttle payloads. Theoretical models of the sheaths and fields around

typical geosynchronous satellites are available, ranging from the simple thick sheath probe models to the NASCAP code which is capable of handling complex geometries. Advanced wake and sheath models are also available for shuttle studies. Rapid time variations on the order of the plasma frequency have even been modeled.

In the area of engineering design, finite element models have been applied in the design of a number of geosynchronous and interplanetary missions. NASCAP has been applied to several systems and proven useful in designing vehicles so as to avoid the worst effects of charging. On the practical side, effort is beginning to be expended in developing charge-reducing materials and techniques. More importantly, the techniques learned on small spacecraft are beginning to be applied in the design of the next generation of large, high voltage vehicles.

Despite this impressive growth of spacecraft charging technology since 1965, there are still a number of areas in need of study. These can be grouped under the headings of material properties, geometry, magnetic fields, wakes, arcs, large size/high potential, and charge deposition in dielectrics. Although work is under way in each of these areas, much still remains to be accomplished. Even so, more than enough has been accomplished in this fourth stage of spacecraft charging analysis so that, as a scientific discipline, spacecraft charging can be said to have come of age.

APPENDIX—Table 7-3 Explanation

The simple probe models of Section 7.4.1. can give first order estimates of the satellite to space potential under a variety of conditions. Given the plasma parameters listed in Table 7-3, this potential has been estimated using approximations to Equations (7.22) and (7.24). As only a first order estimate is desired, a conducting spherical satellite (~ 1 m in diameter) has been assumed. Secondary and backscatter terms are ignored (these would tend to make the potential more positive). For the planar, thin sheath assumption (1-D in Table 7-3): [T is in eV.]

$$V = \frac{-T_E}{q} \text{Ln}\left(\frac{I_{EO}}{I_{IO} + I_{PH}}\right) \qquad V < 0$$

$$V = \frac{-T_I}{q} \text{Ln}\left(\frac{I_{EO} - I_{PH}}{I_{IO}}\right) \qquad V > 0$$

For the spherical, thick sheath assumption (3-D in Table 7-3):

$$V = \frac{-T_E}{q} \text{Ln}\left[\frac{I_{EO}}{I_{IO}\left(1 - q\frac{V}{T_I}\right) + I_{PH}}\right] \qquad V < 0$$

$$V = \frac{-T_I}{q} \text{Ln}\left[\frac{I_{EO}\left(1 + q\frac{V}{T_E}\right) - I_{PH}}{I_{IO}}\right] \qquad V > 0$$

where:

$$I_{EO} = 4\pi r_s^2 \cdot \frac{qn_{EO}}{2}\left(\frac{2T_E}{\pi m_E}\right)^{1/2}$$

$$I_{IO} = 4\pi r_s^2 \cdot \frac{qn_{IO}}{2}\left(\frac{2T_I}{\pi m_I}\right)^{1/2}$$

$$I_{PH} \simeq \pi r_s^2 \cdot J_{PHO}/(1 + V/0.7)^2 \qquad V > 0$$

$$\simeq \pi r_s^2 \cdot J_{PHO} \qquad V < 0$$

r_s = satellite radius

For the "RAM" case and assuming a thin sheath for the electrons:

$$V = \frac{-T_E}{q} \text{Ln}\left(\frac{I_{EO}}{I'_{IO} + I_{PH}}\right) \qquad V < 0$$

$$V = \left[1 - \left(\frac{I_{EO} - I_{PH}}{I'_{IO}}\right)\right] 1/2\, m_I v_s^2 q^{-1} \qquad V > 0$$

Where: $I_{IO} = \pi r_s^2 \cdot qn_{IO}v_s$

v_s = satellite velocity

This assumes, for ion repulsion, that the ion ram current, I_{IO}, is reduced by a factor $(1 - qV/{}^1/_2 m_I v_s^2)$ and that the ions have ~ 0 thermal velocity (see Whipple [1965], p. 28).

The assumed environmental parameters have been adopted from many sources. They should be treated at best as rough approximations as the actual environments can vary by factors of × 10 to × 100. The Jupiter data are from Scudder et al. [1981] and J. Sullivan [private communication]. The solar wind data for less than 1 AU are from Schwenn et al. [1977]. Values greater than 1 AU are estimated.

ACKNOWLEDGMENTS

This work represents the conclusion of 5 years of work at the Air Force Geophysics Laboratory. During that time, my main source of advice and direction was C.P. Pike. Grateful appreciation is also extended to A. Rubin, P. Rothwell, L.W. Parker, and N. Saflekos, who helped me learn the intricacies of spacecraft charging. N.J. Stevens and C. Purvis have provided much practical assistance and advice. T. Gindorf, P. Robinson, E.C. Whipple, Jr., and P. Leung helped in the final revision. M. Spanos and B. Short prepared the transcript. Finally, K. Garrett provided much needed physical support through the long hours expended in producing the manuscript. The work described in this chapter was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under NASA contract NAS7-100.

REFERENCES

Ahmed, M. and R.C. Sagalyn, "Thermal Positive Ions in the Outer Ionosphere and Magnetosphere from OGO 1," *J. Geophys. Res.*, **77:** 1205–1220, 1972.

Albers, N., "Computer Simulation of a Spherical Langmuir Probe," *SU-IPR Report* **449,** Stanford University, Calif., 1973.

Allen, J.E., R.L.F. Boyd, and P. Reynolds, "The Collection of Positive Ions by a Probe Immersed in a Plasma," *Proc. Phys. Soc. London*, **70:** 297–304, 1957.

Al'pert, J.L., Wave-Like Phenomena in the Near-Earth Plasma and Interactions with Man-Made Bodies," in *Handbuch Der Physik*, edited by S. Flugge, *Geophysics III, Part V*, 217–350, Springer-Verlag, Berlin 1976.

Al'pert, J.L., A.V. Gurevich, and L.P. Pitaevskii, *Space Physics with Artificial Satellites*, Consultants Bureau, New York, 186–210, 1965.

Balmain, K.G., "Surface Discharge Effects," in *Space Systems and Their Interactions with the Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, *71*) edited by H.B. Garrett and C.P. Pike, 276–298, AIAA Press, New York, 1980.

Balmain, K.G., M. Cuchanski, and P.C. Kremer, "Surface Micro-Discharges on Spacecraft Dielectrics," in *Proceedings of the Spacecraft Charging Technology Conference*, edited by C.P. Pike and R.R. Lovell, AFGL TR-77-0051/NASA TMX-73537, ADA045459, 519–526, 1977.

Baragiola, R.A., E.V. Alonso, and A.O. Florio, "Electron Emission from Clean Metal Surfaces Induced by Low-Energy Light Ions," *Phys. Rev. B*, **19:** 121–129, 1979.

Beard, D.B. and F.S. Johnson, "Charge and Magnetic Field Interaction with Satellites," *J. Geophys. Res.*, **65:** 1–7, 1960.

Beard, D.B. and F.S. Johnson, "Ionospheric Limitations on Attainable Satellite Potential," *J. Geophys. Res.*, **66:** 4113–4122, 1961.

Beattie, J.R. and R. Goldstein, "Active Spacecraft Potential Control System for the Jupiter Orbiter with Probe Emission," in *Proceedings of the Spacecraft Charging Technology Conference*, edited by C.P. Pike and R.R. Lovell, AFGL TR-77-0051/NASA TMX-73537, ADA045459, 143–165, 1977.

Bernstein, I.B. and I.N. Rabinowitz, "Theory of Electrostatic Probes in a Low-Density Plasma," *Phys. Fluids*, **2:** 112, 1959.

Bernstein, W., et al., "Electron Beam Experiments: The Beam Plasma Discharge at Low Pressures and Magnetic Field Strengths," *Geophys. Res. Lett.*, **5:** 127–130, 1978.

Bernstein, W., H. Leinbach, P.J. Kellogg, S.J. Monson, and T. Hallinan, "Further Laboratory Measurements of the Beam-Plasma Discharge," *J. Geophys. Res.*, **84:** 7271–7278, 1979.

Beers, B.L., et al., "First Principles Numerical Model of Avalanche-Induced Arc Discharges in Electron-Irradiated Dielectrics," *NASA-C-168* (Rev. 10-75), 1979.

Besse, A.L. and A.G Rubin, "A Simple Analysis of Spacecraft Charging Involving Blocked Photoelectron Currents," *J. Geophys. Res.*, **85:** 2324, 1980.

Bohm, D., "Minimum Ionic Kinetic Energy for a Stable Sheath," in *The Characteristics of Elecrical Discharges in Magnetic Fields*, edited by A. Guthrie and R.K. Wakerling, 77–86, McGraw-Hill, New York, 1949.

Bohm, D., H.S. Burhop, and H.S.W. Massey, "The Use of Probes for Plasma Exploration in Strong Magnetic Fields," in *Characteristics of Electrical Discharges in Magnetic Fields*,, edited by A. Guthrie and R.K. Wakerling, McGraw-Hill, New York, 1949.

Bourdeau, R.E., "On the Interaction Between a Spacecraft and an Ionized Medium," *Space Sci. Rev.*, **1:** 719–728, 1963.

Bourdeau, R.E., J. L. Donley, G.P. Serbu, and E.C. Whipple Jr., "Measurement of Sheath Current and Equilibrium Potential on the Explorer VIII Satellite," *J. Astron. Sci.*, **8:** 65–73, 1961.

Brundin, C.L., "Effects of Charged Particles on the Motion of an Earth Satellite," *AIAA J.*, **1:** 2529–2538, 1963.

Cambou, F., R.Z. Sagdeev, and I.A. Zhulin, "Reviewing Artificial Radiation and Aurorae Between Kerguelen and Soviet Union," *Space Sci. Instrument.*, **4:** 117–121, 1978.

Cauffman, D.P., "Ionization and Attraction of Neutral Molecules to a Charged Spacecraft," SAMSO TR-73-263, 1973.

Cauffman, D.P., "The Effects of Photoelectron Emission on a Multiple-Probe Spacecraft Near the Plasmapause," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 153–162, D. Reidel, Dordrecht, Holland, 1973.

Cauffman, D.P. and N.C. Maynard, "A Model of the Effect of the Satellite Photosheath on a Double Floating Probe System," *J. Geophys. Res.*, **79:** 2427–2438, 1974.

Cernuschi, F., "The Physics of Cosmic Grains," *Astrophys. J.*, **105:** 241–254, 1947.

Chang, K.W. and G.K. Bienkowski, "Effects of Electron Emission on Electrostatic Probes at Arbitrary Pressures," *Phys. Fluids*, **13:** 902, 1970.

Chang, H.H.C. and M.C. Smith, "On the Drag of a Spherical Satellite Moving in a Partially Ionized Atmosphere," *J. Brit. Interplanet. Soc.*, **17:**199–205, 1960.

Chang, J.S., S.M.L. Prokopenko, R. Godard, and J.G. Laframboise, "Prediction of Ion Drift Effects on Spacecraft Floating Potentials," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Piek, NASA CP-2071/AFGL TR-79-0082, ADA084626, 179–187, 1979.

Chen, F.F., "Electric Probes," in *Plasma Diagnostic Techniques*, edited by R.H. Huddlestone and S.L. Leonard, 113–200, Academic Press, New York, 1965.

Chopra, K.P., "Interactions of Rapidly Moving Bodies in Terrestrial Atmosphere," *Rev. Mod. Phys.*, **33:** 153–189, 1961.

Chung, M.S. and T.E. Everhart, "Simple Calculation of nergy Distribution of Low-Energy Secondary Electrons Emitted from Metals Under Electron Bombardment," *J. Appl. Phys.*, **45:** 707, 1974.

Cipolla, J.W. and M.B. Silevitch, "Analytical Study of the Time Dependent Spacecraft-Plasma Interaction," in *Spacecraft Charging Technology—1978*, edited by R.C.

Finke and C.P. Pike, NASA CP-2071/AFGL TR-79-0082, ADA084626, 197–208, 1979.

Cohen, H.A., et al., "Design, Development, and Flight of a Spacecraft Charging Sounding Rocket Payload," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-2071/AFGL TR-79-0082, ADA084626, 80–90, 1979.

Davis, A.H. and I. Harris, "Interaction of a Charged Satellite with the Ionosphere," in *Rarefied Gas Dynamics*, edited by L. Talbot, 691–699, Academic Press Inc., New York, 1961.

DeForest, S.E., "Spacecraft Charging at Synchronous Orbit," *J. Geophys. Res.*, **77:** 651, 1972.

DeForest, S.E., "Electrostatic Potentials Developed by ATS 5," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 263–276, D. Reidel, Dordrecht, Holland, 1973.

DeForest, S.E. and C.E. McIlwain, "Plasma Clouds in the Magnetosphere," *J. Geophys. Res.*, **76:** 3587–3611, 1971.

Dekker, A.J. "Secondary Electron Emission," in his *Solid State Physics*, 418–445, Prentice-Hall, Englewood Cliffs, N.J., 1958.

deLeeuw, J.H., "A Brief Introduction to Ionospheric Aerodynamics," in *Rarefied Gas Dynamics*, edited by C.L. Brundin, **2:**1561–1587, Academic Press, New York, 1967.

Deutsch, M.-J.C., "Worst Case Earth Charging Environment," *J. Spacecraft*, **19:** 473–477, 1982.

Durrett, J.C. and J.R. Stevens, "Description of the Space Test Program P78-2 Spacecraft and Payloads," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-2071/AFGL TR-79-0082, ADA084626, 4-10, 1979.

Evdokimov, O.B. and N.P. Tubalov, "Stratification of Space Charge in Dielectrics Irradiated with Fast Electrons," *Sov. Phys. Solid State*, Engl. Trans., **15:** 1869–1870, 1974.

Fahleson, U., "Plasma-Vehicle Interactions in Space -- Some Aspects on Present Knowledge and Future Development," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 563–570, D. Reidel, Dordrecht, Holland, 1973.

Feuerbacher, B., R.F. Willis, and B. Fitton, "Electrostatic Charging and Formation of Composite Interstellar Grains," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 415–428, D. Reidel, Dordrecht, Holland, 1973.

Finke, R.C. and C.P. Pike (eds.), *Spacecraft Charging Technology—1978*, NASA CP-2071/AFGL TR-79-0082, ADA084626, 1979.

Frederickson, A.R., "Electric Fields in Irradiated Dielectrics," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-2071/AFGL TR-79-0082, ADA084626, 554–569, 1979.

Frederickson, A.R., "Radiation Induced Dielectric Charging," in *Space Systems and Their Interactions with the Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, **71**) edited by H.B. Garrett and C.P. Pike, 386–412, AIAA Press, New York, 1980.

Freeman, J.W. Jr., M.A. Fenner, and H.K. Hills, "The Electric Potential of the Moon in the Solar Wind," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 363–368, D. Reidel, Dordrecht, Holland, 1973.

Garrett, H.B., "Review of Quantitative Models of the 0 to 100 KeV Near-Earth Plasma," *Rev. Geophys.*, **17:** 397–417, 1979.

Garrett, H.B., "The Charging of Spacecraft Surfaces," *Rev. Geophys. Space Phys.*, **19:** 577–616, 1981.

Garrett, H.B. and S.E. DeForest, "Time-Varying Photoelectron Flux Effects on Spacecraft Potential at Geosynchronous Orbit," *J. Geophys. Res.*, **84:** 2083–2088, 1979.

Garrett, H.B. and J.M. Forbes, "A Model of Solar Flux Attenuation During Eclipse Passage and Its Effects on Photoelectron Emission from Satellite Surfaces," *Planet. Space Sci.*, **29:** 601–607, 1981.

Garrett, H.B. and A.G. Rubin, "Spacecraft Charging at Geosynchronous Orbit—Generalized Solution for Eclipse Passage," *Geophys. Res. Lett.*, **5:** 865, 1978.

Garrett, H.B., E.G. Mullen, E. Ziemba, and S.E. DeForest, "Modeling of the Geosynchronous Plasma Environment—Part 2, ATS 5 and ATS 6 Statistical Atlas," AFGL TR-78-0304, ADA067018, 1978.

Garrett, H.B., A.G. Rubin, and C.P. Pike, "Prediction of Spacecraft Potentials at Geosynchronous Orbit," in *Solar-Terrestrial Prediction Proceedings*, edited by R.F. Donnelly, Vol. II, 104–118, U.S. Government Printing Office, 1979.

Gibbons, D.S., "Secondary Electron Emission," in *Handbook of Vacuum Physics*, 2, edited by A.H. Beck, 301, Pergamon Press, Oxford, U.K., 1966.

Goldan, P.D., E.J. Yadlowsky, and E.C. Whipple Jr., "Errors in Ion and Electron Temperature Measurements due to Grid Plane Potential Nonuniformities in Retarding Potential Analyzers," *J. Geophys. Res.*, **78:** 2907–2916, 1973.

Goldstein, H., *Classical Mechanics*, Addison-Wesley, Reading, Mass, 63–69, 1965.

Goldstein, R. and S.E. DeForest, "Active Control of Spacecraft Potentials at Geosynchronous Orbit," in *Spacecraft Charging by Magnetospheric Plasmas*, (*Prog Astronaut. and Aeronaut.*, **47**) edited by A. Rosen, 159–168, MIT Press, Cambridge, Mass., 1976.

Gonfalone, A., A. Knudsen, U.V. Fahleson, C.G. Fälthammar, F.S. Mozer, and R.B. Tobert, "Spacecraft Potential Control on ISEE-1," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-2071/AFGL TR-79-0082, ADA084626, 256–267, 1979

Grard, R.J.L. (ed.), *Photon and Particle Interactions with Surfaces in Space*, D. Reidel, Dordrecht, Holland, 1973.

Grard, R.J.L., "Properties of the Satellite Photoelectron Sheath Derived from Photoemission Laboratory Measurements," *J. Geophys. Res.*, 78: 2885–2906, 1973b.

Grard, R.J.L., "The Multiple Applications of Electron Emitters in Space," in *Proceedings of the Spacecraft Charging Technology Conference*, edited by C.P. Pike and R.R. Lovell, AFGL TR-0051/NASA TMX-73537, ADA045459, 203–221, 1977.

Grard, R.J.L. and J.K.E. Tunaley, "Photoelectron Sheath Near the Planar Probe in Interplanetary Space," *J. Geophys. Res.*, 76: 2498–2505, 1971.

Grard, R.J.L., K. Knott, and A. Pedersen, "The Influence of Photoelectron and Secondary Electron Emission on Electric Field Measurements in the Magnetosphere and Solar Wind," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, D. Reidel, Dordrecht, Holland, 163–192, 1973.

Grard, R.J.L., S.E. DeForest, and E.C. Whipple Jr., "Comment on Low Energy Electron Measurements in the Jovian Magnetosphere," *Geophys. Res. Lett.*, **4:** 247, 1977.

Grebowski, R. and T. Fischer, "Theoretical Density Distribution of Plasma Around a Cylinder," *Planet. Space Sci.*, 23: 287–304, 1975.

Gringauz, K.I. and M.Kh. Zelikman, "Measurement of the Concentration of Positive Ions Along the Orbit of an Artificial Satellite," *Uspekhi Fiz. Nauk*, 63: 239–252, 1957; (translated in *The Russian Literature of Satellites, Part II*, International Physical Index, New York, 133–147, 1958.)

Gringauz, K.I., V.V. Bezrukikh, and V.D. Ozerov, "Measurements of the Concentration of Positive Ions in the Ionosphere by Ion Traps on the Third Soviet Earth Satellite," *Iskusstv. Sputniki Zemli*, **6:** 63–100, 1961.

Gross, B. and S.V. Nablo, "High Potentials in Electron-Irradiated Dielectric," *J. Appl. Phys.*, **38:** 2272–2275, 1967.

Guernsey, R.L. and J.H.M. Fu, "Potential Distributions Surrounding a Photo-Emitting Plate in a Dilute Plasma," *J. Geophys. Res.*, **75:** 3193–3199, 1970.

Gurevich, A.V. and Ya.S. Dimant, "Flow of a Rarefied Plasma Around a Disk," *Geomagn. Aeron.*, Engl. Transl., **15:** 183–190, 1975.

Gurevich, A.V. and L.P. Pitaevskii, "Non-Linear Dynamics of Rarefied Gas," *Prog. Aerospace Sci.*, **16:** 227, 1975.

Gurevich, A.V., L.V. Pariiskaya, and L.P. Pitaevskii, "Self-Similar Motion of Rarefied Plasma," *Sov. Phys. JETP*, Engl. Transl., **22:** 449, 1966.

Gurevich, A.V., L.V. Pariiskaya, and L.P. Pitaevskii, "Self-Similar Motion of Low Density Plasma," *Sov. Phys. JETP*, Engl. Transl., **27:** 476, 1968.

Gurevich, A.V., L.V. Pariiskaya, and L.P. Pitaevskii, "Ion Acceleration Upon Expansion of a Rarefied Plasma," *Sov. Phys. JETP*, Engl. Transl., **36:** 274–281, 1973.

Gurevich, A.V., L.P. Pitaevskii, and V.V. Smirnov, "Ionospheric Aerodynamics," *Sov. Phys. Usp.*, **99:** 595, 1970.

Hachenberg, O. and W. Brauer, "Secondary Electron Emission from Solids," in *Advances in Electronics and Electron Physics*, edited by L. Martin, **11:** 413–499, Academic Press, New York, 1959.

Hanson, W.B., D.D. McKibbin, and G.W. Sharp, "Some Ionospheric Measurements with Satellite-Borne Ion Traps," *J. Geophys. Res.*, **69:** 2747–2763, 1964.

Hanson, W.B., S. Sanatani, D. Zuccaro, and T.W. Flowerday, "Plasma Measurements with the Retrading Potential Analyzer on OGO 6," *J. Geophys. Res.*, **75:** 5483–5501, 1970.

Henderson, C.L. and U. Samir, "Observations of the Disturbed Region Around an Ionospheric Spacecraft," *Planet. Space Sci.*, **15:** 1499–1513, 1967.

Hendrickson, R.A., "The Electron Echo Experiment, Observations of the Charge Neutralization of the Rocket and Analysis of the Echo from Electrons Artifically Injected into the Magnetosphere," *Tech. Report CR-160*, School Physics and Astronomy, University of Minnesota, Minneapolis, 1972.

Hess, W.N., "Generation of an Artificial Aurora," *Science*, **164:** 1512, 1969.

Hinteregger, H.E., "Combined Retarding Potential Analysis of Photoelectrons and Environment Charged Particles up to 234 km," *Space Res.*, **1:** 304–327, 1960.

Inouye, G.T., "Spacecraft Potentials in a Substorm Environment," in *Spacecraft Charging by Magnetospheric Plasma*, (*Prog. Astronaut. Aeronaut.*, **42**) edited by A. Rosen, 103–120, MIT Press, Cambridge, Mass., 1976.

Jacobsen, T.A. and N.C. Maynard, "Evidence for Significant Spacecraft Charging by an Electron Accelerator at Ionospheric Altitudes," *Planet. Space Sci.*, **28:** 291–307, 1980.

Jastrow, R and C.A. Fearse, "Atmospheric Drag on the Satellite," *J. Geophys. Res.*, **62:** 413–423, 1957.

Jemiola, J.M., *Proceedings of the USAF/NASA International Spacecraft Contamination Conference*, AFML TR-78-190/NASA CP-2039, 1978.

Jemiola, J.M., "Spacecraft Contamination: A Review," in *Space Systems and Their Interactions with Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, **71**) edited by H.B. Garrett and C.P. Pike, 680–706), AIAA Press, New York, 1980.

Jew, H., "Numerical Studies of the Rarefied Plasma Interaction at Meso-Thermal Speeds," Ph.D. Thesis, University of Michigan, Ann Arbor, Mich., 1968.

Jew, H., "Reply," *J. Geophys. Res.*, **78:** 6829, 1973.

Johnson, B., J. Quinn, and S.E. DeForest, "Spacecraft Charging on ATS-6," in *Effect of the Ionosphere on Space and Terrestrial Systems*, edited by Jonn Goodman, U.S. Government Printing Office, Washington, D.C., 322–327, 1978.

Johnson, C.Y. and E.B. Meadows, "First Investigation of Ambient Positive-Ion Composition to 219 km by Rocket-Borne Spectrometer," *J. Geophys. Res.*, **60:** 193–203, 1955.

Jung, B., "The Origin of Solid Particles in Interstellar Space," *Astron. Nach.*, **263:** 426, 1937.

Kasha, M.A., *The Ionosphere and its Interaction with Satellites*, Gordon and Breach, New York, 1969.

Katz, I., D.E. Parks, S. Wang, and A. Wilson "Dynamic Modeling of Spacecraft in a Collisionless Plasma," in *Proceedings of the Spacecraft Charging Technology Conference*, edited by C.P. Pike and R.R. Lovell, AFGL TR-77-0051/TMX-73537, ADA045459, 319–330, 1977.

Katz, I., J.J. Cassidy, M.J. Mandell, G.W. Schnuelle, P.G. Steen, and J.C. Roche, "The Capabilities of the NASA Charging Analyzer Program," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-2071/AFGL TR-79-0082, ADA045459, 101–122, 1979.

Kennerud, K.L., "High Voltage Solar Array Experiments," *NASA CR-131280*, Boeing Co., Seattle, March 1974.

Knott, K., "The Equilibrium Potential of a Magnetospheric Satellite in an Eclipse Situation," *Planet. Space Sci.*, **20:** 1137–1146, 1972

CHAPTER 7

Knott, K., "Payload of the 'GEOS' Scientific Geostationary Satellite," *ESA Sci. Tech. Rev.*, **1:** 173-196, 1975.

Knudsen, W.C., "Evaluation and Demonstration of the Use of Retarding Potential Analyzers for Measuring Several Ionospheric Quantities," *J. Geophys. Res.*, **71:** 4669, 1966.

Knudsen, W.C. and G.W. Sharp, "Ion Temperatures Measured Around a Dawn-Dusk Auroral-Zone Satellite Orbit," *J. Geophys. Res.*, **72:** 1061–1072, 1967.

Koons, H.C., "Characteristics of Electrical Discharges on the P78-2 Satellite (SCATHA)," Paper 80-0333, AIAA 18th Aerospace Sciences Meeting, AIAA Press, New York, 1980.

Krassovsky, V.I., "Exploration of the Upper Atmosphere with the Help of the Third Soviet Sputnik," *Proc. IRE*, **47:** 289–196, 1959.

Kunemann, B., "The Collisionless Flow of Unmagnetized Plasma Around Bodies, 1," *J. Plasma Phys.*, **20:** 17, 1978.

Kurt, P.G. and V.I. Moroz, "The Potential of a Metal Sphere in Interplanetary Space," *Iskusstvennye Sputniki Zemli*, 7: 78–88, 1961; (translated in *Planet. Space Sci.*, **9:** 259–268, 1962).

Laframboise, J.G., "Theory of Spherical and Cylindrical Langmuir Probes in a Collisonless Maxwellian Plasma at Rest," *Report 100*, University of Toronto Institute of Aerospace Studies, Canada, 1966.

Laframboise, J.B. and E.C. Whipple Jr., Comments on paper by Uri Samir and Howard Jew, "Comparison of Theory with Experiment for Electron Density Distribution in the Near Wake of an Ionospheric Satellite," *J. Geophys. Res.*, **78:** 6827–6828, 1973.

Laframboise, J.G., R. Godard, And S.M.L. Prokopenko, "Numerical Calculations of High-Altitude Differential Charging: Preliminary Results," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-2071/AFGL TR-79-0082, ADA084626, 188–196, 1979.

Langmuir, I., "The Interaction of Electron and Positive Ion Spacecharge in Cathode Sheaths," *Phys. Rev.*, **33&34**, 1929 (reprinted in *Collected Works of Irving Langmuir*, edited by G. Suits, **5**, MacMillan, New York, 1961).

Langmuir, I., and K.B. Blodgett "Currents Limited by Space Charge Between Concentric Spheres," *Phys Rev.*, **24:** 49, 1924.

Lehnert, B., "Electrodynamic Effects Connected with the Motion of a Satellite of the Earth," *Tellus*, **3:** 408–409, 1956.

Liemohn, H.B., "Induced Charging of Shuttle Orbiter by High Electron-Beam Currents," in *Proceedings of the Spacecraft Charging Technology Conference*, edited by C.P. Pike and R.R. Lovell, AFGL TR-77-0051/NASA TMX-73537, ADA045459, 271–286, 1977.

Linson, L.M., "Current-Voltage Characteristics of an Electron-Emitting Satellite in the Ionosphere," *J. Geophys. Res.*, **74:** 2368–2375, 1969.

Liu, V.C. and H. Jew, "Near Wake of the Rarefield Plasma Flows at Mesothermal Speeds," *Paper 68-169*, AIAA Press, New York, 1968.

Liu, V.C., "Ionospheric Gas Dynamics of Satellites and Diagnostic Probes," *Space Sci. Rev.*, **9:** 423, 1969.

Lucas, A.A., Fundamental Processes in Particle and Photon Interactions with Surfaces," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 3–21, D. Reidel, Dordrecht, Holland, 1973.

Manka, R.H., "Plasma and Potential at the Lunar Surface," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, D. Reidel, Dordrecht, Holland, 347–367, 1973.

Massaro, M.J., T. Green, and D. Ling, "A Charging Model for Three-Axis Stabilized Spacecraft," *Proceedings of the Spacecraft Charging Conference*, edited by C.P. Pike and R.R. Lovell, AFGL TR-77-0051/NASA TMX-73537, ADA045459, 237-270, 1977.

Mazzella, A., E. Tobenfeld, and A.G. Rubin, "AFSIM—An Air Force Satellite Interactions Model," AFGL TR-79-0138, ADA078032, 1979.

McCoy, J.E., A Konradi, and O.K. Garriott, "Current Leakage for Low Altitude Satellites," in *Space Systems and Their Interactions with Earth's Space Environment, Prog. Astronaut. Aeronaut.*, edited by H.B. Garrett and C.P. Pike, **71:** 523–553, AIAA Press, New York, 1980.

McPherson, D.A. and W.R. Schober, "Spacecraft Charging at High Altitudes: The SCATHA Satellite Program," in *Spacecraft Charging by Magnetospheric Plasmas, Prog. Astronaut. Aeronaut.*, edited by A. Rosen, **47:** 15–30 MIT Press, Cambridge, Mass., 1976.

Medicus, G., "Theory of Electron Collection of Spherical Probes," *J. Appl. Phys.*, **32:** 2512–2520, 1961.

Meulenberg, A. Jr., "Evidence for a New Discharge Mechanism for Dielectrics in a Plasma," in *Spacecraft Charging by Magnetospheric Plasmas, Prog. Astronaut. Aeronaut.*, edited by A. Rosen, **47:** 237–246, MIT Press, Cambridge, Mass., 1976.

Mizera, P.F., "Natural and Artificial Charging" Results from the Satellite Surface Potential Monitor Flown on P78-2," *Paper 80-0334*, AIAA 18th Aerospace Sciences Meeting, AIAA Press, New York, 1980.

Mizera, P.F., E.R. Schnauss, R. Vandre, and E.G. Mullen, "Description and Charging Results from the RSPM," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-2071/AFGL TR-79-0082, ADA084626, 91–100, 1979.

Montgomery, M.D., J.R. Ashbridge, S.J. Ba and E.W. Hones, "Low-Energy Electron Measurements and Spacecraft Potential: Vela 5 and Vela 6," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 247–262, D. Reidel, Dordrecht, Holland, 1973.

Mott-Smith, H.M. and I. Langmuir, "The Theory of Collectors in Gaseous Discharges," *Phys. Rev.*, **28:** 727–763, 1926.

Mullen, E.G. and M.S. Gussenhoven, "SCATHA Environmental Atlas," AFGL TR-83-0002, ADA131456, 1983.

Mullen, E.G., M.S. Gussenhoven, and H.B. Garrett, "A 'Worst Case' Spacecraft Environment as Observed by SCATHA on 24 April 1979," AFGL TR-81-0231, ADA108680, 1981.

Narcisi, R.S., A.D. Bailey, and L. Della Lucca, "Composition Measurements of Negative Ions in the D and

Lower E Regions (Abstract)," *Eos Trans. AGU*, **49:** 149, 1968.

Nanevicz, J.E. and R.C. Adamo, "Occurrence of Arcing and Its Effects on Space Systems," in *Space Systems and Their Interactions with Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, **71**) edited by H.B. Garrett and C.P. Pike, 252–275, AIAA Press, New York, 1980.

Norman, K., and R.M. Freeman, "Energy Distribution of Photoelectron Emitted from a Surface on the OGO-5 Satellite and Measurements of Satellite Potential," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 231–246, D. Reidel, Dordrecht, Holland, 1973.

Olsen, R.C., "Differential and Active Charging Results from the ATS Spacecraft," Ph.D. Thesis, University of California, San Diego, 1980.

O'Neil, R.R., F. Bien, D. Burt, J.A. Sandock, and A.T. Stair, "Summarized Results of the Artificial Auroral Experiment, Precede," *J. Geophys. Res.*, **83:** 3273, 1978a.

O'Neil, R.R., W.P. Reidy, J.W. Carpenter, R.N. Davis, D. Newell, J.C. Ulwick, and A.T. Stair, "Excede 2 Test, An Artificial Auroral Experiment: Ground-based Optical Measurements," *J. Geophys. Res.*, **83:** 3281, 1978b.

Opik, E.J., "Interplanetary Dust and Terrestrial Accretion of Meteoric Matter," *Irish Astron. J.*, **4:** 84, 1956.

Opik, E.J., "Particle Distribution and Motion in a Field of Force," in *Interactions of Space Vehicles with an Ionized Atmosphere*, edited by S.F. Singer, 3–60, Pergamon Press, New York, 1965.

Parker, L.W., "Computer Solutions in Electrostatic Probe Theory," Mt. Auburn Res. Associates, Inc., Newton, Mass., AFAL TR-72-222, 1973.

Parker, L.W., "Computer Method for Satellite Plasma Sheath in Steady-State Spherical Symmetry," Lee W. Parker, Inc., Concord, Mass., AFCRL-TR-75-0410, ADA015066, July 1975.

Parker, L.W., "Computation of Collisionless Steady State Plasma Flow Past a Charged Disk," *NASA CR-144159*, 1976a.

Parker, L.W., "Theory of Electron Emission Effects in Symmetric Probe and Spacecraft Sheaths," Lee W. Parker, Inc., Concord, Mass., AFGL-TR-76-0294, A037538, 1976b.

Parker, L.W., "Calculation of Sheath and Wake Structure about a Pillbox-Shaped Spacecraft in a Flowing Plasma," in *Proceedings of the Spacecraft Charging Technical Conference*, edited by C.P. Pike and R.R. Lovell, AFGL-TR-77-0051/NASA TMX-73537, ADA045459, 331–366, 1977.

Parker, L.W., "Differential Charging and Sheath Asymmetry of Nonconducting Spacecraft Due to Plasma Flow," *J. Geophys. Res.*, **83:** 4873–4876, 1978a.

Parker, L.W., "Potential Barriers and Asymmetric Sheaths Due to Differential Charging of Nonconducting Spacecraft," AFGL-TR-78-0045, ADA053618, 1978b.

Parker, L.W., "Plasma Sheath Effects and Voltage Distributions of Large High Power Satellite Solar Arrays," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-2071/AFGL TR 79-0082 ADA084626, 341–375, 1979.

Parker, L.W., "Plasmasheath-Photosheath Theory for Large High-Voltage Space Structures," in *Space Systems and Their Interactions with Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, **71**) edited by H.B. Garrett and C.P. Pike, 477–522, AIAA Press, New York, N.Y., 1980.

Parker, L.W., and B.L. Murphy, "Potential Buildup on an Electron-Emitting Ionospheric Satellite," *J. Geophys. Res.*, **72:** 1631, 1967.

Parker, L.W., and E.C. Whipple, Jr., "Theory of a Satellite Electrostatic Probe," *Ann. Physics*, **44:** 126–161, 1967.

Parker, L.W., and E.C. Whipple, Jr., "Theory of Spacecraft Sheath Structure, Potential, and Velocity Effects on Ion Measurements by Traps and Mass Spectrometers," *J. Geophys. Res.*, **75:** 4720, 1970.

Parks, D.E., and I. Katz, "Charging of a Large Object in Low Polar Earth Orbit," in *Proceedings USAF/NASA Spacecraft Charging Technical Conference, III*, 1981.

Pike, C.P. and M.H. Bunn, *Spacecraft Charging by Magnetospheric Plasmas*, (*Prog. Astronaut. Aeronaut.*, **47**) edited by A. Rosen, 45–60, MIT Press, Cambridge, Mass., 1976.

Pike C.P., and R.R. Lovell (eds.), *Proceedings of the Spacecraft Charging Technology Conference*, AFGL-TR-77-0051/NASA TMX-73537, 1977.

Prokopenko, S.M.L., and J.G. Laframboise, "Prediction of Large Negative Shaded-side Spacecraft Potentials," in *Proceedings of the Spacecraft Charging Technology Conference*, edited by C.P. Pike and R.R. Lovell, AFGL-TR-77-0051/NASA, TMX-73537, ADA045459, 369–387, 1977.

Prokopenko, S.M.L., and J.G. Laframboise, "High-Voltage Differential Charging of Geostationary Spacecraft," *J. Geophys. Res.*, **85:** 4125–4131, 1980.

Purvis, C.K., "Configuration Effects on Satellite Charging Response," Paper 80-0040, *AIAA 18th Aerospace Science Conference*, 1980.

Purvis, C.K., and R.O. Bartlett, "Active Control of Spacecraft Charging," in *Space Systems and Their Interactions With the Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, **71**) edited by H.B. Garrett and C.P. Pike 299–317, AIAA Press, New York, N.Y., 1980.

Purvis, C.K., H.B. Garrett, A. Whittlesey, and N.J. Stevens, *Handbook of Design Guidelines for Assessing and Controlling Spacecraft Charging Effects*, NASA Lewis Research Center, NASA TP 2361, September 1984.

Purvis, C.K., N.J. Stevens, and J.C. Oglebay, "Charging Characteristics of Materials: Comparison of Experimental Results with Simple Analytical Models " in *Proceedings Spacecraft Charging Conference*, edited by C.P. Pike and R.R. Lovell, AFGL-TR-77-0051/NASA TMS-73537 ADA045459, 459–486, 1977.

Reasoner, D.L., W. Lennartsson, and C.R. Chappell, "Relationship Between ATS-6 Spacecraft Charging Occurrences and Warm Plasma Encounters," in *Spacecraft Charging by Magnetospheric Plasmas*, (*Prog. Astronaut. Aeronaut.*, **47**) edited by A. Rosen, 89–102, MIT Press, Cambridge, Mass., 1976.

Reddy, B.M., L.H. Brace, and J.A. Findlay, "The Ionosphere at 640 Kilometers on Quiet and Disturbed Days," *J. Geophys. Res.*, **72:** 2709–2727, 1967.

Reiff, P.H., J.W. Freeman, and D.L. Cooke, "Environmental Protection of the Solar Power Satellite," in *Space Systems and Their Interactions With the Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, **71**) edited by H.B. Garrett and C.P. Pike, 554–576, AIAA Press, New York, 1980.

Robinson, P.A., Jr., and A.B. Holman, "Pioneer Venus Spacecraft Charging Model," in *Proceedings Spacecraft Charging Technical Conference*, edited by C.P. Pike and R.R. Lovell, AFGL-TR-77-0051/NASA TMX-73537, 297–308, ADA045459, 1977.

Roche, J.C., and C.P. Purvis, "Comparison of NASCAP Predicitions with Experimental Data," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-1072/AFGL-TR-79-0082, ADA084626, 144–157, 1979.

Rosen, A. (ed.), *Spacecraft Charging by Magnetospheric Plasmas*, (*Prog. Astronaut. Aeronaut.*, **47**) MIT Press, Cambridge, Mass., 1976.

Rosen, A., R.W. Fredericks, Inouye, N.L. Saunders, F.L. Scarf, W.M. Greenstadt, J.L. Vogl, and J.M. Sellen, "RGA Analysis: Findings Regarding Correlation of Satellite Anomalies with Magnetospheric Substorms and Laboratory Test Results," *Report 09760-7020-R0-00*, Aug 1972, TRW Systems Group, Redondo Beach, Calif.

Rothwell, P.L., A.G. Rubin, A.L. Pavel, and L. Katz, "Simulation of the Plasma Sheath Surrounding a Charged Spacecraft," in *Spacecraft Charging by Magnetospheric Plasmas*, (*Prog. Astronaut. Aeronaut.*, **47**) edited by A. Rosen, 121–134, MIT Press, Cambridge, Mass., 1976.

Rothwell, P.L., A.G. Rubin, and G.K. Yates, "A Simulation Model of Time-Dependent Plasma-Spacecraft Interactions," *Proc. Spacecraft Charging Conf.*, edited by C.P. Pike and R.R. Lovell, AFGL-TR-77-0051/NASA TMS-73537, ADA045459, 381–412, 1977.

Rubin, A.G., P.L. Rothwell, and G.K. Yates, "Reduction of Spacecraft Charging Using Highly Emissive Surface Materials," *Proceedings of the 1978 Symposium on the Effects of the Ionosphere on Space and Terrestrial Systems*, Naval Research Laboratory, Washington, D.C., 1978.

Rubin, A.G. et al., I. Katz, M. Mandell, G. Schnuelle, P. Steen, D. Parks, J. Cassidy, and J. Roche, "A Three-Dimensional Spacecraft Charging Computer Code," in *Space Systems and Their Interactions With the Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, **71**) edited by H.B. Garrett and C.P. Pike, 318–336, AIAA Press, New York, 1980.

Sagalyn, R.C., and W.J. Burke, "INJUN 5 Observations of Vehicle Potential Fluctuations at 2500 km," in *Proceedings of Spacecraft Charging Technology Conference*, edited by C.P. Pike and R. R. Lovell, AFGL TR-77-0051, NASA TMS-73537, 67–80, 1977.

Sagalyn, R.C., M. Smiddy, and J. Wisnia, "Measurement and Interpretation of Ion Density Distributions in the Daytime F Region," *J. Geophys. Res.*, **63:** 199–211, 1963.

Samir, U., "Charged Particle Distribution in the Nearest Vicinity of Ionospheric Satellites - Comparison of the Main Results from the Ariel I, Explorer 31, and Gemini-Agena 10 Spacecraft," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 193–220, D. Reidel, Dordrecht, Holland, 1973.

Samir, U., and H. Jew, "Comparison of Theory with Experiment for Electron Density Distribution in the Near Wake of an Ionospheric Satellite," *J. Geophys. Res.*, **77:** 6819–6827, 1972.

Samir, U., and G.L. Wrenn, "The Dependence of Charge and Potential Distribution Around a Spacecraft on Ionic Composition," *Planet. Space Sci.*, **17:** 693, 1969.

Samir, U., L.H. Brace, and H.C. Brinton, "About the Influence of Electron Temperature and Relative Ionic Composition on Ion Depletion in the Wake of the AE-C Satellite," *Geophys. Res. Lett.*, **6:** 101–104, 1979a.

Samir, U., R. Gordon, L. Brau, and R. Theis, The Near-Wake Structure of the Atmosphere Explorer C. (AE-C) Satellite: A Parametric Investigation," *J. Geophys. Res.*, **84:** 513–525, 1979b.

Samir, U., Y.J. Kaufmann, L.H. Brace, and H.C. Brinton, "The Dependence of Ion Density in the Wake of the AE-C Satellite on the Ratio of Body Size to Debye Length in an (O^+)-Dominated Plasma," *J. Geophys. Res.*, **85:** 1769–1772, 1980.

Sanders, N.L. and G.T. Inouye, "Secondary Emission Effects on Spacecraft Charging: Energy Distribution Considerations," in *Spacecraft Charging Technology—1978*, edited by R.J. Finke and C.P. Pike, NASA CP-2071/AFGL TR-79-0082, 747–755, 1979.

Schnuelle, G. W. et al., D.E. Parks, I. Katz, M.J. Mandell, P.G. Steen, J.J. Cassidy, and A. Rubin, "Charging Analysis of the Scatha Satellite," in *Spacecraft Charging Technology—1978*, edited by R.C. Finke and C.P. Pike, NASA CP-2071/AFGL TR-0087, 123–143, 1979.

Schnuelle, G.W., P.R. Stannard, I. Katz, and M.J. Mandell, "Stimulation of the Charging Response of the SCATHA (P78-2) Satellite," in *Proceedings of the USAF/NASA Spacecraft Charging Technical Conference III*, 1981.

Schroder, H., "Spherically Symmetric Model of the Photoelectron Sheath for Moderately Large Plasma Debye Lengths," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 51–58, D. Reidel, Dordrecht, Holland, Mass., 1973.

Schwenn, R., H. Rosenbauer, and K.H. Muhlhauser, "The Solar Wind during STIP II Interval: Stream Structures, Boundaries, Shocks, and Other Features as Observed by the Plasma Instruments on Helios-1 and Helios-2," *Contributed Papers to the Study of Travelling Interplanetary Phenomena/1977*, edited by M.A. Shea, D.F. Smart and S.T. Wu, AFGL TR-77-0309, ADA056931, 1977.

Scudder, J D., E.C. Sittler, Jr., and H.S. Bridge, "A Survey of the Plasma Electron Environment of Jupiter: A View from Voyager," *J. Geophys. Res.*, **86:** 8157, 1981.

Shaw, R.R., J.E. Nanevicz, and R.C. Adamo, "Observations of Electrical Discharges Caused by Differential Satellite Charging," *Spacecraft Charging by Magnetospheric Plasmas*, AIAA (*Prog. Astronaut. Aeronaut.*, **47**) edited by A. Rosen, 61–76, 1976.

Shawhan, S.D., R.F. Hubbard, G. Joyce, and D.A. Gurnett, "Sheath Acceleration of Photoelectrons by Jupiter's

Satellite Io," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 504–414, D. Reidel, Dordrecht, Holland, 1973.

Singer, S.F. (ed.), *Interactions of Space Vehicles with an Ionized Atmosphere*, Pergamon Press, New York, 1965.

Soop, M., "Report on Photosheath Calculations for the Satellite GEOS," *Planet. Space Sci.*, **20:** 859, 1972.

Soop, M., "Numerical Calculations of the Perturbation of an Electric Field Around a Spacecraft," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, 127, D. Reidel, Dordrecht, Holland, 1973.

Spencer, N.W., L.H. Brace, G.R. Corignan, D.R. Taeusch, and H. Nieman, "Electron and Molecular Nitrogen Temperature and Density in the Thermosphere," *J. Geophys. Res.*, **70:** 2665–2698, 1965.

Spitzer, L., Jr., "The Dynamics of the Interstellar Medium," *Astrophys. J.*, **93:** 369–379, 1941.

Spitzer, L., Jr., "The Temperature of Interstellar Material," *Astrophys. J.*, **107:** 6–33, 1948.

Spitzer, L., Jr., and M.P. Savedoff, "The Temperature of Interstellar Matter, III," *Astrophys. J.*, **111:** 593–608, 1950.

Sternglass, E.J., "Backscattering of Kilovolt Electrons from Solids," *Phys. Rev.*, **95:** 345, 1954.

Sternglass, E.J., "Theory of Secondary Electron Emission by High Speed Ions," *Phys. Rev.*, **108:** 1–12, 1957.

Stevens, N.J., "Space Environment at Interactions with Biased Spacecraft Surfaces," in *Space Systems and Their Interactions With the Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, **71**) edited by H.B. Garrett and C.P. Pike 455–476, AIAA Press, New York, N.Y., 1980.

Stevens, N.J., and C.K. Purvis, "NASCAP Modeling Computations on Large Optics Spacecraft in Geosynchronous Substorm Environments," *NASA TM-81395*, 1980.

Stevens, N.J., J.V. Staskus, J.C. Roche, and P.F. Mizera, "Initial Comparison of SSPM Ground Test Results and Flight Data to NASCAP Simulations," *NASA TM-81394*, 1980a.

Stevens, N.J., J.V. Staskus, and J.C. Roche, "Initial Comparison of SSPM Ground Test Results and Flight Data to NASCAP Simulations," *NASA TM-81314*, 1980c.

Suh, P.K., and M.C. Stauber, "Photoelectron Drift and Multiplication due to Surface Potential Gradient: Application to Solar Power Arrays," Paper 80-0044, AIAA 18th Aerospace Sciences Conference, AIAA, New York, 1980.

Taylor, H.A., Jr., H.C. Brinton, and C.R. Smith, "Positive Ion Composition in the Magnetionosphere Obtained from the OGO-A Satellite," *J. Geophys. Res.*, **70**: 5769–5781, 1965.

Treadway, M.J., C.E. Mallon, T.M. Flanagan, R. Denson, and E.P. Wenaas, "The Effects of High-Energy Electrons on the Charging of Spacecraft Dielectrics," *IEEE Trans. Nuc. Sci.*, **NS-26**, (6), 5102–5106, 1979.

Tsien, H.S., "Superaerodynamics, Mechanics of Rarefied Gases," *J. Aero. Sci., 13*, 653–664, 1946.

Tsipouras, P., and H.B. Garrett, "Spacecraft Charging Model 2 Maxwellian Approximation," *AFGL TR-79-0153*, ADA077907, 1979.

Walker, E.H., "Plasma Sheath and Screening Around a Stationary Charged Sphere and a Rapidly Moving Charged Body," in *Interactions of Space Vehicles with an Ionized Atmosphere*, edited by S.F. Singer, 61–162, Pergamon Press, New York, 1965.

Walker, E.H., "Plasma Sheath and Screening of Charged Bodies," in *Photon and Particle Interactions with Surfaces in Space*, edited by R.J.L. Grard, D. Reidel, Dordrecht, Holland, 73–89, 1973.

Whipple, E.C., Jr., The Ion-Trap Results in the "Exploration of the Upper Atmosphere with the Help of the Third Soviet Sputnik," *IRE Trans.*, **47:** 2023–2024, 1959.

Whipple, E.C., Jr., "The Equilibrium Electric Potential of a Body in the Upper Atmosphere," *NASA X-615-65-296*, 1965.

Whipple, E.C., Jr., "Theory of the Spherically Symmetric Photoelectron Sheath: A Thick Sheath Approximation and Comparison with the ATS-6 Observations of a Potential Barrier," *J. Geophys. Res.*, **81:** 601–607, 1976a.

Whipple, E.C., Jr., "Observation of Photoelectrons and Secondary Electrons Reflected from a Potential Barrier in the Vicinity of ATS-6," *J. Geophys. Res.*, **81:** 715, 1976b.

Whipple, E.C., Jr., "Modeling of Spacecraft Charging," in *Proceedings of Spacecraft Technical Conference*, edited by C.P. Pike and R.R. Lovell, AFGL TR-77-0051/NASA TMX-73537, ADA045459, 225–236, 1977.

Whipple, E.C., Jr., and L.W. Parker, "Theory of an Electron Trap on a Charged Spacecraft," *J. Geophys. Res.*, **74:** 2962–2971, 1969a.

Whipple, E.C., Jr., and L.W. Parker, "Effects of Secondary Electron Emission on Electron Trap Measurements in the Magnetosphere and Solar Wind," *J. Geophys. Res.*, **74:** 5763, 1969b.

Whipple, E.C., Jr., J.M. Warnock, and R.H. Winkler, "Effect of Satellite Potential on Direct Ion Density Measurement through the Plasmapause," *J. Geophys. Res.*, **79:** 179–186, 1974.

Wildman, P.J.L., "Dynamics of a Rigid Body in the Space Plasma," in *Space Systems and Their Interactions With Earth's Space Environment*, (*Prog. Astronaut. Aeronaut.*, **71**) edited by H.B. Garrett and C.P. Pike, 633–661, AIAA Press, New York, 1980.

Williamson, P.R., W. Denig, P.M. Banks, and W.J. Raitt, "Vehicle Potential Measurements Using a Mother-Daughter Tethered Rocket (Abstract)," *Eos Trans. AGU*, **61:** 1068, 1981.

Willis, R.F., and D.K. Skinner, "Secondary Electron Emission Yield Behaviors of Polymers," *Solid State Commun.*, **13:** 685, 1973.

Winckler, J.R., "A Summary of Recent Results Under the 'Echo' Program for the Study of the Magnetosphere by Artificial Electron Beams," *Cosmic Phys. Tech. Report 168*, School of Physics and Astronomy, University of Minnesota, Minneapolis, September 1976.

Winckler, J.R., "The Application of Artificial Electron Beams to Magnetospheric Research," *Rev. Geophys. Space Phys.*, **18:** 659–682, 1980.

Zonov, Yu. V., "On the Problem of the Ineraction Between a Satellite and the Earth's Atmosphere and the Earth's Magnetic Field," *Iskusstvenyye Sputniki Zemli*, **3**, 1959, (NASA Tech. Transl. TTF-37).

Chapter 8

MAGNETOSPHERIC AND HIGH LATITUDE IONOSPHERIC ELECTRODYNAMICS

W.J. Burke, D.A. Hardy, and R.P. Vancour

This chapter deals with the volume of space that is bounded externally by the magnetopause and internally by the plasmapause and the high latitude ionosphere at an altitude of 300 km. The magnetopause separates regions of space dominated by the earth's magnetic field (magnetosphere) and by the shocked solar wind (magnetosheath). Earthward of the plasmapause, dynamics are generally controlled by corotation rather than by solar wind driven convection. The arbitrarily chosen, low altitude boundary in the ionosphere represents a transition below which the effects of the earth's neutral atmosphere are dominant. From the viewpoints of both cause and effect, the chapter is something less than self contained. Without the geomagnetic field (Chapter 4) and the solar wind (Chapter 3), there would be no magnetosphere and no magnetospheric electrodynamics; without solar irradiance (Chapter 2), there would be much less of an ionosphere. Without magnetospheric electrodynamics, there would be no aurora (Chapter 12), no high-latitude currents (Chapter 4), no radiation belts (Chapter 5), and no problems with spacecraft charging (Chapter 7).

The term "electrodynamics" encompasses a complex of processes by which charged particles move about in the magnetosphere-ionosphere system. The nature of the processes varies from region to region within the system. Magnetic merging at the magnetopause and field-aligned potential drops above the auroral ionosphere are examples of localized, electrodynamic processes. They are unified as electrodynamic processes in that they emerge, with appropriate boundary conditions, as solutions of the Vlasov-Maxwell equations. General solutions of the Vlasov-Maxwell equations over the entire magnetosphere-ionosphere system are well beyond present capabilities. Some success, however, has been achieved by considering elements of the system in relative isolation. This provides insight into how system elements evolve in response to external inputs. Since the entire system is electrically coupled, the isolated element approach is self-limiting. As one element evolves it affects processes in other elements of the system. The main goals of this chapter are to describe the various system elements and indicate, in a qualitative sense, how they are electrically coupled.

In dealing with the earth's magnetosphere three things quickly impress the mind. First, there is almost nothing there. Particle densities in the plasma sheet range up to about $1/cm^3$. With present technology, laboratory vacuum systems are able to get down to densities of $10^{10}/cm^3$. Second, the volume of space occupied by the magnetosphere is considerable. Typical magnetospheric dimensions are of the order of 10 R_E (1 R_E = 6.4 × 10^3 km). Third, when concentrated to global scales the effects of magnetospheric processes are impressive. The third point which is illustrated in Figure 8-1 provides a convenient point of departure for this survey of magnetospheric and high-latitude ionospheric processes. The figure exemplifies the spatial distribution of visible radiation observed by means of an optical imaging system on a Defense Meteorological Satellite Program (DMSP) satellite. City lights provide an easily recognized map of the western half of North America. The total energy emitted by auroral forms over the northern tier of Canadian provinces rivals or exceeds the combined ground emissions from the United States and Canada. Auroral emissions are largely due to plasma sheet electrons with energies of a few keV impacting the E layer of the ionosphere. The instantaneous locus of plasma sheet electron precipitation is called the auroral oval. Global imagery from satellites such as DMSP have shown that the auroral oval may be approximated by circular bands surrounding the geomagnetic poles. The centers of the circles are offset by ~3° to the night sides of the magnetic poles. The radii of the circle, the widths of the bands and the intensity of emissions vary with the level of geomagnetic activity. However, at all times the auroral oval exists and acts as a major sink for magnetospheric particles and energy. The particles and energy lost by the magnetosphere due to auroral precipitation ultimately come from the solar wind. Thus, an estimate of global precipitation loss is also an estimate of the efficiency of solar wind/magnetospheric interactions required to maintain the auroras

During periods of moderate geomagnetic activity the auroral oval can be approximated as a circular band extending from 75° to 65° magnetic latitude. The area of such a band is $10^{11}/cm^2$. The mean flux of electrons into the

CHAPTER 8

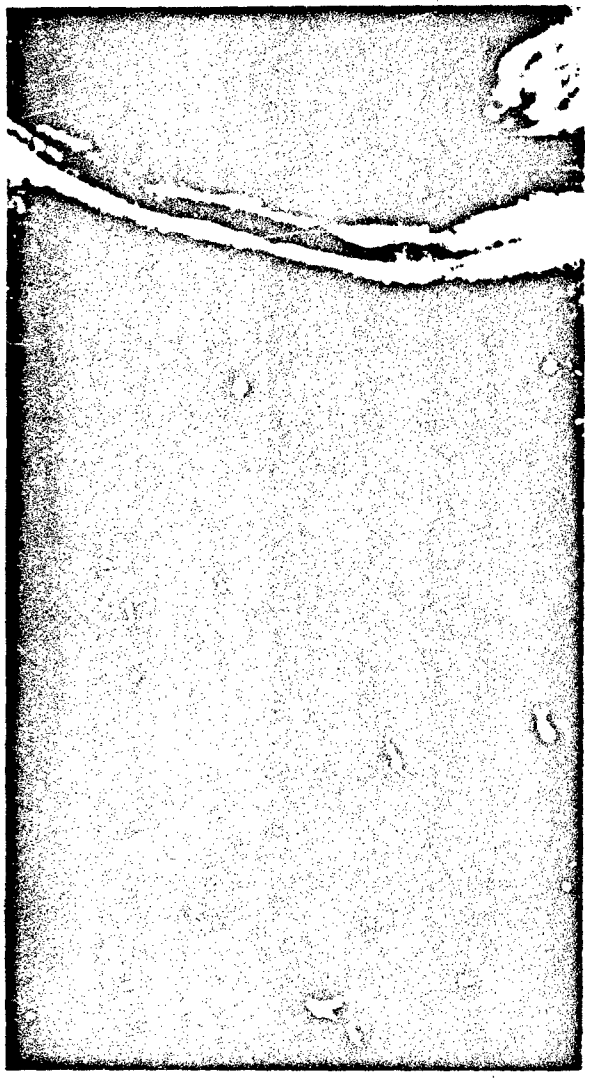


Figure 8-1 DMPS imagery from over western North America

auroral oval is $\sim 10^9$ cm^{-2}s^{-1}. Thus under steady state conditions the solar wind must supply electrons to the magnetosphere at a rate of 10^{26}/s. The average energy of precipitating electrons is of the order of 1 keV. The energy loss due to electron precipitation alone is $\sim 10^{10}$ W. Similar or larger amounts of solar wind energy must be supplied to account for ionospheric Joule heating and for maintaining the ring current. The central focus of this chapter is to outline the present understanding of how 10^{26} electrons per second and tens of billions of watts are extracted from the solar wind to drive magnetospheric and ionospheric electrodynamic processes.

8.1 MAGNETOSPHERIC BOUNDARY INTERACTIONS

In describing interactions between the solar wind and the earth's magnetosphere two coordinate systems are usefully employed: geocentric solar-ecliptic (GSE) and solar magnetospheric (GSM) coordinates. Both coordinate systems, which are described in detail in Chapter 4, have their origins at the center of the earth with the X axes positive toward the center of the sun; that is, $X_{se} = X_{sm}$. The Z_{se} axis is normal to the ecliptic plane and positive toward the north. The Y_{se} that completes the right hand system is positive toward local dusk. The Z_{sm} axis is coplanar with the earth's magnetic moment vector (M) and the X_{sm} axis. It is positive toward ecliptic north. The Y_{sm} axis, which always lies in the SM equatorial plane, completes the right hand coordinate system. For a radially flowing solar wind ($\mathbf{V}_s = -V_s\hat{X}_{se} = -V_s\hat{X}_{sm}$, where $\hat{X}_{se}$ and $\hat{X}_{sm}$ are unit vectors along X_{se} and X_{sm}, respectively) the angle between M and Z_{sm} gives the magnetic latitude of the magnetospheric subsolar point. Note that due to 11° offset between M and the earths rotational axis and to the 23.5° angle between the equatorial and ecliptic planes the magnetic latitude of the subsolar point is subject to ±34.5° combined seasonal and diurnal variations. The GSM is superior to the GSE system for ordering data relevant to interactions between the solar wind and the magnetosphere.

8.1.1 The Magnetopause

The shape of the "steady state" magnetosphere is determined from the force balance equation

$$\nabla \cdot [\underline{\underline{P}} + \underline{\underline{T}}] = 0 \qquad (8.1)$$

where $\underline{\underline{P}}$ and $\underline{\underline{T}}$ are the total pressure and the Maxwell stress tensors, respectively. The total pressure tensor is made up of two parts due to the dynamic and thermal pressures of the solar wind components:

$$\underline{\underline{P}} = 2n_s m_p V_s V_s + p_{si} + p_{se} \qquad (8.1)$$

where n_s is the solar wind density, m_p the mass of a proton; p_{si} and p_{se} are the thermal pressures of solar wind ions and electrons, respectively. The factor of 2 accounts for specular reflection of incoming particles. The shape of the magnetopause on the dayside can be calculated by numerical means using a simplified force balance

$$2n_s m_p V_s^2 (-\hat{X}_{se} \cdot \hat{n}_M)^2 = \frac{B_T^2}{2\mu_o}, \qquad (8.3)$$

where μ_o is the permitivity of free space, $\hat{n}_M$ is an outward directed unit vector normal to the magnetopause, and B_T

the total magnetic field at the magnetopause. $\mathbf{B}_T$ is a superposition of fields due to the earth's dipole $\mathbf{B}_D$, to the currents flowing on the magnetopause $\mathbf{B}_M$, and to other currents distributed in the magnetosphere. Beyond the magnetopause $\mathbf{B}_M$ exactly cancels the internal fields. To a very good approximation at the subsolar point of the magnetopause

$$|\mathbf{B}_T| = 2|\hat{n}_M \times \mathbf{B}_D|. \tag{8.4}$$

In the magnetic equatorial plane

$$\mathbf{B}_D = \frac{B_o}{L^3}, \tag{8.5}$$

where $B_o = 3 \times 10^4$ nT is the strength of the earth's field at the surface on the magnetic equator and L is the distance from the center of the earth in earth radii (R_E). Substitution of Equations (8.4) and (8.5) into (8.3) gives the distance to the magnetopause near the subsolar point

$$L_M = \left(\frac{B_o^2}{\mu_o n_s m_p V_s^2}\right)^{1/6}. \tag{8.6}$$

For a solar wind density and velocity of 5/cm^3 and 400 km/s, $L_M = 9$. The shape of the dayside magnetopause was calculated by Mead and Beard [1964] and by Olson [1969] using iterative numerical techniques in which the tilt of the dipole was ignored and included, respectively. Figure 8-2 shows a meridional cross section of the magnetosphere calculated with $L_M = 10$ in the Mead and Beard model. The locus of dipole field lines (dashed lines) in comparison with the calculated total field strikingly illustrates the effects of the solar wind on the overall magnetic topology. Magnetic field lines on the dayside are compressed while those on the nightside are elongated. Note that in this model, field lines intersecting the earth at magnetic latitudes greater than 83° are swept back to the nightside by the solar wind. There are a pair of singular points on the magnetopause at separatrices between field lines closing on the day and night sides. These points correspond to the dayside cusps.

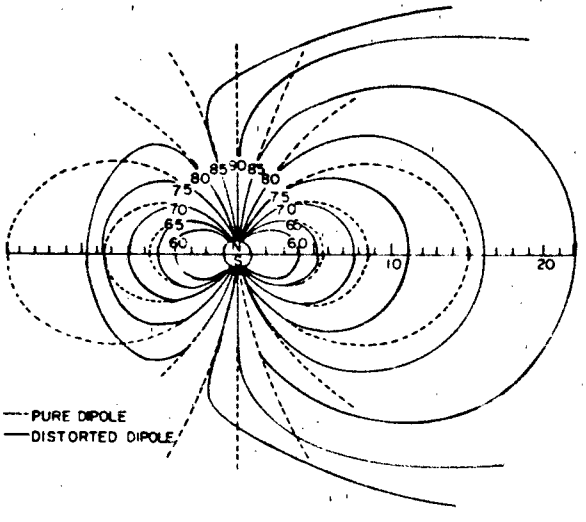


Figure 8-2. Model of earth's magnetic field distorted by the solar wind [Mead and Beard, 1964]

The models just discussed do not describe the nightside of the magnetotail adequately. One reason is apparent from a consideration of Equation (8.2). On the dayside of the magnetosphere, the dynamic pressure of the solar wind dominates over the thermal pressures. On the nightside, with plasma flow almost tangential to the magnetopause ($\mathbf{V}_s \cdot \hat{n}_M \approx 0$), the converse is true. An early model [Johnson, 1950] of the magnetosphere had a tear drop shape with the closing distance determined by the solar wind Mach number. Piddington [1963] suggested that, in flowing past the magnetosphere, the solar wind exerts tangential stresses at the boundary. Such stresses draw the nightside of the magnetosphere into an elongated magnetotail. In the absence of significant plasma within the magnetotail, the tangential force exerted by the solar wind on the magnetosphere is

$$F_T = \frac{B_{MT}^2}{2\mu_o} \pi R_{MT}^2 \tag{8.7}$$

where B_{MT} and R_{MT} are the field strength and radius of the magnetotail, respectively. There are, however, distributed currents in the inner magnetosphere (the ring current) and in the magnetotail (the neutral sheet) currents, whose effects must be included in realistic stress calculations.

8.1.2 Convection

In many cases, the magnetosphere-solar wind interaction is well described by steady state equations such as Equation (8.1). The equilibrium represented by these equations is, however, dynamic rather than static. Only a dynamic situation is consistent with existing high latitude current systems. These currents result from ionospheric convection which is driven by magnetospheric convection [Gold, 1959]. Magnetospheric convection is in turn driven by the solar wind. That is, energy is extracted from the solar wind by the magnetosphere, and at least some of that energy is dissipated in the ionosphere. Two mechanisms for transferring energy to the magnetosphere have been developed over the last two decades: viscous interactions [Axford and Hines, 1961] and magnetic merging [Dungey, 1961]. Both models explain many qualitative features of magnetospheric convection and auroral particle energization. Recent satellite observations suggest that both mechanisms are operative,

but in more complex ways than envisaged by early proponents.

The Axford-Hines model postulates that the magnetosheath plasma exerts a viscous force on a layer of unspecified thickness inside the magnetopause. Magnetic field lines threading this layer are dragged in the antisolar direction and are stretched to great distances in the magnetotail. As elongated flux tubes move out of the viscous interaction layer they snap back to a more dipolar configuration. In the rest frame of the earth this motion of magnetic field lines appears as an electric field, $\mathbf{E} = -\mathbf{V} \times \mathbf{B}$. A magnetospheric equatorial projection of the convection pattern generated in the viscous interaction model is given in Figure 8-3. When mapped to ionospheric altitudes, assuming that

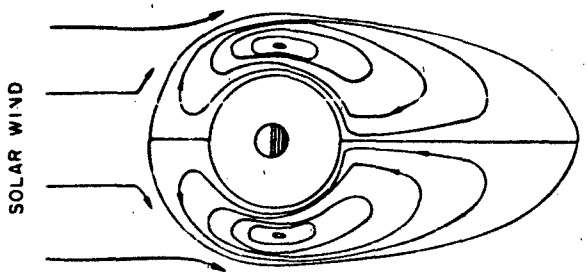


Figure 8-3. Equatorial projection of convection pattern in viscous interaction model.

$\mathbf{E} \cdot \mathbf{B} = 0$, the model reproduces the general features of the polar/auroral current system (Chapter 4). Note that plasma trapped on elongated flux tubes is adiabatically heated as the flux tubes convect earthward and shrink in volume

The second model postulates that the dynamic interaction between the solar wind and the magnetosphere proceeds by means of a magnetic merging process. The simplest features of this phenomenon are illustrated in Figure 8-4.

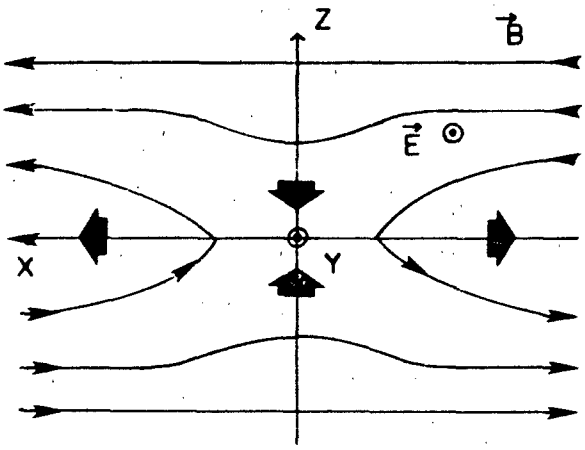


Figure 8-4. Magnetic field geometry and electric field required for magnetic merging.

Consider a magnetic field that at great distances above (below) the X-Y plane points in the + (−) X direction. In the presence of an electric field E_Y, magnetic field lines convect toward the X-Y plane. At the neutral line (X = 0, Z = 0), magnetic field lines from the upper half space merge with field lines of the opposite polarity from the lower half space. To the left (right) of the neutral line, merged magnetic field lines cross the X-Y plane with a + (−) Z component, and $\mathbf{E} \times \mathbf{B}$ convect away from the neutral line in the + (−) direction. Two necessary conditions for magnetic merging are magnetic field of opposite polarity across some plane and an electric field component that is tangent to the plane.

Before considering how magnetic merging might apply to the magnetosphere, it is useful to distinguish between several possible magnetic topologies. It is well known that a weak interplanetary magnetic field (IMF) is carried by the solar wind. Except for a small correction term in the force balance equation, the IMF plays no obvious role in the viscous interaction model. The magnetic merging model assigns important roles to the IMF because this model requires three types of magnetic field lines: (1) IMF lines with both "feet" in the interplanetary medium, (2) closed field lines with both "feet" in the earth, and (3) open field lines with one "foot" on earth and the other in the solar wind. Dungey [1961] pointed out that when the IMF has a southward component, magnetic merging can occur near the subsolar point of the magnetopause. The idea is illustrated in Figure 8-5 which can be viewed either as a snap shot or as

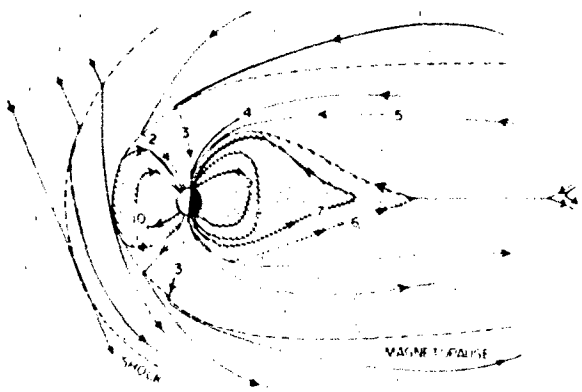


Figure 8-5. Snapshot of magnetic merging between southward IMF and the earth's magnetosphere

a time history of an individual field line. As southward directed IMF lines are convected up against compressed dipolar field lines, merging occurs at time ①. Because one foot of a newly merged field line is imbedded in the solar wind, the whole field line is dragged in the antisolar ($-X_{SM}$) direction. In an earth-stationary frame of reference the motion of the ionospheric foot of the field line appears to result from a dawn to dusk electric field. Times ② through ⑤ show the various stages of antisunward motion of an open

field line. At time ⑥, a portion of the field line has convected to the magnetic equatorial plane where it reconnects with an open field line from the conjugate ionosphere. Under the influence of the dawn to dusk electric field, the field line then convects earthward (times ⑦ through ⑨). Eventually, reconnected field lines move to the dayside (time ⑩) where they are in position to continue the merging-reconnection cycle.

The magnetospheric convection patterns predicted by the viscous interaction and magnetic merging models with a southward IMF are quite similar. Empirical evidence from near the magnetopause indicates that both viscous and magnetic merging processes occur [Eastman et al., 1976; Russell and Elphic, 1979; Mozer et al., 1979]. That many observational studies have shown high correlations between southward turnings of the IMF and the onset of magnetic activity indicates a dominant role for energy transfer by a merging process. As discussed in Section 8.5.1, this surmise is supported by modifications observed in polar cap (open field line) convection patterns with variations in the Y_{SM} component of the IMF and when the Z_{SM} component is northward. Finally, we mention that currently merging is thought to occur sporadically rather than as a steady state process [Haerendel et al., 1978], and in the vicinity of the dayside cusps rather than near the subsolar point [Crooker, 1977; Crooker, 1979].

8.1.3 Boundary Layers

Before considering the dynamics of the closed portion of the magnetosphere, let us return briefly to our original question of how 10^{26} particles/s gain entry to the plasma sheet. Between 10^{28} and 10^{29} solar wind particles/s impact the dayside magnetopause. Thus, an entry efficiency of less than 1% is sufficient to maintain the plasma sheet. Our understanding of how magnetosheath plasma gains entry to the magnetosphere and influences its interior dynamics has been evolving rapidly over the last decade. Under such circumstances, it is not unusual to encounter a multiplicity of nomenclatures that will probably be simplified as relationships between various boundary plasma regimes become more evident. Vasyliunas [1979] has defined magnetospheric boundary layers as regions of space threaded by magnetic field lines of the magnetosphere but populated by plasma similar to that found in the magnetosheath. The four regions satisfying this definition are (1) the plasma mantle, (2) the interior cusp, (3) the low latitude boundary layer (LLBL), and (4) the plasma sheet boundary layer. Figure 8-6a is a schematic representation of the magnetospheric loci of these regimes. Somewhat speculative representations of their ionospheric projections and their "source-relationships" to the plasma sheet are given in Figures 8-6b and 8-7, respectively. The plasma mantle is found on open magnetic field lines; the remaining three regions are found in closed field line portions of the magnetosphere.

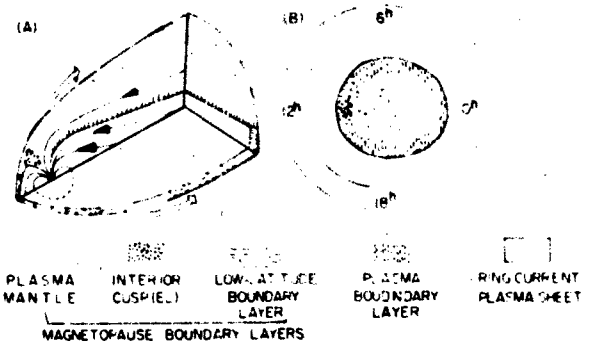


Figure 8-6. (a) Schematic diagram of various observed magnetospheric boundary layers; (b) their mapping down to the ionosphere along magnetic field lines [Vasyliunas, 1979].

The plasma mantle was first identified as a magnetosheath-like plasma flowing along magnetic field lines inside the magnetopause in the near-earth lobes of the magnetotail [Rosenbauer et al., 1975]. The plasma density and the spatial thickness of the mantle are greatest during periods of southward IMF [Sckopke et al., 1976]. The mantle is also observed in the lobes of the magnetotail at lunar distance ($X_{SM} = -60\ R_E$) near the ecliptic plane [Hardy et al., 1975]. Mantle particles are believed to enter the magnetosphere near the dayside cusp. Dawn to dusk electric fields cause particles to convect in the antisunward direction so that particles that mirror at low altitudes find themselves on open field lines as they rise from their mirror points. The same dawn to dusk electric field causes the mantle particles to drift toward the equatorial plane as they move away from the earth. Pilipp and Morfill [1978] suggested that mantle particles may be one source of plasma sheet particles.

As the name suggests, the "interior cusp" refers to the region of closed magnetic field lines passing through or just equatorward of the cusp. Magnetosheath plasma diffuses into this region. As opposed to the mantle or the LLBL, the plasma of this region, which is also called the entry layer, is relatively stagnant. However, depending on the strength and direction of convective electric fields in the cusp, it is possible to think of the interior cusp plasma as a partial source of both the mantle and the LLBL.

Along the dawn and dusk meridians near the magne-

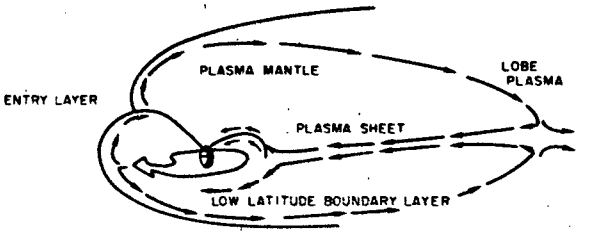


Figure 8-7. A model of magnetospheric circulation for filling the plasma sheet [Freeman, 1979].

tospheric equatorial plane, the LLBL is characterized by magnetosheath-like plasma flowing in the antisunward direction. The thickness of this layer ranges up to ~1 R_E. As shown in Figure 8-6a, the LLBL has been observed to great distances in the antisolar direction. The density of plasma within the layer is about a factor of four less than that of the adjacent magnetosheath [Sckopke et al., 1981]. Electrons within the layer have trapped pitch angle distributions [Eastman and Hones Jr., 1979]. Whether the LLBL is maintained through a diffusive and/or an impulsive [Lemaire and Roth, 1978] entry process is currently a matter of debate. Sckopke et al. [1981] estimate that a diffusion coefficient of $10^9 m^2/s$ is required to maintain the observed LLBL. Figure 8-6a shows the plasma boundary layer together with the LLBL as forming a continuous envelope surrounding the hot plasma contained in the central plasma sheet. The physical processes which connect the LLBL and/or the mantle with the plasma boundary layer and with the central plasma sheet are not known at this time. It is currently believed that discrete arcs in the auroral oval map to the boundary rather than to the central plasma sheet. Within the plasma sheet boundary, rapidly flowing plasmas are observed. These flowing plasmas come from spatially limited acceleration regions called "magnetospheric fireballs" [Frank et al., 1976]. Whether the energization process in fireballs results from magnetic reconnection or some other process is still another open question.

8.2 THE PLASMA SHEET

The earth's plasma sheet is the highly dynamical region of the earth's magnetosphere that acts as a depository for auroral particles. It is a region of closed magnetic field lines. Before being first detected by the Soviet satellites Luna 1 and Luna 2 [Gringanz et al., 1961], somewhat strangely, this important region of the magnetosphere was not anticipated theoretically. Equatorial and noon-midnight meridional projections of the plasma sheet are given in Figures 8-8 and 8-9, respectively [Vasyliunas, 1972]. Both projections show that the plasma sheet extends for great distances in the $-X_{SM}$ direction. The plasma sheet has a distinct inner edge that varies as a function of local time and the level of geomagnetic activity [Vasyliunas, 1968; Frank, 1971]. The dynamics of the inner edge of the plasma sheet are well understood theoretically and are discussed in Section 8.6.2. The equatorial thickness of this boundary is ~1 R_E and is marked by a cooling of electron temperatures. Just tailward of this boundary plasma sheet electrons have an average energy of ~1 keV. During periods of substorm injections, the temperature of electrons may rise to ~10 keV. Temperatures of plasma sheet ions tend to be higher than those of electrons by a factor of 2 or more.

Figure 8-10 is a cross sectional view of the magnetotail portion of the plasma sheet. It has a minimum thickness in the mid-tail region and flares to a maximum thickness near

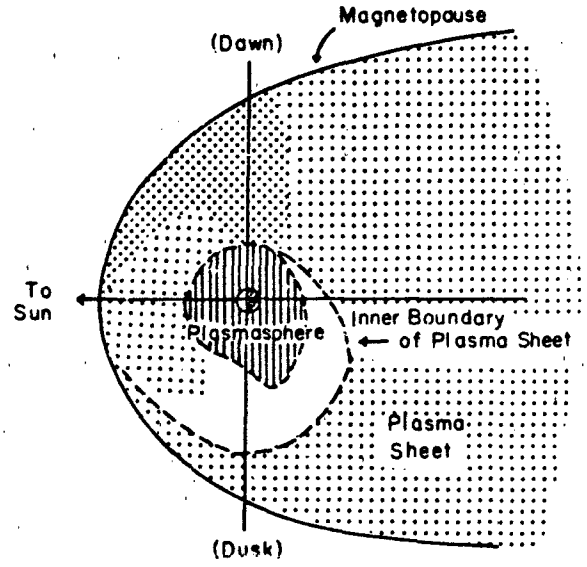


Figure 8-8. The distribution of plasma sheet electrons in the equatorial plane [Vasyliunas, 1972].

the dawn and dusk flanks of the tail. At lunar distance ($X_{SM} = -60\ R_E$) the tail radius is ~25 R_E. The average half-thickness of the mid-tail plasma sheet is ~3 R_E. During the expansion phase of substorms, the thickness of the plasma sheet in the tail decreases and then expands during the recovery phase [Hones et al., 1973]. At lunar distance the average density of the plasma sheet is ~0.1 cm^3. The electron and proton temperatures are ~0.25 and 2.5 keV, respectively [Rich et al., 1973].

The remainder of this subsection is concerned with the physical mechanisms responsible for particle energization and precipitation in the plasma sheet. Energization processes are classified as either adiabatic or non-adiabatic. Examples of non-adiabatic energization are neutral sheet acceleration [Speiser, 1967], stochastic wave-particle acceleration, and heating derived from magnetic field reconnection and annihilation. Although these are undoubtedly important sources of particle energy, we limit ourselves to describing adiabatic energization in some detail. Particle precipitation is main-

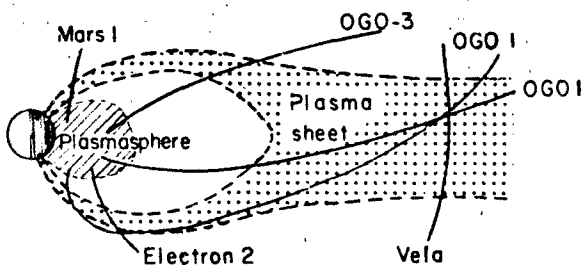


Figure 8-9. The distribution of nightside, plasma sheet electrons in meridional plane [Vasyliunas, 1968].

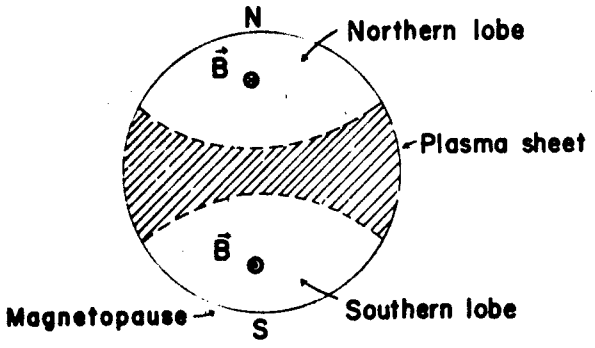


Figure 8-10. Cross-sectional view of magnetotail looking earthward.

tained or enhanced either by magnetic-field aligned electric fields ($E_{\parallel}$) or by pitch angle diffusion. $E_{\parallel}$ is very important for discrete auroral arc formation. Pitch angle diffusion results from wave particle interactions. Here we summarize briefly the collective plasma modes responsible for these phenomena.

8.2.1 Adiabatic Motion in the Plasma Sheet

The adiabatic energization of plasma sheet particles involves the related concepts of adiabatic invariance and guiding center motion. The general concept of adiabatic invariance comes from classical Hamilton-Jacobi theory [TerHaar, 1964]. If a system executes a periodic motion in a force field (F) that changes slowly in time with respect to the period (T)

$$\frac{1}{F}\frac{dF}{dt} \ll \frac{1}{T}$$

then the quantity

$$I = \oint p\, dq,$$

where p and q are canonical momentum and coordinate variables, is a constant of the motion known as an adiabatic invariant. Charged particles moving in the earth's magnetic field may have as many as three periodicities due to their gyration, bounce, and drift motions. The three adiabatic invariants associated with these periodicities are

(1) the magnetic moment

$$\mu = \frac{mv_{\perp}^2}{2B}, \qquad (8.8)$$

where $v_{\perp}$ is the component of velocity perpendicular to the magnetic field,

(2) the longitudinal invariant

$$J = \oint p_{\parallel} ds, \qquad (8.9)$$

where $p_{\parallel}$ and ds are momentum component and distance along B, and

(3) the flux invariant

$$\Phi = \oint \mathbf{A} \cdot d\ell \qquad (8.10)$$

where A is the magnetic vector potential and $d\ell$ a distance element along a particle drift trajectory [Northrop, 1963; Rossi and Olbert, 1970]. Most plasma sheet particles either precipitate or follow drift trajectories that intersect the magnetopause before they can drift all the way around the earth. Thus, in the plasma sheet only the first and second invariants are of interest. The third invariant is important for understanding the ring current and radiation belts (Chapter 5).

At this point it is useful to introduce the related concepts of pitch angle and magnetic mirroring. The pitch angle of a particle (α) is defined as the angle between its instantaneous velocity and the magnetic field

$$\alpha = \cos^{-1}\left(\frac{\mathbf{v}\cdot\mathbf{B}}{|\mathbf{v}|\,|\mathbf{B}|}\right).$$

Magnetic mirroring results from the constancy of a particle's magnetic moment and total energy. The total energy of a non-relativistic particle of mass m and charge q moving with a velocity v in combined magnetic and electric fields is

$$\varepsilon = 1/2\, mv_{\parallel}^2 + \mu B + q\Psi \qquad (8.11)$$

where Ψ is the electrical potential. The component of force exerted along **B** is

$$m\frac{dv_{\parallel}}{dt} = -\mu\frac{\partial B}{\partial s} - q\frac{\partial \Psi}{\partial s}. \qquad (8.12)$$

The second term on the right hand side of Equation (8.12) is due to field-aligned electric field components that are discussed regarding auroral arc formation. The first term on the right hand side of Equation (8.12) is the magnetic mirror force. A particle at the magnetic equator ($s = 0$) with pitch angle α_{eq} can move earthward along B until its pitch angle reaches 90° ($v_{\parallel} = v_{\perp}\cos\alpha = 0$). At this point, it is reflected by the mirror force toward the magnetic equator. The strength of the magnetic field at the mirror point is designated B_M. The total kinetic energy of a particle at its mirror point is $1/2mv^2 = \mu B_M$. Since in general $\mu = 1/2mv^2 (\sin^2\alpha)/B$, the magnetic mirroring condition, in the absence of $E_{\parallel}$, is often written

$$\sin^2\alpha = \frac{B}{B_M}. \qquad (8.13)$$

The line integration for the longitudinal invariant proceeds from the magnetic equator to the mirror distance (s_M).

CHAPTER 8

The notion of guiding center motion is more general than that of adiabatic motion. Under many circumstances charged particle motions are well-approximated superpositions of motions of guiding centers and gyrational motion about the guiding center. For example, in an uniform magnetic field the equation of motion

$$m \frac{dv}{dt} = q (v \times B) \qquad (8.14)$$

describes a particle that gyrates with a circular frequency $\Omega = qB/m$ about a field line and moves along the field line (its guiding center) with a constant velocity. In the presence of an external force field **F** that is perpendicular to **B**, the equation of motion

$$m \frac{dv}{dt} = q (v \times B) + F \qquad (8.15)$$

can be reduced to the form of Equation (8.14) by transforming to a coordinate system moving with a drift velocity

$$V_D = \frac{F \times B}{qB^2}.$$

In this frame of reference the motion of the particle about **B** is purely gyrational.

The most important drifts within the magnetosphere are caused by electric fields

$$V_E = \frac{E \times B}{B^2}, \qquad (8.16)$$

magnetic field gradients

$$V_G = \frac{\mu B \times \nabla B}{qB^2}, \qquad (8.17)$$

and magnetic field line curvature

$$V_C = \frac{mv_\parallel^2}{R^2} \frac{R \times B}{qB^2} = mv_\parallel^2 \frac{B \times (B \cdot \nabla) B}{qB^4} \qquad (8.18)$$

where **R** is the magnetic field line radius-of-curvature vector. Note that V_G and V_C depend on both the particle's energy and charge; V_E depends on neither. Protons (electrons) gradient and curvature drift toward the west (east). In the nearly dipolar part of the inner plasma sheet, V_G and V_C are of comparable magnitudes. Due to sharp magnetic field line curvature across the neutral sheet, $V_C > V_G$ in the magnetotail.

With the expressions for the adiabatic invariants [Equations (8.8) and (8.9)] and guiding center drifts [Equations (8.16), (8.17), (8.18)] we may understand (1) the adiabatic heating of magnetospheric plasma, (2) the existence of a ring current, and (3) the position of the inner boundary of the plasma sheet.

Both the viscous interaction and the magnetic merging models require that particles in the nightside plasma sheet be "adiabatically" heated as they convect earthward under the influence of a dawn to dusk electric field. In this motion, each particle's magnetic moment $1/2mv_\perp^2/B$ is a constant. Consider a particle mirroring in the equatorial plane, $J = 0$. As it drifts from, say, 30 R_E in the magnetotail where $B \sim 20$ nT to a distance of 6 R_E where $B \sim 140$ nT, its kinetic energy must increase seven fold. Particles with J 0 gain kinetic energy, not only from the conservation of μ, but also from the apparent motion of magnetic mirror points (conservation of J). In moving from equatorial crossings of 30 R_E to 6 R_E, particles find themselves on shorter and shorter field lines. Since $\oint p_\parallel ds$ must be a constant, $p_\parallel$ must increase as particles convect earthward.

From the conservation of energy, it is seen that an increase in particle kinetic energy must be due to a decrease in potential energy. The potential energy is electrical; that is,

$$\frac{d}{dt} (1/2\ mv^2) = q \langle V_D \cdot E \rangle . \qquad (8.19)$$

The brackets in equation (8.19) are used to represent time averaging over a gyroperiod, and V_D is the total drift velocity. Since V_E is perpendicular to **E**, only V_G and V_C contribute. Hines [1963] has shown that the energy gained by "adiabatically compressing" a magnetospheric plasma is equivalent to the kinetic energy gained by gradient and curvature drifting in the direction of an electrostatic potential gradient.

Guiding center motion is the simplest basis for understanding the earth's ring current. The existence of a westward current encircling the earth can be inferred directly from decreases in surface values of the horizontal component of the earth's field during the main and recovery phases of magnetic storms (Chapter 4). The general expression for current density is given by a sum over plasma species

$$j = \sum_r n_r q_r V_{Dr}$$

Assuming, for simplicity, that the ring current is made up of a single ion species (H^+) then the current density is

$$j_R = nq (V_{D_+} - V_{D_-}) \qquad (8.20)$$

Since V_E is independent of charge, only V_G and V_C contribute to Equation (8.20). Recall that for protons (electrons), V_G and V_C are westwards (eastwards). In an equiv-

alent fluid description, j_R is driven by magnetospheric pressure gradients. During the main phases of magnetic storms and the expansion phases of substorms, particles are energized and injected into the inner magnetosphere by intense electric fields. With the onset of recovery phase, the electric fields decrease in intensity and/or are shielded from the inner magnetosphere. Injected particles find themselves on closed, stably trapped orbits in which they gradient and curvature drift around the earth. During recovery, phase these particles are slowly removed from the ring current by precipitation or by charge exchange with low energy neutrals (Chapter 5).

The position and shape of the inner edge of the plasma sheet is determined by the drift motions of plasma sheet particles. $\mathbf{V_E}$ decomposes into drifts due to "convective" ($\mathbf{V_{E_C}}$) and "corotational" ($\mathbf{V_{E_\Omega}}$) electric fields. In the following discussion, we use the symbols E_C to represent the "convective" electric field imposed by the solar wind on the magnetosphere and E_Ω to represent the corotation electric field. The direction of ($\mathbf{V_{E_\Omega}}$) is eastward for all particles. For simplicity let us consider the drift motions of charged particles having pitch angles in the equatorial plane of 90°. The conservation of energy Equation (8.11) immediately tells us that cold ($\mu \simeq 0$) particles are constrained to $\mathbf{E} \times \mathbf{B}$ drift along equipotentials. Particles with non-zero μ drift along surfaces of constant ($\Psi + \mu B/q$).

In the magnetotail, particles predominately drift earthward under the influence of a dawn to dusk E-field. As they approach the earth, electrons acquire significant eastward drifts due to both $\mathbf{V_G}$ and $\mathbf{V_{E_\Omega}}$. Since both of these drifts are eastward, cold electrons with $\mathbf{V_G} = 0$ drift closer to the earth before their eastward drifts dominate over their earthward drifts. For this reason outbound satellites encounter cold before hot electrons at the plasma sheet's inner edge [Kivelson et al., 1979].

The boundary between cold electrons that drift along equipotentials from the tail and those that corotate on closed trajectories is called the zero-energy Alfvén layer or the inner boundary of the plasma sheet. Under steady convective electric field conditions, cold electrons that drift in from the magnetotail without precipitating eventually cross the dayside magnetopause. Thus, the inner boundary of the plasma sheet is the boundary between closed (corotation dominated) and open (convection dominated) equipotentials.

Before calculating the shape of the last closed equipotential, we note that for protons, $\mathbf{V_G}$ and $\mathbf{V_{E_\Omega}}$ are oppositely directed. This leads to more complex drift paths for protons than electrons. Protons with $\mu \neq 0$ can drift earthward of the zero-energy Alfvén boundary in the evening local time sector. The different drift paths of protons and electrons eventually lead to the build up of polarization electric fields near the inner edge of the plasma sheet. The main effects of the polarization field are to shield E_C from the inner magnetosphere and to distort its dawn to dusk orientation in the plasma sheet.

As a function of distance R and local time ϕ in the equatorial plane of the magnetosphere, the form of the electric potential is

$$\Psi(R, \phi) = \frac{-\Omega B_o R_E^3}{R} + C R^\gamma \sin \phi \qquad (8.21)$$

where $\Omega = 7.3 \times 10^{-5} s^{-1}$ is the angular spin velocity of the earth, C is a constant to be determined, and γ is a parameter that reflects the level of electrical shielding. The case $\gamma = 1$ corresponds to a uniform, dawn to dusk E_C that completely penetrates the inner magnetosphere. Best empirical values of γ are in the range of 2 to 3 [Ejiri et al., 1978; Gussenhoven et al., 1981]. The value of C is determined by noting that E_Ω is directed radially inward and E_C is mostly the Y_{SM} direction. In the dusk sector E_Ω and E_C are oppositely directed. Depending on the strength of E_C, there is a stagnation point along the 1800 LT ($\phi = 3\pi/2$) axis of symmetry where the two fields exactly cancel.

$$\left.\frac{\partial \Psi\left(R, \frac{3\pi}{2}\right)}{\partial R}\right|_{R = R_S} = 0. \qquad (8.22)$$

Substitution of Equation (8.21) into (8.22) gives

$$C = \frac{\Omega B_o R_E^3}{\gamma R_S^{(\gamma + 1)}}. \qquad (8.23)$$

Thus,

$$\Psi(L, \phi) = -\frac{\Omega B_o R_E^2}{L_S}\left[\left(\frac{L_S}{L}\right) - \gamma\left(\frac{L}{L_S}\right)^\gamma \sin \phi\right] \qquad (8.24)$$

where we have made the substitution $L = R/R_E$ and $L_S = R_S/R_E$. The common term $\Omega B_o R_E^2$ is ~90 kV. The potential of the stagnation point is

$$\Psi\left(L_S, \frac{3\pi}{2}\right) = -\frac{90}{L_S}[1 + 1/\gamma] \text{ kV}. \qquad (8.25)$$

Since along the $\phi = 3\pi/2$ line $E_C < (>) E_\Omega$ for $L < (>) L_S$. Equation (8.25) gives the potential of the zero-energy Alfvén boundary. By setting the term in brackets on the right hand side of Equation (8.24) equal to $[1 + 1/\gamma]$, we arrive at the equation for the zero-energy Alfvén boundary in the equatorial plane as a function of distance from the center of the earth (L_A) and local time

$$\left(\frac{L_A}{L_S}\right)^{\gamma + 1} \sin \phi + (\gamma + 1)\frac{L_A}{L_S} - \gamma = 0. \qquad (8.26)$$

Southwood and Kaye [1979] have shown that to an excellent approximation

$$L_A \approx L_S\left(1 + \frac{\sqrt{2S}}{\gamma}\right) \qquad (8.27)$$

where $S = |\cos(\phi - \pi/2)/2|$. The solution is exact when $\gamma = 1$. Figure 8-11 shows that the shape of the last equi-

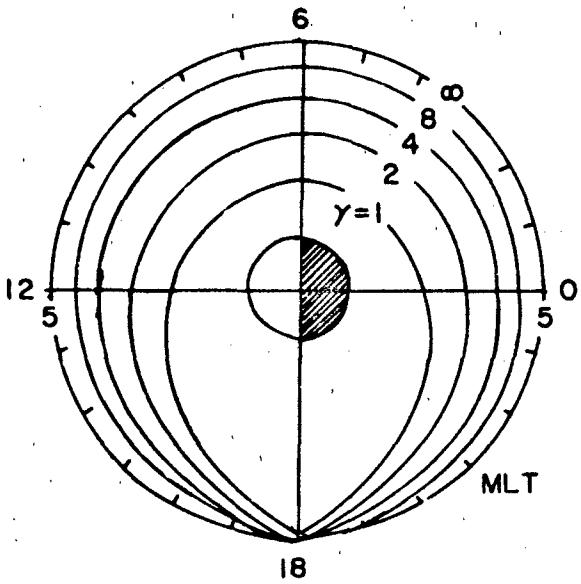


Figure 8-11. Shapes of last closed equipotential for various values of γ [Ejiri et al., 1978].

potential varies from an elongated tear drop for $\gamma = 1$ to a circle for $\gamma = \infty$. Note that Equations (8.26) or (8.27) only allow calculations of the shape, but not the distance to the Alfvén boundary. L_S depends on E_K, which varies with condition in the solar wind and with the level of magnetic activity. It is convenient to defer further comment on the Alfvén boundary until we have discussed its ionospheric projection, the equatorward boundary of diffuse auroral precipitation.

8.2.2 Pitch Angle Diffusion of Plasma Sheet Particles

The final topic to be considered under the heading of general magnetospheric processes is particle precipitation. To anticipate our discussion of the ionosphere, we note that instrumentation on satellites passing through the diffuse auroral ionosphere measure fluxes of electrons and protons that are isotropic over the downcoming hemisphere. The continuous precipitation of plasma sheet electrons and protons into the auroral ionosphere cannot be explained in terms of the individual particle model that we have been using. The problem is illustrated simply in Figure 8-12 where we sketch sequential isocontours of particle distribution functions in the magnetospheric equatorial plane. Figure 8-12a

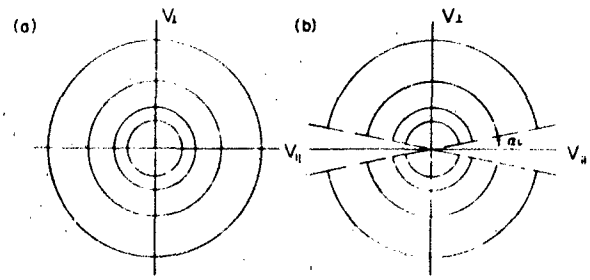


Figure 8-12. Isocontours of distribution functions with (a) isotropic and (b) loss cone distributions.

represents an isotropic population at some initial time. Particles with $\alpha_{eq} \neq 90°$ ($v_\parallel \neq 0$) move along field lines toward their mirror points. If their mirror points are sufficiently deep in the atmosphere, typically altitudes < 200 km, the particles are lost. Otherwise they return to the magnetosphere. If we designate the strength of the magnetic field at an altitude of 200 km as B(200), then Equation (8.13) shows that particles initially with equatorial pitch angles

$$\alpha_{eq} < \alpha_L = \sin^{-1}\left[\frac{B_{eq}}{B(200)}\right]^{1/2} \qquad (8.28)$$

are lost after a few bounce periods. Such particles are said to be in the atmospheric loss cone. For plasma sheet particles α_L is ~2°. Figure 8-12b shows the distribution after several bounce periods as made up of the initial population minus an empty loss cone. Particle motion that conserves μ and J allows no further loss.

Kennell and Petschek [1966] pointed out that above certain particle flux levels, loss cone distribution functions such as shown in Figure 8-12b are unstable to the growth of whistler waves. The waves grow in energy by causing particles to diffuse into the loss cone. To produce strong pitch angle diffusion, that is, maintain isotropy over the loss cone, a resonant condition must be fulfilled. Resonant scattering occurs for particles whose energy is equal to that of the magnetic energy per particle

$$\varepsilon_{Res} = \frac{B^2}{2\,\mu_o n}. \qquad (8.29)$$

This model successfully explains flux limits observed for stably trapped ring current particles. The requirement for resonant pitch angle diffusion given in Equation (8.29) was

empirically verified by observations of the proton ring current from Explorer 45 during the magnetic storm of December 1971 [Williams and Lyons, 1974].

An examination of the resonance condition given in Equation (8.29) shows that a whistler mode instability cannot be responsible for strong pitch angle scattering in the plasma sheet. At geostationary altitude (6.7 R_E) in the plasma sheet, B ~ 100 nT and n ~ 1 cm^3. This gives a magnetic field energy density in eV/cm^3 of 2.5 B^2(nT). A resonant energy of ~25 keV is well in excess of mean thermal energies for either electrons or protons in the plasma sheet.

Realizing that whistler mode interactions could not explain the isotropic precipitation of plasma sheet particles, investigators in the 1970s concentrated on sources of electrostatic wave energy. There are two important developments from the decade of which we take note. The first concerns direct observations of broadband electrostatic noise all along magnetic field lines connecting the auroral ionosphere to the equatorial plasma sheet [Gurnett and Frank, 1977]. The amplitudes of observed waves are of sufficient intensity to drive strong pitch angle scattering. The second development concerns the theoretical recognition of the role played by cold plasma for making available free energy contained in anisotropic pitch angle distributions. Cold plasma, of ionospheric origin, in the plasma sheet can produce velocity space gradients in the total distribution functions ($\partial f / \partial v_\perp > 0$) that are unstable to the growth of (N + 1/2) electron cyclotron waves [Young et al., 1973] and lower hybrid ion waves [Ashour-Abdalla and Thorne, 1978]. Numerical studies show that such waves are unstable over restricted ranges of parameter space. Measuring low density, cold plasma embedded in a hot plasma sheet is experimentally difficult. A successful measurement of the cold plasma component is a critical requirement for our understanding plasma sheet processes.

8.3 ELECTRICAL COUPLING OF THE MAGNETOSPHERE AND IONOSPHERE

Before discussing the electrodynamics of the high latitude ionosphere, it is useful to review briefly the theory of magnetosphere-ionosphere coupling. Vasyliunas [1970] has developed a theoretical model that illustrates the physical laws describing how magnetospheric convention couples with the ionosphere. The model is presented in Figure 8-13 in the form of a closed loop of equations (straight lines) and of quantities to be determined (boxes). External sources of particles, cross-magnetospheric potentials, neutral winds in the ionosphere, and so forth, are imposed boundary conditions. [illegible] loop of equations is self-consistently closed, it can be [illegible] at any point. Let us assume that we have an initial idea about the distribution of magnetospheric electric fields and particles

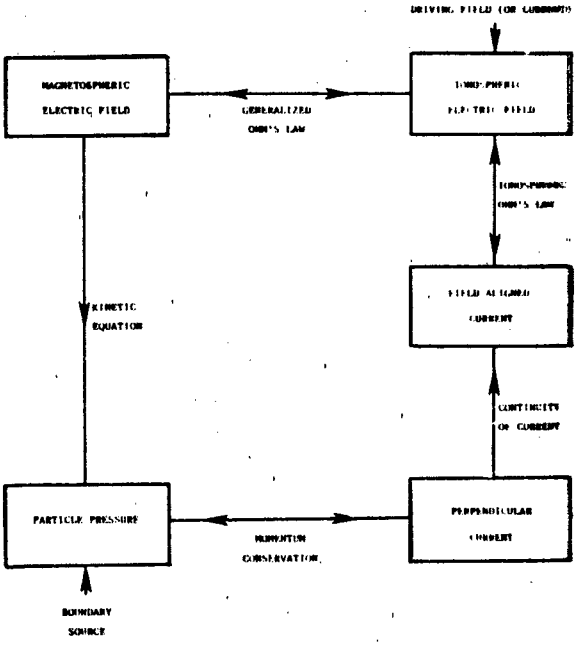


Figure 8-13. Outline of self consistent calculation of magnetospheric convection [Vasyliunas, 1970].

(a) First link: with knowledge of the electric field, we calculate the motion and distribution of protons and electrons in the magnetosphere, and hence the total plasma pressure at any point.

(b) Second link: from the plasma pressure gradients, we calculate the components of the electric current perpendicular to the magnetic field. That is, from the force balance equation

$$\nabla p = \mathbf{j} \times \mathbf{B} \tag{8.30}$$

we calculate

$$j_\perp = \frac{B \times \nabla p}{B^2} \tag{8.31}$$

For simplicity we have assumed that the pressure is isotropic.

(c) Third link: by calculating the divergence of the perpendicular current and averaging over each flux tube, we obtain ($j_\parallel$) the field aligned currents (FAC) flowing between the magnetosphere and the ionosphere.

$$\frac{\partial}{\partial s}\left(\frac{j_\parallel}{B}\right) = -\frac{1}{B}\nabla_\perp \cdot \mathbf{j}_\perp \tag{8.32}$$

(d) Fourth link: from the requi[illegible] that these field aligned currents be closed by perpendi[illegible] ohmic currents in the ionosphere, we obtain the configuration of the electric

field in the ionosphere. The continuity of ionospheric current requries that

$$\nabla \cdot \mathbf{I} = j_{\parallel} \sin\chi \qquad (8.33)$$

where $\mathbf{I}$ is the height integrated current and χ the inclination of magnetic field lines. In the ionosphere

$$\mathbf{I} = \Sigma\sum \cdot (\mathbf{E} + \mathbf{V}_n \times \mathbf{B}) \qquad (8.34)$$

where Σ is the height integrated conductivity tensor and $\mathbf{V}_n$ is the neutral wind velocity.

(c) Fifth and final link: the ionospheric electric field can be mapped into the magnetosphere, and the requirement that it agree with the magnetospheric electric field assumed at the outset determines the field, and thus closes the system of equations. Except near discrete arcs, the mapping may be done by assuming that magnetic field lines are equipotentials. There is empirical evidence suggesting a functional relationship between $j_{\parallel}$ in discrete arcs and field-aligned potential drops [Lyons et al., 1979]. We note in passing that the Rice University group has successfully simulated the ionospheric features of a magnetospheric substorm using this model [Harel et al., 1981].

From Equations (8.33) and (8.34), it is clear that measurements of $\mathbf{E}$ and $j_{\parallel}$ are critical for understanding the magnetosphere-ionosphere circuit. Measurements of precipitating particle fluxes are needed to (1) understand spatial variation in Σ, (2) identify the dominant carriers of $j_{\parallel}$, (3) calculate field-aligned potential drops, and (4) help distinguish between topologically different regions.

8.4 HIGH LATITUDE ELECTRIC FIELDS

Electric fields at ionospheric altitudes are measured from potential differences between the ends of extended booms on satellites [Fahleson, 1967] and from the $\mathbf{E} \times \mathbf{B}$ drifts of cold plasma [Hanson and Heelis, 1975]. In principle, $j_{\parallel}$ can be determined from particle fluxes. With present technology, full distribution functions of ions and electrons cannot be measured with sufficient accuracy. For this reason, highly sensitive triaxial fluxgate magnetometers [Armstrong and Zmuda, 1973] are used to determine $j_{\parallel}$ from magnetic deflections. Particle fluxes in approximately the energy range 10 eV to 30 keV are measured by means of continuous channel electron multipliers (channeltrons) placed behind electrostatic energy-analyzers [Gold, 1959].

It is useful to explain the format of high latitude data presented in the following subsections by providing an example of the simplest kinds of electric fields and magnetic field deflections expected to be measured by instrumentation on a polar orbiting satellite. Figure 8-14a shows the trajectory of a satellite in circular polar orbit with the ascending node at the dusk (1800 LT) meridian. We define a satellite

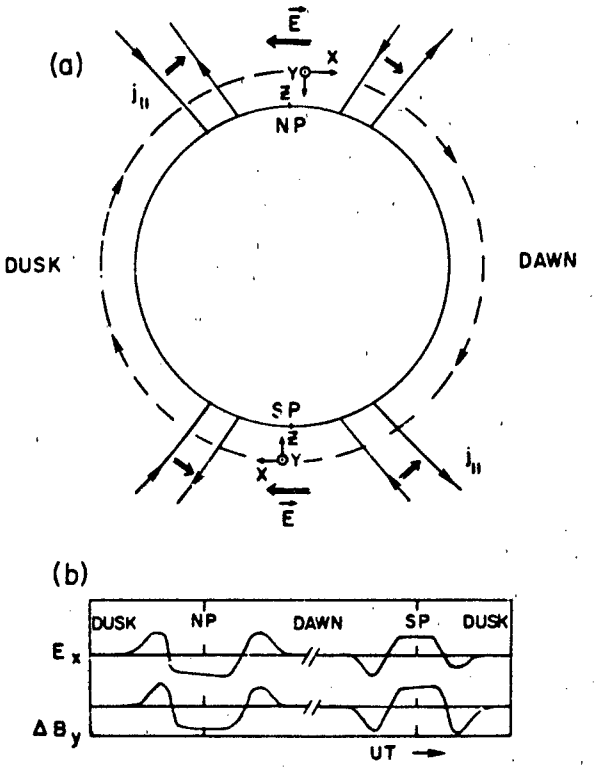


Figure 8-14. (a) Field-aligned currents and electric fields as seen by polar orbitting satellite in dawn-dusk meridian. (b) Idealized electric field and magnetic perturbation measurements

centered coordinate system; $\hat{X}$ is positive along the satellite velocity; $\hat{Z}$ is positive toward local nadir; $\hat{Y}$ completes the right-hand system. In the dawn-dusk meridian Y is positive in the antisunward direction. At high latitudes in the northern (southern) hemisphere we approximate $\mathbf{B}$ as being along the $+(-)Z$ axis. Figure 8-14a also shows dusk to dawn electric fields in the auroral ovals and uniform dawn to dusk electric fields across the polar caps. The convective electric field reverses directions near the poleward boundary of the auroral oval and goes to zero at the equatorward boundary.

The governing equations are current continuity [Equation (8.33)], Ohm's law [Equation (8.34)] and the Maxwell equation

$$\nabla \times \mathbf{B} = \mu_o \mathbf{j} \qquad (8.35)$$

In the infinite current sheet approximation, these equations may be combined to eliminate $j_{\parallel}$ and reduce to

$$\frac{\partial}{\partial x}\left[\Delta B_y - \mu_o \Sigma_p E_x\right] = 0 \qquad (8.36)$$

where ΔB_y is the deflection of the magnetic field due to $j_{\parallel}$ and Σ_p is the height-integrated Pedersen conductivity

[Smiddy et al., 1980]. Except near local noon at the dayside cusp and near local midnight at the Harang discontinuity where there are significant divergences of the ionospheric Hall current [Rostoker, 1980], Equation (8.36) is a very useful approximation. It tells us that for a uniformly conducting ionosphere, fluctuations in the transverse magnetic field component should track variations in the meridional component of the electric field. Deviations from correlated variations are due to the presence of conductivity gradients. Figure 8-14b is a plot of E_X and ΔB_Y expected over a full orbit assuming a uniform ionospheric conductivity. Positive (negative) slopes in ΔB_Y correspond to regions of current into (out of) the ionosphere. Thus, due to the divergence of ionospheric Pedersen currents, a field-aligned current (FAC) should flow into the ionosphere at the equatorward boundary of the dusk side oval and out at the poleward boundary. The morning side oval currents have the opposite polarity. If there are significant Pedersen currents across the polar cap, the poleward FAC system should be of greater intensity than the equatorward system.

In the following subsections, ΔB_Y is given as a function of time rather than distance. Since satellites at ionospheric altitudes travel at speeds of ~7 km/s, Equation (8.35) can be transformed to give a convenient expression for $j_\parallel$

$$j\left(\frac{\mu A}{m^2}\right) = 0.113\,\frac{\partial\,\Delta B_y(nT)}{\partial\,t(s)}. \qquad (8.37)$$

A current of 1 $\mu A/m^2$ corresponds to a locally unbalanced flux of $\sim 10^9/cm^2 s$.

Finally, in both the polar cap and the auroral oval we refer to small- and large-scale structures. Small-scale structures have latitudinal dimensions of a few tens of km or less. They are traversed by satellites in a few seconds. Discrete arcs and inverted-V's are examples of small-scale structures. Somewhat arbitrarily, we define large-scale systems as having latitudinal dimensions greater than 100 km. The E field and FAC systems shown in Figure 8-14b are of large scale.

8.5 POLAR CAP ELECTRODYNAMICS

This section treats three topics: (1) large-scale electric field patterns, (2) electron precipitation morphologies, and (3) characteristics of discrete, sun-aligned arcs in the polar cap. Here we use the term "polar cap" to designate the portion of the high latitude ionosphere containing only open magnetic field lines. Precipitating particles on these field lines should be of direct magnetosheath origin. Except possibly during periods of northward interplanetary magnetic field (IMF), cold ionospheric plasma should convect in the antisunward direction under the influence of a dawn to dusk electric field.

8.5.1 Large Scale Electric Fields Patterns

Table 8-1 lists the six polar orbiting satellites launched to date that were capable of measuring ionospheric electric fields. Data from the double-probe experiment on Injun 5 confirmed the existence of convective reversals near the poleward boundaries of the auroral oval. An inclination of 68° allowed the Atmospheric Explorer (AE-C) driftmeter to measure convective drifts in the oval but usually not in the polar cap. The AE-D satellite had an inclination of 90° but failed ~ 4 months after launch. The orbit was initially close to the noon-midnight meridian and precessed toward dawn-dusk. Because of the high altitude of S3-3, its near-apogee data have been most useful for identifying the small-scale features of auroral arc/inverted-V phenomena. The OGO 6 and S3-2 satellites spent sufficient periods of time near the dawn-dusk meridian to identify the main large-scale features of polar cap convection.

Table 8-1 Satellites capable of measuring electric fields

SATELLITE	LAUNCH DATE	INCLINATION	INITIAL APOGEE (km)	TYPE INSTRUMENT
Injun 5	Aug 1968	81°	2550	double probe
OGO 6	June 1969	82°	1600	double probe
AE-C	Dec 1973	68°	4000	drift meter
AE-D	Oct 1975	90°	4000	drift meter
S3-2	Dec 1975	96°	1550	double probe
S3-3	Aug 1976	98°	8050	double probe

Two examples of E_X measured by OGO 6 at northern (summer) high latitudes are given in Figure 8-15. Except for small-scale variations, the main features of the expected E_X patterns are found in the auroral oval. The example in the top trace is consistent with a uniform dawn to dusk electric field across the polar cap. In the bottom trace E_X has relatively high (low) values near the morning (evening) flank of the polar cap.

A simplified summary of large-scale electric field patterns identified by Heppner [1972] is given in Figure 8-16.

- Type A: Uniform E_X across polar cap. These were observed by OGO 6 only in the northern (summer) polar cap.
- Type B: Strong E_X near the morning (evening) flank of the northern (southern) polar cap.
- Type D: Strong E_X near the evening (morning) flank of the northern (southern) polar cap.
- Type F: Strong E_X along flanks of polar cap with weak fields in the central polar cap.
- Type I: Irregular fields across the polar cap.

Types A and D (B) were found when the IMF was in a toward (away) structure.

During the last three months of 1976 the S3-2 orbit was close to the dawn-dusk meridian. Figure 8-17 is a scatter plot of Heppner's patterns observed in S3-2 data as a function of IMF B_X and B_Y. It is seen that A and B types are found only when $B_Y > 0$. For both OGO 6 and S3-2, A types are only found in the summer hemisphere and are associated with the IMF polarity that tends to produce strong E fields along the evening flank of the polar cap. The occurrence ration of A to B was 2:1 in S3-2 data; the A to D ratio was 3:1 in June 1969 OGO 6 data. The dependence of D types on B_Y (upper right plot) agrees with OGO 6. The type F patterns were found when $B_Y \sim 0$. The distribution of B, D, and F patterns tend to confirm the critical role of B_Y [Friis-Christensen et al., 1972]. A survey of S3-2 measurements shows that A, B, D, and F types are found when $B_Z < 0.5$ nT. Type I shows no correlation with B_X or B_Y. The highly irregular electric fields are found only when $B_Z > 0$. Type I patterns are further discussed below in connection with polar cap arcs.

An example of an E-field pattern that was found in approximately half the summer polar cap passes of S3-2 when $B_Z > 0.7$ nT is shown in Figure 8-18. During S3-2, Rev 5215 southern high-latitude pass B_Z was 4.9 nT [Burke et al., 1979]. As expected for driving sunward convection in the auroral oval, E_X was directed from dusk to dawn. At the poleward boundary of the oval, E_X reversed polarity becoming dawn to dusk near the morning and evening flanks of the polar cap. Within the central polar cap E_X was directed from dusk to dawn. Sunward convection in the central polar cap is inconsistent with a viscous interaction model. It was theoretically anticipated as a consequence of the magnetic merging model

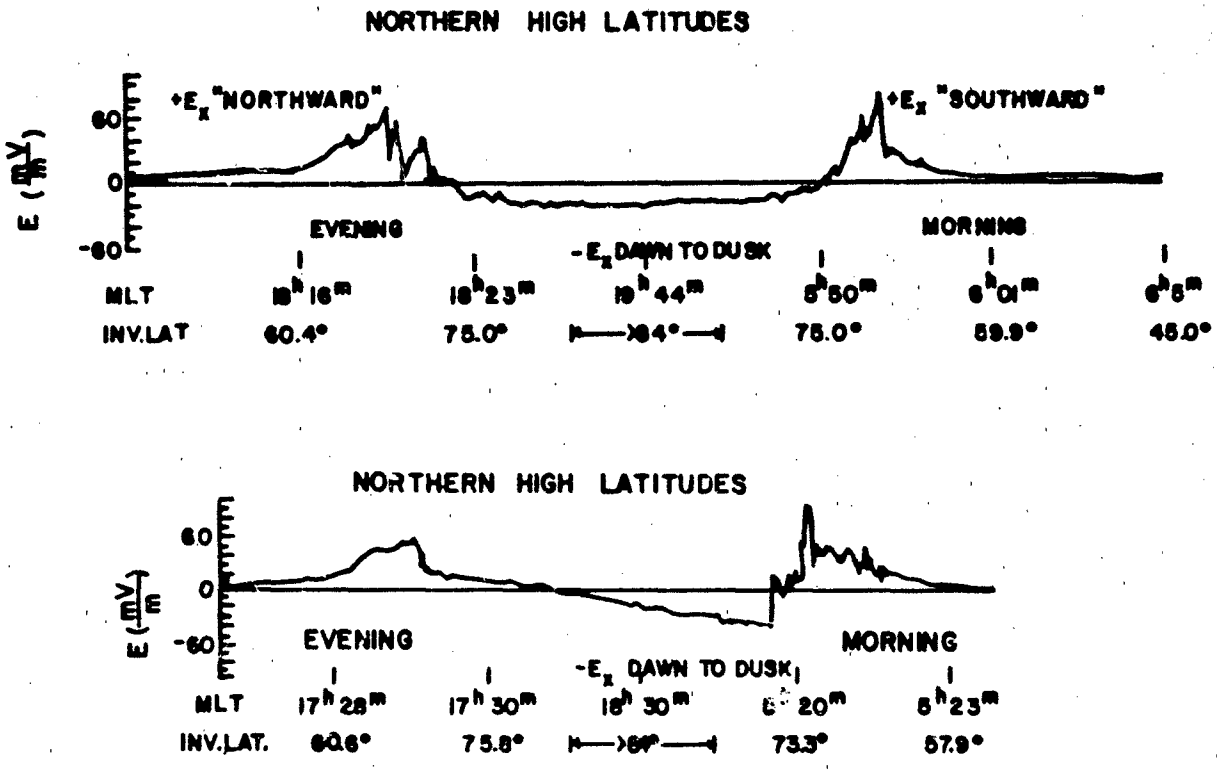


Figure 8-15. Two examples of large scale electric fields measured by OGO 6 [Heppner, 1972].

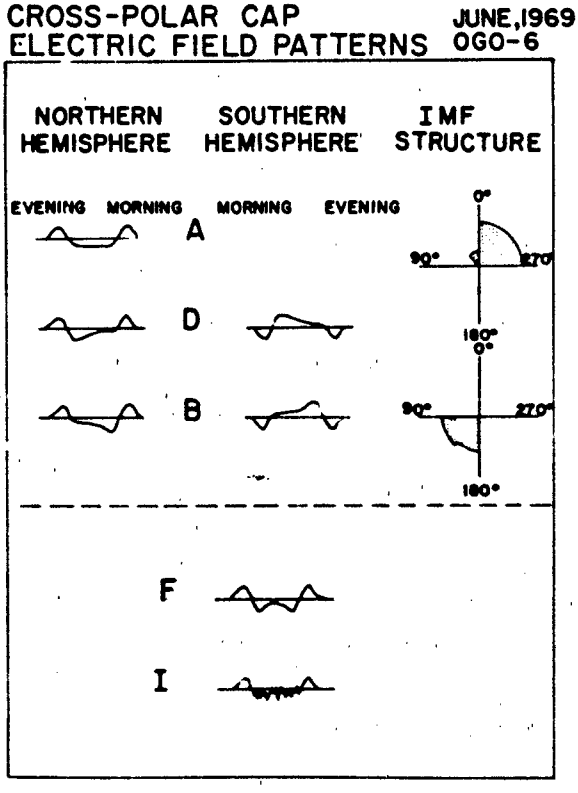


Figure 8-16. Types of electric field patterns observed by OGO 6 and their dependence of IMF B_x and B_y [Heppner, 1972].

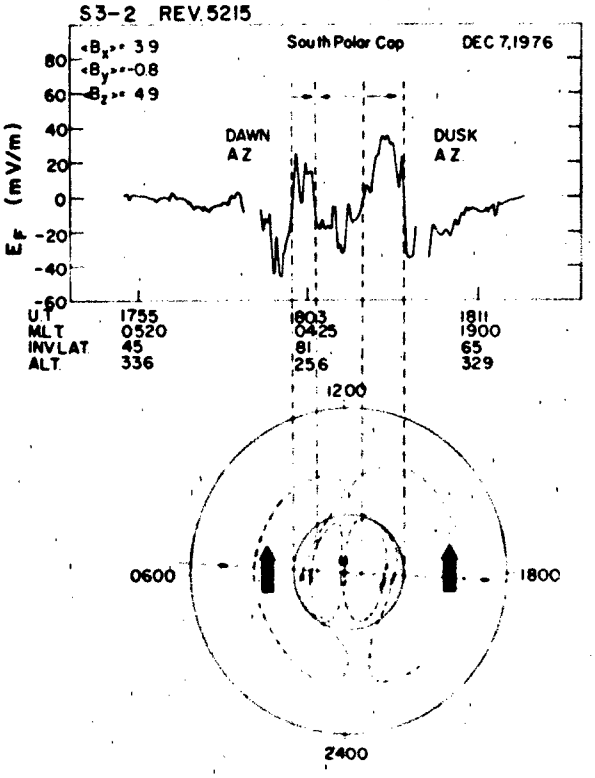


Figure 8-18. Example of sunward convection, dusk to dawn electric field in the central polar cap during a period of northern IMF.

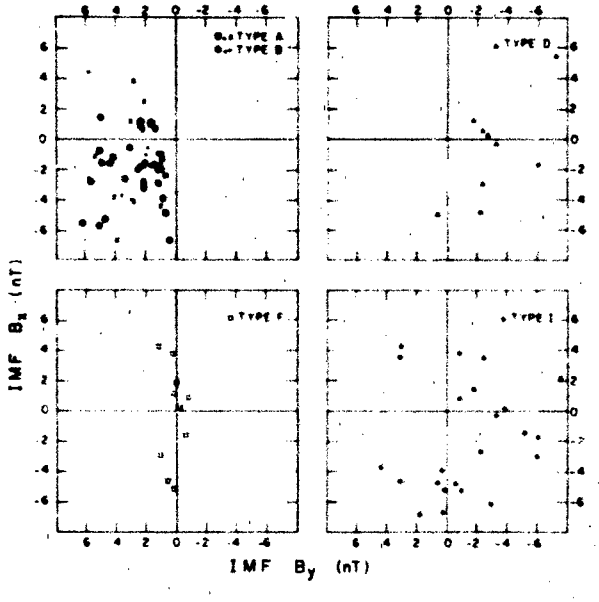


Figure 8-17. Scatter plot of convective electric field patterns observed by S3-2 as a function of IMF B_x and B_y.

[Russell, 1972]. Sunward convection in the summer polar cap during periods of northward B_z is consistent with both ground magnetometer [Maezawa, 1976] and laboratory simulation [Podgorny et al., 1978] results.

The cross polar cap potential $\Delta\Psi_{pc}$ can be derived from satellite electric field measurements and is an important parameter for magnetospheric modeling. This potential (in volts) gives the rate (in webers/s) that magnetic flux is transferred from the day to the night side of the magnetosphere. Based on two weeks of OGO 6 data Heppner [1977] found that on the average $\Delta\Psi_{pc}$ increased from 20 to 100 kV as the magnetic index Kp increased from 0 to 6. There were individual cases in which $\Delta\Psi_{pc}$ significantly exceeded 100 kV.

Reiff et al. [1981] analyzed 32 measurements of $\Delta\Psi_{pc}$ from AE-C and AE-D as a function of various solar wind and IMF parameters. To calculate merging rates that account for compression of the IMF in the magnetosheath, B at the magnetopause was set at the lesser of 8 times its solar wind value or 60 nT. The latter value was taken as typical of the earth's field near the subsolar magnetopause. Although the best agreement was found with theoretical merging rates, a high correlation was found with Akasofu's [1978] ϵ parameter. Figure 8-19 shows that in the AE measurements $\nabla\Psi_{pc}$

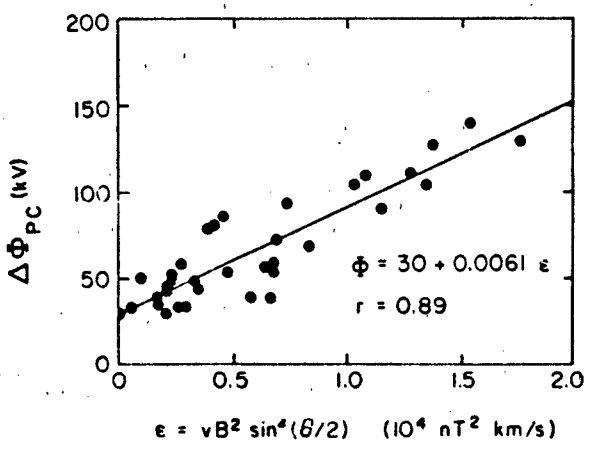


Figure 8-19. Cross polar cap potential as a function of solar wind parameter E [Heppner, 1977].

varied from 30 to 150 kV. This implies that 30 kV cannot be accounted for by merging. Such a potential greatly exceeds the potential across the boundary layer theoretically estimated by Hill [1979] and measured at ionospheric altitudes by Smiddy et al. [1980]. The upper limit of ~150 kV is much less than the cross magnetosphere potential drop ($\Delta\Psi_{cs}$) in the solar wind. With V_x = 400 km/s and B_y = −5nT, the Y component of the electric field in the solar wind is 2 mV/m (= 12.8 kV/R_E). For a magnetospheric diameter of 30 R_E at the dawn-dusk meridian, $\Delta\Psi_{cs}$ = 384 kV.

8.5.2 Polar Cap Precipitation

Particle fluxes into the polar ionosphere are conveniently divided into high and low energy components. Energetic particles from solar flares can seriously disrupt the polar ionosphere; these are important during magnetic storm periods. The flux levels of low energy protons in the polar cap are below the sensitivity levels of existing detectors. Winningham and Heikkila [1974] identified three classes of low energy electron precipitation: polar rain, polar showers, and polar squalls.

Polar rain is a relatively uniform type of precipitation that can fill the entire polar cap. Particles have mean thermal energies of ~ 100 eV and are isotropically distributed outside the atmospheric loss cone. The energy fluxes carried by these particles are in the 10^{-2} to 10^{-1} ergs cm^{-2} s^{-1} range, two to three orders of magnitude less than typical auroral fluxes. The highest energy fluxes for polar rain occur during periods of geomagnetic activity [Meng and Kroehl, 1977]. Figure 8-20 is a plot of precipitating electron spectra measured in the dayside cusp and in the polar rain. The similarity in spectral shape suggests that polar rain particles are of direct magnetosheath origin. Particle fluxes measured in the

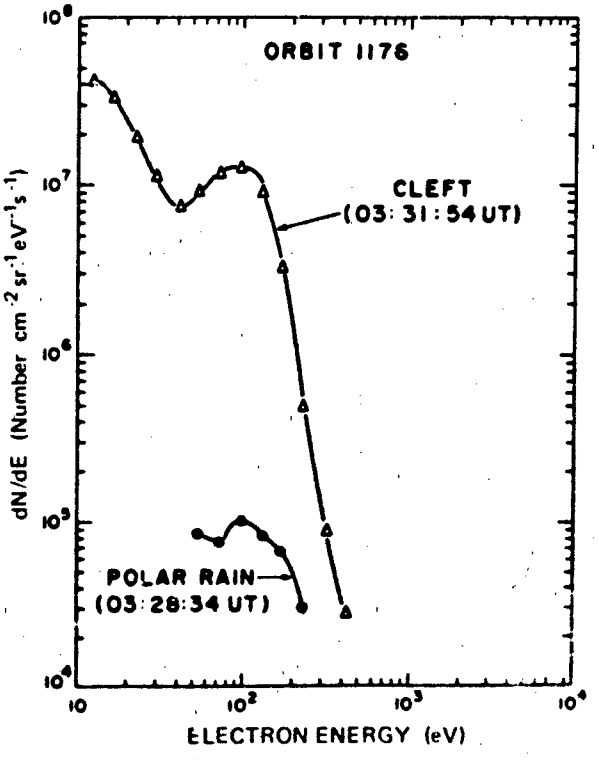


Figure 8-20. Typical differential spectra for the polar rain and cleft precipitation observed on orbit 1176 on 15 May 1969 [Winningham and Heikkila, 1974].

lobes of the magnetotail indicate that polar rain electrons enter the magnetosphere at great distance downstream from the earth. The efficiency of the entry process is modulated by the polarity of the IMF. Yaeger and Frank [1976] found that fluxes of soft electrons in the northern lobe of the tail increased by more than an order of magnitude when the IMF was in an away (B_x < 0) sector. There is also evidence suggesting a B_y influence. Meng et al. [1977] found that the intensity of polar rain fluxes is strongest near the flank of the polar cap along which convection is strongest.

Polar showers are characterized by locally enhanced fluxes of precipitating electrons with mean energies of ~1 keV. These electron structures are embedded in broader regions of polar rain. They are thought to be responsible for sun-aligned arcs in the polar cap and are discussed further in Section 8.5.3.

Polar squalls are described by Winningham and Heikkila [1974] as localized, intense fluxes of electrons that have undergone field-aligned accelerations of several kV. They are found in the polar cap during geomagnetic storms. Foster and Burrows [1976; 1977] have reported observing fluxes of electrons into the polar cap that are spectrally identical to those found in polar squalls. However, these fluxes were observed to be widely and uniformly distributed over the polar cap. Like squall particles, they were observed during the recovery phases of magnetic storms. These fluxes also

appear to be modulated by the polarity of the IMF. While intense fluxes of keV electrons were measured in the northern polar cap, only polar rain was detected in the southern hemisphere. Figure 8-21 is a plot of particle fluxes measured

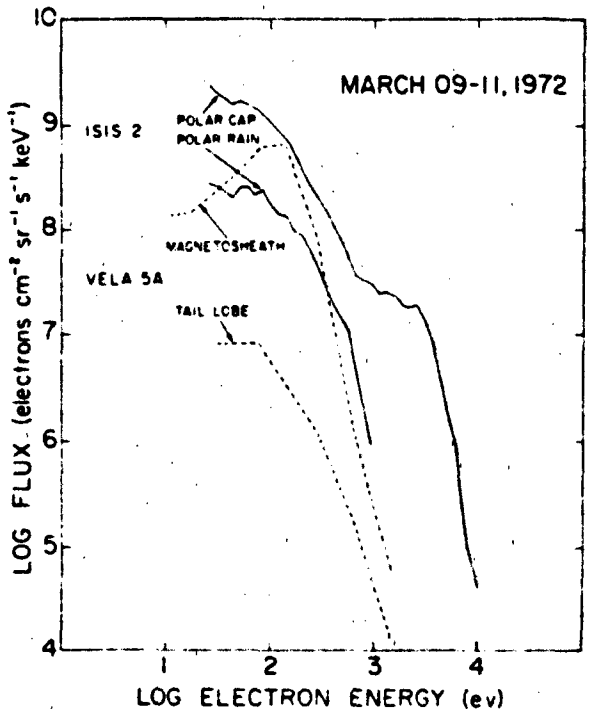


Figure 8-21 Differential electron energy spectra measured in the polar cap (solid curves) and in the tail lobe and magnetosheath (dashed curves). The magnetosheath measurement was made at ~13 hrs on 9 March, the tail lobe at ~9 hrs on 11 March, the polar cap at ~16 hours on 10 March, and the polar rain at ~11 hrs 9 March. The polar cap spectra display a more or less pronounced high-energy tail relative to the magnetosheath [Foster and Burrows, 1976]

in the polar rain and extended squalls by ISIS 2 and nearly simultaneous measurements of electron fluxes in the magnetosheath and the tail lobe from VELA 5. Foster and Burrows [1977] argue from the near isotropy of the keV particles in the polar cap and the simultaneous absence of keV electrons in the magnetosheath that the electrons were accelerated along magnetic field lines at great distance from the earth in the magnetosphere.

8.5.3 Polar Cap Arcs

Investigations of discrete arcs in the polar cap have shown that polar cap arcs tend to be sun-aligned and are most frequently observed during periods of magnetic quieting when the IMF has a northward component [Ismail et al., 1977]. Visible arcs are caused by precipitating electrons with energies of ≤ 2 keV. Another class of subvisual arcs are produced at F-layer altitudes by electrons with energies of a few hundred eV [Weber and Buchau, 1977]. Here we illustrate many of the known characteristics of polar cap arcs using data from the USAF satellites S3-2 and DMSP (Defense Meteorological Satellite Program). At the times of interest, both satellites were in orbits close to the dawn-dusk meridian. S3-2 measured E_X, ΔB_Y (described in Section 8.5.1), and fluxes of electrons with energies between 50 eV and 17 keV. DMSP satellites are three-axis stabilized and are in circular, sun-synchronous orbit at an altitude of 840 km. All DMSP satellites are equipped with scanning, optical imagers [Eather, 1979]. Some, but not all, are also equipped with spectrometers that look toward local zenith and measure fluxes of electrons with energies between 50 eV and 20 keV.

Figure 8-22 is a cartoon that represents, in magnetic

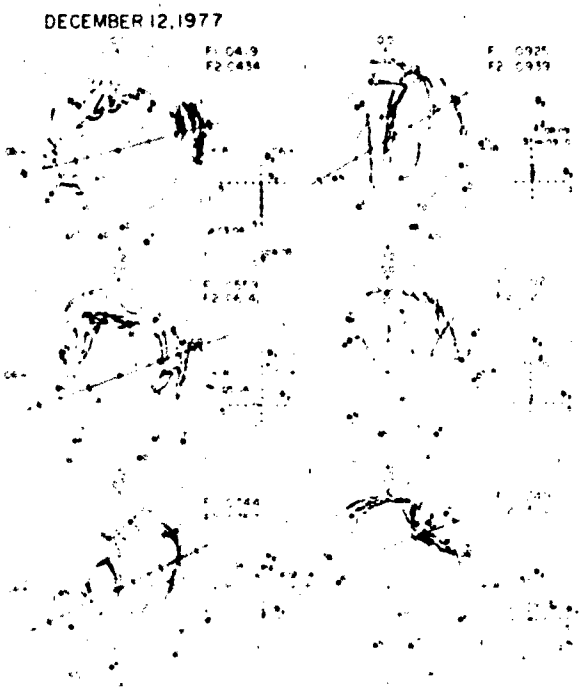


Figure 8-22 Cartoon representation of sequential DMSP high latitude imagery on 12 December 1977

latitude and local time, composites of visible imagery from DMSP/F1 and DMSP/F2 taken over the northern hemisphere during a period of magnetic quieting on 12 December 1977. Straight solid lines give the portions of F2 trajectories during which electron data were taken. To the right of each cartoon, the hourly average values of IMF B_Y and B_Z are represented. During the initial period of southward B_Z, the polar cap was clear of visible emissions and only uniform, polar rain fluxes were detected. Approximately one hour after the IMF turned northward sun-aligned arcs were found in the polar cap. Polar cap arcs persisted until the IMF again turned southward. An hour after a second northward turning

of B_z, arcs returned to the polar cap [Hardy et al., 1981]. The sun-aligned arcs were embedded in a region of high density (~0.1 cm $^{-3}$) polar rain. Within the arc the up-looking DMSP spectrometer detected three spectral components, a cold (100 eV) high density (1.5 cm $^{-3}$) population, a peaked primary distribution with a temperature of 350 eV that had been accelerated through a potential drop of ~750 volts and a secondary and/or degraded primary population. Burch et al. [1979] found that over polar showers the low-energy component was highly field-aligned. The secondary and accelerated primary populations were nearly isotropic in pitch angle.

Figure 8-23 gives a plot of E_X, ΔB_Y, the directional flux of electrons (cm^2 s sr) $^{-1}$ and electron pitch angles measured during S3-2 Rev 5231 as functions of invariant latitude, magnetic local time, and altitude. The pass occurred while the satellite was near apogee over the north polar cap where it passed within 1° of the magnetic pole along the dawn-dusk meridian. The IMF X, Y, and Z components were −3.7, 3.8 and 7.4 nT, respectively. As compared with the idealized measurements of Figure 8-14b, E_X and ΔB_Y were highly irregular. Recall that in the northern hemisphere E_X positive corresponds to sunward convection; $j_\parallel$ is out of the ionosphere in regions where ΔB_Y has a negative slope. I[illegible] regions of negative slope in ΔB_Y accompanied by enhanced electron fluxes are noted in Figure 8-23. As evidenced by their being embedded in polar rain, Events 3 through 7 lie in the polar cap.

Event 6 has been analyzed in detail by Burke et al. [1982]. It was shown that the FAC out of the ionosphere had an intensity of 2.8 μ A/m^2 that was carried by electrons with a temperature of 200 eV and had been accelerated through a potential drop of ~1 kV. A nearly isotropic pitch angle distribution of electrons across event 6 suggests a field-

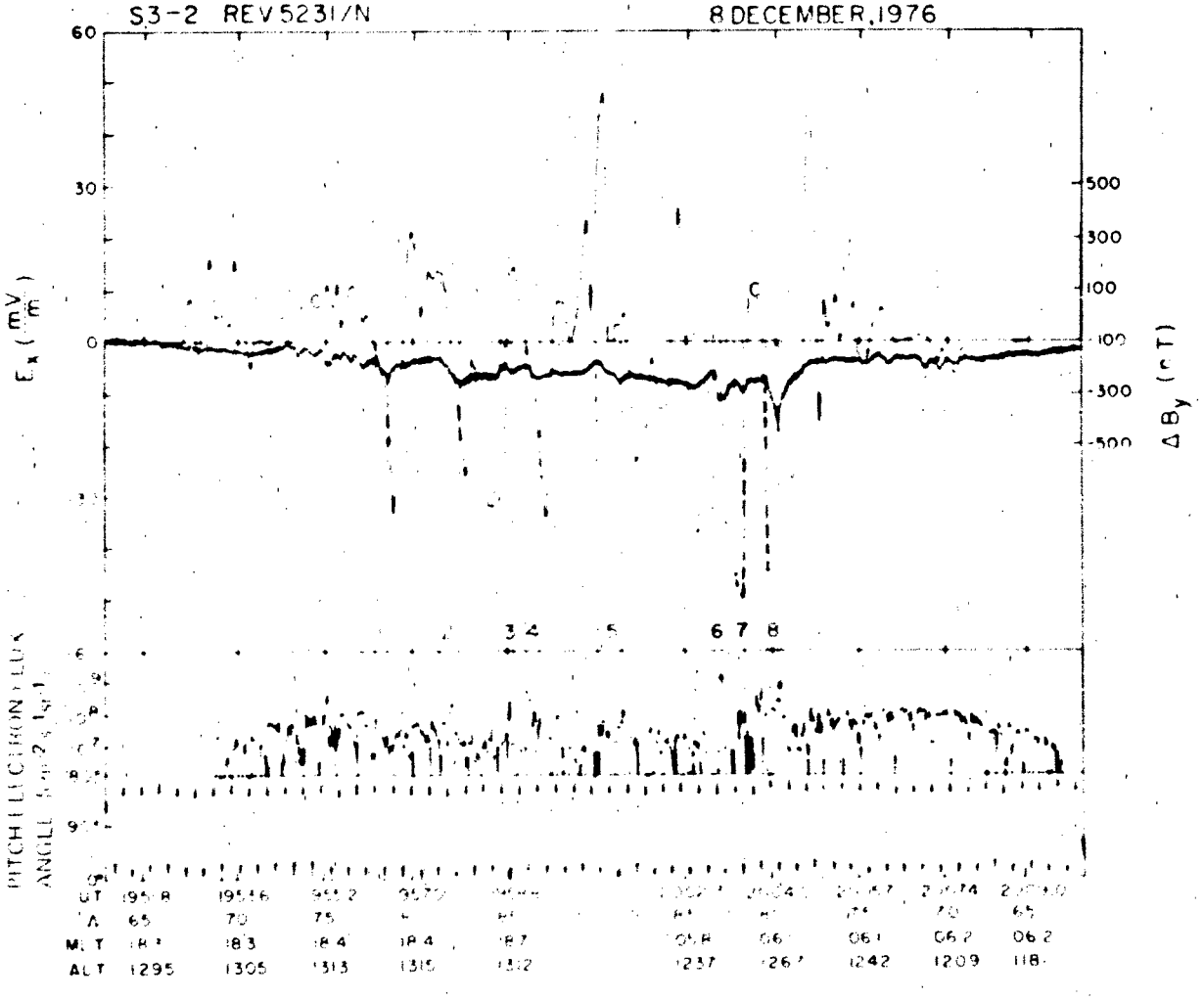


Figure 8-23 The top panel gives the dawn to dusk electric field component and the transverse magnetic field deflection (heavy line). The bottom panels give the directional electron flux and pitch angles. Data were taken over winter polar cap with IMF B_z northward

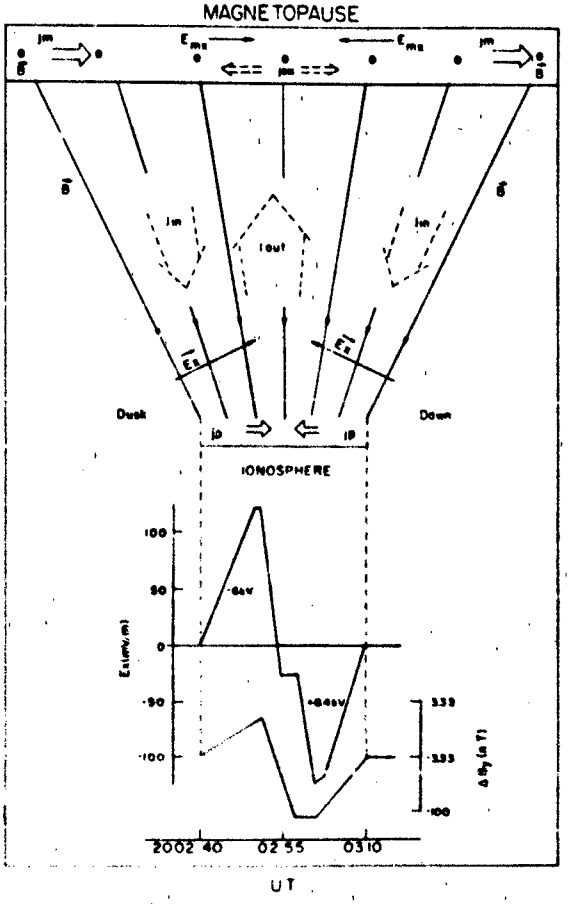


Figure 8-24. Idealized two dimensional projection of electric fields and currents in vicinity of Event 6 in Figure 8-23.

aligned potential drop extending for large distances along **B**. The measured electron energy fluxes of 2.5 ± 0.5 ergs $cm^{-2} s^{-1} sr^{-1}$ were sufficient to produce a visible arc. Figure 8-24 gives an idealized, two-dimensional projection of the electric fields and currents associated with Event 6. As predicted by Lyons [1980;1981], the arc is in a region of negative electric field divergence. Similar, recently reported measurements from the Dynamics Explorer satellites suggest that such polar cap arcs may result from a bifurcation of the distant plasma sheet. That is, some polar cap arcs are on closed magnetic field lines [Frank et al., 1982].

8.6 AURORAL OVAL ELECTRODYNAMICS

The auroral oval is a region of closed magnetic field lines mapping through the plasma sheet. Ionospheric plasma convection should be mostly in the sunward direction. Near local midnight (noon) strong equatorward (poleward) convection components are expected.

The stock-in-trade vocabulary of a contemporary auroral physicist is replete with terms such as "inverted-Vs", "double-layers", "diffuse aurora", and "beams and conics" [Mozer et al., 1980]. These terms describe distinct auroral processes first reported in the 1970s. These new phenomena defy adequate summation in this section. Here we consider three main topics: (1) global auroral morphology, (2) phenomena associated with inverted-V events, and (3) substorms. Under the heading of global morphology we treat the systematics of large-scale, field-aligned current systems and the equatorward boundary of diffuse auroral precipitation.

8.6.1 Global Field Aligned Currents

The average global field-aligned current (FAC) system for periods of high and low magnetic activity are shown in Figure 8-14. Iijima and Potemra [1978] define Region 1 (2) as a region of FAC near the poleward (equatorward) portion of the auroral oval. In the evening sector current flows into the ionosphere in Region 2 and out in Region 1. The polarity of current flow is reversed in the morning sector. The intensities of these currents are of the order 1 $\mu A/m^2$. Currents out of the ionosphere are mostly carried by precipitating electrons. It is believed that cold electrons moving from the ionosphere to the magnetosphere carry the current into the ionosphere. Klumpar [1979] found that the equatorward boundary of Region 2 is coterminus with diffuse auroral electron precipitation in the post-midnight sector. In the evening sector Region 2 extends ~2° equatorward of the electron precipitation boundary.

Near local midnight and noon large-scale FACs are the most complex. Figure 8-25 shows that near midnight the

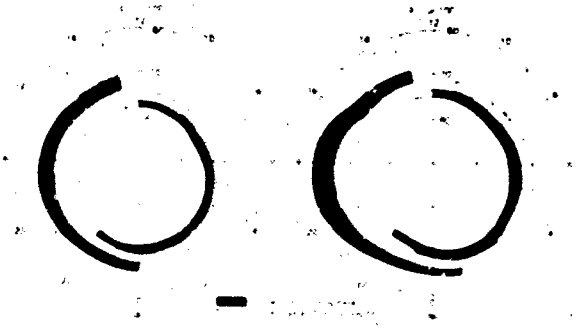


Figure 8-25. A summary of the distribution and flow directions of large scale field-aligned currents determined from (a) data obtained from 430 passes of Triad during weakly disturbed conditions and (b) data obtained from 366 Triad passes during active periods [Iijima and Potemra, 1978].

morning side Region 1 current overlaps the evening side Region 1. No simultaneous electric field and magnetic field measurements have yet been reported from this region. However, it is expected that the morphology of these FACs can be understood in terms of latitudinal [illegible] of the electric field just pri[illegible] [1974]

showed that in the late evening sector the large scale convective electric field is directed poleward in the equatorward part of the oval. It rotates through west across the Harang discontinuity to equatorward in the poleward part of the oval. Pedersen currents driven in the ionosphere converge from both sides on the Harang discontinuity. To maintain an overall divergence-free current system, current must flow into the ionosphere at both the equatorward and poleward boundaries of the oval and out of the ionosphere near the Harang discontinuity. These are the essential features found near midnight in Figure 8-25.

In the vicinity of the dayside cusp an extra FAC system has been observed poleward of the Region 1. Its polarity is opposite to that of the nearby Region 1 current. In the northern (southern) hemisphere it appears only on the afternoon (morning) side of noon when IMF $B_y < 0$, and only on the morning (afternoon) side when IMF $B_y > 0$ [McDiarmid et al., 1979]. Simultaneous electric and magnetic field measurements from the USAF satellite S3-2 in the region of the dayside cusp suggest that the extra FAC system lies entirely on open, newly merged magnetic field lines that are being dragged toward the dusk or dawn flank of the polar cap [Doyle et al., 1981].

8.6.2 Equatorward Boundary of the Oval

In the previous section it was pointed out that the inner edge of the plasma sheet, or zero-energy Alfvén boundary, maps to the equatorward boundary of diffuse auroral precipitation. More than four thousand crossings of this boundary have been analyzed using data from an upward-looking electron spectrometer on the polar orbiting USAF satellites DMSP F2 and DMSP F4 [Gussenhoven et al., 1981]. The corrected geomagnetic latitudes of the boundaries ($\Lambda_{c,o}$) were studied as functions of magnetic local time (MLT) and Kp. Kp is a three-hour index of magnetic activity compiled from a world-wide network of midlatitude magnetometer stations. Results of linear correlation analyses

$$\Lambda_{c,o} = \Lambda_o + \alpha\,Kp$$

from available MLT sectors along with correlation coefficients are given in Tables 8-2 and 8-3. Using the magnetic field model of Fairfield and Mead [1975] the auroral boundaries were projected to the magnetospheric equatorial plane. Figure 8-26 is a plot of projected boundary positions (open circles) in comparison with predictions of the Volland-Stern electric field model [Ejiri et al., 1978] (solid line) and the injection boundary of Mauk and McIlwain [1974] (dashed lines). Best fits are obtained for a shaping factor $\gamma = 2$. Note that the stagnation points are offset from the dusk meridian. The offset angle ϕ_o varies from 23° toward evening for Kp = 0 to 45° toward afternoon for Kp = 5. The potential distribution in the equatorial plane (Equation (8.24) of the previous chapter) takes the form

$$\Psi(L, \phi) = \Omega B_o R_I^2 \left[\frac{L^\gamma}{\gamma L_s^{\gamma+1}} \sin(\phi - \phi_o) - \frac{1}{L}\right] \tag{8.38}$$

For $\gamma = 2$ the term $1/\gamma (L_s^{\gamma+1})$ is empirically related to Kp

$$\frac{1}{\gamma L_s^{\gamma+1}} = (1.6 + 2.4\,Kp) \cdot 10^{-4}$$

with a correlation coefficient of 0.97. Thus,

$$\Psi(L, \phi) = \Omega B_o R_I^2 \left[(1.6 + 2.4\,Kp)\,10^{-4} L^2 \sin(\phi - \phi_o) - \frac{1}{L}\right] \tag{8.39}$$

A further analysis [Hardy et al., 1982] has been performed correlating the boundary data and hourly averaged solar wind and IMF measurements. The most significant results were obtained when $\Lambda_{c,o}$ was correlated with the interplanetary electric field VB_z for the hour preceeding the boundary measurement subject to the condition $B_z \leq 1$ nT. Correlation parameters are summarized in Table 8-4. To compare the results of the two studies a correlation analysis was performed between Kp and VB_z.

With $B_z \leq 1$ nT

$$Kp = 2.09 - 0.91\,VB_z \tag{8.40}$$

Table 8-2. Regression coefficients for auroral boundaries in the morning sector

	North				South			
MLT	Λ_o	α	N	cc	Λ_o	α	N	cc
0300–0500	67.4	−1.35	171	−0.58				
0500–0600	67.8	−1.87	265	−0.75				
0600–0700	68.5	−1.96	403	−0.82	67.4	−1.67	376	−0.74
0700–0800	70.2	−2.15	367	−0.83	68.3	−1.97	411	−0.81
0800–0900					68.7	−1.88	302	−0.72
0900–1000					69.1	−1.64	217	−0.67

Table 8-3. Regression coefficients for auroral boundaries in the evening sector.

	North				South			
MLT	Λ_o	α	N	cc	Λ_o	α	N	cc
1600–1700					71.3	−1.19	107	−0.65
1700–1800					70.7	−1.20	256	−0.69
1800–1900	71.6	−2.00	103	−0.90	70.6	−1.60	327	−0.80
1900–2000	71.2	−1.96	426	−0.89	70.0	−1.82	447	−0.87
2000–2100	69.4	−1.85	452	−0.82	69.5	−1.89	345	−0.84
2100–2200	68.7	−1.66	556	−0.83				
2200–2300	68.3	−1.79	184	−0.63				

with VB, in mV/m. This allows us to express the magnetospheric potential as a function of VB_Y.

$$\Psi(L, \phi) = \Omega B_o R_E^2 \left[(6.6 + 2.2\ VB_Y)\ 10^{-4} L^2 \sin(\phi - \phi_o) - \frac{1}{L}\right] \quad (8.41)$$

To test the validity of this equation it is possible to calculate the cross-magnetospheric potential drop ($\Delta\Psi_m$) and compare it with measured cross polar cap potential drops ($\Delta\psi_{pc}$).

To estimate $\Delta\Psi_m$ we use the position and shape of the magnetopause in the equatorial plane given by McIlwain [1972]. In this representation the magnetopause lies at a distance of $11R_E$ near the subsolar point and flares to a distance of $15R_E$ at the dawn-dusk meridian. We ignore stormtime compressions of the magnetosphere in the calculations. Two methods of calculating $\Delta\Psi_m$ are used. The first method sets $\phi_o = 0$ in Equation (8.40) making the axis of symmetry the

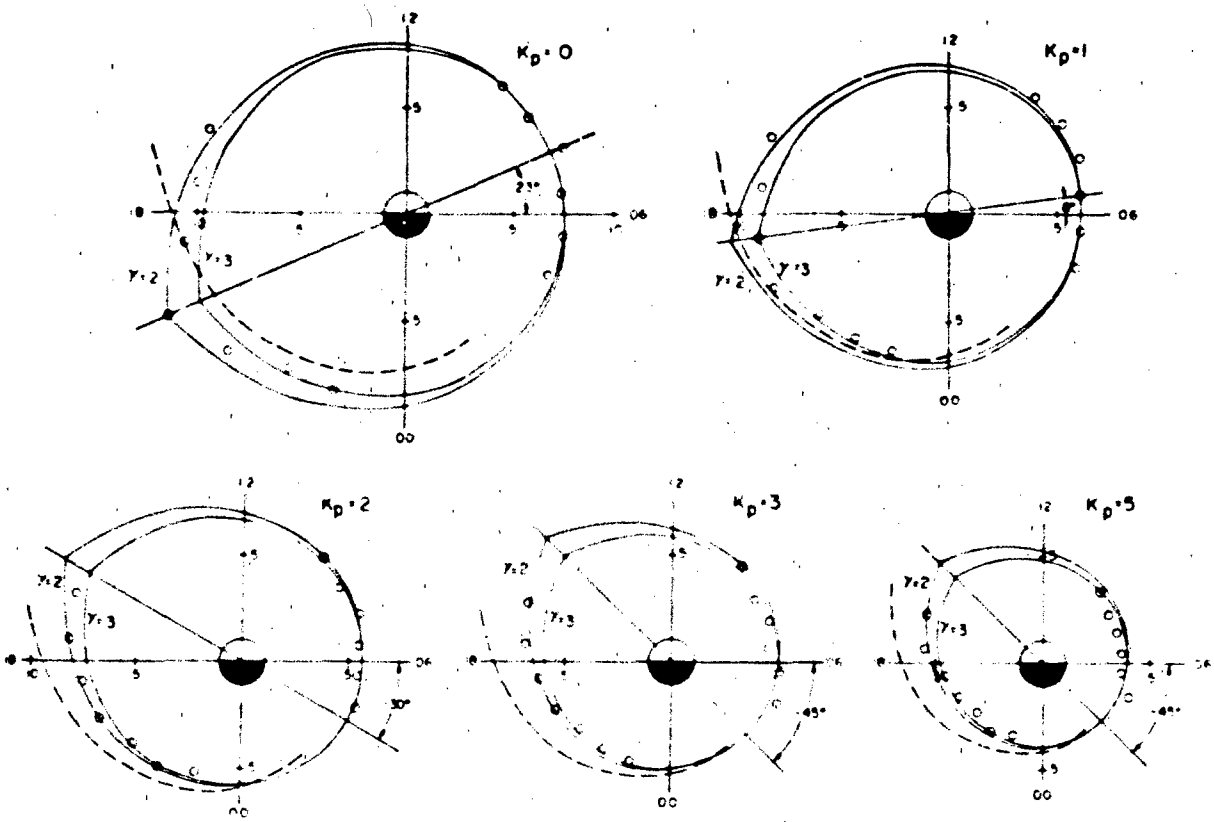


Figure 8-26. Volland-Stern injection boundaries for $\gamma = 2$ and $\gamma = 3$, rotated to fit the inner edge of the plasma sheet as determined by the DMSP F2 auroral boundaries, for various Kp. The Mauk-McIlwain [1974] injection boundary is also shown (dashed line).

Table 8-1. The intercepts (Λ_o), slopes, correlation coefficients (cc) and sample sizes (N) for the linear regression of the boundary location with VB_Z in each magnetic local time zone, with a 1 h delay in the value of the interplanetary magnetic field (IMF) used.

	North Pole					South Pole		
MLT	Λ_o	Slope	cc	N	Λ_o	Slope	cc	N
0400–0500	65.1	3.28	0.52	43				
0500–0600	63.6	1.94	0.61	110				
0600–0700	64.2	1.80	0.70	105	64.4	1.70	0.56	124
0700–0800	66.1	2.56	0.68	85	63.7	1.95	0.61	94
0800–0900					65.1	2.27	0.69	68
0900–1000					65.5	1.74	0.74	38
1000–1100					65.8	0.44	0.59	19
1600–1700					69.5	1.89	0.82	26
1700–1800					68.6	1.36	0.74	57
1800–1900	68.6	2.49	0.81	44	67.2	1.60	0.70	83
1900–2000	67.3	1.76	0.68	123	66.6	1.81	0.70	141
2000–2100	65.2	1.71	0.70	130	65.8	2.01	0.64	44
2100–2200	65.5	1.91	0.71	132				

dawn-dusk meridian. The second method uses empirical values of ϕ_o, given in Figure 8-26. For Kp = 0 the axis is tilted with the dusk stagnation point 23° to the nightward side of the dawn-dusk meridian. For higher values of Kp the stagnation point rotates sunward reaching a constant value of −45° for Kp > 2. In Figure 8-27 we have sketched the

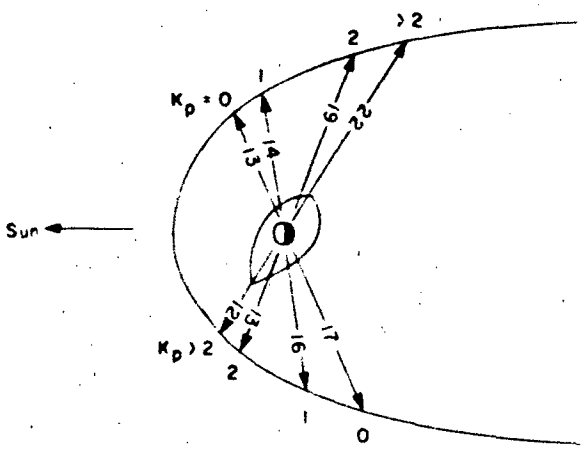


Figure 8-27. Equatorial magnetopause with distances along axes of symmetry for various values of Kp.

axes of symmetry indicating approximate distances to the dawn and dusk sides of the magnetopause for Kp = 0, 1, 2 and >2.

In Figure 8-28 $\Delta\Psi_m$ is plotted as a function of Kp, VB_Y and B_Y. VB_Y is positive when the interplanetary electric field (IEF) is directed from dawn to dusk (positive Y in standard geocentric, solar-magnetospheric coordinates). Values of B_Y are derived from VB_Y using a solar wind speed of 400 km/s. The solid line calculations assume $\phi_o = 0$, that is, the dawn-dusk meridian is the axis of symmetry. For this case, $\Delta\Psi_m$ increases linearly from 6.5 kV for Kp = 0 to 65 kV for Kp = 6. The dot-dash line indicates values of $\Delta\Psi_m$ using empirical values of ϕ_o and axes of symmetry indicated in Figure 8-27. In these calculations $\Delta\Psi_m$ ranges from 5.5 kV for Kp = 0 to 95 kV for Kp = 6. We note that in these calculations there is an abrupt transition in the trend of $\Delta\Psi_m$ near Kp = 2. This is close to our independently derived "breakpoint" in B_Z.

For reference, in Figure 8-28 we have also plotted average values of $B\Psi_{pc}$ (dashed line) as a function of Kp based on two weeks of OGO 6 electric field measurements. Heppner [1977] found that $\Delta\Psi_{pc}$ increased linearly from 20 kV for

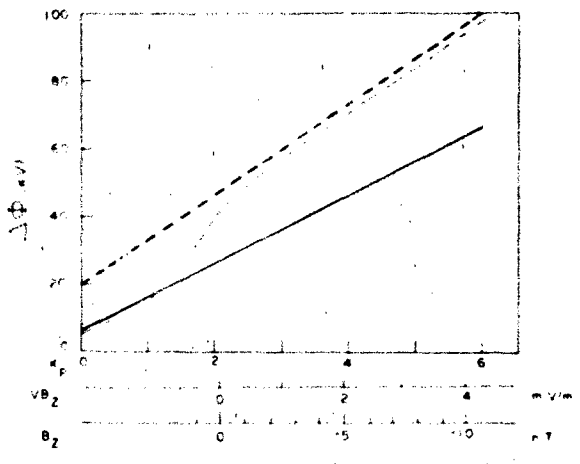


Figure 8-28. Polar cap potential, from Heppner [1977] (dashed line), and from Equation (8.12) with $\phi_o = 0$ (solid line) and ϕ_o with empirical values (dot-dash line).

Kp = 0 to 100 kV at Kp = 6. At low values of Kp the DMSP values of $\Delta\Psi_m$ are not sensitive to the value of ϕ_o and are considerably less than Heppner's $\Delta\Psi_{pc}$. With values of $\phi_o = -45°$ for Kp > 2 $\nabla\Psi_m$ and $\Delta\Psi_{pc}$ are of comparable magnitudes. In a time-average sense $\Delta\Psi_{pc}$ and $\nabla\Psi_m$ are coupled through the process of magnetic reconnection in the magnetotail. The two quantities must be roughly the same since on average, the amount of magnetic flux transferred to the nightside magnetosphere must equal the amount being returned to the dayside magnetosphere.

We emphasize that these are "average" results in the sense that they represent least square fits to the data. Heppner [1977] pointed out the $\Delta\Psi_{pc}$ frequently exceeds 100 kV during period of high Kp. Similarly, in both DMSP boundaries for cases of large negative B_Z and high Kp respectively, the latitude of the boundary is often much lower than the average. Since such cases correspond to the Alfvén boundary being closer than average to the earth, they also correspond to potentials exceeding 100 kV.

8.6.3 Inverted-V Phenomenology

Diffuse auroral electron precipitation is fairly uniform and isotropically distributed. In the poleward portion of the oval, electron fluxes are more spatially structured and are often field aligned. Discrete optical arcs are the most striking manifestation of the structuring process. The dominant structural features of electron fluxes observed with polar orbiting satellites are the so-called inverted-V structures. These structures are latitudinally narrow (~ 1°) bands of electron precipitation that increase in average energy from a few hundred eV to several keV, then return to a few hundred eV [Frank and Ackerson, 1971]. On energy-time spectragrams they have the shapes of inverted-Vs. Although individual inverted-V structures have been identified with ground-observed discrete arcs [Ackerson and Frank, 1972], the general relationship between the two phenomena is not clear. Inverted-Vs have latitudinal dimensions of several hundred km, whereas optical arcs have typical widths of 1 to 10 km [Maggs and Davis, 1968].

Lin and Hoffman [1979] have studied the global distribution of inverted-V structures and the pitch angle distribution of electrons within them. Figure 8-29 is an event occurrence map of 280 inverted-V structures observed with the AE-D satellite. This map shows that inverted-Vs are found throughout the high-latitude region. They appear predominantly in the late evening sector and are absent from the late pre-noon MLT sector.

The pitch angle distribution of electrons measured over an inverted-V event is given in Figure 8-30. This spectrogram gives the ratio of field-aligned flux at a given energy to the flux at pitch angles of 60°. The heavy black line gives the energy at which the maximum differential flux was measured. Note that the energy of the peak flux rises to ~3 keV at 1008:05 UT, then decreases to ~0.5 keV 20 seconds later. At the energy of the peak flux the electrons are field aligned whereas those with higher energies are isotropic over the downcoming hemisphere. This distribution is consistent with the electrons having been accelerated through a field-aligned potential drop at an altitude above the point of observation [Evans, 1974]. Electrons with energies less than that of the peak show highly complex pitch angle distributions. Some of these electrons are secondaries and degraded primaries trapped between a magnetic mirror point and an electrostatic potential barrier. Some of the low-energy electrons are highly field-aligned suggesting that they are accelerated electrons of ionospheric origin. If so, the complexity of their distribution suggests that the field-aligned potential drop is varying temporally and/or spatially.

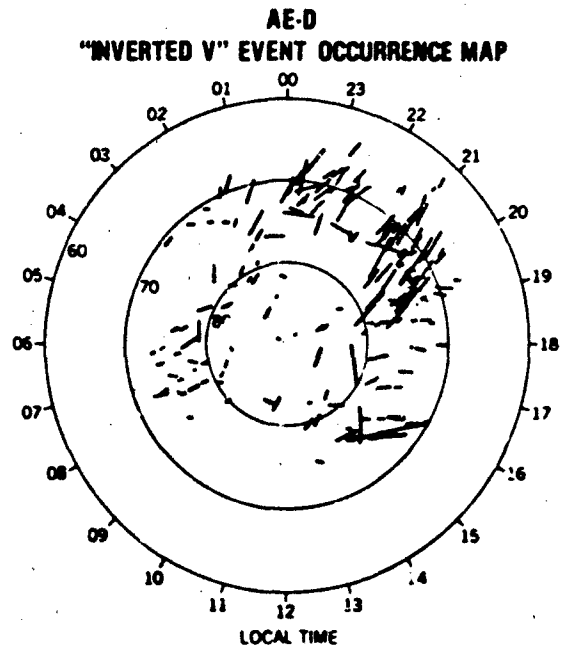


Figure 8-29. Spatial occurrence map of 280 inverted-V events. Local time and circles of constant invariant latitude are shown as reference [Lin and Hoffman, 1979].

One of the most exciting developments of the 1970s was the development of direct evidence for the existence of field-aligned potential drops from measurements by instrumentation on the S3-3 satellite. The observational work of the University of California, Berkeley group is summarized by Mozer et al. [1980].

The highly eccentric orbit of S3-3 carried it to altitudes of ~ 8000 km above the auroral ionosphere. This is a hitherto unexplored region. At these altitudes very intense electric field structures were observed. An example given in Figure 8-31 shows electric fields at 8000 km reaching 400 mV/m and undergoing rapid reversals. If mapped, assuming $\mathbf{E} \cdot \mathbf{B} = 0$, to auroral arc altitudes the electric field intensities would be of the order of 1000 mV/m and the latitudinal width of the structure would be <10 km. This is a

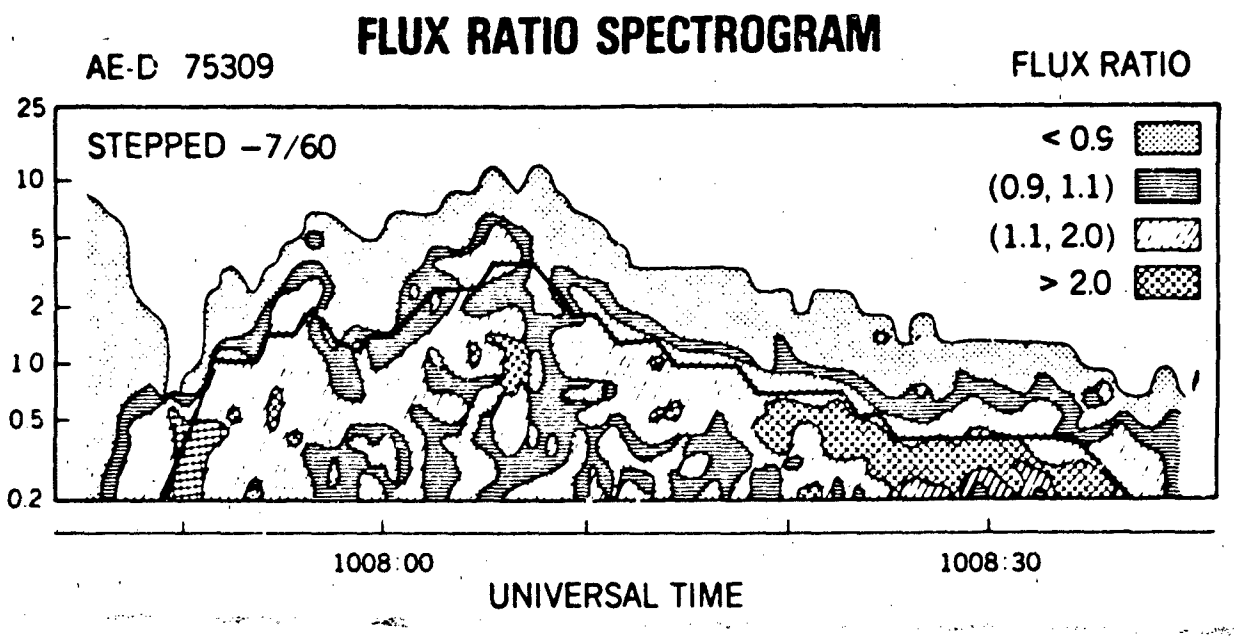


Figure 8-30. Spectrogram displaying the directional differential flux ratios (−7°) and (60°) over an inverted-V event. The flux ratios are separated into four categories as shown in the upper right corner [Lin and Hoffman, 1979].

typical width of a discrete optical arc. However, the electric field intensity greatly exceeds that measured in the lower ionosphere. For this reason it is argued that there must be a potential drop along magnetic field lines that accelerate plasma sheet electrons to form arcs.

These electric field structures have been observed at all altitudes down to 1000 km. They are mostly found above 4000 km. Comparisons of S3-3 electric field measurements with simultaneous measurements from other instruments on the spacecraft show that the electric field structures are wave turbulence that are collocated with inverted-V precipitation structures. Within the electric field reversals we find fluxes of field-aligned electrons going into, and H^+ and O^+ ions going out of the ionosphere. The upgoing ions are decided according to pitch angle distributions into "beams" and "conics." Ion beams are highly directed along magnetic field lines. They have only been observed at altitudes greater than 4000 km and in the region of discrete rather than diffuse aurora. They are thought to be accelerated upward by the

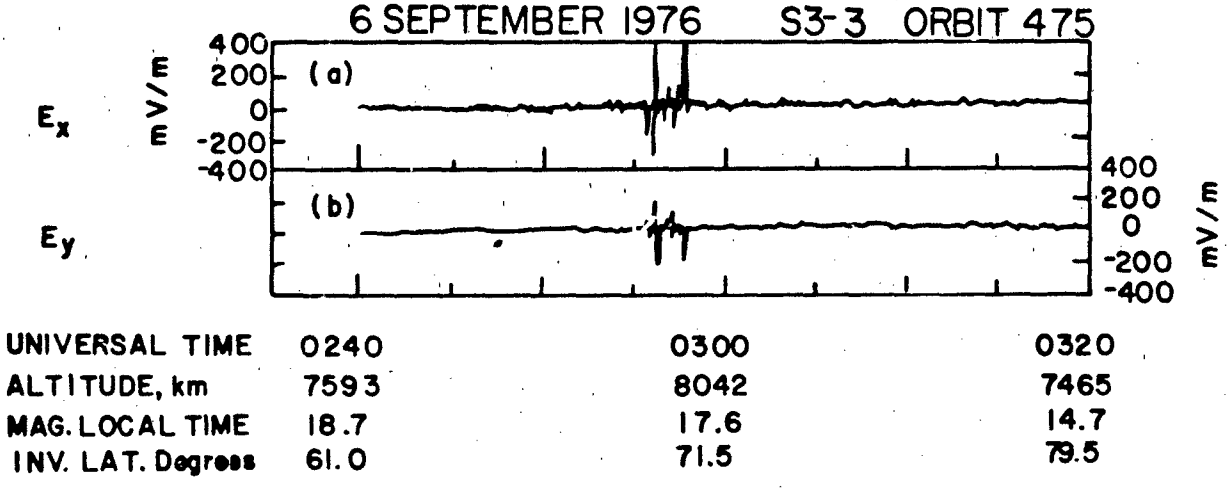


Figure 8-31. An example of a rapid electric field reversal observed at 8000 km by S3-3 [Mozer et al., 1980].

same parallel electric fields that accelerate plasma sheet electrons downward to form discrete auroral arcs. Ion conics have pitch angle distributions that peak at some intermediate angles between 90° and 0°. They have been detected at all latitudes in the auroral oval and at all altitudes greater than 1000 km. They are thought to be accelerated through resonant interactions with upper hybrid and ion cyclotron waves [Gorney et al., 1981].

8.6.4 Substorms

No overview of high-latitude electrodynamics would be complete without some comment, no matter how cursory, on substorms. It is during substorms that the dynamic coupling between the magnetosphere and ionosphere is most striking. Despite intensive studies of substorm processes over the last fifteen years, the richness of the observations has made total agreement on what constitutes the essential elements of a substorm elusive.

The evolution of discrete auroral arcs in the late evening sector as synopsized from all-sky camera data by Akasofu [1964] is given in Figure 8-32. Under pre-substorm conditions, homogeneous arcs extending for thousands of kilometers in the east-west direction are found in the oval and sun-aligned arcs are found in the polar cap. Substorm onset is announced by a brightening of the most equatorward arc. The discrete arcs expand poleward and westward traveling surges develops in the evening sector. After expanding to some maximum latitude, the arcs slowly retreat toward their pre-substorm condition.

To resolve differences between various schools of thought, nine active investigators met in August 1978 to specify substorm signatures and unambiguously define words commonly (often differently) used to describe substorm processes [Rostoker et al., 1980]. Major points of agreement were

1. During extended periods of northward IMF, the magnetosphere quiets and asymptotically approaches a ground state.
2. As the IMF turns southward, magnetospheric convection increases. This enhanced convection can exist for some time prior to substorm onset.
3. Substorm onset is signaled by an explosive increase in luminosity of the most equatorward arc, an intensification of the auroral electrojet, and a burst of Pi2 micropulsation. The burst of micropulsations testifies to the explosive nature of the onset process in the magnetospheric source region.
4. The expansion phase occurs from onset to the time when the midnight sector arcs have undergone their most poleward excursions. Note that in this definition, multiple intensifications of the substorm process, each marked by a micropulsation burst, are allowed.
5. The recovery phase coincides with the period in which midnight sector arcs retreat equatorward.

The substorm signatures dealt with in the August 1978 meeting are observable from ground-based instrumentation. Except for the micropulsation bursts, the signatures are of ionospheric effects whose causes lie in the magnetosphere. A key observation for understanding how substorms may be triggered comes from numerous satellites at geostationary altitude. During pre-substorm periods the nightside magnetic field at 6.7 R_E takes on a tail-like configuration. At substorm onset the magnetic field rapidly recovers its normal, nearly dipolar configuration. Simultaneously hot plasma sheet particles, with no dispersion, are injected.

To explain the observations the following scenario has been pieced together.

1. At a southward turning of the IMF, magnetic flux is transferred from the day to the night side of the magnetosphere. This process proceeds for about half an hour in which potential energy, in the form of stored magnetic flux, builds up in the tail. During this period the neutral sheet current moves earthward to ~10 R_E leading to a tail like field geometry at 6.7 R_E. Recall that discrete arcs map to the boundary plasma sheet rather than the central plasma sheet. Thus, the most equatorward arc maps to a region near the inner edge of the neutral sheet current.
2. At substorm onset, the neutral sheet current near the inner edge of the boundary plasma sheet is diverted via field-aligned currents through the ionosphere. This leads to a collapse of the inner portion of the tail. In the ionosphere, part of the energy released in the collapse of the tail appears as an explosive brightening of the most equatorward arc. As the inner-tail field lines snap back to dipolar, plasma sheet electrons are rapidly accelerated by inductive electric fields and are injected to the vicinity of geostationary distance. The process continues while B_Z remains southward.
3. When the IMF turns northward the rate of flux transfer decreases abruptly. If the IMF maintains a northward component for considerable time, the potential energy stored in the tail is slowly dissipated and the magnetosphere relaxes toward a "ground state."

8.7 CONCLUSIONS

This chapter provides an introductory summary of what is known about the electrodynamics of the magnetosphere and the high latitude ionosphere prior to the launch of the Dynamics Explorer satellites. Our knowledge is an amalgam of theoretical models and *in situ* observations. In a qualitative sense, we are able to explain the gross features of magnetospheric processes. For example, a set of MHD equations has been used successfully to calculate the equilibrium shape of

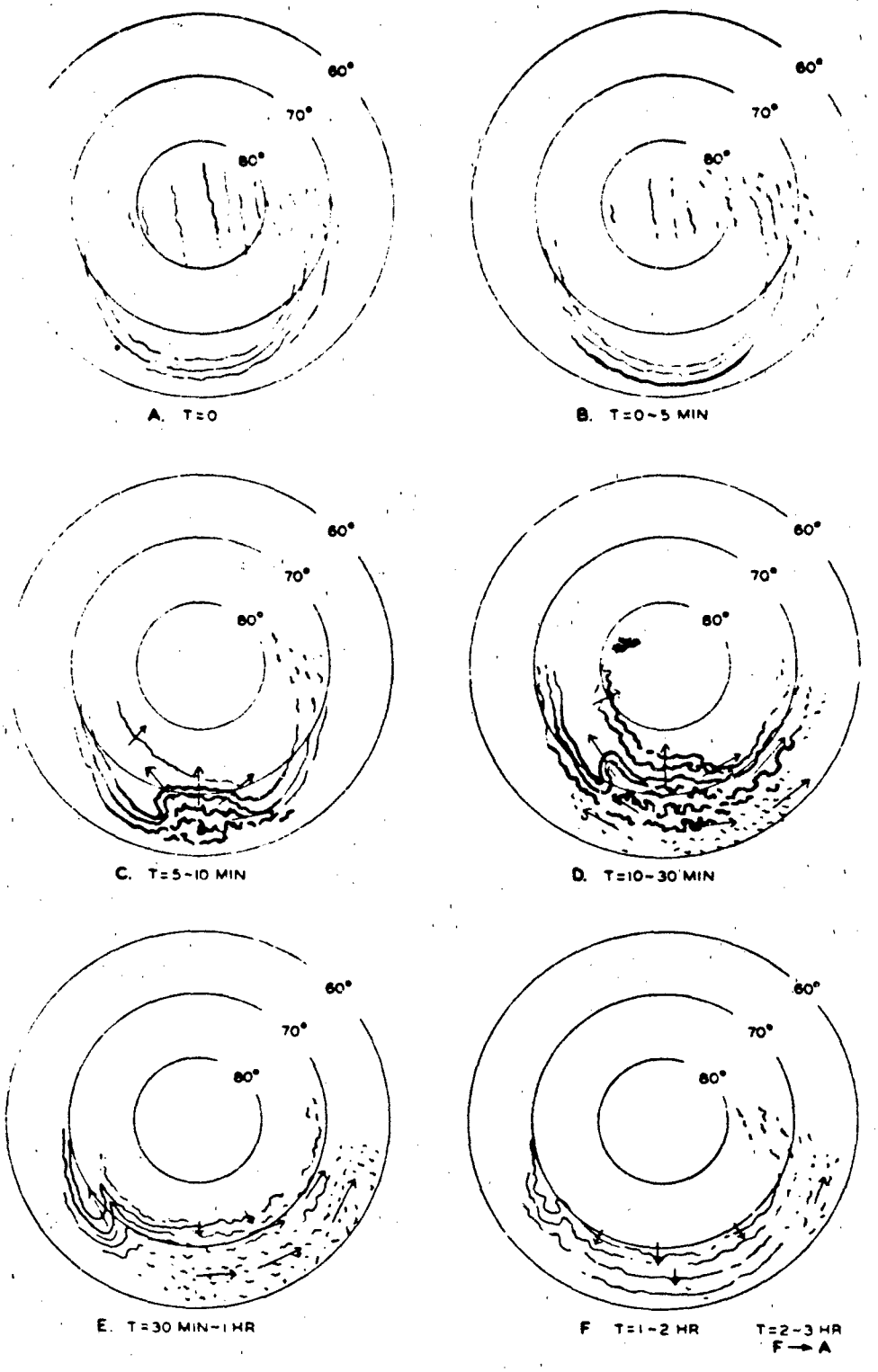


Figure 8-32. Schematic diagram to illustrate the development of the auroral substorm. The center of the concentric circles in each state is the north geomagnetic pole, and the sun is toward the top of the diagram [Akasofu, 1964]. (Reprinted with permission from Pergamon Press, Ltd. © 1964.)

the magnetopause [Mead and Beard, 1964]. However, the microphysics of how (if) magnetic merging occurs at the magnetopause and how particles cross the magnetopause to form the boundary layer are not yet understood. An essentially MHD model for coupling between the ionosphere and magnetosphere [Ashour-Abdalla and Thorne, 1978] has been applied to successfully simulate substorm effects in the inner plasma sheet and low latitude portion of the auroral oval [Lyons et al., 1979; Harel et al., 1981]. Thus, adiabatic energization is an important and well understood magnetospheric process. Other energization processes associated with reconnection in the magnetotail are not understood. The role of parallel electric fields for the formation of auroral arcs was experimentally established in the 1970s. Competing theoretical models of how parallel electric fields form in a collisionless plasma fill the literature.

With the gift of hindsight, it is interesting to examine the work of cartographers from the mid-seventeenth century. The mixed fruits of explorations spanning two hundred years are evident. Europe, the Middle East, Africa, and the eastern shores of the Americas are mapped with exquisite detail. Shapes assigned to the American west coasts and to the nations of East Asia range from vague to fanciful. Our own position twenty years into the age of satellite exploration is analogous to that of seventeenth-century map-makers. Some parts of the magnetosphere-ionosphere system are well explored. Our understanding of other important parts is based on fragmentary data samples and leaves much to be desired. Undoubtedly, readers a few decades hence will find many of our ideas as quaintly amusing as the seventeenth-century image of Japan. To those who view cartography or space physics as quasi-static processes, this is a scandalous situation. To those involved in the process, it crystalizes the excitement of exploration.

CHAPTER 8

REFERENCES

Ackerson K.L. and L.A. Frank, "Correlated Satellite Measurements of Low-Energy Electron Precipitation and Ground-Based Observations of a Visible Auroral Arc," *J. Geophys. Res.*, **77:** 1128, 1972.

Akasofu, S.-I., "The Development of the Auroral Substorm," *Planet Space Sci.*, **12:** 273, 1964.

Akasofu, S.-I., "Interplanetary Energy Flux Associated with Magnetospheric Substorms," *Planet Sp. Sci.*, **27:** 425, 1978.

Armstrong, J. and A.J. Zmuda, "Triaxial Magnetic Measurements of Field Aligned Currents at 800 km in the Auroral Region: Initial Results," *J. Geophys. Res.*, **78:** 6802, 1973.

Ashour-Abdalla M. and R.M. Thorne, "Toward a Unified View of Diffuse Auroral Precipitation," *J. Geophys. Res.*, **83:** 4755, 1978.

Axford W.I. and C.O. Hines, "A Unifying Theory of High-Latitude Geophysical Phenomena and Geomagnetic Storms," *Can. J. Phys.*, **39:**1433, 1961.

Burch, J.T., S.A. Fields and R.A. Heelis, "Polar Cap Electron Acceleration Regions," *J. Geophys. Res.*, **84:** 5863, 1979.

Burke, W.J., M.C. Kelley, R.C. Sagalyn, M. Smiddy, and S.T. Lai, "Polar Cap Electric Field Structure with a Northward Interplanetary Magnetic Field," *Geophys. Res. Lett.*, **6:** 21, 1979.

Burke, W.J., M.S. Gussenhoven, M.C. Kelley, D.A. Hardy, and F.J. Rich, "Electric and Magnetic Field Characteristics of Discrete Arcs in the Polar Cap," *J. Geophys. Res.*, **87:** 2431–2443, 1982.

Crooker, N.U., "Dayside Merging and Cusp Geometry," *J. Geophys. Res.*, **82:** 3629, 1977.

Crooker, N.U., "The Magnetospheric Boundary Layers: A Geometrically Explicit Model," *J. Geophys. Res.*, **84:** 951, 1979.

Doyle, M.A., F.J. Rich, W.J. Burke and M. Smiddy, "Field Aligned Currents and Electric Fields in the Region of the Dayside Cusp," *J. Geophys. Res.*, **86:** 5656, 1981.

Dungey, J.W. "Interplanetary Magnetic Field and the Auroral Zones," *Phys. Rev. Lett.*, **6:** 47, 1961.

Eastman T.E. and E.W. Hones Jr., "Characteristics of the Magnetospheric Boundary Layer and Magnetopause Layer as Observed by IMP 6," *J. Geophys. Res.*, **84:** 2019, 1979.

Eastman, T.E., E.W. Hones J- S.J. Bame, and J.R. Asbridge, "The Magnetospheric Boundary Layer: Site of Plasma Momentum and Energy Tranfer from the Magnetosheath into the Magnetosphere," *Geophys. Res. Lett.*, **3:** 685, 1976.

Eather, R.E., "DMSP Calibration," *J. Geophys. Res.*, **84:** 4134, 1979.

Ejiri, M., R.A. Hoffman, and P.H. Smith, "The Convection Electric Field Model for the Magnetosphere Based on Explorer 45 Observations," *J. Geophys Res.*, **83:** 4811, 1978.

Evans, D.S., "Precipitating Electron fluxes formed by a magnetic field-aligned potential difference," *J. Geophys. Res.*, **79:** 2853, 1974.

Fahleson, U.V., "Theory of Electric Field Measurements Conducted in the Magnetosphere with Electric Probes," *Space Sci Rev.*, **7:** 238, 1967.

Fairfield, D.H. and G.D. Mead, "Magnetospheric Mapping with a Quantitative Geomagnetic Field Model," *J. Geophys. Res.*, **80:** 535, 1975.

Foster, J.C. and J.R. Burrows, "Electron Fluxes Over the Polar Cap 1, Intense keV Fluxes During Poststorm Quieting," *J. Geophys. Res.*, **81:** 6016, 1976.

Foster, J.C. and J.R. Burrows, "Electron Fluxes Over the Polar Cap 2, Electron Trapping and Energization on Open Field Lines," *J. Geophys. Res.*, **82:** 5165, 1977.

Frank, L.A., "Initial Observations of Low-Energy Electrons in the Earth's Magnetosphere with OGO3," *J. Geophys. Res.*, **72:** 185, 1967.

Frank, L.A., "Relationship of the Plasma Sheet, Ring Current, Trapping Boundary, and Plasmapause Near the Magnetic Equator and Local Midnight," *J. Geophys. Res.*, **76:** 2265, 1971.

Frank L.A. and K.L. Ackerson, "Observations of Charged Particle Precipitation into the Auroral Zone," *J. Geophys. Res.*, **76:** 3612, 1971.

Frank, L.A., K.L. Ackerson and L.P. Lepping, "On Hot Tenuous Plasmas, Fireballs, and Boundary Layers in the Earth's Magnetotail," *J. Geophys. Res.*, **81:** 5859, 1976.

Frank, L.A., J.D. Cravens, J.L. Burch, and J.D. Winningham, "Polar View of Earth's Aurora with Dynamics Explorer," *Geophys. Res. Lett.*, **9:** 1001, 1982.

Freeman, J.W., "Internal Magnetospheric Plasma Flow," in *International Sydney Chapman Conference on Magnetospheric Boundary Layers, Alpbach, Austria, 1979* (ESA-SP-148), European Space Agency, Paris, 1979.

Friis-Christensen, E., K. Lassen, J. Willjelm, J.M. Wilcox, W. Gonzalez, and D.S. Colburn, "Critical Component of the Interplanetary Magnetic Field Responsible for Large Geomagnetic Events," *J. Geophys. Res.*, **77:** 3371, 1972.

Gold, T., "Motions in the Magnetosphere of the Earth," *J. Geophys. Res.* **64:** 1219, 1959.

Gorney, D.T., A. Clarke, D. Croley, J. Fennell, J. Luhmann, and T. Mizera, "The Distribution of Ion Beams and Conics Below 8000 km," *J. Geophys. Res.*, **86:** 83, 1981.

Gringanz, K.I., V.V. Bezrukikh, V.D. Ozerov and E.E. Rybchinskii, "A Study of the Interplanetary Ionized Gas, High Energy Electrons, and Corpuscular Radiation from the Sun by Means of the Three-Electrode Trap for Charged Particles on the Second Soviet Cosmic Rocket," *Sov. Phys. Dokl.* **5:** 361, 1961.

Gurnett, D.A. and L.A. Frank, "A Region of Intense Plasma Wave Turbulence on Auroral Field Lines," *J. Geophys. Res.*, **82:** 1031, 1977.

Gussenhoven, M.S., D.A. Hardy and W.J. Burke, "DMSP/F2 Electron Observations of Equatorward Auroral Boundaries and Their Relationship to Magnetospheric Electric Fields," *J. Geophys. Res.* **86:** 768, 1981.

Haerendel, G., G. Paschmann, N. Sckopke, H. Rosenbauer and P.C. Hedgecock, "The Frontside Boundary Layer of the Magnetosphere and the Problem Reconnection," *J. Geophys. Res.* **83:** 3195, 1978.

Hanson, W.B. and R.A. Heelis, "Techniques for Measuring Bulk Gas Motions from Satellites," *Space Sci Inst.*, **1:** 493, 1975.

Hardy, D.A., H.K. Hills, and J.W. Freeman, "A New Plasma Regime in the Distant Geomagnetic Tail," *Geophys. Res. Lett.* **2:** 169, 1975.

Hardy, D.A., M.S. Gussenhoven, E. Holman, W.J. Burke, and N. Heinemann, "DMSP/F2 Electron Observations of Equatorial Auroral Boundaries and Their Relationship to the Solar Wind Velocity and the North-South Component of the Interplanetary Magnetic Field," *J. Geophys. Res.*, **86:** 9961–9974, 1981.

Hardy, D.A., W.J. Burke, and M.S. Gussenhoven, "DMSP Optical and Electron Measurements in the Vicinity of Polar Cap Arcs," *J. Geophys. Res.*, **87:** 2413–2430, 1982.

Harel, M., R.A. Wolf, P.N. Reiff, R.W. Spiro, W.J. Burke, F.J. Rich, and M. Smiddy, "Quantitative Simulation of a Magnetospheric Substorm, 1, Model Logic and Overview, and 2, Comparison with Observations," *J. Geophys. Res.*, **86,** 2217, 1981.

Heppner, J.P., "Polar Cap Electric Field Distributions Related to the Interplanetary Magnetic Field Direction," *J. Geophys. Res.*, **77:** 4877, 1972.

Heppner, J.P., "Empirical Models of High-Latitude Electric Fields," *J. Geophys. Res.*, **82:** 1115, 1977.

Hill, T.W., "Generation of the Magnetospheric Electric Field," in *Quantitative Modelling of Magnetospheric Processes, AGU Monograph 21*, edited by W.P. Olson, American Geophysical Union, Washington D.C., p. 297.

Hines, C.O., "The Energization of Plasma in the Magnetosphere: Hydromagnetic and Particle Drift Approaches," *Planet. Space Sci.* **10:** 239, 1963.

Hones, Jr., E.W., J.R. Asbridge, S.J. Bame, M.D. Montgomery, and S. Singer, "Substorm variations of the magnetotail plasma sheet from $X_{SM} \approx -6\ R_E$ to $X_{SM} \approx -60\ R_E$," *J. Geophys Res.*, **78:** 1109, 1973.

Iijima, T. and T.A. Potemra, "Large-scale characteristics of field aligned currents associated with substorms," *J. Geophys. Res.*, **83:** 599, 1978.

Ismail, S., D.D. Wallis, and L.L. Cogger, "Characteristics of Polar Sun-Aligned Arcs," *J. Geophys. Res.*, **82:** 4741, 1977.

Johnson, F.S., "The Gross Character of the Geomagnetic Field in the Solar Wind," *J. Geophys. Res.*, **65:** 3049, 1960.

Kennel C. F. and H.E. Petschek, "Limit on Stably Trapped Particle Fluxes," *J. Geophys. Res.*, **71:** 1, 1966.

Kivelson, M.G., S.M. Kaye and D.J. Southwood, "The Physics of Plasma Injection Events," in *Dynamics of the Magnetosphere*, edited by S.-I. Akasofu, 385, D. Reidel, Dordrecht, Holland, 1979.

Klumpar, D.M., "Relationships Between Auroral Particle Distributions and Magnetic Field Perturbations Associated with Field Aligned Currents," *J. Geophys. Res.*, **84:** 6524, 1979.

Lemaire, J. and M. Roth, "Penetration of solar wind plasma elements into the magnetosphere," *J. Atmos. Ter. Phys.*, **40:** 331, 1978.

Lin, C.S. and R.A. Hoffman, "Characteristics of the Inverted-V Event," *J. Geophys. Res.*, **84:** 1514, 1979.

Lyons, L.R., "Generation of Large-Scale Regions of Auroral Currents, Electric Potentials, and Precipitation by the Divergence of the Convection Electric Field," *J. Geophys. Res.*, **85:** 17, 1980.

Lyons, L.R., "Discrete Aurora as the Direct Result of an Inferred High-Altitude Generating Potential Distribution," *J. Geophys. Res.*, **86:** 1, 1981.

Lyons, L.R., D.S. Evans, and R. Lundin, "An Observed Relationship Between Magnetic Field Aligned Electric Fields and Downward Electron Energy Fluxes in the Vicinity of Auroral Forms," *J. Geophys. Res.*, **84:** 457, 1979.

Maezawa, K., "Magnetospheric Convection Induced by the Positive and Negative Z Components of the Interplanetary Magnetic Field: Quantitative Analysis Using Polar Cap Magnetic Records," *J. Geophys. Res.*, **81:** 2289, 1976.

Maggs, J.E. and T.N. Davis, "Measurement of the Thickness of Auroral Structures," *Planet. Space Sci.*, **16:** 205, 1968.

Mauk, B.H. and C.E. McIlwain, "Correlations of K_p with Substorm Injected Plasma Boundary," *J. Geophys. Res.*, **79:** 3193, 1974.

Maynard, N.C., "Electric Field Measurements Across the Harang Discontinuity," *J. Geophys. Res.*, **79:** 4620, 1974.

McDiarmid, I.B., J.R. Burrows, and M.D. Wilson, "Large Scale Magnetic Field Perturbations and Particle Measurements at 1400 km on the Dayside," *J. Geophys. Res.*, **84:** 1431, 1979.

McIlwain, C.E., "Plasma Convection in the Vicinity of Geosynchronous Orbit" in *Earth's Magnetospheric Processes*, edited by B.M. McCormac, 268, D. Reidel, Dordrecht, Holland, 1972.

Mead, G.D. and D.B. Beard, "Shape of the Geomagnetic Field Solar Wind Boundary," *J. Geophys. Res.*, **69:** 1169, 1964.

Meng, C.-I. and H.W. Kroel, "Intense Uniform Precipitation of Low-Energy Electrons Over the Polar Cap," *J. Geophys. Res.*, **83:** 2305, 1977.

Meng, C.-I., S.-I. Akasofu and K.A. Anderson, "Dawn-Dusk Gradient of the Precipitation of Low-Energy Electrons Over the Polar Cap and its Relation to the Interplanetary Magnetic Field," *J. Geophys. Res.*, **82:** 5271, 1977.

Mozer, F.S., R.B. Torbert, U.V. Fahleson, C.G. Falthammer, A. Gonfalone, A. Pedersen, and C.T. Russell, "Direct Observation of a Targential Electric Field Component at the Magnetopause," *Geophys. Res. Lett.*, **6:** 305, 1979.

Mozer, F.S., C.A. Cattell, M.K. Hudson, R.L. Lysak, M. Temerin, and R.B. Torbert, "Satellite Measurements and Theories of Low Altitude Auroral Particle Acceleration," *Space Sci. Rev.*, **27:** 155, 1980.

Northrop, T.G., *The Adiabatic Motion of Charged Paticles*, Chap. 3, Wiley, New York, 1963.

Olson, W.P., "The Shape of the Tilted Magnetopause," *J. Geophys. Res.*, **74:** 5642, 1969.

Pilipp, W.G. and G. Morfill, "The formation of the plasma sheet resulting from plasma mantle dynamics," *J. Geophys. Res.*, **83:** 5670, 1978.

CHAPTER 8

Piddington, J.H., "Theories of the Geomagnetic Storm Main Phase," *Planet Space Sci.*, **11:** 1277, 1963.

Podgorny, I.M., E.M. Dubinin, and Yu. N. Potanin, "The Magnetic Field on the Magnetospheric Boundary from Laboratory Simulation Data," *Geophys. Res. Lett.*, **4:** 207, 1978.

Reiff, P.H., R.W. Spiro, and T.W. Hill, "Dependence of Polar-Cap Potential Drop on Interplanetary Parameters," *J. Geophys. Res.*, **86:** 7639, 1981.

Rich, F.J., D.L. Reasoner, and W.J. Burke, "Plasma Sheet at Lunar Distance: Characteristics and Interactions with the Lunar Surface," *J. Geophys. Res.* **78:** 8097, 1973.

Rosenbauer, H., H. Grunwalt, M.D. Montgomery, G. Paschmann, and N. Sckopke, "HEOS 2 Plasma Observations in the Distant Polar Magnetosphere: The Plasma Mantle," *J. Geophys. Res.*, **80:** 2723, 1975.

Rossi, B. and S. Olbert, *Introduction to the Physics of Space*, Chapter 5, McGraw-Hill, New York, 1970.

Rostoker, G., "Magnetospheric and Ionospheric Currents in the Polar Cusp and Their Dependence on the B_y Component of the Interplanetary Magnetic Field," *J. Geophys. Res.*, **85:** 4167, 1980.

Rostoker, G., S.-I. Akasofu, J. Foster, R.A. Greenwald, Y. Kamide, K. Kawasaki, A T.Y. Lui, R.L. McPherron, and C.T. Russell, "Magnetospheric Substorms—Definitions and Signatures," *J. Geophys. Res.*, **85:** 1663, 1980.

Russell, C.T., "The Configuration of the Magnetosphere," in *Critical Problems of Magnetospheric Physics*, edited by E.R. Dyer, 1, National Academy of Science, Washington, D.C., 1972.

Russell, C.T. and R.C. Elphic, "ISEE Observations of Flux Transfer Events at the Dayside Magnetopause," *Geophys. Res. Lett.*, **6:** 33, 1979.

Sckopke, N., G. Paschmann, H. Rosenbauer, and D.H. Fairfield, "Influence of the Interplanetary Magnetic Field Oneral and Thickness of the Plasma Mantle," *J. Geophys. Res.*, **81:** 2687, 1976.

Sckopke, N., G. Paschmann, G. Haerendel, B.U.O. Sonnerup, S.J. Bame, T.G. Forbes, E.W. Hones, Jr., and C.T. Russell, "Structure of the Low-Latitude Boundary Layer," *J. Geophys. Res.*, **86:** 2099, 1981.

Smiddy, M., et al., "Effects of High Latitude Conductivity on Observed Convection Electric Fields and Birkeland Currents," *J. Geophys. Res.*, **85:** 6811, 1980.

Southwood, D.J. and S.M. Kaye, "Drift Boundary Approximations in Simple Magnetospheric Convection Models," *J. Geophys. Res.*, **84:** 5773, 1979.

Speiser, T.W., "Particle Trajectories in Model Current Sheets, II: Application to Auroras Using a Geomagnetic Tail Model," *J. Geophys. Res.* **72:** 3919, 1967.

TerHaar, D., *Elements of Hamiltonian Mechanics*, Chapter 6, North Holland, Amsterdam, 1964.

Vasyliunas, V. M., "A Survey of Low-Energy Electrons in the Evening Sector of the Magnetosphere with OGO 1 and OGO 3," *J. Geophys. Res.*, **73:** 2839, 1968.

Vasyliunas, V.M., "Mathematical Models of Magnetospheric Convection and its Coupling to the Ionosphere," in *Particles and Fields in the Magnetosphere*, edited by B.M. McCormac, D. Reidel, Dordrecht, Holland, 1970.

Vasyliunas, V.M., *Magnetospheric Plasma, in Solar Terrestrial Physics*, edited by E. Dyer, D. Reidel, Dordrecht, Holland, p. 192, 1972.

Vasyliunas, V.M., "Interaction Between the Magnetospheric Boundary Layers and the Ionosphere," *International Sydney Chapman Conference on Magnetospheric Boundary Layers, Alpbach, Austria, 1979* (ESA-SP-148), European Space Agency, Paris, 1979.

Weber, E.J. and J. Buchau, "Polar Cap F-Layer Auroras," *Geophys. Res. Lett.*, **8:** 125, 1981.

Williams, D.J. and L.R. Lyons, "The Proton Ring Current and Its Interaction with the Plasmapause: Storm Recovery Phase," *J. Geophys. Res.*, **79:** 4195, 1974.

Winningham, J.D. and W.J. Heikkila, "Polar Cap Auroral Electron Fluxes Observed with Isis 1," *J. Geophys. Res.*, **79:** 949, 1974.

Yaeger, D.M. and L.A. Frank, "Low Energy Electron Intensities at Large Distances Over the Earth's Polar Cap," *J. Geophys. Res.*, **81:** 3966, 1976.

Young, T.S.T., J.D. Callen and J.E. McCune, "High Frequency Electrostatic Waves in the Magnetosphere," *J. Geophys. Res.*, **78:** 1082, 1973.

Chapter 9

IONOSPHERIC PHYSICS

Section 9.1 F.J. Rich
Section 9.2 F.J. Rich
Section 9.3 Su. Basu

9.1 STRUCTURE OF THE IONOSPHERE

9.1.1 Ionospheric Layers

The ionized atmosphere of the earth is composed of a series of overlapping layers. In each layer there is an altitude of maximum density, above and below which the ionization density tends to drop off. The total ionization profile with the layers indicated is shown in Figure 9-1. Characteristics of the layers are shown in Table 9-1. This profile is valid only for midlatitudes. In the equatorial region, the profile is distorted by the geomagnetic field, and in the polar region, the profile is distorted by ionization by energetic particles, magnetospheric coupling, and other effects.

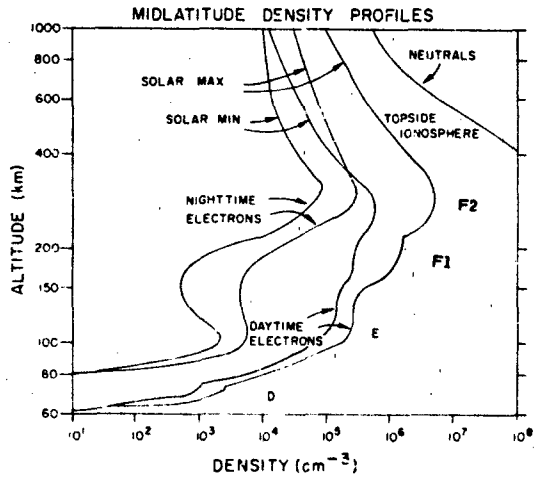


Figure 9-1. Total ionization profile with ionospheric layers.

The *D region* is present only during daylight hours. The altitude of the peak density is normally around 90 km, but this may decrease considerably to ~ 78 km when the solar x-ray flux is enhanced. The *E region* peak density occurs at a peak altitude of 110 km. At sunset, the E Region electron density drops by a factor of 10 or more in a short period (tens of minutes) before reaching a nighttime equilibrium density. At night, the region of low density near 150 km between the E region and the F region may have a sharply lower density than shown in Figure 9-1 or the density may be great enough so that there is no depletion region depending on geophysical conditions. The *F region* is a combination of two somewhat different regions. The *F1 region* has an altitude peak near 200 km, but is absent at night. The *F2 region* has a peak near 300 km during the day and at higher altitudes at night. Shortly after sunset, the absolute density near the peak of the F region often increases due to plasma transport processes before decreasing to a night time value.

The *Topside Ionosphere* is the name given to the rest of the ionosphere above the F region peak. In a simple model of the ionosphere, the density of the topside ionosphere decreases exponentially with height with some char-

Table 9-1. Layers of daytime midlatitude ionosphere.

Layer	Altitude(km)	Major Component	Production Cause
D	70–90 km	NO^+, O_2^+	Lyman Alpha, x-rays
E	95–140 km	O_2^+, NO^+	Lyman Beta, Soft x-rays, UV Continuum
F1	140–200 km	O^+, NO^+	He II, UV Continuum (100–800Å)
F2	200–400 km	O^+, N^+	He II, UV Continuum (100–800Å)
Topside F	> 400 km	O^+	Transport from Below
Plasmasphere	> 1200 km	H^+	Transport from Below

acteristic scale height until the ionization density is below detectable levels. In the mid and low latitude ionosphere, the geomagnetic field tends to trap ions, especially hydrogen ions, that would otherwise drift off into deep space. Thus, somewhere between 800 and 2000 km altitude the scale height increases to a very large value (≥ one earth radius). If one follows a geomagnetic field line out to the equatorial plane and back into the conjugate ionosphere, the density would change by less than 10^2 and in some cases less than 10. This region of trapped ionospheric ions is the *Plasmasphere* [Carpenter and Park, 1973]. The outer edge of the plasmasphere where ionospheric ions are not trapped is the *Plasmapause*. The plasmapause is located approximately along the geomagnetic field line that maps down to 60° magnetic latitude. A representation of the topside ionosphere and plasmasphere is shown in Figure 9-2.

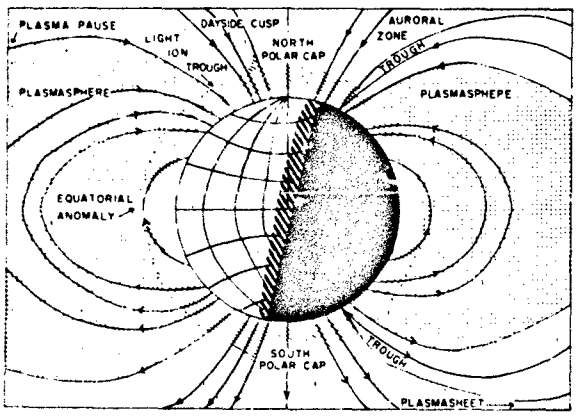


Figure 9-2. Representation of topside ionosphere and plasmasphere

For further information about ionospheric layers see Banks and Kockarts [1973]; Kohnlein [1978]; and Chapters 12 and 21.

9.1.2 Chapman Theory for Ionospheric Layers

The layering of the ionosphere was first discussed by Sydney Chapman in the 1920s. Ionizing photons from the sun will produce more and more ions as they penetrate deeper and deeper into an atmosphere with rapidly increasing density. As photoionization occurs, the flux of photons is attenuated until a depth is reached that photoionization production drops. Thus, a layer of ionization near the altitude of maximum production is created.

The process of quantitatively determining the height profile of ionization is outlined as follows. We assume that direct photoionization and recombination are the only sources of production and loss respectively. We can find the ion density at a height h due to a single frequency of ionizing photons ν by setting the production rate equal to the loss rate for a quasi-equilibrium ionosphere. The production rate is

$$Q_\nu(h) = \beta_\nu N(h) F_\nu(h) \text{ ionizations cm}^{-3}\text{ s}^{-1} \quad (9.1)$$

where β_ν is the absorption cross section for photoionization, N(h) is the neutral density at the height h, and $F_\nu(h)$ is the photon flux of frequency ν. The recombination rate is approximately

$$L_\nu(h) = \alpha_R N_i N_e \quad (9.2)$$

where α_R is the recombination coefficient and N_i and N_e are the ion and electron densities. If multiply charged and negative ions are not important, $N_i = N_e$. For an appropriate set of assumptions, the density near the peak of the layer is

$$N_e(h) \approx \left(\frac{Q^M}{\alpha_R}\right)^{1/2}\left[1 - \frac{(h-h_m)^2}{4H^2}\right] \quad (9.3)$$

where Q_M is the production at the peak of the layer, h_m is the height of the layer and H is the scale height of the neutral atmosphere. The parabolic variation of density with h around the peak gives the layers of the ionosphere their characteristic shape.

In principle, the above equation needs to be summed over all appropriate frequencies of incident photons to obtain the total ionospheric profile. In fact, many details about ionospheric chemistry (especially in the D and E layers) and about plasma transport (especially in the F layer) must be considered. For example, NO is not a significant component of the neutral atmosphere, but NO^+ is a significant ion component due to a chain of chemical reactions induced by photoionization. Details about computing an ionospheric model with all factors included is given by Banks and Kockarts [1973].

9.1.3 Ionization Production, Loss and Vertical Transport

Ionization production is generally a two-step process. The first step is the creation of ions from the neutrals by solar photons in the ultraviolet and x-ray spectrum, and to a lesser degree by collisions with energetic particles. Most of the ionization is produced by solar radiation with wavelengths less than 1026 Å, which ionizes O, O_2, and N_2. There are a vast array of minor constituents, especially molecular and metallic ions, that are important in understanding the ionosphere. Some of these minor constituents play important roles in the absorption of solar UV radiation (especially in the D region), in the production of airglow, in the chemistry of the ionosphere, and in the role of tracers to indicate ionospheric and atmospheric transport. A review

of photoionization is given by Hudson [1971] and Stolarski and Johnson [1972].

The second step in the creation of the ionosphere is the reaction between ions, neutrals, and electrons to create different forms of ionization than that created by direct ionization. Ionospheric chemistry explains why NO^+ is an important ionospheric ion despite the low abundance of NO in the neutral atmosphere. Two of the major reactions in the ionosphere which create NO^+ are

$$N_2 + O^+ \rightarrow NO^+ + N$$
$$N_2^+ + O \rightarrow NO^+ + N \qquad (9.4)$$

All of the major reactions are summarized in Figure 9-1 of Torr [1979]. The reactions that result in minor constituents are quite numerous and not all of them are well understood. Reactions that should be mentioned involve metastable atomic states, negative ions, ionization by photoelectrons, energetic neutrals and vibrational states of molecules. Further details of ionospheric chemistry are given in Chapter 21.

After recombination of ions and electrons to form neutrals, the atoms are often in an excited state that is metastable. On the ground, such states are de-excited through collisions, but with the lower collision frequency in the thermosphere, forbidden atomic transitions occur which release photons not generally seen in the laboratory such as the 5577 Å and 6300 Å emissions of atomic oxygen.

Negative ions are generally found in the D and lower E regions. They are formed primarily through electron-neutral molecule collisions. Photodissociation of electrons from negative ions provides the major source of D-region electrons shortly after sunrise. See Ferguson [1971] and Turco [1974] for a discussion of D-region chemistry.

After electrons are removed from an atom by photoionization, they have an energy that depends upon the work function of the atom and the energy of the photon. The energy distribution function of photoelectrons is a complex function [Jasperse, 1977 and Doering et al., 1976] that can crudely be approximated by a combination of 2 eV and a 20 eV Maxwellian distribution. In the D and E regions, the photoelectrons lose their energy through collisions close to the location of their formation. In the F region, the photoelectrons can travel significant vertical distances before losing all of their energy. As a rule of thumb, photoelectrons traveling upward at altitudes above 300 km are considered as escaping because of the low collision frequency. Except in the polar regions, these photoelectrons do not truly escape but follow magnetic field lines into the opposite hemisphere. Some of their energy is lost to plasma in the plasmasphere and the rest is used to heat and populate the opposite ionosphere [Mantas et al. 1978]. Photoelectrons from the opposite hemisphere are most important when one hemisphere is in darkness and the other in sunlight.

Charge exchange reactions such as

$$O^+ + H \leftrightarrows O + H^+ \qquad (9.5)$$

are important to the dynamics of the ionosphere, especially above the density peak of the F region. Energetic ions can undergo a charge exchange process, cross magnetic field lines as a neutral and be re-ionized by a second charge exchange. The result is a transfer of energy across field lines that otherwise would be impossible.

The principal loss of ionization comes from recombination. This is simply the reaction of positive ions and electrons to form neutrals. The most important recombination reactions are

$$NO^+ + e^- \rightarrow N + O$$
$$O_2^+ + e^- \rightarrow O + O. \qquad (9.6)$$

The charge exchange reaction

$$O^+ + O_2 \rightarrow O_2^+ + O \qquad (9.7)$$

is important to facilitate the recombination process. The major problem related to loss of ionization is not a lack of knowledge of the reactions but a lack of precise knowledge about the reaction rates. Reaction rates depend upon the density of the ions and neutrals, the atomic state of the incident and resultant ions, the vibrational state of molecules, and the temperature of the ions and electrons. If one reaction rate for any related part of the recombination process is not known precisely, it is difficult to determine the reaction rate of the complete recombination process. Solar eclipse data have been used to empirically determine some of the effective recombination rates. The knowledge of reaction rates for both production and loss covers a vast array of atomic and molecular reactions [Torr, 1979; Ferguson, 1971; and *DNA Rate Book*, 1972] known with varying degrees of precision.

Vertical transport of ions and electrons is principally governed by collision frequency and gravity. Below 130 km, the collision frequency of ions and electrons is so large that to a good approximation no ions or electrons enter or leave a unit volume. The unit volume may move up or down under the influence of pressure gradients. Between 130 km and 300 km the mean free path of electrons, and to a lesser degree ions, becomes comparable or larger than the scale height or layer thickness of the ionosphere. Diffusion becomes an important factor in plasma transport and energy exchange. Above an altitude of 300–400 km, the plasma can be treated as collisionless for many purposes. At high and midlatitudes, vertical transport is approximately along magnetic field lines and particles will tend to diffuse upward or downward according to their mass. Because the ions and electrons are electrically coupled, ambipolar diffusion must be considered. The ions and electrons diffuse upward together at a rate that must be less than a neutral particle with the mass of an electron and slightly greater than a neutral particle with the mass of the ion. Near the equator, vertical transport is impeded by the magnetic field lines. In the daytime, however, due to electrodynamic effects ionization

is transported upwards at the magnetic equator which subsequently diffuses down magnetic field lines at ± 15° magnetic latitudes. This results in the Appleton anomaly as discussed in greater detail in Section 9.2.2.4.

9.1.4 Neutral Winds and Horizontal Transport

Below 130 km altitude, the ion-neutral collision frequency is so high that ions will freely flow across field lines with the neutrals but the electrons are relatively fixed with respect to the magnetic field. As a result, neutral winds tend to cause ionospheric currents. These ionospheric currents are detected on the ground by the diurnal variations in the magnetic field at mid and low latitudes (+60° to −60°). The Sq (solar quiet) current system and the equatorial electrojet are the major current systems related to the neutral winds. The separation of ions and electrons in the E region produce polarization electric fields that can cause E × B drift of plasma in both the E- and F-regions [Evans, 1978].

Neutral winds are predominantly atmospheric tides with periods of 24, 12, 8, 6, . . . hours. The zero order neutral wind consists of a steady flow away from ~ 1400 local time toward ~ 0200. The latitude of these high and low pressure zones shifts seasonally with the sun. At high latitudes the neutral winds are strongly affected by geomagnetic activity [Roble et al., 1981]. The ring current, which causes the negative Dst midlatitude deflection is dissipated in the ionosphere near 60° which, in turn, heats the thermosphere. The Joule heating of the auroral zones and polar caps and the heating from precipitating particles also cause the high latitude thermosphere to be heated. As a result of the heating of the thermosphere, the neutral winds at high and midlatitudes can be strongly affected and even reverse direction.

The convection electric fields imposed upon the high latitude ionosphere from the magnetosphere also cause horizontal plasma transport. In the winter ionosphere, the drifts driven by the convection electric field are a major source of ionization in the polar cap drawing from the auroral zone and the dayside [Sojka et al., 1981]. The convection electric field can also decrease ionization by enhancing the recombination rates as ions are driven through the neutrals.

9.2 HIGH LATITUDE PHENOMENA

9.2.1 Total Ionization Trough and Light Ion Trough.

The total ionization trough is a region of decreased F-region ionization and/or total electron content found in a latitudinally narrow band near 60°–65° magnetic latitude [Ahmed et al., 1979]. The ends of the trough are typically found an hour after sunset and an hour before sunrise. The contour of the trough follows the contours of the auroral oval instead of paralleling the magnetic latitudes. As the auroral oval moves poleward, the trough moves poleward and expands in width. As the auroral oval moves equatorward, the trough moves equatorward and decreases in width. The depth of the trough (the ratio of minimum density to the density a few degrees equatorward and/or poleward) is typically a factor of ten in density or total electron content (TEC) but may vary from barely discernible to a factor of 10^3.

The equatorward wall of the trough tends to be found on the same field line or 1°–2° poleward of the plasmapause. From this observation has come the suggestion that the trough is formed by plasma flow upwards into evacuated flux tubes. The region equatorward lies on filled flux tubes that can replenish the plasma, and the region poleward of the trough is replenished by ionization induced by precipitating energetic particles. An alternate explanation is that the trough is created by the cancellation of the corotation and convection electric field in the evening sector. Plasma remains nearly stationery for several hours; depletion is due to recombination. After leaving the stagnation zone, the trough is sustained through the night by the lack of production in the region of the trough.

The light ion trough is a sharp drop in H^+ and He^+ density near the peak of F region and in the topside ionosphere. On the night side, the light ion trough or density gradient is collocated with the equatorward wall of the total ion trough. On the dayside, the light ion trough continues to be found near 60° magnetic latitude while the total ion trough either ceases to exist or moves to higher latitudes. For a satellite traveling in the altitude range from ~800 to ~1500 km on the dayside and ~600 to ~1500 km on the night side, the light ion trough is seen as a rapid transition between H^+ and O^+ as the dominant ion [Titheridge, 1976]. Above ~1500 km, the O^+ density is so low even when it is the dominant ion that the total density drops by a factor of ~10 to >10^3 at all local times at the same latitude as the light ion trough.

The plasma temperature in the F region near the trough and the equatorward wall of the auroral zone is increased substantially from the temperature in adjacent regions. This is partly related to energy from the ring current in the equatorial plane of the magnetosphere being transferred to the ionosphere. When this energy deposition is large enough, the airglow is enhanced to form stable auroral red (SAR) arcs. SAR arcs are generally subvisual. They tend to be most intense following a major geomagnetic storm.

9.2.2 The Ionosphere in the Auroral Oval.

We have known of the visual displays called auroras for centuries and we have known for the past century that auroras are associated with electromagnetic disturbances. However, it has only been since the start of the space age

we have known auroras are global and are the projection of activity at great altitudes in the magnetosphere and solar wind (Chapters 3 and 8). The global nature of auroras is generally acknowledged by referring to the auroral oval, a band a few degrees in latitude and around both magnetic poles where auroral phenomena are found [Feldstein and Starkov, 1967; Meng, 1977; Gussenhoven et al., 1981 and 1983; see also Chapter 12]. The optical emissions are a result of the energy from precipitating energetic particles (mostly electrons) being deposited in the ionosphere. Optical emissions visible to the naked eye are mostly in narrow latitudinal bands (1 to 10 km in width) called auroral arcs, but there are precipitating particles and optical emissions throughout the auroral zone. In the equatorward portion of the auroral oval the optical emissions are spatially uniform and are known as the diffuse aurora. Along the poleward portion of the oval, the precipitating particles tend to be grouped into bands about 1° wide. The intensity and maximum energy of the precipitating particles are greatest in the center of the band and fall off near the edges of the band. These have been called "inverted-V" events due to their signature in the records of polar orbiting satellites. If the maximum intensity is great enough, a visible auroral arc will appear at the center of the inverted-V event.

The precipitating energetic particles lose energy to the atmosphere by ionizing neutrals in a manner similar to the ionization caused by protons. If the precipitating particles all had the same energy, a thin layer of ionization would be formed. The altitude of the ionization layer is determined from the energy of the precipitating particles: 10 keV electrons produce an ionization layer near 110 km altitude; 500 eV electrons produce an ionization layer near 180 km altitude. The density of the ionization layer is determined by the intensity of the particle flux.

In the nighttime auroral zone, the electron density profile can be accurately estimated if the spectrum of precipitating particles is known, or conversely the spectrum of precipitating particles can be estimated from the electron density profile [Vondrak and Baron, 1977]. This is especially true in the E region where the ionization lifetimes are short. In the F region where ionization lifetimes are long, the ionization present at any given moment is influenced by the precipitation over the past few minutes to tens of minutes as well as the instantaneous precipitation. Also, vertical and horizontal transport has a major effect upon the structure of the F region. Regions of enhanced F region ionization can drift many degrees from the production region [Vickrey et al., 1980].

In the sunlit auroral zone, photoionization dominates the production of ionization, but ionization from particles has a major effect on plasma irregularities. Also, in limited altitude regions the ionization from particles can occasionally dominate photoionization. Even where the photoionization is the dominant source, the heat from precipitating particles increases the scale height of the ionosphere and the neutrals.

An important feature of the high latitude ionosphere is that the plasma density is irregular on the scale of meters to kilometers vertically and on the scale of meters to hundred of kilometers horizontally. The small scale irregularities cause scintillation of radio signals passing through the ionosphere. The large scale irregularities or density gradients provide a necessary condition for various plasma turbulence mechanisms that result in small scale irregularities (see Chapter 10). The causes of irregularities in the auroral zone are numerous [Fejer and Kelley, 1980] but are related to factors such as particle precipitation and $E \times B$ drifts. The region of auroral zone scintillations extends equatorward of the optical auroral [Martin and Aarons, 1977], but are roughly collocated with the region of particle precipitation. Some of the strongest scintillations are observed when the ray path of the radio signal is aligned with the magnetic L-shell in the F region. This has been analyzed to indicate that irregularities are in the form of sheets extended along the magnetic field and in the magnetic E-W direction [Rino et al., 1978].

9.2.3 Substorm Effects.

In the ionosphere, a substorm is an intensification of processes and structures that are normally observed in the quiet time auroral zone ionosphere. See Chapter 8 for a complete description. During a substorm the flux and average energy of the precipitating particles increase rapidly and substantially. This causes more visible features to appear. The increase in the energy of the particle causes ionization in the lower E layer that is not present during quiet times. The increased activity causes an increase in the scintillation producing irregularities. The increased $E \times B$ drift rates heat the ionosphere and increase the scale heights.

During the early phases of a substorm, the auroral zone moves equatorward several degrees in 10–20 minutes. The trough either moves equatorward with the auroral zone movement or ceases to exist during the substorm depending on geomagnetic conditions and local time. During the late phases of a substorm the auroral oval contracts poleward and precipitation induced ionization tends to die away by recombination.

9.2.4 Polar Cap Structure.

The polar cap ionosphere is relatively placid compared to the auroral zone, but soft particle precipitation known as polar rain [Winningham and Heikkila, 1974] does affect the polar cap ionosphere. In the summer months, the polar cap ionosphere is dominated by photoionization similar to the midlatitude ionosphere. In the winter months, ionization is generally maintained by the polar rain and by convection from the day side to the night side of the polar cap. In the winter months, the F layer can be sunlit while the E layer is in darkness. This can lead to He^+ being the dominant

ion in parts of the topside ionosphere. During times of weak convection and precipitation, the winter polar ionosphere near the midnight sector can decay to very low levels of ionization; this area is called the polar hole [Brinton et al., 1978].

Since the magnetic field lines in the polar cap diverge effectively toward infinity, the H^+ in the polar ionosphere escapes rapidly in a process known as the polar wind. Unlike the midlatitude ionosphere, O^+ is generally the dominant ion at all altitudes of the F-region.

9.3 EQUATORIAL PHENOMENA

9.3.1 Sq Current System

See Chapter 4.

9.3.2 Equatorial Electrojet

The intense eastward ionospheric current that flows by day over a narrow latitudinal strip along the magnetic equator is known as the *equatorial electrojet* [Matsushita and Campbell, 1967]. The electrojet causes the large daily variations of the horizontal component of magnetic field intensity recorded by ground magnetometers near the magnetic equator. The ionospheric current system is a result of a dynamo action of the horizontal wind system and the electrical conductivity of the ionosphere in the presence of the electrons and ions. The concentration of ionospheric current near the magnetic equator is a result of the high value of electrical conductivity of the upper atmosphere at the dip equator, which arises from an inhibition of Hall current due to the horizontal configuration of the earth's magnetic field and the horizontal stratification of the ionosphere.

The electrical conductivity of the ionosphere is not only governed by the concentration of the charged particles but by the neutral particles and the earth's magnetic field as well [Chapman, 1956]. Through collisions the neutral particles restrict the motion of charged particles under the action of any impressed electric field. The presence of the magnetic field, on the other hand, restricts the motion of charged particles across the magnetic field and therefore makes the conductivity anisotropic.

When a dc electric field is impressed parallel to the magnetic field, the longitudinal electrical conductivity that exists parallel to the magnetic field is given by

$$\sigma_0 = ne^2\left[\frac{1}{m_e\nu_e} + \frac{1}{m_i\nu_i}\right], \qquad (9.8)$$

where n is the electron or ion concentration, e is the electronic charge, ν_e and ν_i are the electron and ion collision frequencies, and m_e and m_i are the masses of an electron and an ion respectively. The longitudinal conductivity is independent of magnetic field intensity and is identical to the conductivity obtained in the absence of any magnetic field.

When an electric field is applied perpendicular to the magnetic field, the conductivity in the direction of the electric field is called the Pedersen conductivity and is given by

$$\sigma_1 = ne^2\left[\frac{\nu_e}{m_e\left(\nu_e^2 + \Omega_e^2\right)} + \frac{\nu_i}{m_i\left(\nu_i^2 + \Omega_i^2\right)}\right], \qquad (9.9)$$

where Ω_e and Ω_i are the electron and ion gyrofrequencies respectively.

In such cases of a crossed electric and magnetic field, a Hall current usually flows perpendicular to both the electric and magnetic fields and the resulting conductivity, called the Hall conductivity, is given by

$$\sigma_2 = ne^2\left[-\frac{\Omega_e}{m_e\left(\nu_e^2 + \Omega_e^2\right)} + \frac{\Omega_i}{m_i\left(\nu_i^2 + \Omega_i^2\right)}\right]. \qquad (9.10)$$

At the magnetic equator, an eastward electric field is developed by the dynamo action of the horizontal wind system, which gives rise to a motion of charged particles in the east-west direction (X) due to the Pedersen conductivity σ_1 and in the vertical direction (Z) due to the Hall conductivity σ_2. In view of the horizontal stratification of the ionosphere, the flow of Hall current in the vertical direction (Z) is totally inhibited and a polarization electric field (E_Z) develops. In such a case the current densities J_X, J_Z in the X and Z directions can be expressed as

$$J_X = \sigma_1 E_X + \sigma_2 E_Z$$

$$J_Z = -\sigma_2 E_X + \sigma_1 E_Z = 0$$

or,
$$J_X = \left(\sigma_1 + \frac{\sigma_2^2}{\sigma_1}\right) E_X$$

or,
$$J_X/E_X = \sigma_1 + \frac{\sigma_2^2}{\sigma_1}$$

$$= \sigma_3 . \qquad (9.11)$$

The resulting conductivity in the east-west direction is called the Cowling conductivity σ_3.

The variations of Pederson σ_1, Hall σ_2, and Cowling σ_3 conductivities with altitude at the magnetic equator are

shown in Figure 9-3a for an assumed variation of electron density N and temperature T shown in Figure 9-3b. The enhanced value of the Cowling conductivity in the dynamo region is sufficient to account for the intensity of the equatorial electrojet. Away from the equator, the geomagnetic

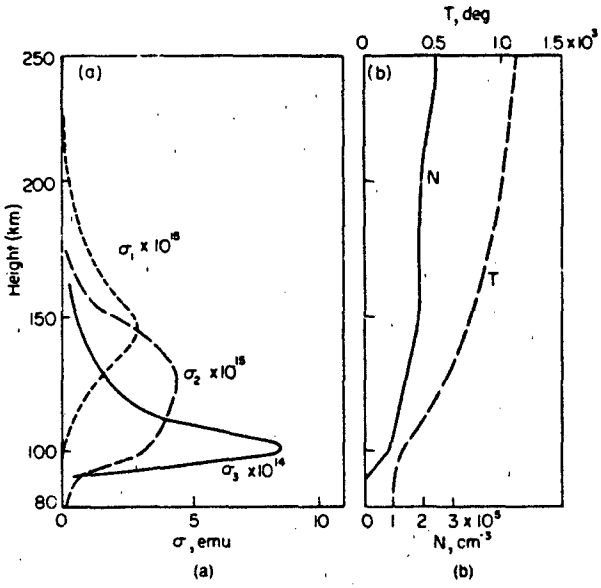


Figure 9-3. (a) Variation of Pedersen σ_1, Hall σ_2, and Cowling σ_3 conductivities with altitude for an ionosphere in which electron density N and temperature T vary as shown at (b). Note that scale for σ_3 is smaller than σ_1 and σ_2 by a factor of 10 [Chapman and Raja Rao, 1965, based on Chapman, 1956].

field is no longer horizontal, which allows the Hall field to leak away. Baker and Martyn [1952] estimated that the half-width of the strip of enhanced east-west conductivity around the dip equator is about 3° in latitude. The equatorial electrojet corresponds to an east-west electric field of 0.5 mV/m and a vertical polarization field of about 10 mV/m. This gives rise to an eastward current or westward electron drift of several hundred m/s. The electron drift is westward by day and eastward at night. This electrojet model was studied in detail by Sugiura and Cain [1966]. More complete models allowing for vertical currents have been discussed by Untiedt [1967], Sugiura and Poros [1969], Richmond [1973], and others.

The equatorial electrojet current has been observed to reverse its normal direction during day or night and during magnetically quiet or disturbed condtions; this reverse current system has been termed the *counter electrojet* [Gouin and Mayaud, 1967; Hutton and Oyinloye, 1970; Rastogi, 1973; Fejer et al., 1976; Fejer and Kelley, 1980]. The rapid reversals during disturbed conditions have been related to magnetospheric and high-latitude phenomena [Matsushita and Balsley, 1972], whereas the reversals during quiet conditions have been related to lunar tides [Rastogi, 1974].

9.3.3 Electrojet Irregularities

Ionosondes first detected the existence of a distinct type of sporadic E near the magnetic equator which is patchy and transparent to radio waves reflected from higher layers. The intensity of equatorial E_s (or E_{sq}) is strongly correlated with the strength of the electrojet current discussed in the last section [Matsushita, 1951]. A typical ionogram showing E_{sq} echoes is shown in Figure 9-4. VHF forward scatter

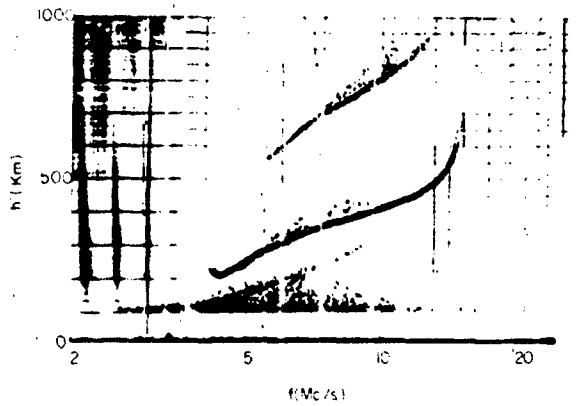


Figure 9-4. The typical equatorial sporadic E configuration on an ionogram recorded at Huancayo, Peru at 1229 hr (75°W) on 19 April 1960 [Cohen et al., 1962].

experiments established that these echoes arise as a result of scattering from field aligned irregularities of electron density immersed in the equatorial electroject [Bowles and Cohen, 1962].

The important characteristics of the electrojet irregularities as related to the physics of the scattering region have been probed by the VHF radar measurements performed at the Jicamarca Radio Observatory. Radar spectral studies have shown the existence of two classes of irregularities called Type 1 and Type 2, associated with the electrojet [Balsley and Farley, 1973; Farley and Balsley, 1973; Fejer and Kelley, 1980].

Type 1 irregularities have a very narrow spectrum with a Doppler shift corresponding approximately to the ion acoustic velocity of about 360 m/s. Farley [1963] and Buneman [1963] have explained the Type 1 irregularities by showing that a plasma is unstable to waves when the relative electron-ion drift velocity in the direction of the wave exceeds the ion acoustic velocity. As such, Type 1 irregularities are also called two-stream irregularities. Type 2 irregularities on the other hand have phase velocities smaller than the ion-acoustic velocity and are observed even when the eastward drift velocity is very small during the day. Type 2 irregularities are identified with E_{sq} echoes in ionograms and both disappear under counter electrojet condi-

tions during the daytime. At night the Type 2 irregularities are almost always observed except for momentary disappearance when the electrojet electric field reverses sign. The Type 2 irregularities are explained by the gradient drift instability mechanism. This is because the horizontal polarization field arising from a relative electron-ion drift in the electrojet region can, in the presence of the earth's magnetic field, develop a drift in the direction of the vertically oriented density gradient and give rise to these irregularities [Fejer and Kelley, 1980].

9.3.4 Equatorial Anomaly and Fountain Effect

During the equinox the sun is overhead at the equator, and in terms of solar control the ionization density is expected to be maximum in that region. Instead, the daytime ionization density at the F_2 peak shows a pronounced trough at the magnetic equator and crests at about 30°N and 30°S magnetic dip. This anomalous latitude variation of F_2 ionization near the magnetic equator obtained from bottomside ionograms illustrated in Figure 9-5, was first recognized by

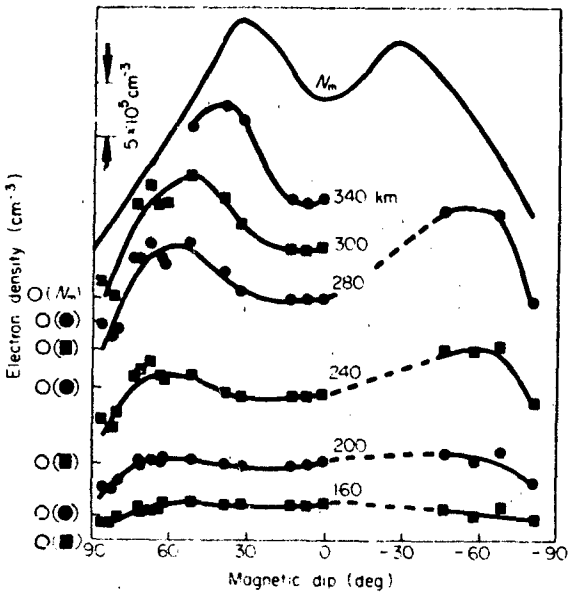


Figure 9-5. Variation of N_mF_2 and of electron density (electron concentration) at fixed heights with magnetic dip, for noon on magnetically quiet days in September 1957. The zero level for each curve is indicated on the left [Croom et al., 1959]. Reprinted with permission from MacMillan Journals Ltd., © 1959.)

Appleton [1946] and is known as the equatorial anomaly or *Appleton* anomaly.

The equatorial anomaly is explained in terms of a *fountain effect* caused by vertical electrodynamic drift at the equator and plasma diffusion along geomagnetic field lines [Martyn, 1959]. Figure 9-6 illustrates how the eastward E-region dynamo electric field at locations slightly off the magnetic equator maps to F-region altitude over the equator. The eastward electric field in conjunction with the northward

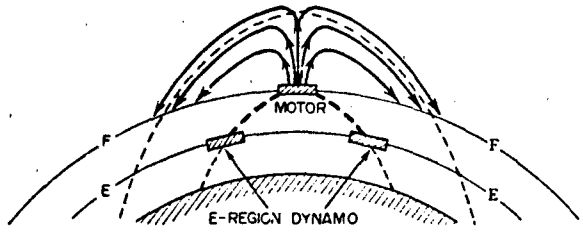


Figure 9-6. The F region geomagnetic anomaly. Near the equator the electric fields of the atmospheric dynamo in the E layer are conveyed upwards along geomagnetic lines of force to the *motor* in the F layer where they produce an upwards movement of the plasma during the day. The raised plasma then diffuses down lines of force to produce enhanced concentration at places on each side of the equator and decreased concentration at the equator itself [Ratcliffe, 1972].

geomagnetic field gives rise to a vertically upward plasma motion. At high altitudes over the equator, the plasma encounters field lines that connect to the F_2 peak at 30°N and 30°S magnetic dip along which the plasma diffuses under the action of gravity. Such plasma transport depletes the F_2 ionization at the equator and increases the density at locations 30°N and 30°S. Theoretical studies of the equatorial anomaly on a more rigorous basis have been performed by many workers [see Hanson and Moffett, 1966 and references therein].

The transport processes involved in the formation of the equatorial anomaly are best illustrated by the Alouette I topside sounder results. Figure 9-7 shows the variation of ionization density as a function of height and latitude in the daytime topside ionosphere. At high altitudes over the magnetic equator, the density shows a dome-like structure following the shape of a magnetic field line. At lower altitudes, below 700 km, a field-aligned double humped structure is obtained, with the maxima being closer together at the greater heights.

The diurnal development of the equatorial anomaly has been studied from ground based as well as topside sounders. Ground based data indicate that during years of sunspot minimum the anomaly is most pronounced at about 1400 LT and then declines steadily until it disappears around 2000 LT [Rastogi, 1959]. However, during the period of sunspot maximum, the anomaly after an initial decay in the afternoon hours shows substantial redevelopment in the late evening hours [Martyn, 1959; Appleton, 1960; Wright, 1960] and often the ionization density at the crests in the evening period exceeds the daytime values. Latitudinal asymmetry of the equatorial anomaly in the northern and southern hemisphere as a function of season and longitude has also been studied

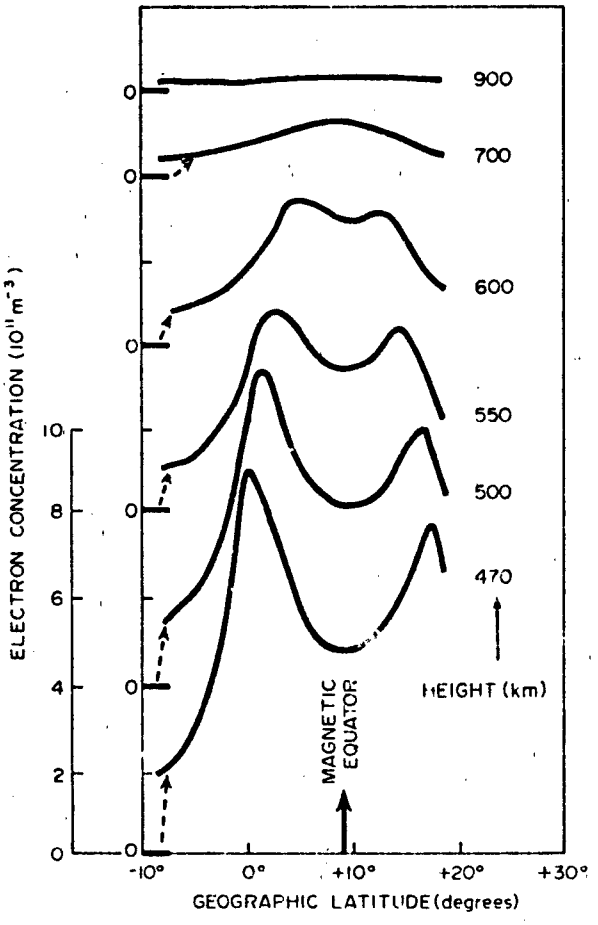


Figure 9-7. Latitudinal variation of electron density across the equatorial anomaly at various altitudes above h_{max} from topside ionograms [Eccles and King, 1969] (Reprinted with permission from JATP; © 1969.)

[Lyon, 1963; Lyon and Thomas, 1963]. Interhemispheric neutral wind and variation of magnetic declination with longitude have been invoked in theoretical models [Hanson and Moffett, 1966] to account for such asymmetry.

9.3.5 Equatorial F Region Irregularities

Historically, the signature of equatorial nighttime F-region irregularities was first obtained from the spread-F signature on ionograms [Booker and Wells, 1938]. Equatorial spread-F has since been divided into two types, range and frequency spread [Calvert and Cohen, 1961; Rastogi, 1980], the former type being attributed to strongly scattering irregularities.

The advent of orbiting and geostationary beacons in the early 1960s provided another technique for monitoring equatorial irregularities, that of measuring the phase, amplitude, and plane of polarization of the trans-ionospherically propagated signal. The bulk of the information on magnitude and occurrence of these irregularities came from amplitude fluctuation or scintillation measurements at a host of equatorial stations (see Chapter 10 for further details).

During the last decade a determined effort has been made to understand the nature and occurrence of nighttime equatorial F-region irregularities since the largest propagation effects extending up to S-band frequencies are observed in this region. Our insight has come from multi-technique observations comprising satellite and rocket *in situ* measurement of irregularity amplitude and spectra, coherent and incoherent radar backscatter measurements, total electron content and ground based and airborne multi-frequency scintillation, and all-sky imaging photometer measurements. Together these techniques measure irregularities over scale lengths of 5 to 6 orders of magnitude from hundreds of kilometers to tens of centimeters, and given the right conditions the post-sunset equatorial F-region is indeed found to contain irregularities over this enormous scale size range [Basu and Basu, 1981]. A brief description of the different techniques and their results are given below.

The rocket and satellite *in situ* measurements have detected large scale irregular biteouts of ion concentration (N_i) in the nighttime equatorial spread-F region associated with small scale irregularities in N_i [Hanson and Sanatani, 1973; Kelley et al., 1976; Morse et al., 1977]. A comprehensive study of such depleted regions by McClure et al., [1977] indicates the presence of very sharp electron density structures (see bottom panel of Figure 9-8) and the existence of ionic species near the F-peak that are normally obtained in the bottomside and valley region between the E and F layers. This study also revealed the existence of a highly structured upward velocity within these depleted regions on the order of 100 m/s (hence the name *bubbles*), and sometimes, in addition, a westward velocity of about 20 m/s as shown in

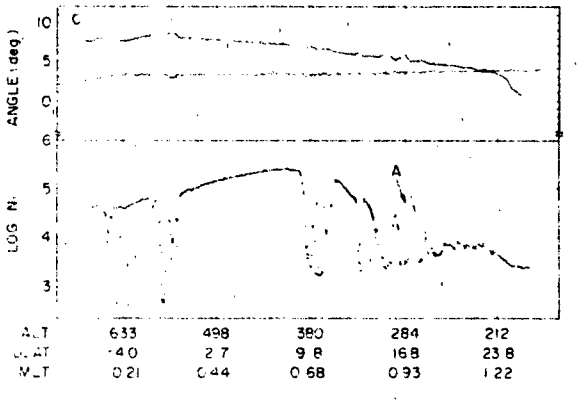


Figure 9-8. Ion drift meter data, orbit 2282. Satellite altitude, dip latitude, and magnetic local time are indicated on the figure. The satellite longitude was −50° to −65°. The observed pitch and yaw angles are shown in the upper and lower curves, respectively. Positive angles correspond to ions moving up or left with respect to the spacecraft [McClure et al., 1977].

Figure 9-8. Woodman and LaHoz [1976] using the radar technique at Jicamarca observed plume-like structures in backscatter power maps of equatorial irregularities at 3 m wave length. The maps they obtained are similar to that shown in the top panel of Figure 9-9. They interpret the plumes as being due to vertically rising bubbles and their wakes. Evidence for the fact that the plasma bubbles, most probably initiated by the Rayleigh-Taylor instability, are probably extended in altitude has been obtained from conventional polarimeter observations of the total electron content (TEC). These observations [Yeh et al., 1979; DasGupta et al., 1982] show that scintillation patches in the early evening hours occur in association with depletions of TEC which may be as large as 40% of the ambient value.

From topside sounder observations, Dyson and Benson [1978] have shown that plasma bubbles are confined within magnetic field tubes. By the use of an all sky imaging photometer on board an aircraft, Weber et al., [1978, 1980] have detected 6300 Å airglow depletions in the nighttime equatorial F-region in association with radar backscatter and scintillation patches. The 6300 Å airglow depletions signify depletions of integrated ionization density and the two-dimensional maps of these structures show that the depletions observed near the magnetic equator have typical E-W dimension of 100 km and are as large as 1200 km in the magnetic N-S direction. These observations establish that the bubbles are open at the bottom and confined within magnetic field tubes. From measurements of both incoherent scatter and coherent backscatter using a steerable radar at 155 MHz, Tsunoda [1980a] located the 1 m field aligned irregularities at the top edge of a plasma bubble and mapped the bubble along the magnetic field tube.

By performing careful coordinated studies of radar backscatter, high resolution satellite and rocket *in situ* and ground scintillation measurements, the spatial and temporal coexistence of kilometer and meter scale irregularities have been studied [Basu et al., 1978; 1980; Rino et al., 1981]. Figure 9-9, taken from Basu and Basu [1981], shows that the 3 m irregularities causing the radar backscatter and kilometer to several hundred meter irregularities causing scintillations are simultaneously (within the limit of time resolution of the experiment ~ min) generated in the onset phase, but the short scale irregularities are outlived by the large scale ones by several hours. A considerable effort has been made

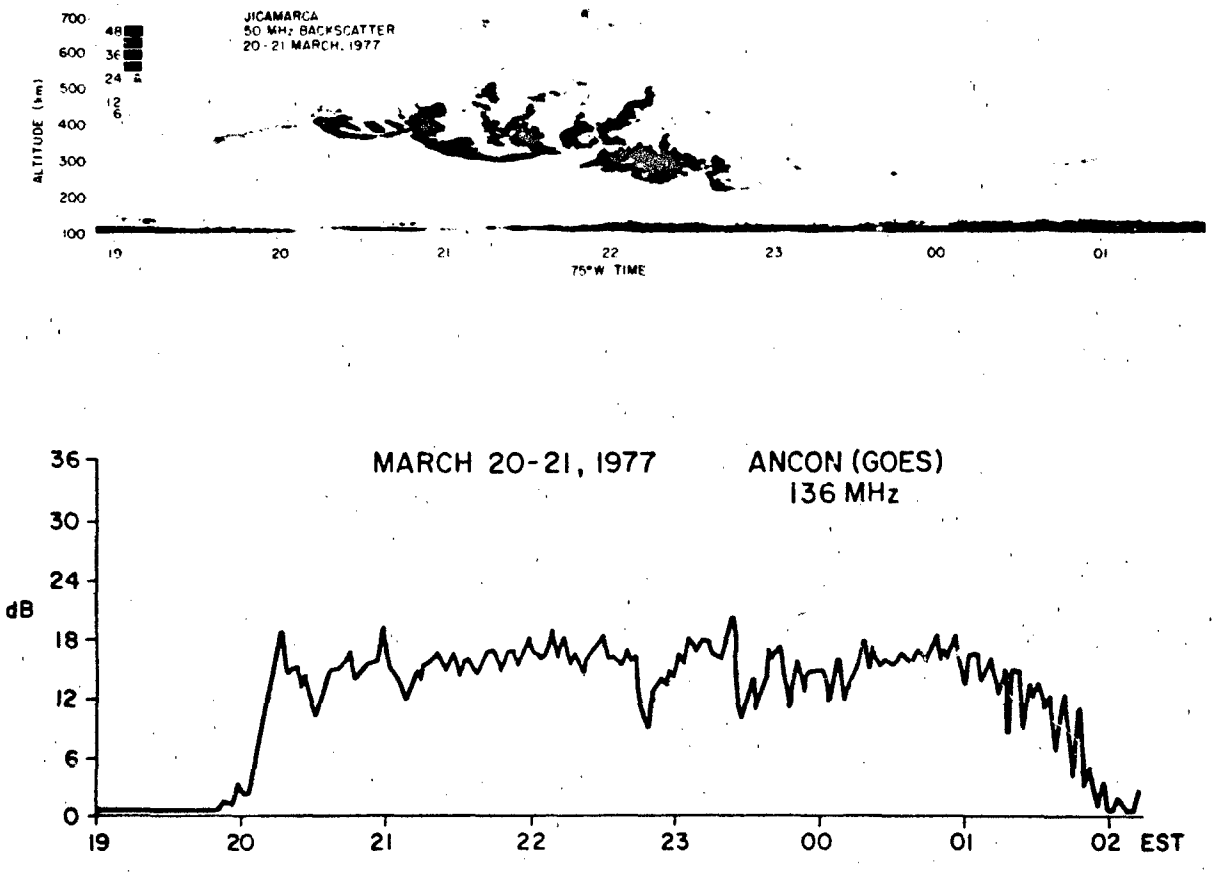


Figure 9-9 Temporal variation of range and intensity (different grey tones) of 50 MHz backscattered power at Jicamarca on 20-21 March 1977 (top panel) and 137 MHz scintillations (bottom panel) over a nearly common ionospheric volume [Basu et al., 1980]

over the past few years to detect irregularities shorter than the ion gyroradius which is ~5 m in the topside equatorial ionosphere [Woodman and Basu, 1978] for radar system applications and for the understanding of the complex plasma processes in equatorial spread F. By the use of the ALTAIR radar at Kwajalein, Marshall Islands, irregularities with spatial wavelengths of 1 m and 36 cm have been detected [Tsunoda et al., 1979; Towle, 1980; Tsunoda, 1980b] and more recently the TRADEX radar at Kwajalein has been used to detect 11 cm irregularities which are approximately 3 times the electron gyroradius and 30 times the Debye length [Tsunoda, 1980c]. Thus equatorial spread F is found to encompass irregularity wavelengths extending over 5–6 orders of magnitude.

CHAPTER 9

REFERENCES

Ahmed, M., R.C. Sagalyn, and P.J.L. Wildman, "Topside Ionosphere Trough Morphology: Occurrence Frequency and Diurnal, Seasonal and Altitude Variations," *J. Geophys. Res.*, **84:** 489–498, 1979.

Appleton, E.V., "Two Anomalies in the Ionosphere," *Nature* (London), **157:** 691, 1946.

Appleton, E.V., "Two Anomalies in the Behavior of the F2 Layer of the Ionosphere," *Some Ionospheric Results Obtained During the IGY.*, edited by W.J.G. Beynon, p. 3, Elsevier, Amsterdam, 1960.

Baker, W.G. and D.F. Martyn, "Conductivity of the Ionosphere," *Nature*, **170:** 1090, 1952.

Balsley, B.B. and D.T. Farley, "Radar Observations of Two-Dimensional Turbulence in the Equatorial Electrojet," *J. Geophys. Res.*, **78:** 7174, 1973.

Banks, P.M. and G. Kockarts, *Aeronomy, Part B*, Academic Press, New York, 1973.

Basu, S. and Su. Basu, "Equatorial Scintillations—A Review," *J. Atmos. Terr. Phys.*, **43:** 473, 1981.

Basu, S., Su. Basu, J. Aarons, J.P. McClure, and M.D. Cousins, "On the Coexistence of Kilometer- and Meter-Scale Irregularities in the Nighttime Equatorial F Region," *J. Geophys. Res.*, **83:** 4219, 1978.

Basu, S., J.P. McClure, S. Basu, W.B. Hanson, and J. Aarons, "Coordinated Study of Equatorial Scintillation and In Situ and Radar Observations of Nighttime F Region Irregularities," *J. Geophys. Res.*, **85:** 5119, 1980.

Booker, H.G. and H.W. Wells, "Scattering of Radio Waves by the F-Region of the Ionosphere," *Terr. Magn. Atmos. Elect.*, **43:** 249, 1938.

Bowles, K.L. and R. Cohen, "A Study of Radio Wave Scattering from Sporadic E. Near the Magnetic Equator," in *Ionospheric Sporadic E.* edited by E.K. Smith and S. Matsushita, Pergamon Press, New York, 1962.

Brinton, H.C., J.M. Grebowsky, and L.H. Brace, "The High Latitude Winter F-Region at 300 km: Thermal Plasmas Observations from AE-C," *J. Geophys. Res.*, **83:** 4767–4776, 1978.

Buneman, O., "Excitation of Field Aligned Sound Waves by Electron Streams," *Phys. Res. Lett.*, **10:** 285, 1963.

Calvert, W. and R. Cohen, "The Interpretation and Synthesis of Certain Spread-F Configurations Appearing on Equatorial Ionograms," *J. Geophys. Res.*, **66:** 3125, 1961.

Carpenter, D.L. and C.G. Park, "On What Ionospheric Workers Should Know About the Plasmapause-Plasmasphere," *Rev. Geophys. Space Sci.*, **11:** 133–154, 1973.

Chapman, S., "The Electrical Conductivity of the Ionosphere: A Review," *Nuovo Cimento*, **4**(Suppl): 1385, 1956.

Chapman, S. and K.S. Raja Rao, "The H and Z Variations Along and Near the Equatorial Electrojet in India, Africa and the Pacific," *J. Atmos. Terr. Phys.*, **27:** 559–581, 1965.

Cohen, R., K.L. Bowles, and W. Calvert, "On the Nature of Equatorial Slant Sporadic E," *J. Geophys. Res.* **67:** 965–972, 1962.

Cormier, R.J., "Thule Riometer Observations of Polar Cap Absorption Events, (1962–1972)," AFCRL TR-73-0060, 1973.

Croom, S., A. Robbins, and J.O. Thomas, "A Review of Topside Sounder Studies of the Equatorial Ionosphere," *Nature* (London), **184:** 2003, 1959.

DasGupta, A., J. Aarons, J.A. Klobuchar, S. Basu, and I. Bushby, "Ionsopheric Electron Content Depletions Associated with Amplitude Scintillations in the Equatorial Region," *Geophys. Res. Lett.*, **9:** 147–150, 1982.

DNA Ratebook, Defense Nuclear Agency, U.S. Government Printing Office, Washington, D.C., 1972.

Doering, J.P., T.A. Potemra, W.K. Petersen, and C.O. Bostrom, "Characteristic Energy Spectra of 1 to 500 eV Electrons Observed in the High Latitude Ionosphere from Atmospheric Explorer C," *J. Geophys. Res.*, **81:** 5507, 1976.

Donnelly, R.F., "Contribution of X-Ray and EUV Bursts of Solar Flares to Sudden Frequency Deviations," *J. Geophys. Res.*, **74:** 1973, 1969.

Donnelly, R.F., "Extreme Ultraviolet Flashes of Solar Flares Observed Via Sudden Frequency Deviations: Experimental Results," *Sol. Phys.*, **29:** 188, 1971.

Dyson, P.L. and R.F. Benson, "Topside Sounder Observations of Equatorial Bubbles," *Geophys. Res. Lett.*, **5:** 795, 1978.

Eccles, D. and J.W. King, "A Review of Topside Sounder Studies of the Equatorial Ionosphere," *Proc. IEEE*, **57:** 1012, 1969.

Evans, J.V., "Incoherent Scatter Contributions to Studies of the Dynamics of the Lower Thermosphere," *Rev. Geophys. Space Sci.*, **16:** 195–216, 1978.

Farley, D.T., "A Plasma Instability Resulting in Field-Aligned Irregularities in the Ionosphere," *J. Geophys. Res.*, **68:** 6083, 1963.

Farley, D.T. and B.B. Balsley, "Instabilities in the Equatorial Electrojet," *J. Geophys. Res.*, **78:** 227, 1973.

Fejer, B., D.T. Farley, B.B. Balsley, and R.F. Woodman, "Radar Studies of Anomalous Velocity Reversals in the Equatorial Ionosphere," *J. Geophys. Res.*, **81:** 4621, 1976.

Fejer, B.G. and M.C. Kelley, "Ionospheric Irregularities," *Rev. Geophys. Space Sci.*, **18:** 401–454, 1980.

Feldstein, Y.I. and G.V. Starkov, "Dynamics of Auroral Belts and Polar Geomagnetic Disturbances," *Planet. Space Sci.*, **15:** 209, 1967.

Ferguson, E.E., "D-Region Ion Chemistry," *Rev. Geophys. Space Phys.*, **9:** 997–1008, 1971.

Gouin, P. and P.N. Mayaud, "A Propos de Existence Possible d'un Contre-electrojet aux Latitudes Magnetiques Equatoriales," *Ann. Geophys.*, **23:** 41, 1967.

Gussenhoven, M.S., D.A. Hardy, and W.J. Burke, "DMSP/F2 Electron Observations of Equatorward Auroral Boundaries and Their Relationship to Magnetospheric Electric Fields," *J. Geophys. Res.*, **86:** 768, 1981.

Gussenhoven, M.S., D.A. Hardy, N. Heinemann, and E. Holeman, "1978 Diffuse Auroral Boundaries and a De-

rived Auroral Boundary Index," AFGL-TR-82-0398, ADA130175, 1982.

Hanson, W.B. and R.J. Moffett, "Ionization Transport Effects in the Equatorial F Region," *J. Geophys. Res.*, **71:** 5559, 1966.

Hanson, W.B. and S. Sanatani, "Large N, Gradients Below the Equatorial F- Peak," *J. Geophys. Res.*, **78:** 1167, 1973.

Hudson, R.D., "Critical Review of Ultraviolet Photoabsorption Cross Section for Molecules of Astrophysical and Aeronomic Interest," *Rev. Geophys. Space Phys.*, **9:** 305–406, 1971.

Hutton, R. and J.O. Oyinloye, "The Counter-Electrojet in Nigeria," *Ann. Geophys.*, **26:** 921, 1970.

Jasperse, J.R., "Electron Distribution Function and Ion Concentrations in the Earth's Lower Ionosphere from Boltzman-Fokker-Planck Theory," *Planet. Space Sci.*, **25:** 743, 1977.

Kelley, M.C., G. Haerendel, H. Kappler, A. Valenzuela, B.B. Balsley, D.A. Carter, W.L. Eklund, C.W. Carlson, B. Hausler, and R. Torbert, "Evidence for a Rayleigh-Taylor Type Instability and Upwelling of Depleted Density Regions during Equatorial Spread-F," *Geophys. Res. Lett.*, **3:** 448, 1976.

Kohnlein, W., "Electron Density Models of the Ionosphere," *Rev. of Geophys. Space Phys.*, **16:** 341–354, 1978.

Lyon, A.J., *The Ionosphere*, p. 88, Physical Society, London, 1963.

Lyon, A.J. and L. Thomas, "The F2-Region Equatorial Anomaly in the African, American and East Asian Sectors during Sunspot Maximum," *J. Atmos. Terr. Phys.*, **25:** 373, 1963.

Mantas, G.P., H.C. Carlson, and V.B. Wickwar, "Photoelectron Flux Build-up in the Plasmasphere," *J. Geophys. Res.*, **83:** 1–15, 1978.

Martin, E. and J. Aarons, "F-Layer Scintillations and Aurora," *J. Geophys. Res.*, **82:** 2717-2722, 1977.

Martyn, D.F., *The Physics of the Ionosphere*, p. 254, Physical Society, London, 1963.

Martyn, D.F., "The Normal F Region of the Ionosphere," *Proc. IRE N.Y.*, **47:** 147, 1959.

Matsushita, S., "Intense E, Ionization Near the Magnetic Equator," *J. Geomagn. Geoelect.*, **3:** 44–46.

Matsushita, S. and B.B. Balsley, "A Question of DP-2," *Planet. Space Sci.*, **20:** 1259, 1972.

Matsushita, S. and W.H. Campbell (eds.), *Physics of Geomagnetic Phenomena*, Academic Press, New York, 1967.

McClure, J.P., W.B. Hanson, and J.H. Hoffman, "Plasma Bubbles and Irregularities in the Equatorial Ionosphere," *J. Geophys. Res.*, **82:** 2650, 1977.

Meng, C.-I., R.H. Holzworth, and W.-I. Akasofu, "Auroral Circle Delineating the Poleward Boundary of the Quiet Auroral Oval," *J. Geophys. Res.*, **82:** 164, 1977.

Morse, F.A., B.C. Edgar, H.C. Koons, C.J. Rice, W.J. Heikkila, J.H. Hoffman, B.A. Timsley, J.D. Winningham, A.B. Christiansen, R.F. Woodman, J. Pomalaza, and R.N. Teixeiro, "Equion: An Equatorial Ionospheric Irregularity Experiment," *J. Geophys. Res.*, **82:** 578, 1977.

Rasmussen, J.E., P.A. Kossey, and E.A. Lewis, "Evidence of an Ionospheric Reflecting Layer Below the Classical D Region," *J. Geophys. Res.*, **85:** 3037–3044, 1980.

Rastogi, R.G., "The Diurnal Development of the Anomalous Equatorial Belt in the F2 Region of the Ionosphere," *J. Geophys. Res.*, **64:** 727, 1959.

Rastogi, R.G., "Counter Electrojet Currents in the Indian Zone," *Planet. Space Sci.*, **21:** 1355, 1973.

Rastogi, R.G., "Lunar Effects in the Counter-Electrojet Near the Magnetic Equator," *J. Atmos. Terr. Phys.*, **36:** 167, 1974.

Rastogi, R.G., "Seasonal and Solar Cycle Variations of Equatorial Spread-F in the American Zone," *J. Atmos. Terr. Phys.*, **42:** 593, 1980.

Ratcliffe, J.A., *An Introduction to the Ionosphere and Magnetosphere*, Cambridge University Press, 1972.

Richmond, A.D., "Equatorial Electrojet, I. Development of a Model Including Winds and Instabilities," *J. Atmos. Terr. Phys.*, **35:** 1083, 1973.

Richmond, A.D., and S.V. Venkateswaran, "Geomagnetic Crochets and Associated Ionospheric Current Systems," *Radio Sci.*, **6:** 139, 1971.

Rino, C.L., T. Tsunoda, J. Petriceks, R.C. Livingston, M.C. Kelley, and K.D. Baker, "Simultaneous Rocket-Borne Beacon and in situ Measurements of Equatorial Spread-F—Intermediate Wavelength Results," *J. Geophys. Res.*, **86:** 2411, 1981.

Rino, C.L., R.C. Livingston, and S.J. Matthews, "Evidence for Sheet-Like Auroral Ionospheric Irregularities," *Geophys. Res. Lett.*, **5:** 1039, 1978.

Roble, R.G., R.E. Dickinson, E.C. Ridley, and Y. Kamiue, "Thermospheric Response to the November 8–9, 1969 Disturbance," *J. Geophys. Res.*, **84:** 4207–4216, 1979; "Thermosphere Circulation Models," *Eos, Trans. AGU*, **62:** 3, 19, 1981.

Sojka, J.J., W.J. Raitt, and R.W. Schunk, "A Theoretical Study of High Latitude Winter F Region at Solar Minimum for Low Magnetic Activity," *J. Geophys. Res.*, **86:** 609–621, 1981.

Stolarski, R.S. and N.P. Johnson, "Photoionization and Photoabsorption Cross Sections for Ionospheric Calculations," *J. Atmos. Terr. Phys.*, **34:** 1691, 1972.

Strascio, M.A. and B. Sellers, "The Calculation of Riometers Absorption and Approximate Connection Between Riometer Absorptions and Solar Proton Fluxes during Night-Time PCA Events," AFCRL-TR-75-0469, ADA019656, 1975.

Sugiura, M. and J.C. Cain, "A Model Equatorial Electrojet," *J. Geophys. Res.*, **71:** 1869, 1966.

Sugiura, M. and D.J. Poros, "An Improved Model Equatorial Electrojet with a Meridional Current System," *J. Geophys. Res.*, **74:** 4025, 1969.

Titheridge, J.E., "Ion Transition Heights from Topside Electron Density Profiles," *Planet. Space Sci.*, **24:** 229–245, 1976.

Torr, D.G., "Ionospheric Chemistry," *Rev. Geophys. Space Phys.*, **17:** 510–521, 1979.

Towle, D.M., "VHF and UHF Radar Observations of Equatorial Ionospheric Irregularities and Background Densities," *Radio Sci.*, **15:** 71, 1980.

Tsunoda, R.T., "Magnetic-Field-Aligned Characteristics of Plasma Bubbles in the Nighttime Equatorial Ionosphere," *J. Atmos. Terr. Phys.*, **42:** 743, 1980a.

Tsunoda, R.T., "On the Spatial Relationship of 1-m Equatorial Spread-F Irregularities and Plasma Bubbles," *J. Geophys. Res.*, **85:** 185, 1980b.

Tsunoda, R.T., "Backscatter Measurements of 11-cm Equatorial Spread-F Irregularities," *Geophys. Res., Lett.*, **7:** 848, 1980c.

Tsunoda, R.T., M.J. Baron, J. Owen, and D.M. Towle, "Altair. An Incoherent Scatter Radar for Equatorial Spread-F Studies," *Radio Sci.*, **14:** 1111, 1979.

Turco, R.P., "A Discussion of Possible Negative Ion Detachment Mechanisms in the Sunrise D Region," *Radio Sci.*, **9:** 655, 1974.

Untiedt, J., "A Model of the Equatorial Electrojet with a Meridional Current System," *J. Geophys. Res.*, **72:** 5799, 1967.

Vickrey, J.F., C.L. Rino, and T.A. Potemra, "Chatanika/TRIAD Observations of Unstable Ionization Enhancements in the Auroral F-Region," *Geophys. Res. Lett.*, **7:** 798–792, 1980.

Vondrak, R.R. and M.J. Baron, "Radar Measurements of the Latitudinal Variation of Auroral Ionization," *Radio Sci.*, **11:** 939–946, 1976; "A Method of Obtaining the Energy Distribution of Auroral Electrons from Incoherent Scatter Radar Measurements," *Radar Probing of the Auroral Plasma*, edited by A. Brekke, pp 315–330, Scandinavian University Books, Oslo, Norway, 1977.

Weber, E.J., J. Buchau, R.H. Eather, and R.S. Mende, "North-South Aligned Equatorial Airglow Depletions," *J. Geophys. Res.*, **83:** 712, 1978.

Weber, E.J., J. Buchau, and J.G. Moore, "Airborne Studies of F Layer Ionospheric Irregularities," *J. Geophys. Res.*, **85:** 4631, 1980.

Winningham, J.D. and W.J. Heikkila, "Polar Cap Auroral Electron Fluxes Observed with ISIS-I," *J. Geophys. Res.*, **79:** 949–958, 1974.

Woodman, R.F. and C. LaHoz, "Radar Observations of F Region Equatorial Irregularities," *J. Geophys. Res.*, **81:** 5447, 1976.

Woodman, R.F. and S. Basu, "Comparison Between in situ Spectral Measurements of Equatorial F Region Irregularities and Backscatter Observations at 3-m Wavelength," *Geophys. Res. Lett.*, **5:** 869, 1978.

Wright, J.N., "A Model of the F2 Region Above h_{max} F2," *J. Geophys. Res.*, **65:** 185, 1960.

Yeh, K.C., H. Soicher, and C.H. Liu, "Observations of Equatorial Ionospheric Bubbles by the Radio Propagation Method," *J. Geophys. Res.*, **84:** 6589, 1979.

Chapter 10

IONOSPHERIC RADIO WAVE PROPAGATION

Section 10.1	**S. Basu, J. Buchau, F.J. Rich and E.J. Weber**
Section 10.2	**E.C. Field, J.L. Heckscher, P.A. Kossey, and E.A. Lewis**
Section 10.3	**B.S. Dandekar**
Section 10.4	**L.F. McNamara**
Section 10.5	**E.W. Cliver**
Section 10.6	**G.H. Millman**
Section 10.7	**J. Aarons and S. Basu**
Section 10.8	**J.A. Klobuchar**
Section 10.9	**J.A. Klobuchar**
Section 10.10	**S. Basu, M.F. Mendillo**

The series of reviews presented is an attempt to introduce ionospheric radio wave propagation of interest to system users. Although the attempt is made to summarize the field, the individuals writing each section have oriented the work in the direction judged to be most important.

We cover areas such as HF and VLF propagation where the ionosphere is essentially a "black box", that is, a vital part of the system. We also cover areas where the ionosphere is essentially a nuisance, such as the scintillations of trans-ionospheric radio signals.

Finally, we have included a summary of the main features of the models being used at the time of writing these reviews. [J. Aarons]

10.1 MEASURING TECHNIQUES

10.1.1 Ionosonde

For more than four decades, sounding the ionosphere with ionospheric sounders or ionosondes has been the most important technique developed for the investigation of the global structure of the ionosphere, its diurnal, seasonal and solar cycle changes, and its response to solar disturbances. Even the advent of the extremely powerful incoherent scatter radar technique [Evans, 1975], which permits measurement of the complete electron density profile, electron and ion temperatures, and ionospheric motions, has not made the relatively inexpensive and versatile ionosonde obsolete. On the contrary, modern techniques of complex ionospheric parameter measurements and data processing [Bibl and Reinisch, 1978a; Wrigh, and Pitteway, 1979; Buchau et al., 1978] have led to a resurgence of interest in ionospheric sounding as a basic research tool, while a renewed interest in HF communications is leading to a rejuvenation of the global ionosonde network.

10.1.1.1 Ionogram. Ionospheric sounders or ionosondes are, in principle, HF radars that record the time of flight or travel of a transmitted HF signal as a measure of its ionospheric reflection height. By sweeping in frequency, typically from 0.5 to 20 MHz, an ionosonde obtains a measurement of the ionospheric reflection height as a function of frequency. A recording of this reflection height measurement as a function of frequency is called an ionogram. Ionograms can be used to determine the electron density distribution as a function of height, $N_e(h)$, from a height that is approximately the bottom of the E layer to generally the peak of the F2 layer, except under spread F conditions or under conditions when the underlying ionization prevents measurement of the F2 layer peak density. More directly, ionosondes can be used to determine propagation conditions on HF communications links.

Two typical ionograms produced by a standard analog ionospheric sounder using film recording techniques are shown in Figure 10-1. The frequency range is 0.25 to 20 MHz (horizontal axis), and the displayed height range is 600 km, with 100 km height markers. The bottom ionogram is typical for daytime, showing the signatures of reflections from the E, F1 and F2 layers. The cusps, seen at various frequencies (where the trace tends to become vertical) indicate the so-called critical frequencies, foE, foF1, and foF2. The critical frequencies are those frequencies at which the ionospheric sounder signals penetrate the respective layers. These frequencies are a measure of the maximum electron densities of the respective layers. Since the densities vary with time, ionospheric sounding is used to obtain information on changes in the critical frequency and other parameters of the electron density vs height profile.

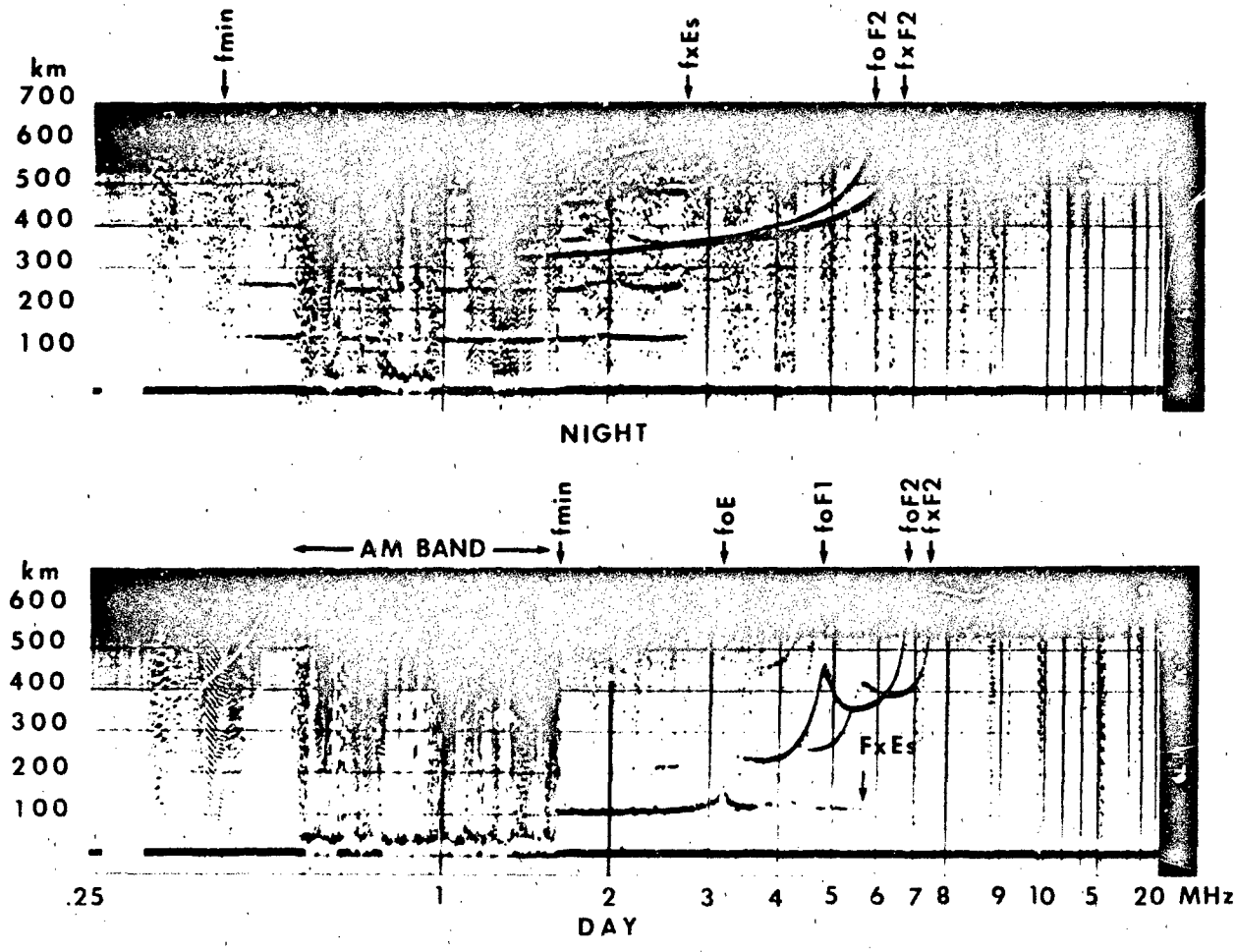


Figure 10-1. Typical midlatitude day and nighttime ionograms, recorded by a C-4 ionosonde at Boulder, Colorado. The daytime ionogram shows reflections from E, Es, F1 and F2 layers; the nighttime ionogram those from Es and F2 layers.

The ionogram (Figure 10-1) shows signatures of various phenomena that complicate the process of ionospheric sounding or the ionogram analysis. Superimposed on the primary F layer echo trace is a similar but not identical trace, shifted up in frequency: the so-called extraordinary or X component. The primary trace is called the ordinary or O component. The echo trace is split into two traces due to effects of the earth's magnetic field. A second trace similar to the primary trace is seen at twice the range, a multiple reflection. Only a small fraction of the wave energy is received by the antenna after it has returned from the ionosphere. Most of the returned energy is reflected back from the ground and provides the first multiple (second order echo) at twice the range. If the ionosphere is a good reflector, and losses in the D region are low, additional reflections can be observed. Figure 10-1 (Night) shows a second multiple (third order echo) for part of the Es trace. It is easy to see that slopes increase by a factor that corresponds to the order of the echo.

Finally, we see vertical bands in the frequency range from 0.5 to 1.7 MHz, the signature of radio frequency interference (RFI) in an ionogram, here from the AM band. RFI can become severe enough to prevent the recording of ionospheric echoes; for example, interference masks part or all of the E layer trace below 4.7 MHz.

The top of Figure 10-1 shows a typical nighttime ionogram. The E and F1 traces have disappeared because these layers dissipate after sunset. (Residual nighttime E region ionization of low density can be observed in the absence of low sporadic E layers at stations with low RFI and large antennas.) Echoes from a sporadic E layer (Es) and F2 layer echoes and their multiples are clearly visible in Figure 10-1. At times a brushlike spreading of the F2 layer cusps is observed. It is called spread F and is caused by small scale irregularities embedded in the ionosphere and ripples in the equidensity contours on the order of hundreds of meters to kilometers. For a detailed discussion of spread F see Davies [1966] and Rawer and Suchy [1967]; for a dis-

cussion of the occurrence and global distribution see Herman [1966]. The nighttime ionogram also shows increased RFI bands at higher frequencies. Because the D layer disappears at night, HF propagation over large distances is possible. This long distance propagation is heavily used for broadcasting by commercial users and for shortwave radio communications by government services and radio amateurs. Fortunately the ionosonde's own echoes also increase in amplitude due to the disappearance of the D layer, reducing to some extent the effect of increased propagated noise on the systems overall signal-to-noise ratio.

10.1.1.2 Principles of Ionospheric Sounding. The concept of ionospheric sounding was born as early as 1924, when Breit and Tuve [1926] proved the existence of an ionized layer with the reception of ionospheric echoes of HF pulses transmitted at 4.3 MHz from a remote transmitter (distance 13.8 km). This, during the next decade, led to the development of monostatic ionospheric sounders by the National Bureau of Standards and the Carnegie Institution. Even today the principles used by Breit and Tuves constitute the principles on which most ionospheric sounders are based. These are the transmission of HF pulses and the measurement of their time of flight to the reflection level. For a short historical review of the development of ionospheric sounders see Villard [1976].

Ionospheric sounding takes advantage of the refractive properties of the ionosphere. A radio wave propagating into the ionospheric plasma encounters a medium with the refractive index (in the absence of the earth's magnetic field B, and ignoring collisions between electrons and the neutral atmosphere)

$$\mu^2 = 1 - X = 1 - \left(\frac{f_N}{f}\right)^2 \qquad (10.1)$$

where

$$X = \frac{N_e e^2}{4\pi^2 \varepsilon_0 m f^2}, \qquad (10.2)$$

e, ε_0 and m are natural constants, N_e is the electron density, and f is the wave frequency. Below the ionosphere, $N_e = 0$, and $\mu = 1$. Within the ionosphere, $N_e > 0$, and $\mu < 1$. At a level where $X = 1$,

$$f_N^2 = \frac{N_e e^2}{4\pi^2 \varepsilon_0 m} = f^2 \qquad (10.3)$$

the refractive index μ becomes zero. The wave cannot propagate any farther and is reflected. The quantity f_N, which relates the electron density to the frequency being reflected, is called the plasma frequency. Inserting the natural constants into Equation (10.3) permits us to deduce the useful relation between electron density and plasma frequency (which is identical to the probing frequency being reflected)

$$f_N = 0.009 \sqrt{N_e} \qquad (10.4)$$

$$N_e = 1.24 \times 10^4 f_N^2 \qquad (10.5)$$

where f_N is in MHz and N_e in electrons/cm^3. The plasma frequency is the natural frequency of oscillation for a slab of neutral plasma with the density N_e after the electrons have been displaced from the ions and are allowed to move freely. For further discussions of the relation of μ to the wave propagation see Davies [1966].

Peak densities of the ionospheric layers vary between 10^4 and $> 10^6$ el/cm^3. Inserting these numbers into Equation (10.4) gives a plasma frequency range from 1 to > 9 MHz, this is the reason for the frequency range (0.5 MHz $\leq f \leq$ 20 MHz) covered by a typical ionosonde. The low densities of the D layer can only be probed with low frequencies < 250 Hz, requiring large antennas and complex processing/analysis techniques and are not directly measurable by the standard ionosondes (for details see Kelso [1964] and references therein). Indirectly the D region ionization is measured by the integral absorption effects that it imposes on the HF waves propagating through it to the E or F region reflection levels (see discussion of fmin).

The inclusion of the magnetic field in the formula for the refractive index leads to the well known Appleton dispersion formula (dispersion means that the refractive index depends on the propagating frequency) for a magnetized plasma, here given for the case of no collisions, generally valid for frequencies > 2 MHz, in the E and F regions.

$$\mu^2 = 1 - \frac{2X(1 - X)}{2(1 - X) - Y_T^2 \pm \sqrt{Y_T^4 + 4(1 - X)^2 Y_L^2}} \qquad (10.6)$$

with

$$Y_{L,T} = \frac{e B_{L,T}}{2\pi m f} \qquad (10.7)$$

and

$$\frac{e B}{2\pi m} = f_H \qquad (10.8)$$

where f_H is the gyrofrequency, the natural frequency at which free electrons circle around the magnetic field lines. $B_{L,T}$ are the components of the magnetic field in the direction of (*l*ongitudinal) or perpendicular to (*t*ransverse) the wave normal. Inserting the constants into Equation (10.8) leads to the useful relation for the gyrofrequency

$$f_H = 2.8 \times 10^4 B, \qquad (10.9)$$

where f_H is in MHz and B in gauss (1 gauss = 10^{-4} tesla).

The refractive index given in Equation (10.6) shows, by the $\pm$ solution to the square root, that in a magnetized

plasma two and only two "characteristic" waves can propagate. These two characteristic waves are called the ordinary or o-component and the extraordinary or x-component seen in the ionogram shown in Figure 10-1. A radio wave with arbitrary (often linear) polarization will split in the ionospheric medium into two characteristic or o-and x-components, which in general propagate independently.

The reflection condition $\mu = 0$ gives two solutions for X; for the + sign (o-component)

$$X = 1 \tag{10.10}$$

as in the no-field case, Equation (10.3); for the − sign (x-component)

$$X = 1 - Y. \tag{10.11}$$

At the reflection level for the O-component the plasma frequency equals the probing frequency, $f_N = f$. The x-component is reflected at a lower level that depends on the local magnetic field strength. It can be shown that the critical frequencies fo and fx, for $fo \gg f_H$, are related by

$$fx - fo \approx \frac{f_H}{2}, \tag{10.12}$$

that is, the magneto-ionic splitting (due to the presence of the magnetic field in the ionospheric plasma) depends on the local magnetic field strength and therefore varies, from station to station. For a typical midlatitude station, B = 0.5G and from Equation (10.9) we determine $f_H = 1.4$ MHz, leading to the fo-fx separation of ~0.7 MHz seen in Figure 10-1. A solution X = 1 + Y exists for frequencies below the gyrofrequency f_H. For details see Davies [1966].

Using ionograms to determine the true height electron density profile $N_e(h)$ is further complicated by the slowing-down effect that the ionization below the reflection level has on the group velocity of the pulse. While the phase velocity v of the wave is

$$v = \frac{c_o}{\mu(h,f)} \tag{10.13}$$

it can be shown that the group velocity u, defined as the propagation velocity of the pulse envelope, is given for the no-magnetic field case by

$$u = \frac{c_o}{\mu'(h,f)} = \mu(h,f)c_o, \tag{10.14}$$

where $\mu'(h,f)$ is the group refractive index. Therefore, while the phase velocity increases above the speed of light in a plasma, the group velocity, the velocity at which the energy propagates, slows down ($\mu < 1$ in a plasma). For a more detailed discussion see Davies [1966] and Chapter 10 of Budden [1961].

As a result, the actual reflection height h is smaller than the so-called virtual height h', which is derived, assuming propagation in the medium with the speed of light from

$$h' = \frac{c_o t}{2} \tag{10.15}$$

with t the round trip travel time of the pulse. Or since

$$u < c_o, \tag{10.16}$$

then

$$h' > h. \tag{10.17}$$

As stated before, one of the main objectives of ionospheric sounding is the determination of h(f), which through the relation between f and N_e, Equations (10.3) and (10.4) represents the desired function $N_e(h)$. Since the group travel time is

$$t = \frac{2}{c_o}\int_o^h \frac{dz}{\mu} = \frac{2}{c_o}\int_o^h \mu' dz, \tag{10.18}$$

the virtual height is related to the group refractive index by

$$h'(f) = \int_o^{h(f)} \mu'\,[f_N(z), f]\,dz. \tag{10.19}$$

If the electron density $N_e(h)$ is considered as a function of the height h above the ground, μ' is also a function of h and the problem is now to solve the integral equation (10.19), for given values of h'(f) obtained from the ionogram. The techniques used to solve this equation are known as true height analysis for which in general numerical methods are used; they are discussed in detail in a 1967 special issue of *Radio Science*.

10.1.1.3 Analog Ionosonde. The general principle of an ionospheric pulse sounder is shown in Figure 10-2 [Rawer and Suchy, 1967]. A superheterodyne technique is used to both generate the transmitted pulse of frequency f_T and to mix the received signals back to an intermediate frequency or IF for further amplification. Tuning the receiver mixer stage so that its output frequency is equal to the frequency of the fixed frequency (pulsed) oscillator f_c, and using a common variable local oscillator f_v, ensures that the receiver and transmitter are automatically tuned for every value of

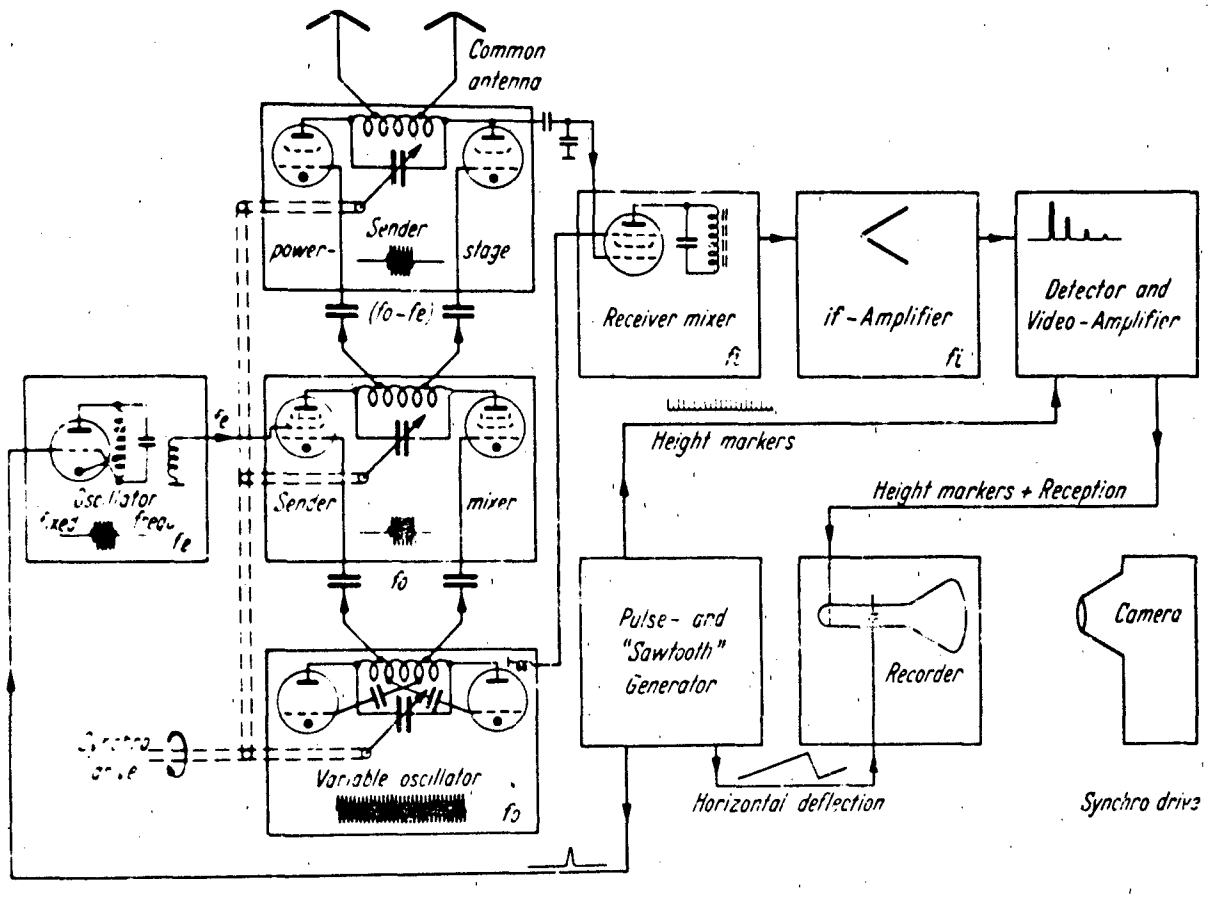


Figure 10-2. Schematic presentation of the major components of an Ionospheric Pulse Sounder.

the oscillator frequency f_o. The transmit pulse is amplified in one or several power stages and transmitted, using a suitable wide-band antenna with a vertical radiation pattern. The same antenna can be used for reception using either a tap (voltage divider) of the transmitter tank circuit or a Transmit/Receive or T/R switch, which protects the receiver input from overloading during transmission of the pulse, especially problematic for transistorized receivers. More recently separate antennas for transmission and reception have become commonplace. This permits the use of smaller and therefore less costly receiver antennas in phased arrays for angle-of-arrival measurements and as polarized antennas for polarization or mode identification [Bibl and Reinisch, 1978a,b; Wright and Pitteway, 1979].

The received signals are mixed down (or up) to the intermediate frequency and amplified in an IF amplifier, that is matched in bandwidth to the pulse width (overall bandwith B = 1/P, with P the width of the transmitted pulse). After detection and amplification, the video signal modulates the intensity of the CRTs electron beam (Z-axis modulation). Deflecting the beam in the Y-axis with a sawtooth voltage in synchronism with the transmission and pulling a film slowly in the direction of the X-axis in the focal plane of an imaging optic results in an ionogram recording such as that shown in Figure 10-1. Since sounders based on the general concepts shown in Figure 10-2 rely completely on analog techniques for signal generation, reception, and processing, they have more recently become known as analog sounders, in contrast to the digital sounders developed in several places during the last decade. Analog sounders are still operated at many ionospheric observatories, especially the well-known C3 and C4 ionosondes, which were developed by NBS and which were distributed worldwide as the primary ionosonde for the International Geophysical Year 1957/58.

10.1.1.4 Analog/Digital Hybrid Ionosonde. While vertical sounding with the transmitter and receiver and their respective antennas collocated made synchronization of transmitter and receiver relatively easy, a much more de-

manding task arose when investigators attempted to sound the ionosphere over paths of varying distances to determine the ionospheric structure and the propagation conditions directly.

If the transmitted signal is to be received within the receiver bandwidth, the systems must be started at a precise time, and must have perfectly aligned frequency scans. This was achieved using linear frequency scans and synchronous motor drives, which derived their A/C voltage from crystal oscillators [Bibl, 1963]. A large step forward was the development of frequency stepping sounders such as the Granger Path Sounder [Gowell and Whidden, 1968] which combined digital and analog techniques. Digital techniques generated ionograms by stepping synthesizer/transmitter and receiver through the desired frequency range, providing selectable frequency spacing (for example, 25, 50, or 100 kHz, linear or linear over octave bands). The frequency synthesis itself and the data processing/recording however, used the standard analog techniques. All digital and hybrid pulse sounders currently available use these frequency stepping techniques.

10.1.1.5 Digital Ionosonde. The rapid development of integrated circuits, microprocessors and especially Read-Only-Memories, and of inexpensive storage of large capacity, has led to the development of digital ionospheric sounders. These systems have some analog components, but use digital techniques for frequency synthesis, receiver tuning, signal processing, recording, and displaying of the ionograms. However, to the modern sounder, the digital control of all sounder functions, the ability to digitally control the antenna configuration, and above all, the immense power of digital real time processing of the data prior to recording on magnetic tape or printing with digital printers are of special importance.

A digital amplitude ionogram, recorded by a Digisonde 128 PS at the AFGL Goose Bay Ionospheric Observatory is shown in Figure 10-3. This system developed at the University of Lowell [Bibl and Reinisch, 1978a,b] uses phase coding, spectral integration, polarized receive antennas for o/x component identification, and fixed angle beam steering of the receive antenna array for coarse angle of arrival measurements to provide a rather complete description of the properties and origin of the reflected echoes. Using a standard set of 128 range bins for each frequency, the sounder integrates the sampled receiver output signals for a selectable number of integrations, improving the signal-to-noise ratio and providing the samples for spectral analysis. Since for each frequency-range-bin or FRB only one return is recorded, a search algorithm determines from the set of separate signals (o, x, several antenna directions, Doppler lines) the signal with the largest amplitude and retains amplitude and STATUS, that is, special signal characteristics. Using a special font [Patenaude et al., 1973], the resulting digital amplitudes are printed out providing the analog presentation essential for the recognition of the detailed structure of an ionogram trace simultaneously with the digital information. Preprocessing has largely eliminated the noise background. The bottom part of Figure 10-3 shows a digital amplitude ionogram, represented by all amplitudes above a noise level determined automatically and separately for each frequency. The noise level on each frequency can be estimated, since the unmodified signals of the lowest four height bins are shown at the bottom of the ionogram. The displayed range starts at 60 km and in 128 height increments with a Δ = 5 km covers the range to 695 km. Each frequency step is in 100 kHz, which covers the range from (nominally) 0 to 13 MHz in 130 frequency steps. Ionograms of this type can be produced in between 30 s and 2 min, depending on the complexity of signal characterization selected. The number of integrations required to achieve an acceptable signal-to-noise ratio, and the desired spectral resolution of the Doppler measurements also affect the duration of the ionogram sweep. The ionogram is similar in structure to the daytime ionogram in Figure 10-1, showing clearly an E-trace (foE = 3.25 MHz), an F1-cusp (foF1 = 5.0 MHz), and the F2 trace (foF2 = 8.2 MHz). The top part of the figure was produced by printing only those amplitudes which had a STATUS indicating o-polarization, vertical signals only. The resulting suppression of the x-component and of the (obliquely received) noise shows the effectiveness of these techniques.

The digital "HF Radar System" developed at NOAA, Boulder, Colorado [Grubb, 1979] is an ionospheric sounder, built around a minicomputer. Appropriate software allows freedom in generating the transmit signal phase coding and sequence, and in processing procedures. However, instruction execution times of the minicomputer limit this flexibility. The sounder with its present software uses an echo detection scheme rather than a fixed FRB grid to obtain the information on the ionospheric returns. This scheme requires that the return has to be identified beforehand, generally using a selectable level above the noise, and an "a hits out of b samples" criterion. This system, by the use of a receive antenna array, then determines on-line for this initially identified return, the echo amplitude, its polarization, Doppler shift, and reflector location [Wright and Pitteway, 1979].

The spectral information available in digital ionograms has been used to track moving irregularities in the equatorial [Weber et al., 1982] and polar ionosphere [Buchau et al., 1983]. An example of a Doppler ionogram recorded in the polar cap is shown in Figure 10-4. The right lower panel shows a heavily spread amplitude ionogram and superimposed two oblique backscatter traces. The right upper panel shows the Status or Doppler ionogram: each FRB displays the Doppler bin number instead of an amplitude. FRB's with an amplitude below an automatically determined noise level show neither amplitude nor status. Separating the ionogram into positive and negative Doppler ionograms permits the identification and subsequent tracking of approaching

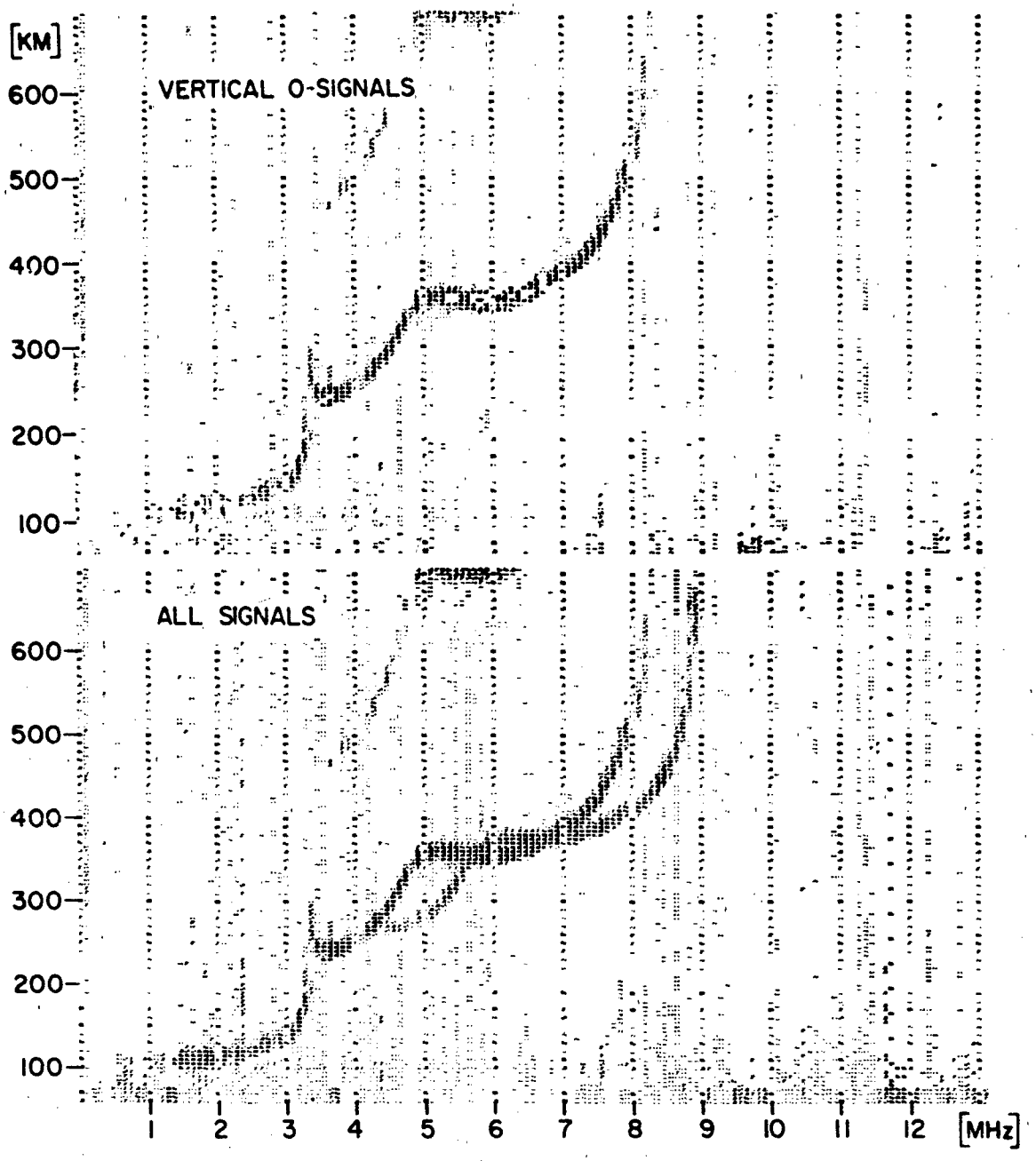


Figure 10-3. Digital daytime amplitude ionogram recorded by a Digisonde 128PS at the AFGL Goose Bay Ionospheric Observatory 16 June 1980 1720 AST. Coarse angle of arrival and polarization information is used to separate the vertical ordinary trace shown in the upper part of the figure.

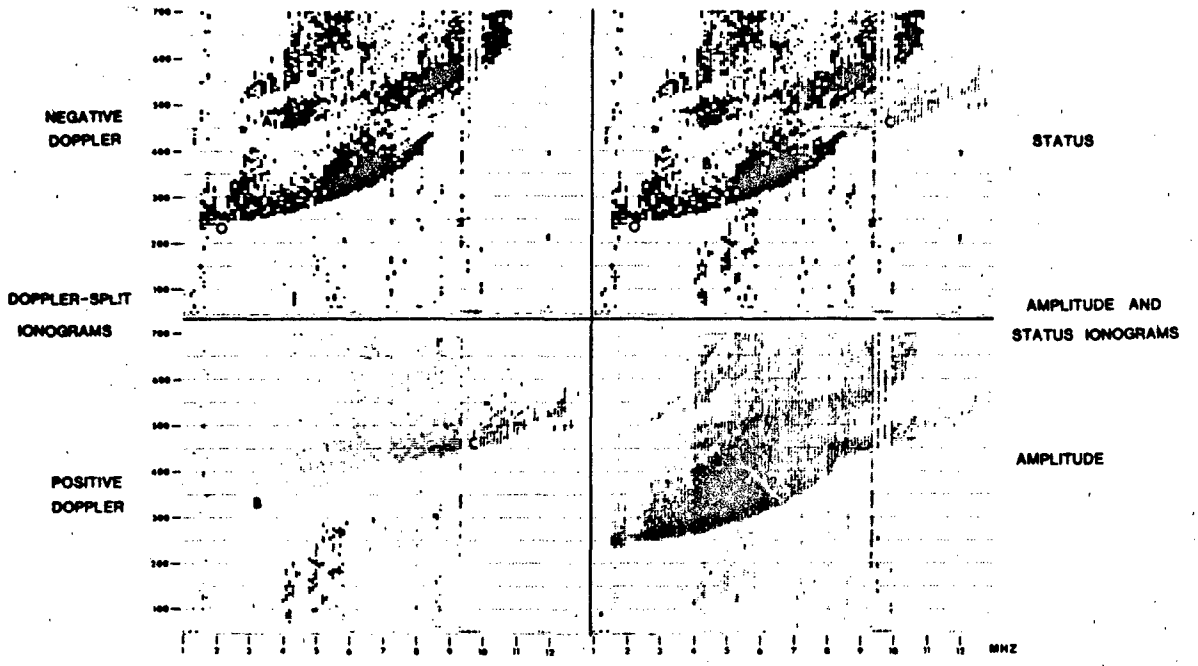


Figure 10-4. Amplitude/status ionogram taken by the AFGL airborne ionospheric observatory with a Digisonde 128PS at Thule, Greenland 9 December 2231 (UT). The lower right panel shows the amplitude ionogram after removal of radio noise. The Doppler ionogram shown in the upper right panel is produced by replacing each amplitude in the ionogram below with a number representing the measured Doppler shift. The separation into positive and negative Doppler traces (approaching and receding reflection regions) is shown in the two panels on the left.

(traces marked B and C) and receding (trace marked A) reflecting or scattering centers. The overhead trace (very low Doppler) is marked 0.

10.1.1.6 Digital Data Processing. The availability of ionograms in digital form has finally provided the basis for successful automatic processing of these complex data. Real time monitoring [Buchau et al., 1978], survey of large data bases [Reinisch et al., 1982a], real time analysis of ionospheric parameters [Reinisch et al., 1982b], automatic trace identification and true height analysis [Reinisch and Xuequin, 1982, 1983] have been made possible by the availability of data in digital form. Analysis concepts for angle-of-arrival determination and other parameters for the NOAA/SEL digital sounder have been presented by Wright and Pitteway [1982] and by Paul [1982].

An example of a data survey presentation using digital ionosonde data from Goose Bay is shown in Figure 10-5. The top row shows the integrated height characteristic, obtained by collapsing each ionogram onto its height axis. This characteristic provides the history of E and F layer (minimum) height variations over the course of three days. The middle panel shows the temporal changes of the F layer returns, with the lower envelope determined by foE (daytime) or fmin (nighttime), while the upper envelope is determined generally by foF2. The bottom panel representing the frequency extent of E and Es traces shows the typical cos X (X = solar zenith angle) pattern of the solar E-layer, maximizing at noon. Sporadic E events observed on all three nights are typically observed at these high latitudes during auroral storms [Buchau et al., 1978].

10.1.1.7 FM/CW or Chirp Sounder. The availability of very linear sweep-frequency synthesizers resulted in the development of FM/CW (frequency modulated continuous wave) or Chirp Sounders, initially for oblique incidence and in the 1970s also for monostatic vertical incidence sounding [Barry, 1971]. A linear waveform with the constant sweep rate df/dt is transmitted. Receiving the waveform after propagation to the ionosphere and back and measuring the time delay of each frequency component against the original waveform permits the determination of the travel time as a function of frequency. This is actually done by mixing the received waveform with the original, resulting in a difference frequency which can be measured by spectrum analysis. The difference frequency as a function of frequency (the "Chirp") is proportional to the travel time of the signal as a function of frequency; therefore, a graph of the difference frequency as a function of time or transmitted frequency, through the known sweep rate df/dt, forms an ionogram. While initially transmitter and receiver were separated by a substantial distance, to avoid overloading of the re-

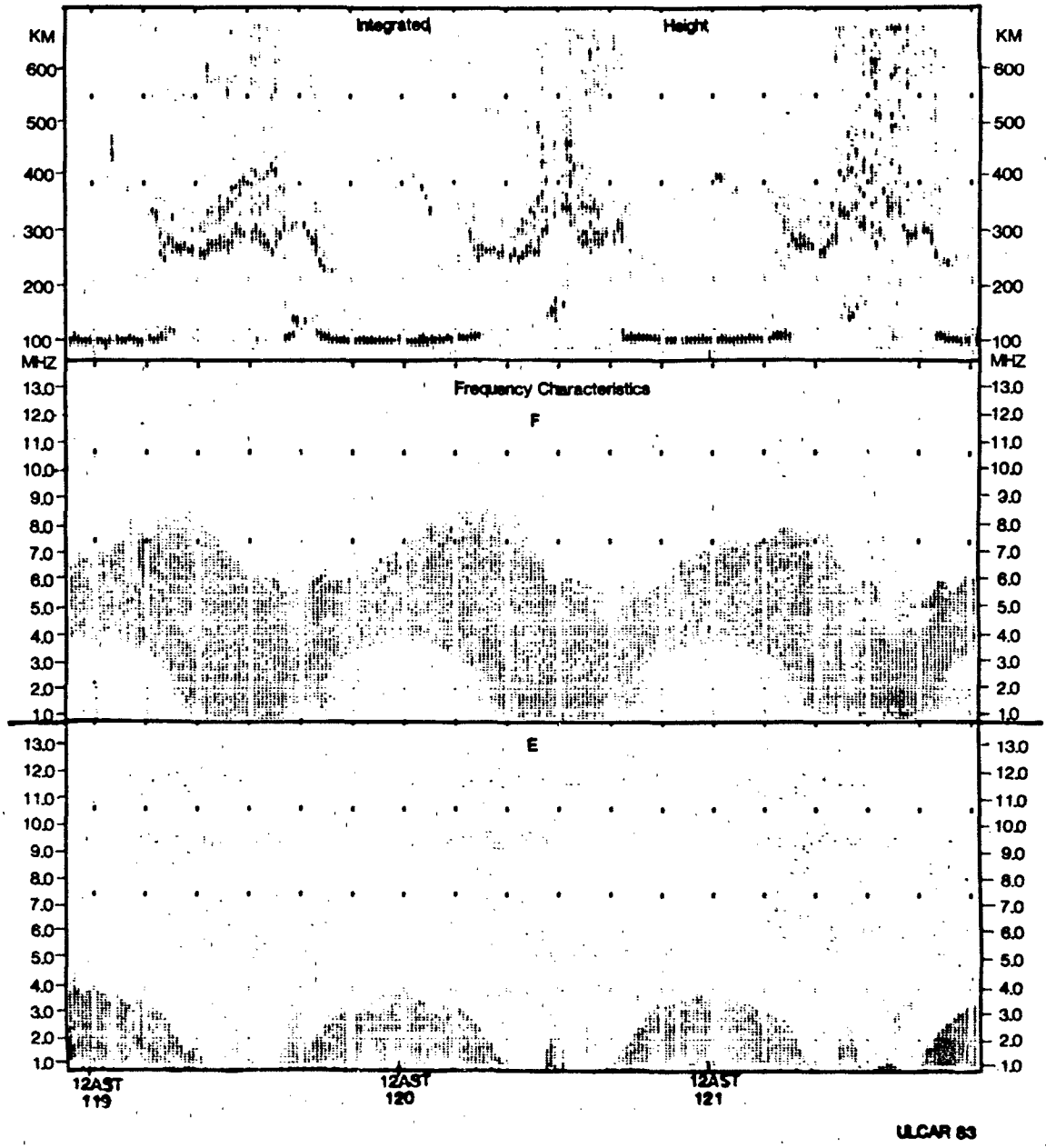


Figure 10.5 Ionospheric characteristics spanning three days produced from digital ionograms recorded at Goose Bay 28–30 April 1980. The integrated height characteristic shows the dynamic changes of the minimum height of the F layer and the appearance of the solar and sporadic E layers. The F and E frequency characteristics show the diurnal variability of these layers as well as evidence of some auroral events.

ceivers with the unwanted direct signal, a monostatic system was developed, using a T/R switch and a quasi-random interruption of the linear waveform transmission. The main advantage of the FM/CW system is the very narrow instantaneous bandwidth of the transmitted signal, allowing a similarly narrow receiver bandwidth (nominal 100 Hz at sweeprates of 20 kHz s). This bandwidth is further reduced by spectrum analysis to an effective bandwidth of the order of 1Hz.

Although the digital integration and spectral analysis used in the modern digital pulsed ionosondes decreases the effective bandwidth of a pulse receiver significantly (by a

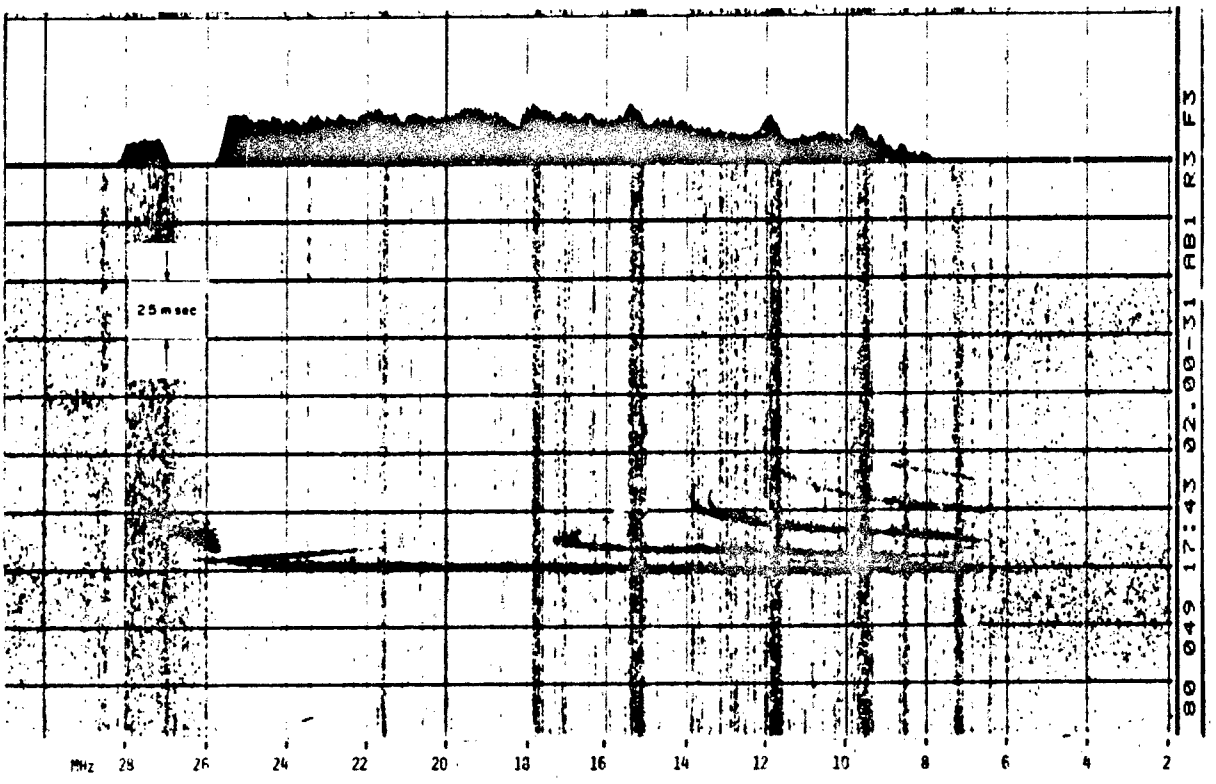


Figure 10-6. FM/CW or chirp-oblique propagation ionogram over a 2400 km path. Transmitted from an aircraft over the North Atlantic and recorded at a receiver site in Maine.

factor between 10 and 100, depending on the integration time and processing procedures) from the original 20 kHz bandwidth for a typical 50 μs pulse and makes the actual signal-to-noise ratio comparable to that of a chirp sounder, the FM/CW system allows a substantial reduction of the peak power of the transmission. Good FM/CW ionograms have been obtained with transmit power as low as 1W (CW). The FM/CW system is definitely a good solution for the already strained occupancy of the HF spectrum. However, several drawbacks have limited its application as a vertical sounder. motion of the reflecting region imposes a Doppler shift, which introduces range errors that are intolerable under arctic conditions. The important Doppler shift of the echoes cannot be obtained. However, due to their low transmit power and superior S/N characteristics, many FM/CW systems are used as backscatter sounders and as oblique incidence or propagation sounders. A typical propagation ionogram showing extensive multihop structure is shown in Figure 10-6. The ionogram was obtained over a 2400 km path with the transmitter located on board an aircraft and the receiver located at the AF OTH-B radar site at Columbia Falls, Maine, U.S.A.

10.1.1.8 Topside Sounders. Shortly after the onset of the satellite era, an ionospheric topside sounder satellite, Alouette I, was launched into space on 29 September 1962. Since then a series of follow-on improved topside sounders were launched, either as orbiting ionospheric sounders (Alouette II, 1965) or integrated with a complement of *in situ* measuring geophysical instruments (ISIS I, 1969 and ISIS II, 1971). For details on these sounders see Franklin and MacLean [1969].

Since groundbased ionosondes obtain ionospheric echoes only up to the peak of the layer with the highest electron density (generally the F2 layer), knowledge about the topside ionosphere was based on theory, the few incoherent scatter facilities, and sporadic high-altitude rocket flights. The topside sounders therefore filled a gap in the techniques available to measure the global ionosphere in its entirety. During the approximately 15 years that these topside sounders have provided continuous data, the many millions of ionograms that have been accumulated have contributed substantially to our picture of the global ionosphere [Chan and Colin, 1969; Thomas and Rycroft, 1970] and of its subregions, such as the high-latitude ionosphere [Jelly and Petrie, 1969; Thomas and Andrews, 1968], the F region trough [Muldrew, 1965; Pike, 1976], and the equatorial ionosphere [Eccles and King, 1969]. The topside sounder provides an h'f trace from the satellite altitude down to the peak of the F layer (except for those rare occasions where

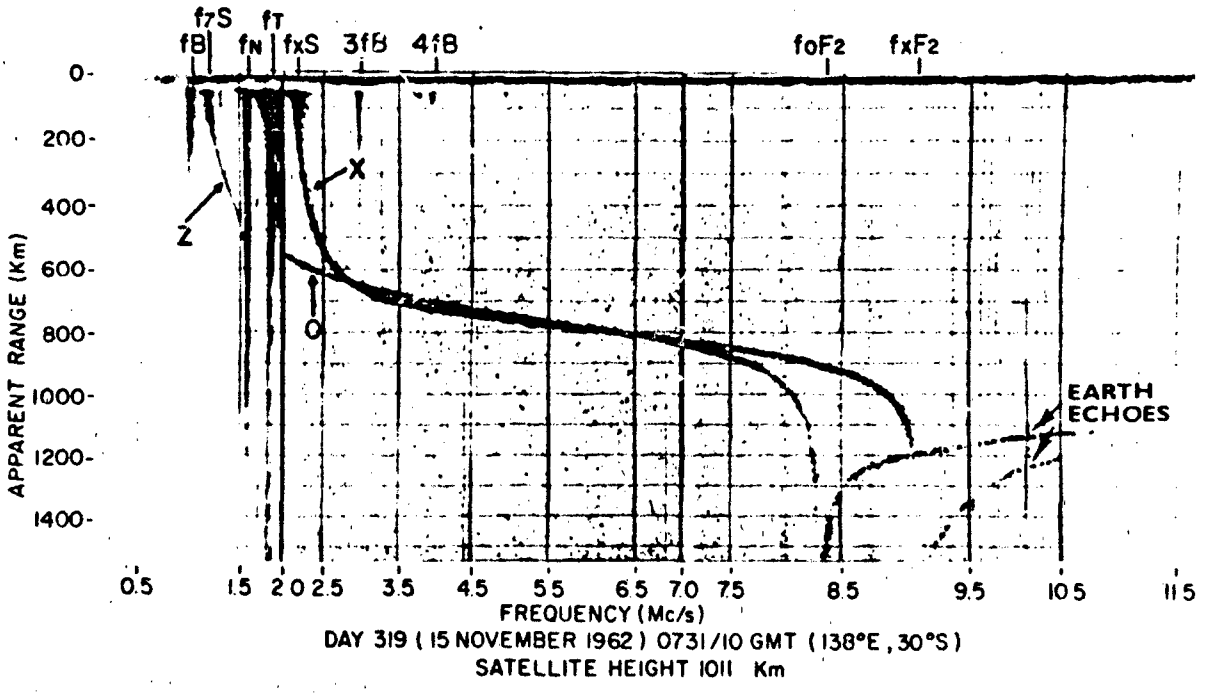


Figure 10-7. An Alouette I topside ionogram illustrating Z-, O- and X-wave traces, cutoffs, resonance spikes, and earth echoes

ionization of lower layers exceeds the maximum density of the F2-layer). A typical topside ionogram is shown in Figure 10-7 from the *URSI Handbook of Ionogram Interpretation and Reduction* [UAG-23, 1972]. A unique phenomenon observed in topside ionograms are the ionospheric resonance spikes due to the excitation of the ambient plasma by the transmissions. The most frequently observed resonance spikes occur at the (local) plasma frequency f_N, at the local gyrofrequency f_H (labeled f_B in Figure 10-7), at the hybrid frequency

$$f_T = \sqrt{f_N^2 + f_H^2} \qquad (10.20)$$

and at certain harmonics of these frequencies [Hagg et al., 1969].

Many of the references given here and a large amount of further material can be found in the special issue on topside sounding of the *Proceedings of the IEEE* [1969].

10.1.1.9 Ionogram Interpretation. The behavior of the ionosphere is often very dynamic. This fact and the large range of electron densities, over which the ionospheric layers change from day to day, from day to night, with season and with solar cycle result in a large variety of ionograms. There are also extreme differences in ionospheric variations and structures from the equators to the poles and in the regular or sporadic appearance and disappearance of the lower layers. Dynamic effects that shape the profile along the ray path and specifically in the vicinity of the reflection region also affect the group delay at each frequency and therefore the final appearance of the h'f-trace. This trace is sometimes further complicated by ionospheric irregularities and oblique returns. All these factors combined ensure an incredible variety of ionograms. To capture their geophysically significant parameters, a large number of rules and definitions have evolved over the decades, which after acceptance by the International Radio Science Union (URSI) have been published as the *URSI Handbook of Ionogram Interpretation and Reduction*, [UAG-23, 1972] governing the analysis of ionograms at all ionospheric stations. This set of rules, resulting from the still continuing or terminated operation of more than 300 ionosonde stations distributed over the whole globe, has produced a rather uniform analyzed data base which is archived at the World Data Centers for Solar Terrestrial Research located at Boulder, Colorado (WDC A), Izmiran, USSR (WDC B), Tokyo, Japan (WDC C1) and Slough, UK (WDC C2). With some exceptions, the individual world data centers store data originating in their respective regions. WDC A stores the data from the western hemisphere and also data from France and India.

To provide special instructions for the analysis of the extremely complex ionograms from high latitude stations, a *High Latitude Supplement to the URSI Handbook on Ionogram Interpretation and Reduction* has been published [UAG-50, 1975].

For special research efforts, it is often essential to go back to the source data, the ionospheric films of a specific station(s). For the western hemisphere, these films are stored

at the World Data Center A for Solar Terrestrial Physics, NOAA/NGSDC, Boulder, Colorado. A *Catalogue of Ionosphere Soundings Data* [UAG-85, 1982] provides access to this data base, which spans the period from 1930 through today. The longest and still continuing operation of an ionosonde station started at Slough, UK in January 1930. Continuous operation starting before 1940 is still ongoing at Canberra, Australia (1937); Heiss Island, USSR (1938); Huancayo, Peru (1937); Leningrad, USSR (1939); Tomsk, USSR (1937); and Tromso, Norway (1932).

To provide an overview of some of the more important ionospheric parameters that can be derived from an ionogram and introduce their geophysical meaning, two ionograms are provided in the form of a sketch (Figure 10-8), and the parameters are identified. Both ionograms depict the same ionospheric conditions (taken from Figure 10-1) with the exception of an Es layer that can suddenly appear, possibly as the result of a windshear at E layer heights. This Es layer can obscure parts of the trace from reflections at higher regions of the ionosphere. A list of parameters and their identification and interpretation is provided here as a general reference and not as a guide for ionogram analysis. For detailed instructions in the evaluation of ionograms please refer to the *URSI Handbooks UAG-23 and UAG-50*.

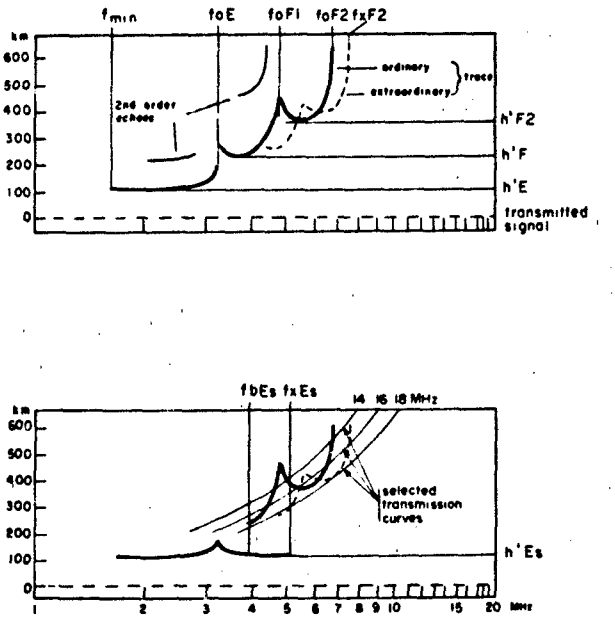


Figure 10-8. Line sketch of daytime ionogram shows definition of important ionogram parameters.

Parameter	*Meaning/Comments*
a) Critical and characteristic frequencies	
foF2	F2 layer ordinary wave critical frequency. A measure of the maximum density Ne_{max} of this layer [see Equation (10.5)].
fxF2	F2 layer extraordinary wave critical frequency. Can be used to infer foF2 using Equation (10.12) if foF2 is obscured by interference.
foF1	F1 layer ordinary wave critical frequency. This layer is often smoothly merging with the F2 layer resulting in the absence of a distinct cusp and in difficulties of determining the exact frequency (L condition).
foE	solar produced E layer ordinary wave critical frequency. *Comment:* Extraordinary wave returns exist for all layers. However, absorption of the extraordinary component is stronger than that of the o-component and the x-trace of the E layer is rarely, that of the F1 layers not always observed.
fbEs	Es layer blanketing frequency. Returns from higher layers are obscured by the Es layer up to this frequency. This frequency corresponds closely to the maximum plasma density in the (thin) Es-layer [Reddy and Rao, 1968].
fxEs	Highest frequency at which a continuous Es trace is observed.
foEs	foEs can be inferred, applying Equation (10.12). If fbEs < foEs, the layer is semitransparent. Es and higher layers are both observable. The determination of foEs and fxEs for all cases is subject to a complex set of rules beyond the scope of this outline (see URSI Handbook on Ionogram Interpretation). Modern Sounders, using polarized receive antennas, permit unambiguous foEs determination.

Parameter	*Meaning/Comments*
f_{min}	Minimum frequency at which returns are observed on the ionogram. Since radio wave energy is absorbed in the D region according to an inverse square law (Absorption ~ $1/f^2$), the variation of f_{min} is often used as a coarse indicator of the variation of D region ionization. f_{min} is not an absolute value (as for example foF2), but depends directly on the transmitted power and the antenna gain. Comparison between stations, therefore, can be only qualitative.
b) Virtual heights	
h'F	The minimum virtual height of the ordinary wave F trace taken as a whole. Due to the effects of underlying ionization and profile shape on the travel time of the pulse, these minimum virtual heights are only useful as coarse and "relative" height classifiers (high, average, or low layer, compared to a reference day). True height analysis must be made to give more meaningful height parameters, such as the height of the layer maximum ($h_{max}F2$).
h'F2	The minimum virtual height of the ordinary wave F2 layer trace during the daytime presence of the F1 layer. When an F1 layer is absent, the minimum virtual height of the F2 layer is h'F, defined above.
h'E	The minimum virtual height of the normal E layer, taken as a whole.
h'Es	The minimum virtual height of the trace, used to determine foEs.
hpF2	The virtual height of the ordinary wave mode F trace at the frequency 0.834 × foF2. For a single parabolic layer with no underlying ionization this is equal to the height of the maximum of the layer, h_{max}. In practice hpF2 is usually higher than the true height of the layer maximum. Useful as a rough estimator of h_{max} but strongly affected by a low foF2/foF1 ratio ($\leq$ 1.3).
MUF(3000)F	A set of "transmission curves" [Davies, 1966 and 1969] developed for a selected propagation link distance (the URSI standard is 3000 km) permits the determination of the *M*aximum *U*sable *F*requency, which the overhead ionosphere will permit to propagate over the selected distance. The MUF is determined from the estimated transmission curve tangential to the F-trace. For this ionogram MUF(3000)F would be 17.0 MHz.

10.1.1.10 Ionosonde Network. Even though the routinely operating ionosondes forming the worldwide network are independent, generally operated as subchains or as individual stations by national or private organizations, their operation is coordinated by the "Ionospheric Network Advisory Group (INAG)", working under the auspices of Commission G (On the Ionosphere), a Working Group of the International Union of Radio Science (URSI). INAG publishes the "Ionospheric Station Information Bulletin" at varying intervals. The Bulletin provides a means of exchanging experiences gained at the various ionospheric stations, discusses in detail difficult ionograms for the benefit of all participants, and disperses information on new systems, new techniques, special events (for example, eclipses), relevant meetings, and general network news. URSI's International Digital Ionosonde Group (IDIG), which provides a forum for the discussion of standardization proposals, for the exchange of software, and for the general exchange of experiences with these rather new and still maturing systems has been incorporated into INAG as of September 1984. The INAG bulletin can be obtained from the World Data Center A, Boulder, Colorado, 80303.

With the advent of modern digital ionosondes and on-site automated processing, a carefully planned network of remotely controllable ionosondes can provide ionospheric data and electron density profiles to a control location for real time monitoring of ionospheric and geophysical conditions. Automatic oblique propagation measurements between stations of the link can increase manyfold the number of ionospheric points that can be monitored. Considerations for the deployment of a modern ionosonde network have recently been presented by Wright and Paul [1981]. Operational and technical information on the individual stations of the world wide network of ionosondes, as well as their respective affiliations and addresses, are available in the *Directory of Solar Terrestrial Physics Monitoring Stations* [Shea et al., 1984]. Figure 10-9, taken from the report in preparation, shows the locations of all ionosondes reported

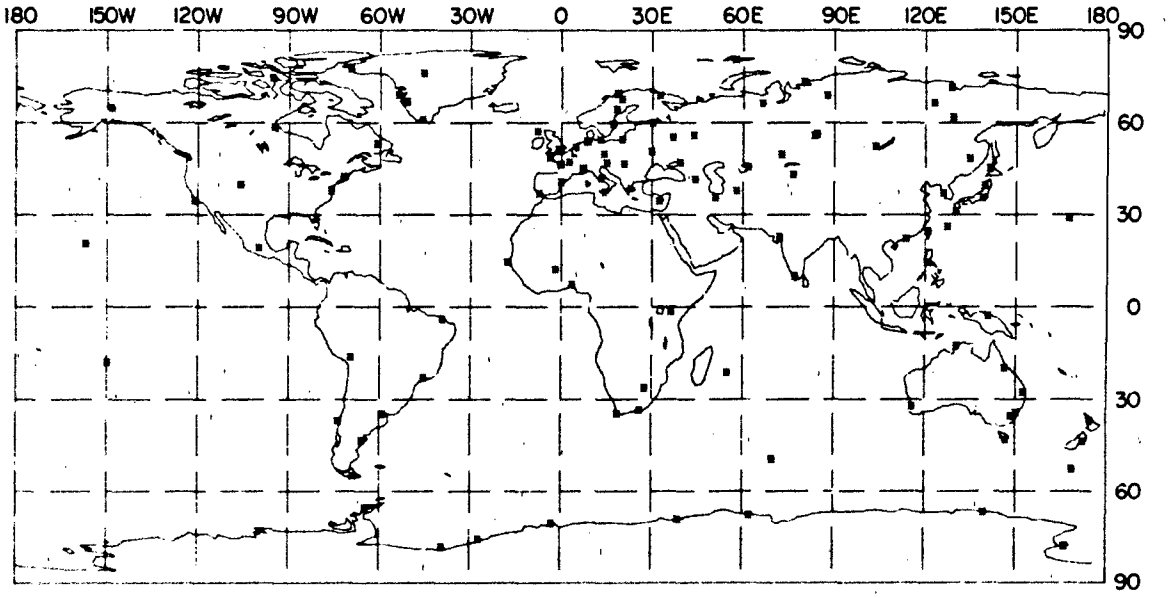


Figure 10-9. Map of vertical incidence ionospheric sounder stations 1984.

as operational or operating in 1984. World Data Center A Report. UAG-85, lists all past and present ionospheric observatories.

10.1.2 Incoherent Scatter

J.J. Thomson [1906] showed that single electrons can scatter electromagnetic waves, and that the energy scattered by an electron into unit solid angle per unit incident flux is given by $(r_e \sin\psi)^2$ where r_e is the classical electron radius $(= e^2/\varepsilon_o m_e c^2 = 2.82 \times 10^{-15}$ m) and ψ is the polarization angle, that is, the angle between the direction of the incident electric field and the direction of the observer. Thus the radar backscatter ($\psi = \pi/2$) cross-section of a single electron will be $\sigma_e = 4\pi r_e^2$. Gordon [1958] first proposed that by the use of a powerful radar operating at a frequency $f \gg$ foF2 where foF2 is the plasma frequency at the peak of the F2 layer, the backscattered power from the electrons in the upper atmosphere should be detectable. The measurement of scattered power and its characteristics as a function of altitude was expected to provide a measurement of the various geophysical parameters both in the bottomside and the topside ionosphere. Gordon assumed that the electrons were in random thermal motion of the same type as the motion executed by neutral particles so that the radar would detect scattering from individual electrons that are random in phase or incoherent. This is known as incoherent scatter or Thomson scatter (for a comprehensive review, see Evans [1969]). Gordon calculated the backscattered power per unit volume to be $N\sigma_e$, where N is the electron number density. He also predicted that the spectrum of the scattered signal will be Doppler broadened by the random electron thermal motion. The spectrum of the scattered signal was expected to be Gaussian with center to half-power width of 0.71 Δf_e where Δf_e is the Doppler shift of an electron approaching the radar at mean thermal speed so that

$$\Delta f_e = \frac{1}{\lambda}\left(\frac{8kT_e}{m_e}\right)^{1/2} \text{ Hz}, \qquad (10.21)$$

where λ is the radar wavelength (m), k is Boltzmann's constant $(= 1.38 \times 10^{-23}$ J/K), T_e is the electron temperature, and m_e is the mass of an electron $(= 9.1 \times 10^{-31}$ kg). At a wavelength $\lambda = 1$ m, and $T_e = 1600$ K, 0.71 $\Delta f_e \approx 200$ kHz. Soon after Gordon [1958] proposed the feasibility of the incoherent scatter radar experiment to study the upper atmosphere, Bowles [1958] was able to detect radar echoes from the ionosphere. The echoes resembled the predicted ionospheric scatter signal except that the bandwidth of the signal was considerably less than the predicted value. The decrease of the bandwidth of the scatter signals contributing a larger signal power per unit bandwidth obviously made it easier to detect the signal. Bowles [1958] correctly surmised that the presence of ions causes a reduction of the bandwidth of the scattered signal. Later theoretical work [Fejer, 1960; Dougherty and Farley, 1960; Salpeter, 1960; Hagfors, 1961] showed that the spectral form of the scattered signal is dictated by the radar wave-

length in relation to the Debye length in the upper atmosphere. The Debye length (D) for electrons is defined as

$$D = (\varepsilon_o kT_e/4\pi N_e^2)^{1/2} \text{ m} \qquad (10.22)$$

where ε_o is the permittivity of free space ($= 8.85 \times 10^{-12}$ F/m), e is the charge on an electron ($= 1.6 \times 10^{-19}$ C), k is the Boltzmann's constant, N_e is the electron density (m^{-3}) and T_e is the electron temperature (K). The Debye length for the electrons in the ionosphere is typically of the order of 1 cm or less below 1000 km and it is not possible to sustain organized motion at scales smaller than these values.

It was shown that, in general, the spectrum of the scattered signal consists of two parts, one due to the ions and the other to electrons. If the radar wavelength is much smaller than the Debye length, the scattered energy is entirely due to the electronic component and the initial predictions of Gordon [1958] for the scattered power ($N\sigma_e$) and its spectrum (Δf_e) are valid. On the other hand, for radar wavelengths much larger than the Debye length, which represents the experimental situation, the electronic component decreases and appear as a single plasma line at a Doppler shift approximately equal to the plasma frequency of the medium. Under this condition, the largest part of the scattered energy resides in the ionic component and the spectral width is controlled by the Doppler shift Δf_i for an ion approaching the radar at the mean thermal speed of the ions, given by

$$\Delta f_i = \frac{1}{\lambda}\left(\frac{8kT_i}{m_i}\right)^{1/2} \text{Hz}, \qquad (10.23)$$

where m_i is the mass of the dominant positive ion and T_i is the ion temperature. Considering $T_e \approx T_i$, and the dominant ion to be O^+, $\Delta f_i \approx 2 \times 10^{-2} \Delta f_e$. The echo energy is, therefore, mainly concentrated in a relatively narrow spectral window rendering the radar investigation feasible with apparatus of much lower sensitivity than initially envisaged. Figure 10-10 shows how the spectral shape depends on a parameter $\alpha = 4\pi D/\lambda$ for the case $T_e = T_i$. For $\alpha \geq 10$, the scattered energy is entirely due to the electronic component, whereas for very small values of α, the electronic component decreases and the energy appears mainly in the ionic component with a much smaller bandwidth. The electronic component now appears as a single line, known as the plasma line, at a Doppler shift approximately equal to the plasma frequency of the medium.

In the ionosphere, the electrons and the ions are at different temperatures and the spectrum of the ionic component changes for different values of the ratio T_e/T_i at a given value of α. This is shown in Figure 10-11 for the case $\alpha = 0.1$ for O^+ ions which illustrates the double-humped form of the spectrum. By measuring the height of the hump at the wing relative to the center of the spectrum it is possible to estimate T_e/T_i and the total scattering cross-section due to the ionic component is simply obtained from the area under the curves. For small values of T_e/T_i, which

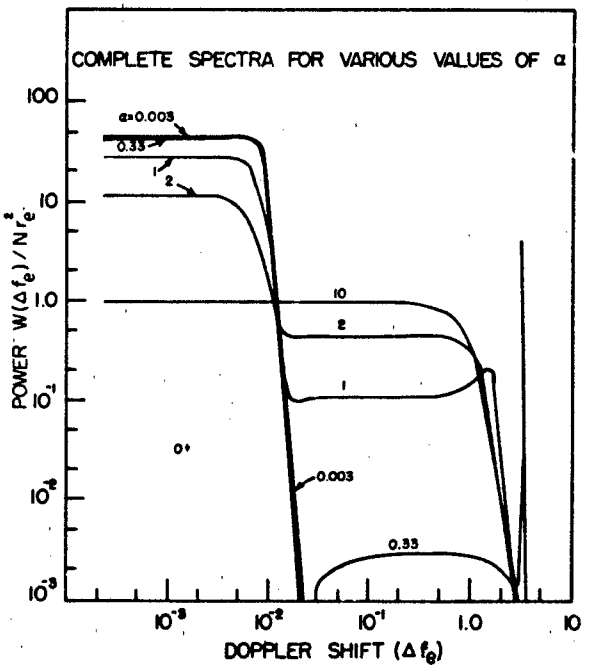


Figure 10-10. The variation of the overall spectrum for different values of the ratio $\alpha (= 4\pi D/\lambda)$. The ion has been assumed O^+. These curves assume that collisions are negligible and that $T_e = T_i$ [Hagfors, 1961].

is encountered in the ionosphere, the total scattering cross-section (σ) may approximately be given by

$$\sigma = \frac{\sigma_e}{(1 + \alpha^2)(1 + \frac{T_e}{T_i} + \alpha^2)}. \qquad (10.24)$$

The incoherent scatter radar technique opened up the possibility of *in situ* sampling of a wide range of upper atmospheric parameters by the use of a powerful ground-

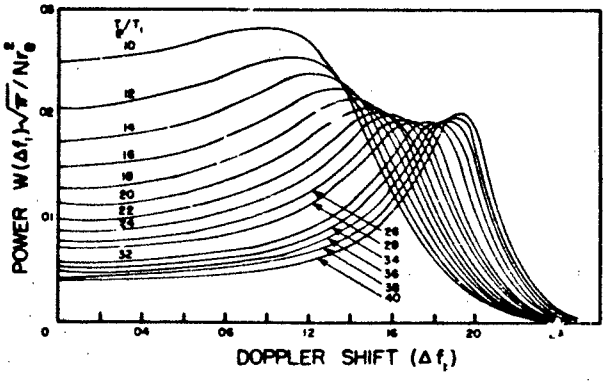


Figure 10-11. Spectra of the ionic component for the case of $\alpha (= 4\pi D/\lambda) = 0.1$ [Evans, 1969] (Reprinted with permission from IEEE © 1969.)

based radar system [Evans, 1969]. The most obvious measurement is the electron density (N) versus altitude (h) profile made by recording the variation of echo power P_r as a function of delay by using a vertically directed pulsed radar. The echo power is given by

$$P_r = C \frac{N(h)\,\sigma(h)}{h^2}, \qquad (10.25)$$

where C is a constant. The constant C can be determined either by a careful determination of the radar parameters or by an absolute determination of N at an altitude by an ionosonde or other techniques. However, as mentioned earlier (Equation 10.24), the scattering cross-section $\sigma(h)$ depends both on α and T_e/T_i which are both functions of altitude. From a measurement of the scattered energy spectrum, these corrections can be introduced and electron density profiles are determined. It has also been possible to obviate this difficulty entirely by the use of Faraday rotation technique. Figure 10-12 illustrates the electron density profile extending to almost one earth radius obtained at Jicamarca by this technique. In addition to the rather straightforward measurement of electron density profiles, electron and ion temperatures, ion composition, and photoelectron flux, the ionospheric electric field and a variety of other upper atmospheric parameters have been successfully measured at various locations extending from the magnetic equator to the auroral zone [Radio Science, special issue, 1974].

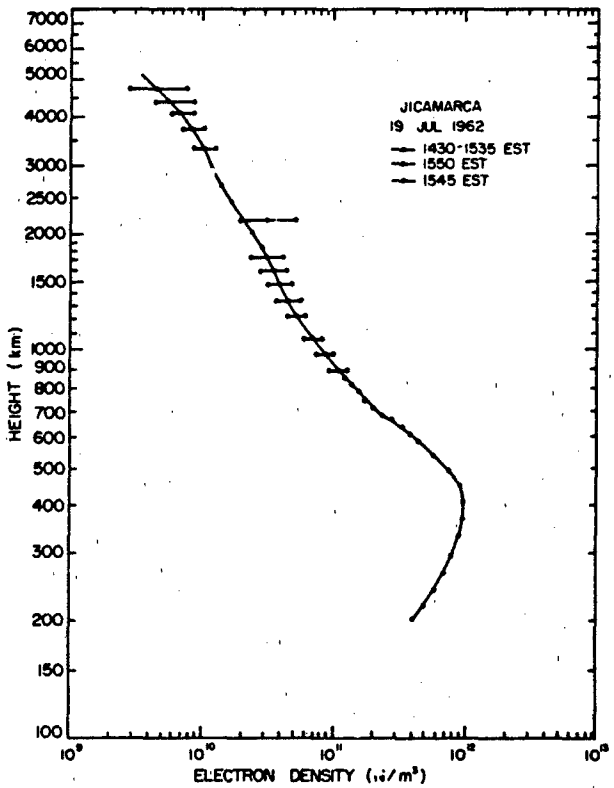


Figure 10-12. An electron density profile obtained at Jicamarca that extends to almost one earth radius [Bowles, 1963].

Table 10-1. Incoherent scatter facilities.

Location	Frequency (MHz)	Power (MW)	Antenna	Dip latitude (°N)
Jicamarca, Peru	50	6 Pulsed	290 m × 290 m array	1
Arecibo, Puerto Rico	430	2 Pulsed	300 m spherical reflector	30
St. Santin, France	935	0.15 Continuous	20 m × 100 m reflector	47
Millstone Hill, USA	440 1300	3 4 Pulsed	68 m 25 m parabola	57
Sondreström, Greenland	1300	5 Pulsed	32 m parabola	71
EISCAT*				
Transmitter:				
Tromso, Norway	224 (monostatic) 933.5 (tristatic)	5 2	30 m × 40 m parabola cylinder 32 m parabola	67
Receiver:				
Tromso, Norway				
Kiruna, Sweden			32 m parabola	
Sodankyla, Finland				

*European Incoherent SCATter facility

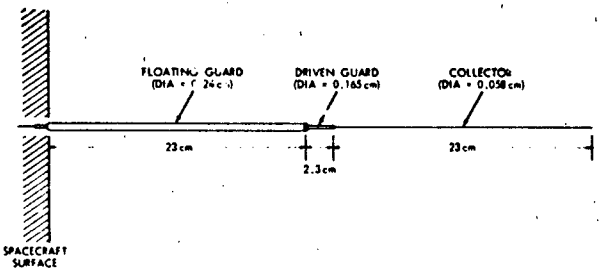


Figure 10-13. Cylindrical Langmuir Probe.

Table 10-1 adopted from Hargreaves [1979] gives a list of the incoherent scatter facilities now in operation and the characteristics of the radar system.

10.1.3 Langmuir Probes

One of the simplest devices used on rockets or satellites to measure the ionospheric density *in situ* is the Langmuir probe, named for Irving Langmuir, who pioneered the method at General Electric in the 1920s. The density is determined from a measurement of electric current passing between two conducting surfaces in contact with the environment. A varying electrostatic potential placed between the two surfaces causes the current to vary. The magnitude of the current indicates the density of the ionospheric plasma, and the change in current with respect to changes in the potential between the surfaces indicates the ion and electron temperatures. The double-floating-probe, which is the closest version to an idealized Langmuir probe, usually consists of a conductor at each end of a dipole antenna flown on a rocket or satellite for other purposes. The major disadvantage of a double-floating-probe is that ion thermal velocity is much lower than the electron thermal velocity and the rocket or satellite speeds. Therefore, the usual Langmuir probe is a single probe to measure only electrons; the other probe is replaced with the entire exposure conducting surface of the rocket or satellite. If the exposed conducting surface of the rocket or satellite is much greater than the area of the probe, the potential of the surfaces will remain fixed as the potential on the probe is swept. As a minimum, the area of the conducting surfaces should be 100 times greater than the area of the probe, and ideally the area should be 10 000 times greater then the area of the probe. By setting the potential of the probe very positive (+1.5v to +20v), all electrons within a few Debye lengths of the probe will be drawn in and measured; this allows a direct measurement of plasma density oscillations which are directly related to plasma turbulence.

The two most common shapes for Langmuir probes are the cylindrical probe (Figure 10-13) and the spherical probe (Figure 10-14). Any shape probe is possible, but these shapes are the easiest to analyze mathematically.

10.1.4 Faraday Cups for Rockets and Satellites

The most commonly used device for measuring the thermal ions is the Faraday cup (see Figure 10-15). It is usually an aperture that is a section of a flat, infinite surface in contact with a plasma. A screen across the aperture shields electrostatic potentials inside the sensor from the outside environment. The arrangement of grids or screens inside the sensor is determined by the function of the sensor. Most Faraday cups use a suppressor screen in front of the collector. This screen has a large negative potential (−10V to −100V) to repel electrons from the environment away from the collector and to drive secondary and photoelectrons from the collector back to the collector.

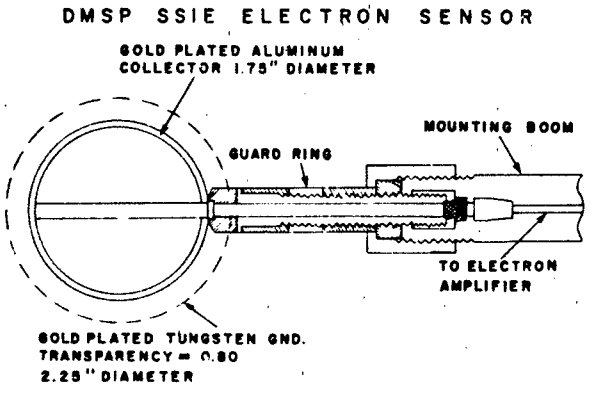


Figure 10-14. The Spherical Langmuir Probe on the DMSP Satellite.

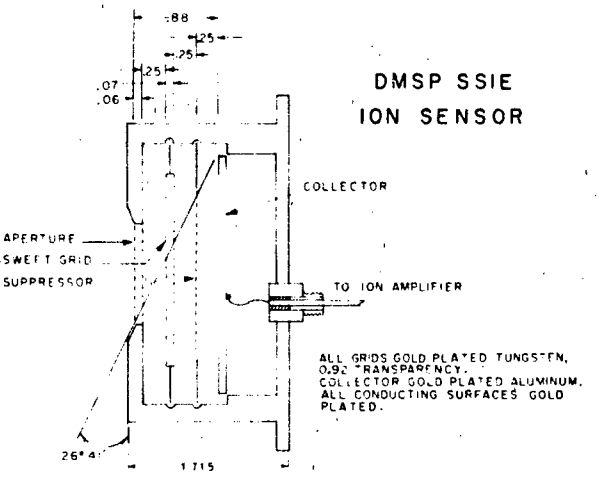


Figure 10-15. The Faraday Cup used on the DMSP Satellite.

10.1.5 Optical Measurements

Ground, airborne, satellite, and rocket based optical measurements are commonly used to determine ionospheric structure and dynamics. While a number of different instruments are employed, all analysis techniques must relate spectral emission features to ionospheric structure and dynamic processes. This is done through a knowledge of the atmosphere/ionosphere chemistry that leads to the measured photon emission. Ionospheric domains are conveniently divided into regions that are produced or influenced by energetic particle precipitation (auroral regions) and those controlled mainly by solar ionizing radiation (equatorial and midlatitude). Optical measurements have played important roles in both regions in defining the spatial and temporal characteristics of ionospheric plasma. Commonly used observing techniques will be discussed followed by a section describing important results.

10.1.5.1 Observing Techniques. Optical instruments can be classified according to spectral resolution as low, medium, and high.

Low Resolution Systems: The all sky camera has historically been used to measure auroral structure. This is perhaps the lowest resolution system, measuring all wavelengths over the sensitivity range for the type of film used (typically Kodak TRI X). The system uses a 160° field of view lens to measure auroras over a circle of 1000 km diameter in the lower ionosphere (110 km altitude). All sky cameras typically measure only bright auroral features, primarily at E region altitudes.

Photometers are low resolution systems. They rely on narrow band interference filters to isolate spectral lines and bands of interest. Meridian scanning photometers use a narrow (0.5° to 2.5°) field of view and ~2 A° Filters to measure absolute intensity of auroral and airglow along a vertical circle, commonly aligned along a magnetic meridian. Tilting filters use the change in transmitted wavelength versus tilt angle to perform a limited wavelength scan. This allows separation of non-spectral continuum from the line or band emission.

More recently, all-sky imaging photometers have been developed to perform all-sky (155°) monochromatic measurements at high sensitivity (20 Rayleighs). These employ slightly wider (~ 20 A°) interference filters because of the lack of convergence of the extreme optical rays at large zenith angles. Image intensifiers are employed to achieve the high sensitivity. Data are recorded either on a photographic image or by using a TV system to produce a video signal. Typical system parameters are shown in Table 10-2.

Medium Resolution Systems: Ebert-Fastie type scanning spectometers are used as medium resolution optical systems. These are effectively used over the (visible) wavelength range of 3800 to 7900 Å with variable spectral resolution from 2-18 Å. They can be scanned across the sky or operated in the zenith. When properly calibrated, spectrometers provide the absolute intensity of auroral and airglow features as well as some measure of spectral character.

High Resolution Systems: Fabry-Perot interferometers use multiple path interference to achieve high spectral resolution. These instruments are primarily used to measure spectral line broadening and Doppler shift. From these parameters, atmospheric temperatures and drift velocities can be derived. Primary spectral features of interest are 6300 Å [OI], 5577 Å [OI] for neutral winds, and 7320 Å [OII] for plasma drift.

10.1.5.2 Ionospheric Structure from Optical Measurements. Ionospheric structure at mid and equatorial latitudes is controlled by solar ionizing radiation, electric fields, and neutral atmosphere dynamics. Airglow observations of equatorial plasma depletions are one example of optical measurements used to define ionospheric processes. A brief review of equatorial airglow chemical production mechanisms is presented to illustrate the techniques used to infer ionospheric plasma density variations from remote optical measurements.

Two primary airglow spectral emission features are of interest for nighttime, F region phenomena 6300 Å [OI] and 7774 Å OI. The 6300Å atomic oxygen emission results from the following sequence of reaction.

$$O^+ + O_2 \xrightarrow{K_1} O_2^+ + e \qquad (10.26)$$

$$O_2^+ + e \xrightarrow{K_2} O + O(^1D) \qquad (10.27)$$

$$O(^1D) \rightarrow O(^3P) + h\nu(6300) \qquad (10.28)$$

$$O(^1D) + N_2 \xrightarrow{K_Q} O(^3P) + N_2. \qquad (10.29)$$

Since $K_2 \gg K_1$, and in regions where O^+ is the dominant ion ($O^+ = N_e$) the 6300Å volume emission rate is given by

$$dI(6300) = \left[0.75\, K_2E\, [N_e][O_2] / \left(1 + \frac{K_Q N_2}{A} \right) \right] dh \qquad (10.30)$$

where $K_2E = 1.4 \times 10^{-11}$ cm^3 s^{-1}

$K_Q = 7.0 \times 10^{-11}$ cm^3 s^{-1}

$A = 10^{-2}$ s^{-1}

(See Weber et al., [1980] and Noxon and Johanson [1970] for a more complete discussion).

The 7774 Å OI results from radiative recombination of O^+:

Table 10-2. Summary of system specifications.

Field of View	155°
Pass Band	25 A at f 1.4; 5 Å at f 8
Resolution	1/2° zenith, 2° horizon
Spectral Response	S-20, exceeding 100 μA/lumen
Picture Storage	No detectable degradation for up to 3 s
Tube Gain	Photon noise granularity visible above tube noise
Threshold Sensitivity	20 R at 2 s exposure 1 kR at 30 frames/s
Dynamic Range	20 R to 10 kR covered by 3 preset HV settings
Flatness of Field	30 percent loss at edge of field
Repetition Rate	Typically 20 s for complete filter cycle
Temporary Storage	Video disc, three video tracks + one sync track
Permanent Storage	Video tape deck, time-lapse type (9 h recording time on a single reel): 16 mm color movie camera
Process Controller	In-field programming capability
Display Systems	Four black and white monitors, 9 in. diagnal; Color monitor, RGB and A-B input, 12 in. diagonal
Real-Time Display	Simultaneous fully registered display of three filter channels. Capability of displaying difference of any two pictures. Display of two or three filters as pseudo-color on RGB monitor
Character Generators	Date/time display on each frame for frame identification
Digital Encoding	Digital encoding of time and housekeeping data for computer-controlled data handling

$$O^+ + e \xrightarrow{K_3} O(^5P) \qquad (10.31)$$

$$O(^5P) \rightarrow O(^5S) + h\nu(7774) \qquad (10.32)$$

and the volume emission rate is given by

$$dI(7774) = K_3 [O^+][N_e]\, dh. \qquad (10.33)$$

To illustrate the altitude dependence, 6300 Å and 7774 Å airglow volume emission rates were calculated for an electron density profile representative of the Appleton anomaly region, and are shown in Figure 10-16. The bottomside profile was obtained from true height analysis of a digital ionogram. This was matched to a modified Chapman function [Tinsley et al. 1973] to represent the topside profile from hmax F_2 to 690 km. Because of the exponentially

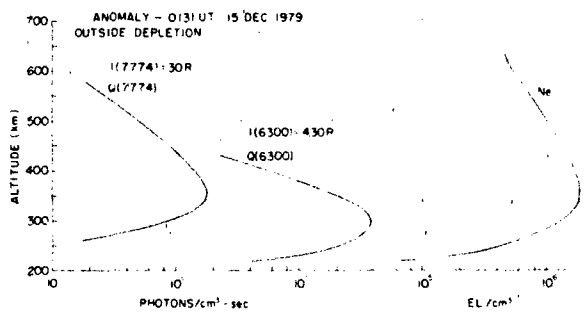


Figure 10-16. Electron density profile derived from true height analysis of bottomside ionogram matched to a modified Chapman function for the topside. Also shown are calculated 6300 and 7774 Å volume emission rates and column intensities in Rayleighs.

decreasing O_2 concentration, the 6300 Å volume emission rate is confined to the bottomside and reaches a maximum value at 300 km, below h_{max} F2 (360 km). The 7774 Å volume emission rate is proportional to $[Ne]^2$ and attains its maximum value at h_{max} F2. Because of the broad altitude exatent of the equatorial electron density profile the 7774 Å volume emission rate displays a similar broad extent and falls to 50% of the maximum value at 300 km and 450 km. Thus the 7774 A airglow measurements provide information over a broad altitude range, with approximately one half of the emission produced above the F layer peak.

All sky imaging photometer measurements conducted within a few degrees of the magnetic equator and near the Appleton Anomaly region (~18° ML) have established the two-dimensional horizontal extent of equatorial plasma depletions. These are also the regions of post-sunset equatorial spread F, VHF radar backscatter plumes, and amplitude and phase scintillation on transionospheric radio propagation. Figure 10-17 shows an example of all sky images at 6300 Å and 7774 Å near the equatorial edge of the Appleton Anomaly. The bright region over the southern two-thirds of the image is airglow from the high-density anomaly region. The North-South aligned dark band is a region of decreased airglow emission. Comparison with simultaneous *in situ* measurements from the Atmosphere Explorer satellite (AE-E) shows this airglow depletion to the collocated with a region of significantly decreased ion density. Having established the relation between airglow emission processes and F layer density, the all sky images provide a two-dimensional map of these depleted regions. In addition, the dynamics of these regions can be monitored over extended periods in the absence of "snapshot" satellite measurements. Optical measurements coordinated with VHF radar, ionosonde, *in situ* density, and satellite radio beacon scintillation observations have helped to provide a detailed description of the development, structure, drift, and decay of these important equatorial ionospheric features.

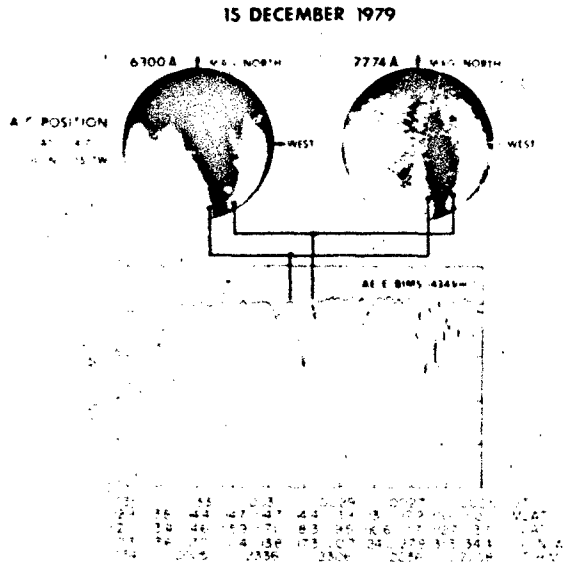


Figure 10-17. All sky images at 6300 and 7774 A recorded near the equatorial edge of the Appleton Anomaly

Optical measurements have improved our understanding of auroral zone and polar cap ionospheric structure and magnetosphere-ionosphere coupling processes. In this region dominated by the effect of precipitating electrons and ions over a wide energy range (few eV to 100's of keV), optical measurements of impact excitation and chemical recombination aid in understanding a wide variety of processes. In this section, several examples of all-sky monochromatic images are shown, primarily to demonstrate the use of optical measurements, especially when coordinated with other ionospheric diagnostics. Figure 10-18 shows a montage of auroral images at 10-min intervals at 6300 Å [OI] and 4278 Å N_2^+. These images were recorded on an aircraft which flew North-South legs along the Chatanika, Alaska Incoherent Scatter Radar magnetic meridian. The images provide a map of the instantaneous particle precipitation patterns separately for the E (4278 Å) and F (6300 Å) layers. Measurements with the radar mapped electron density structure and satellite UHF radio beacon scintillation measurements mapped regions of ionospheric irregularities (from tens of meters to a few kilometers). In this experiment, optical measurements provide a continuous map of particle precipitation regions over a large area (1200 km diameter at F-region altitudes) for interpretation of magnetosphere-ionosphere coupling and ionospheric dynamics.

Measurements in the polar cap have recently clarified local particle precipitation effects from plasma transport (E-field) effects. Local precipitation of low energy (100's of eV) electrons, during IMF Bz north conditions, leads to the production of sun-aligned F region auroras. Some of these auroras are characterized by F region plasma density enhancement and structuring within these auroras leds to amplitude and phase scintillation.

During Bz south conditions, large patches of F region plasma are observed to drift across the polar cap in the anti-sunward direction. Coordinated satellite measurements show that these patches are not locally produced by precipitating particles, but are convected from a source region at or equatorward of the dayside cusp. These patches are also subject to structuring processes that lead to scintillation.

10.2 SOME ASPECTS OF LONG WAVE PROPAGATION

It is convenient to refer to radio waves having frequencies below 3000 kHz as "long waves". These include Extremely Low Frequencies (ELF), Very Low Frequencies (VLF), Low Frequencies (LF), and Medium Frequencies (MF), as outlined in Table 10-3. ELF has had very little

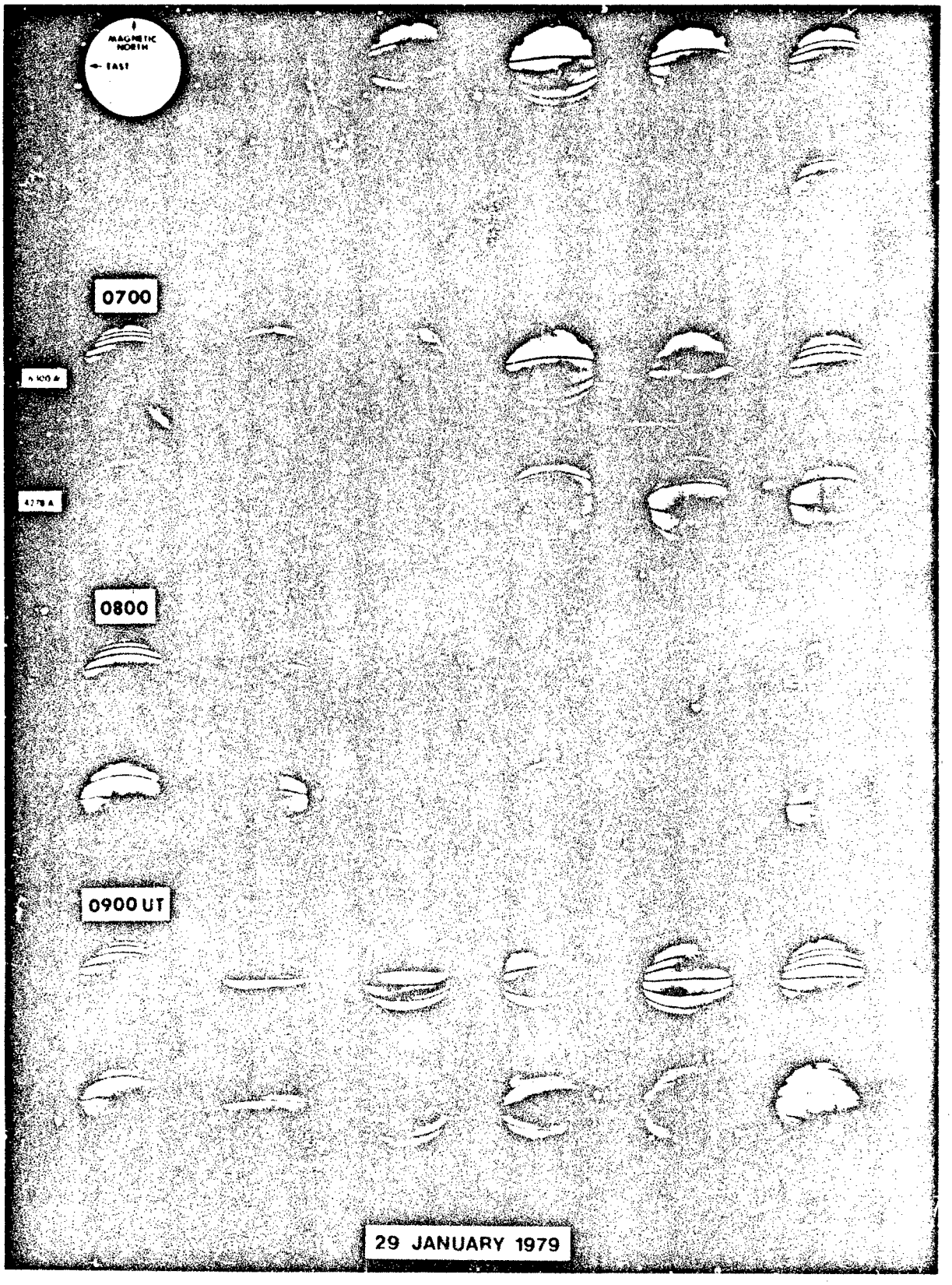


Figure 10-18 Auroral images taken at 10 minute intervals from 0620 to 0950 UT on 29 January 1979. The upper row under each hour shows the 6300 Å images, the lower row, the 4278 Å images

Table 10-3. Long wave frequency bands.

Designation	Abbreviation	Frequency Range	Wavelengths
Extremely Low Frequency	ELF	0.003–3 kHz	10^5–10^2 km
Very Low Frequency	VLF	3–30 kHz	10^2–10^1 km
Low Frequency	LF	30–300 kHz	10–1 km
Medium Frequency	MF	300–3000 kHz	1–0.1 km

use, except for communications that require wave penetration beneath the surface of the ocean or earth. The VLF/LF bands are used extensively for navigation and military communication. The standard AM broadcast systems utilize part of the MF band (535-1606 kHz). Long radio waves are also used in basic ionospheric research, meteorology and thunderstorm study and tracking, standard frequency and time distribution, geological studies, and minerals exploration.

Long waves propagate by a number of different modes. These include propagation over the surface of the earth by diffraction modes, ELF propagation by transmission-line type modes, propagation by ionospheric reflection (or earth-ionosphere waveguide modes) and propagation through the ionosphere by so-called "whistler" modes. Each type of mode requires a separate physical description and mathematical formulation.

10.2.1 Groundwave Propagation

The most general definition of the groundwave is the wave that would be excited by an antenna at or near the air-earth boundary if there were no wave reflections from the upper atmosphere. At long wavelengths ionospheric reflections are important, and for continuous wave (CW) transmissions it is necessary to regard the total wave field as a vector sum of the groundwave and skywaves. If an antenna radiates a very short pulse, however, it may be possible for a distant receiver to resolve the groundwave and skywaves individually. The time interval between the onsets of the groundwave and the first-hop skywave is given by

$$\Delta t = (2\sqrt{h^2 + 4a(a+h)\sin^2(d/4a)} - d)\,(1/c), \quad (10.34)$$

providing

$$d \leq 2a\cos^{-1}\{a/(a+h)\}. \quad (10.35)$$

where d is the distance between the transmitter and receiver, h is the effective height of reflection in the ionosphere, a is the earth's radius (≈6370 km) and c is the wave velocity (3×10^5 km/s). If the transmitted pulse is short enough, Δt may be long enough (for example, Δt = 93 μs for d = 500 km and h = 80 km) to permit very accurate measurements of the groundwave, especially its arrival time or phase. The latter is the basis of the long-range 100 kHz groundwave navigation system, Loran-C.

If the transmitted signal is a continuous wave, the amplitude and phase of the composite signal received at a fixed distance vary with time as the ionosphere changes. On the other hand, at a given moment the signal amplitude is a function of distance [Hollingworth, 1926], having maxima and minima typical of an interference pattern. The groundwave component is stronger than the skywaves out to a distance that depends on the wave frequency, among other factors. This region of groundwave dominance is the most stable, or primary, coverage area of MF broadcast transmitters.

As defined above, the groundwave exists at all radio frequencies, but at wavelengths comparable to the height of the ionosphere or greater, the usefulness of the concept begins to fade. Also, for transmitters high above the ground, or at high frequencies where quasi-optical propagation analysis is appropriate, the term groundwave is seldom used.

The earth often acts as a fairly good conductor for long waves, in which case the electromagnetic boundary conditions permit electric fields perpendicular to the surface, while tending to suppress electric fields tangential to the surface. It follows that groundwave fields near the earth's surface tend to have transverse magnetic (TM) polarization rather than transverse electric (TE) polarization. In common usage the unqualified term "groundwave" implies TM polarization.

10.2.1.1 Idealized Flat-Earth Models. In a simple model the earth's surface is regarded as a flat perfect conductor, and the air is homogeneous with refractive index 1. The most elementary source is a vertically-directed current $I(t) = I_0 \exp(i\omega t)$, at frequency $f = \omega/2\pi$, of infinitesimal length $d\ell$, which has an electric dipole moment $M(t) = I(t)\,d\ell$ (Note: complex antennas can be regarded as distributions of such elementary currents). The fields of such a source may be found readily by the method of images. When the current element is just above the surface, the fields in air are simply twice the homogeneous free-space fields. Because of symmetry the magnetic field is everywhere in the azimuthal direction ϕ (see Figure 10-19), while the electric field on the surface is constrained by the boundary conditions to be strictly in the vertical direction z. In mks units the magnetic and electric fields at a distance, d, on the flat perfectly-conducting surface are, respectively,

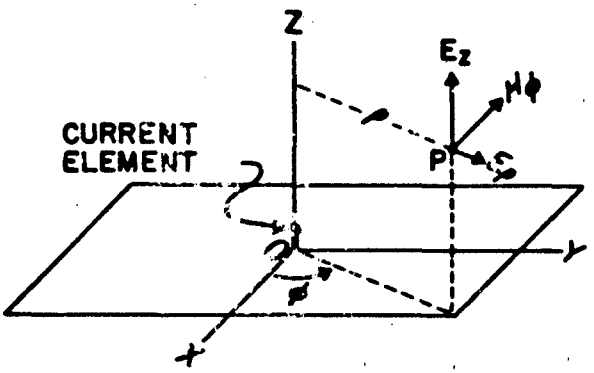


Figure 10-19 Vector field-components at a point P in a cylindrical coordinate system. The plane XOY represents the surface of the earth.

$$H_\phi = \frac{M(t')}{2\pi}\left[\frac{1}{d^2} + \frac{i2\pi}{\lambda d}\right] \text{ At/m.} \qquad (10.36)$$

$$E_z = -\frac{M(t')}{2\pi}\left[\frac{1}{i^2\varepsilon_o\omega d^3} + \frac{120\pi}{d^2} + \frac{i\mu_o\omega}{d}\right] \text{ V/m.} \qquad (10.37)$$

$$E_\rho = 0. \qquad (10.38)$$

where $t' = (t - d/c)$, λ is the vacuum wavelength ($= 3 \times 10^8/f$ m), ε_o is the permittivity of free space ($= 8.854 \times 10^{-12}$ F/m), and μ_o is the free space permeability ($= 4\pi \times 10^{-7}$ H/m). The far-field components are related by

$$E_z = -Z_o H_\phi \qquad (10.39)$$

where Z_o is the impedance of free-space ($\approx 120\pi$ ohm).

In a more realistic model, the plane earth is allowed to have finite conductivity σ and permittivity ε. The solution of this boundary-value problem was given by Sommerfeld [1909] in terms of an infinite complex integral [see Stratton, 1941]. A more complex problem, that of an elevated dipole, was solved by Weyl [1919], who expressed the free-space field as a sum of plane waves that reflected at the earth's surface in accordance with the Fresnel formulas. Norton [1941] and others have calculated numerical values from the formal solutions. Height variations of the fields are shown in Figure 10-20 for a source on a plane earth surface [Heckscher and Tichovolsky, 1981], and curves illustrating groundwave field amplitudes along the surface are given in Figure 10-21 for both plane and spherical earth models.

A type of groundwave, the Zenneck surface wave, has fields expressed exactly in simple closed forms. Although the Zenneck wave is important historically and conceptually, it is generally difficult to excite because of its rather slow decay with height. In fact, the excitation of a pure

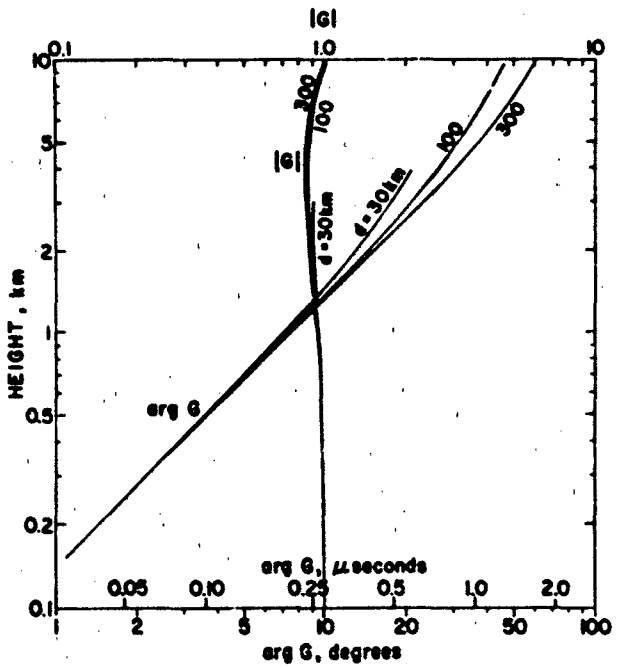


Figure 10-20 Height variations in the amplitudes and phases of 100 kHz groundwave fields for a source on a plane earth. Values are shown at distances of 30, 100, and 300 km, for propagation over fresh water, $\sigma = 10^{-1}$ S/m, $\varepsilon/\varepsilon_o = 80$ [Heckscher and Tichovolsky, 1981].

Zenneck wave, without any radiation field, requires an infinitely long source distribution [Hill and Wait, 1978].

For a finitely conducting earth, Equation (10.39) is still true approximately, but the radial electric field component E_ρ has a finite value related to the loss of wave power into

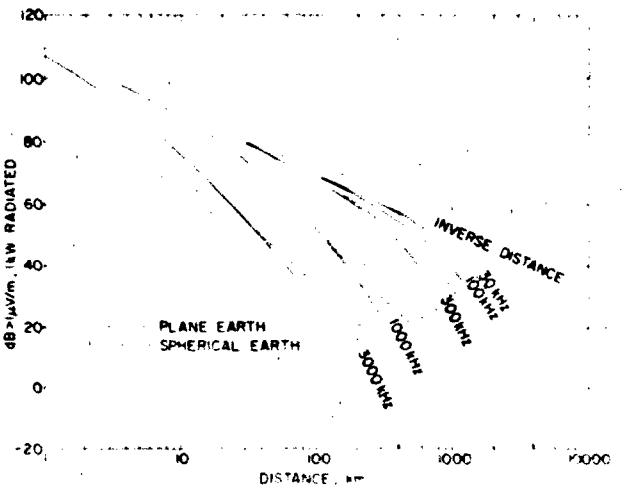


Figure 10-21 Long wave groundwave field amplitudes as a function of distance over plane and spherical earths, for propagation over good earth, $\sigma = 10^{-2}$ S/m, $\varepsilon/\varepsilon_o = 20$.

Table 10-4 Ratio of groundwave radial and vertical electric fields for various earth surfaces and 100 kHz

SURFACE TYPE	σ (S m)	ε/ε₀	E_ρ/E_z
Perfect Conductor	∞	—	0
Sea Water	4	80	0.00118 ∠44.99°
Good Soil	10^{-2}	20	0.0236 ∠44.66°
Fresh Water	10^{-3}	80	0.0713 ∠32.89°
Poor Soil	10^{-3}	10	0.0745 ∠43.25°
Thick Ice	2×10^{-5}	5	0.403 ∠14.76°

the ground. The ratio of the radial and vertical electric fields at the surface is given by

$$E_\rho/E_z \approx \sqrt{(1 - 1/p)/p}, \qquad (10.40)$$

where

$$p = \varepsilon/\varepsilon_0 - i\sigma/\omega\varepsilon_0. \qquad (10.41)$$

Starting just below the surface of the assumed uniform earth, E_ρ and H_ϕ decrease exponentially with depth. The fields are 1/e of their values on the surface at the "skin depth" δ, which may be estimated from

$$\delta = -\frac{c}{\omega}\sqrt{\frac{\varepsilon_0}{\varepsilon}}\frac{1}{\mathrm{Im}(p\sqrt{1+p})} \qquad (10.42)$$

$$\approx 503/\sqrt{\sigma f} \text{ when } \arg p \approx -\pi/2. \qquad (10.43)$$

Sample values of E_ρ/E_z are given in Table 10-4 for various earth surfaces at a frequency of 100 kHz, and Figure 10-22 shows skin depths over the long wave spectrum.

The magnetic field of the groundwave H_ϕ may be sensed with a loop antenna having its axis parallel to the surface and perpendicular to the direction of propagation. E_z may be sensed with a vertical monopole antenna, or a vertical dipole. The radial component E_ρ may be sensed by a horizontal dipole with its axis in the direction of propagation. It follows from the principle of reciprocity that if a horizontal dipole on the earth is driven with RF current, a TM-polarized groundwave is radiated in directions along its axis. The fields produced by horizontal and vertical current elements have been discussed in mathematical detail by Wait [1954, 1957, 1961, 1971].

10.2.1.2 Idealized Spherical-Earth Models. A mathematical treatment of the groundwave on a smooth spherical earth of homogeneous, isotropic material was undertaken by Watson [1919] to determine if an atmospheric reflecting layer (ionosphere) was required to explain the large fields produced by distant transmitters. Such early theoretical analyses were handicapped by the poor convergence of the infinite series contained in the solutions. That difficulty was largely overcome by Van der Pol and Bremmer's [1938] "residue series" solution, which has become a basis for modern numerical analysis of the groundwave.

Figure 10-21 shows examples of field-strength vs distance curves for waves of selected frequencies propagating over "good soil," assuming a vertical source on the surface. The earth-curvature causes the wave amplitude to decrease with distance more rapidly than it would on a flat earth of the same material, but near the source the flat- and round-earth models give essentially the same fields.

One way to present both amplitude and phase data is in terms of the complex factor W, which is the ratio of the actual field component to an idealized one calculated as if the earth were flat and perfectly conducting. The flat earth propagation distance is taken to be the same as the curvilinear distance on the sphere. Figure 10-23 shows curves of W in the complex plane, for 100 kHz propagation over sea water and good soil.

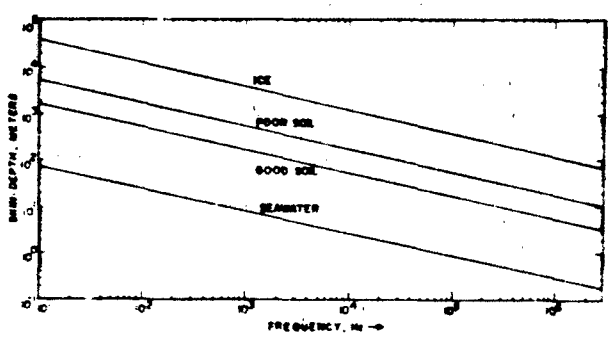


Figure 10-22 Long wave skin depths for various earth surfaces

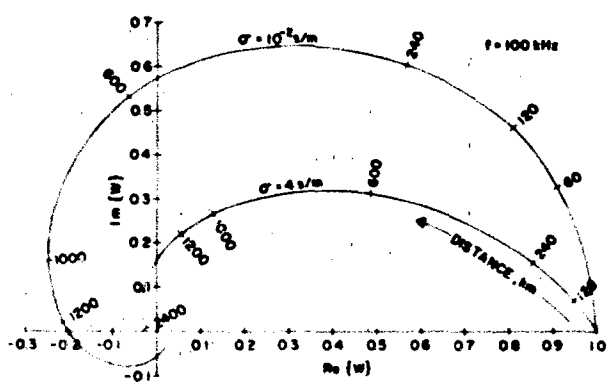


Figure 10-23 Complex values of the ratio W of actual groundwave field components and ones obtained assuming a flat, perfectly conducting earth, for 100 kHz over sea water ($\sigma = 4$ S/m) and good soil ($\sigma = 10^{-2}$ S/m). Derived from Wait and Howe [1956].

10.2.1.3 Models with Earth-Properties Gently Varying Along the Propagation Path. Except for large bodies of water, the earth's surface is too uneven, both in electrical properties and in topology, to be represented well by the idealized models discussed above. However, if the earth's electrical properties and curvature do not vary much in a wavelength, the groundwave amplitude and phase can be approximated by the solution of an integral equation for W [Hufford, 1952]. The two-dimensional integral equation—the most general form—is valid provided, first, impedance boundary conditions can be applied; and second, terrain irregularities are not too severe.

A much simpler, one-dimensional, integral equation is more commonly used. It is derived by a stationary-phase integration that reduces the dimensionality of the general version; being an approximation, it is not valid for all terrain types, particularly at low frequencies where wave lengths and Fresnel zones have sizes comparable with terrain features. Field [1982a] compared solutions of the one- and two-dimensional equations to quantify the errors incurred by using the one-dimensional equation for terrain features narrower than a Fresnel zone. For a frequency of 100 kHz and on-path features narrower than about 10 km, the two-dimensional equation is needed to properly account for the signal's dependence on obstacle width, recovery at long distances, and transverse diffraction patterns. The one- and two-dimensional solutions approach one another far beyond wide terrain features.

Considerable error can be incurred at low frequencies by applying the one-dimensional equation, even for large terrain features. For example, for a path-length of 500 km, that equation overstates by a factor of about four the effect of an obstacle 6 km in diameter. It cannot give accurate results unless the diameter approaches a Fresnel zone width, which for this example is several tens of kilometers. However, because numerical solution of the two-dimensional integral equation is costly, its use has been limited to highly idealized irregularities. An alternate approach given by King and Wait [1976] obtains an equivalent one-dimensional model by averaging the terrain over the Fresnel zone.

If, instead of being homogeneous, the earth were comprised of layers of different conductivity and dielectric constant, it could still be characterized by a (frequency dependent) surface impedance, so that the integral equation for W can be formulated. Some progress has been made in this and other ways of estimating earth-constants for groundwave phase prediction at 100 kHz, but uncertainties in these constants remain a major source of prediction error.

Special models of non-uniform terrain exist for which solution of the integral equation is unnecessary. Groundwave propagation from land to sea, and vice versa, has been modeled with an earth having a sharp discontinuity of conductivity along a horizontal straight line. A useful approximate solution was given by Millington [1949] based on reciprocity arguments, and mathematical solutions have been given by Wait [1956] and Wait and Walters [1963].

Groundwave perturbations caused by hills have been studied with models having semielliptical bosses on otherwise smooth surfaces [Wait and Murphy, 1957, 1958]. In the electrostatic case, the vertical electric field at the summit of a hemispherical boss is exactly 3 times that on a flat plane, and for a semicylindrical ridge the factor is 2 times. This field-enhancement effect carries over (with modifications) to VLF groundwaves, and has been demonstrated experimentally [Harrison et al., 1971].

10.2.2 ELF Propagation

ELF propagation has been the subject of theoretical study for many years. The texts by Wait [1970], Galejs [1972] and Burrows [1978], along with review papers by Bernstein et al. [1974], Wait [1974 and 1977] and Bannister, [1980], provide comprehensive descriptions of the propagation characteristics of ELF waves. Much of the discussion that follows is based on a review paper by Bannister [1982].

ELF waves below a few hundred hertz propagate with little attenuation, penetrate well into lossy media and are very stable compared with higher frequencies. Nevertheless, ELF has limitations relative to conventional radio communication bands. Its restricted bandwidth allows only very low data rates, and because of the great wavelengths, ELF transmitting antennas are very inefficient (less than 0.5% is typical).

The energy of an ELF wave is confined principally to the waveguide that exists between the earth and the ionosphere. At ELF the effective height h of the waveguide is much less than the wavelength λ of the wave and the waveguide is below cut-off for all but the lowest order mode, the transverse electromagnetic (TEM) mode. On the other hand, at VLF the waveguide height exceeds the wavelength and several modes propagate. At LF the number of significant propagating modes may exceed 20.

The principal TEM fields are the vertical electric field E_z, and the horizontal magnetic field, H_θ. Secondary field components arise because the surface impedance η_s of the ground—albeit small—is not zero; hence the term "quasi TEM" mode. The secondary fields are small compared with the principal fields, but are important because horizontal antennas would not radiate if the secondary fields were zero [Burrows, 1978].

Attenuation of the ELF quasi-TEM mode in the earth-ionosphere waveguide is low, on the order of 1 or 2 dB/Mm. The attenuation is caused mainly by power absorption in the ionosphere, since the surface impedance of the ionosphere is typically much larger than the surface impedance of the ground. That effect is evidenced in the expression for the waveguide attenuation rate α, which is inversely proportional to h. That behavior indicates that the rate of power leakage from the guide is proportional to the field intensity at the surface, whereas the rate of power flow along the guide is proportional to the guide's volume. Thus, as

the guide's height decreases, the ratio of power leakage to total power flow increases [Burrows, 1978].

ELF attenuation is low enough to support very long-range propagation, and a planar model of the earth is inadequate if the path length exceeds the earth's radius. The most important effect of earth curvature is the closure of the guide, which allows the field to return to the source point after one complete encirclement. Therefore, the total field is the sum of the field arising from propagation over the shorter great-circle path from the source, and that arising from propagation over the longer one. The local effect of the curvature is small, however, and the wave propagates in the curved guide with nearly the same parameters as it would in a planar one.

The geomagnetic field interacts with the charged particles of the ionosphere to produce an anisotropic conducting medium. However, the electrical mismatch between the atmosphere and the ionosphere is large at ELF, and the transition between them abrupt, so the ionosphere acts much like a perfect reflector. The effect of the anisotropy is therefore small, and ELF attenuation and phase velocity depend only slightly on the direction of propagation.

10.2.2.1 ELF Field-Strength Calculations. The expressions most often employed for calculating ELF fields in the earth-ionosphere waveguide are based upon an idealized model that assumes the earth and the ionosphere to be sharply bounded and homogeneous. Experimental measurements of the waveguide properties at ELF have shown consistently that they can be represented accurately by formulas based on such a simple model. Complicated calculations that account for vertical structure of the ionosphere [Field, 1969 and Pappert and Moler, 1978] also confirm that the simple model is adequate for many purposes.

For the idealized model, the magnitude of the magnetic field of the signal from a horizontal dipole is approximately

$$|H| = Id\ell\, \Lambda \left[\frac{k_o}{2\mu_o}\left(\frac{\omega\mu_o}{2\pi}\right)^{1/2}\right]\left(\frac{\rho/a}{\sin \rho/a}\right)^{1/2} \times \frac{10^{-\alpha\rho/2\pi \times 10}}{(k_o\rho)^{1/2}} \cos\phi, \quad (10.44)$$

where

$$\Lambda = \frac{\eta_s}{h\sqrt{\omega\mu_o c/v}} \quad (10.45)$$

is an excitation factor, η_s is the surface impedance of the ground, c is the speed of light, v is the phase velocity of the TEM mode and a is the radius of the earth.

There are six distinct factors in Equation (10.44): (1) the source strength $Id\ell$; (2) the excitation factor Λ; (3) a collection of free space parameters, all of which are known once the frequency is specified; (4) a spherical focusing factor; (5) the radial propagation loss factor, including both the exponential decay due to absorption and the $\rho^{-1/2}$ decay due to spreading; and (6) the directional dependence of the radiated field. Once the current moment $Id\ell$, angular frequency ω, and coordinates ρ, ϕ of the field point are specified, only two parameters are left undetermined, Λ and the attenuation rate, α. Both depend on the ionosphere.

Greifinger and Greifinger [1978, 1979], derived simple approximate expressions for ELF propagation constants that agree well with full wave numerical calculations. For daytime propagation, the approximate expressions for c/v and α are

$$c/v \sim 0.985\sqrt{h_1/h_o} \quad (10.46)$$

and

$$\alpha \sim 0.143f\sqrt{h_1/h_o}\left(\frac{\zeta_o}{h_o} + \frac{\zeta_1}{h_1}\right). \quad (10.47)$$

where h_o is the altitude where $\sigma_1 = \omega\varepsilon_o$; h_1 is the altitude where $4\omega\mu_o\zeta_1^2 = 1$; ζ_o and ζ_1 are the conductivity scale heights at altitudes h_o and h_1, respectively; and σ_1 is the conductivity of the ionosphere, which varies with altitude.

Equations (10.46) and (10.47) show that the phase constant depends primarily on the two reflecting heights and is nearly independent of the conductivity scale heights. On the other hand, the attenuation rate depends on the scale heights as well as the reflection heights.

A simple exponential fit to the ionospheric conductivity profile is given by Wait [1970] for determining propagation parameters:

$$\sigma_1(z)/\varepsilon_o = 2.5 \times 10^5 \exp[(z - H)/\zeta_o]. \quad (10.48)$$

where H is an (arbitrary) reference height. The corresponding values for h_o and h_1 are

$$h_o = H - \zeta_o \operatorname{Ln}\left(\frac{2.5 \times 10^5}{2\pi f}\right) \quad (10.49)$$

and

$$h_1 = h_o + 2\zeta_o \operatorname{Ln}\left(\frac{2.39 \times 10^4}{f\zeta_o}\right). \quad (10.50)$$

Note that all heights are in kilometers.

10.2.2.2 Theoretical and Measured ELF Propagation Constants. It can be shown that the effective waveguide height of reflection is roughly h_o, rather than the higher reflecting height h_1. This is in excellent agreement with the effective reflection heights inferred from ELF propagation measurements [Bannister, 1975]. The most common values of H and ζ_o employed in interpreting VLF daytime propagation measurements are H = 70 km and ζ_o = 3.33 km. By using these values the values of h_o, h_1, c/v and α can be determined readily at ELF.

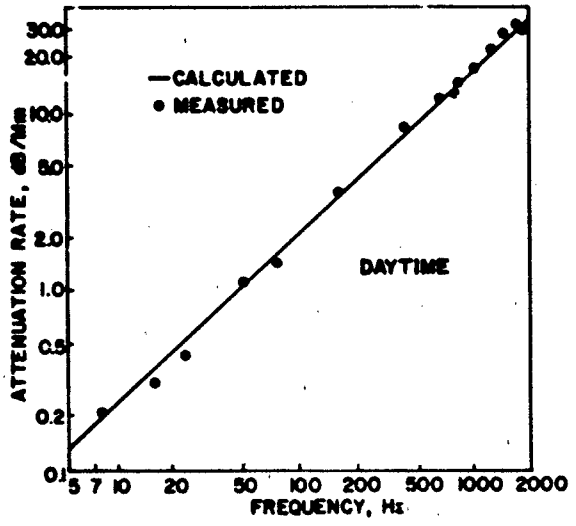


Figure 10-24 Calculated and measured ELF daytime attenuation rates [Bannister, 1982]

Theoretical values of ELF daytime attenuation rates are plotted in Figure 10-24 for frequencies from 5 Hz to 2000 Hz. Also plotted are values of α determined from controlled source measurements [Bannister, 1982], or inferred from Schumann resonance measurements [Chapman et al., 1966]. Figure 10-24 shows that the agreement between the theoretical and measured values of ELF daytime attenuation rates is excellent.

Theoretical values of ELF daytime relative phase velocity are plotted in Figure 10-25 for frequencies from 5 Hz to 1000 Hz. Also plotted are various values of c/v determined from measurements of atmospherics. Figure 10-25 shows excellent agreement between the theoretical and measured values of c/v for frequencies above 50 Hz, and fair agreement for frequencies below 50 Hz.

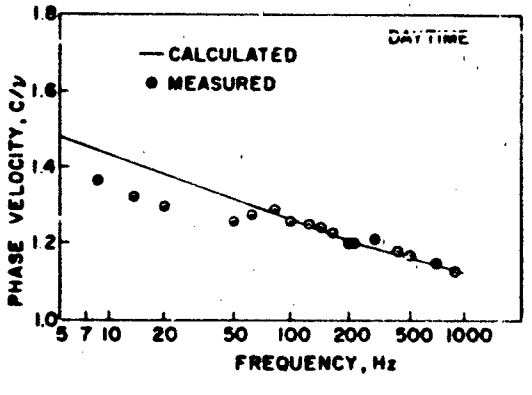


Figure 10-25 Calculated and measured daytime ELF phase velocities [Bannister, 1982]

Under nighttime propagation conditions, an E region bottom where the electron density increases very sharply is usually encountered below the altitude h_1. For a simple model that assumes the density above the E region bottom to vary slowly on the scale of the local wavelength, the propagation constant is [Greifinger and Greifinger, 1979]

$$c/v \sim \sqrt{h_1/h_0} \qquad (10.51)$$

and the attenuation rate is

$$\alpha \sim 0.143 f \sqrt{h_1/h_0} \left(\frac{\zeta_0}{h_0} + \frac{1}{\pi k_0 n_1 h_1} \right), \qquad (10.52)$$

where h_1 is the altitude of the E region bottom and $k_0 n_1$ is the local wave number.

Nighttime ELF attenuation rates are plotted in Figure 10-26 for frequencies from 40 Hz to 1000 Hz. The calculated values were obtained using Wait's nighttime ionospheric conductivity model (with a reference height of 90 km and scale height of 2.5 km) in conjunction with Equations (10.51) and (10.52), and assuming the height of the E region bottom to be 90 km and its conductivity to be 8×10^{-6} S/m. Also plotted are various measured values of α. Figure 10-26 shows that, for frequencies from 45 Hz to 800 Hz, there is excellent agreement between the theoretical and the measured values.

In addition to their very low attenuation rates, ELF radio waves below about 200 Hz can penetrate lossy media and retain usable strengths to substantial depths. Those features make them attractive for communicating over great distances to sub-surface locations. For example, even for sea water, with a conductivity of 4 S/m, the skin depth is about 36 m at a frequency of 50 Hz.

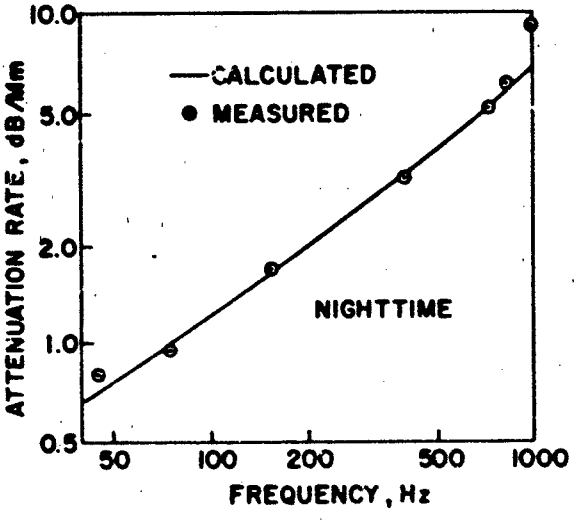


Figure 10-26 Calculated and measured ELF nighttime attenuation rates [Bannister, 1982]

10.2.2.3 Anomalous ELF Propagation Occasionally the nighttime field strengths measured at 40 Hz to 80 Hz have decreased by 4 dB to 8 dB in the northeastern United States [Bannister, 1975, 1980]. Those relatively severe nighttime fades sometimes occurred during the several days following magnetic storms, when similar (but less pronounced) behavior was found to coincide with phase disturbances on VLF paths across the northern United States. These short-path (~1.6 Mm) field strength reductions might have been caused by enhanced ionospheric ionization due to precipitating electrons from the radiation belts. However, attempts to correlate the fades with geomagnetic indices have met with limited success. Simutaneous measurements taken in Connecticut and the North Atlantic area during the magnetically quiet period of early March 1977 have indicated some of these anomalies might have been caused by a moving nocturnal sporadic E layer.

Calculations by Barr [1977] and Pappert and Moler [1978] show that nocturnal sporadic E can produce marked maxima and minima in the propagation characteristics of ELF radio waves. One physical explanation for the effect is interference between waves reflected from the normal E region and from the sporadic E region. Pappert [1980] showed that a sporadic E patch one square megameter in extent could account for the 6-8 dB fades that have been observed. Similarly, patches 0.5 Mm^2 in extent could account for the more commonly observed 3-4 dB fades.

Many other ionospheric disturbances can cause ELF propagation anomalies, including those associated with solar x-ray flares, energetic electrons, protons from solar particle events (SPEs) in the polar cap, and high-altitude nuclear bursts. The attenuation rate can increase easily by 1 or 2 dB/Mm relative to normal daytime conditions during such disturbances, depending on the wave frequency and severity of the disturbance, and certain moderate solar proton events can cause the attenuation rate to approach 4 dB/Mm at 75 Hz [Field, 1982b].

10.2.2.4 Analysis of Laterally Non-Uniform Ionospheric Disturbances. Most predictions of ELF fields in the earth-ionosphere waveguide have used a WKB method described by Pappert and Molar [1978]. To find E-fields along any particular radial from the transmitter, that method assumes the properties of the guide to depend only on distance from the transmitter. However, even such large inhomogeneities as sporadic E patches, the polar cap boundary, and the day/night terminator can cause the properties of the earth-ionosphere waveguide to change markedly over the huge wavelength, or Fresnel zone, of an ELF signal. Such inhomogeneities can cause lateral reflection, diffraction, and focusing of ELF modes. Those phenomena are usually unimportant at higher frequencies where the earth-ionosphere waveguide can be assumed to be vary slowly in the lateral directions.

To handle such situations, Field and Joiner [1979, 1982] employed an integral equation approach for analyzing propagation in the earth-ionosphere waveguide where conditions change over transverse distances comparable with a Fresnel zone. They derived an expression for the relative errors introduced by neglecting transverse ionospheric gradients over the path and found that the WKB method is inaccurate when the width of a disturbance is less than two-thirds of the width of the first Fresnel zone. Further, the WKB approximation significantly overestimates the propagation anomaly when the disturbance is centered near the propagation path and underestimates the anomaly when the disturbance is centered far off path.

Strong localized disturbances behave like a cylindrical lens filling a narrow aperture: Lateral diffraction, focusing, and reflection cause the transverse electromagnetic (TEM) mode to exhibit a transverse pattern of maxima and minima beyond the disturbance and a standing-wave pattern in front of it. The focusing and diffraction diminish when the transverse dimension of the disturbance approaches the width of the first Fresnel zone, typically, several megameters. Reflection from widespread inhomogeneities can be important in two situations: first, for great-circle propagation paths that are nearly tangential to the boundary of the disturbed polar cap; and second, when the TEM mode is obliquely incident on the day/night terminator, in which case a phenomenon analogous to internal reflection can occur.

10.2.3 Long-Range VLF/LF Propagation

Very low and low frequency (VLF/LF) waves are reflected from the lowest regions of the ionosphere (the D region during daylight and the lower E region at night), and apart from the sunrise and sunset periods, exhibit propagation characteristics that are very stable in both phase and amplitude. The LF band (30-300 kHz) is useful for communications to distances ranging from hundreds to several thousand kilometers, shorter than the almost global ranges achievable at VLF but much longer than the groundwave distances normally associated with the MF band.

Beginning in 1911 with Austin, various empirical formulas, deduced from numerous measurements, have been used to estimate the field strengths of these long waves. For example, Pierce derived a semiempirical formula to describe VLF propagation over water during the day. That formula, which gives the vertical electric field strength at a distance d from a transmitter having a radiated power P (kw) at frequency f (kHz), is [Watt, 1967]

$$E(\mu V/m) = \frac{21.3\sqrt{P}}{\sqrt{(\sin d/a)}} \exp\left[-0.2(d/a) f^{0.75}\right], \quad (10.53)$$

where a is the radius of the earth and the absorption term is the exponential.

Modern mathematical approaches for predicting the propagation characteristics of VLF/LF waves are formulated

in terms of a conducting spherical earth surrounded by a concentric electron-ion plasma (the ionosphere) into which waves are launched from a Hertzian dipole source. The application of Maxwell's equations and the appropriate boundary conditions allow the electromagnetic field to be calculated everywhere. In doing this the earth and the ionosphere can be regarded as forming a waveguide (without side-walls) in which propagation may be viewed either as a series of wave reflections (wave-hops), or by the mathematical equivalent—traveling wave modes. Generally, it is more convenient to apply the waveguide mode approach to the VLF and lower LF band, and the wave-hop approach to the higher LF band, as described below. In addition to the references cited under the specific topics that follow, more detailed descriptions of the features of VLF/LF propagation can be found in the works by Budden [1961], Wait and Spies [1964], Watt [1967], Pappert [1968], Wait [1970], Galejs [1972], Field et al. [1976], and the AGARD proceedings on long waves edited by Belrose [1982].

10.2.3.1 Waveguide Modes. In the waveguide-mode formulation for VLF/LF propagation an arbitrary propagating field is regarded as being composed of a series of modal patterns that propagate with characteristic phase velocities, little change in pattern shape, and gradually decreasing amplitudes. Generally, the more complex (or higher-order) modal patterns attenuate more rapidly: at great distances only the simpler (lower-order) modes may be important. The earth-ionosphere cavity exhibits such waveguide features as cut-off frequencies, and reflections from mismatched sections caused by abrupt changes in the electrical properties along its boundaries.

The field at a point in the waveguide depends on how strongly the various modes are excited, and upon their relative amplitudes and phases at the observation point. Although there is evidence of wave focusing in the vicinity of the point antipodal to the transmitter, typical field strength versus distance curves generally show a decreasing trend, upon which are superimposed local variations due to modal interference.

A transmitting antenna modeled by a vertical current-element produces transverse magnetic (TM) polarized waves, which have a magnetic field parallel with the earth's surface, while the electric field is perpendicular to the magnetic field and not quite vertical. For TM waves, the earth-ionosphere waveguide has a quasi cutoff frequency during the daytime at about 2 kHz. Figure 10-27 illustrates idealized TM wave vectors and mode patterns.

Airborne VLF/LF transmitting antennas with horizontal current elements excite transverse electric (TE) waves with horizontal electric fields, and magnetic fields in the vertical plane. Figure 10-27 illustrates idealized TE wave vectors mode and waveguide patterns. In general, TE fields are very small at the ground, and the modes are difficult to excite with groundbased transmitters.

Because of the presence of the geomagnetic field in the ionosphere, the TM and TE modes are not entirely independent, but are coupled. In general, when an electron acquires a velocity from the electric field of the wave the magnetic forces cause it to have a component of motion perpendicular to the electric field, thus causing polarization conversion. The interaction is described by the Appleton-Hartree equation (for example, see Ratcliffe [1959]), and the polarization conversion has been demonstrated experimentally (as reported by Bracewell, et al., [1951] Lewis, et al., [1973] among others). The coupling effect is most pronounced at night when the waves reflect higher in the ionosphere, where the electron-neutral collision frequency is smaller than the electron gyrofrequency.

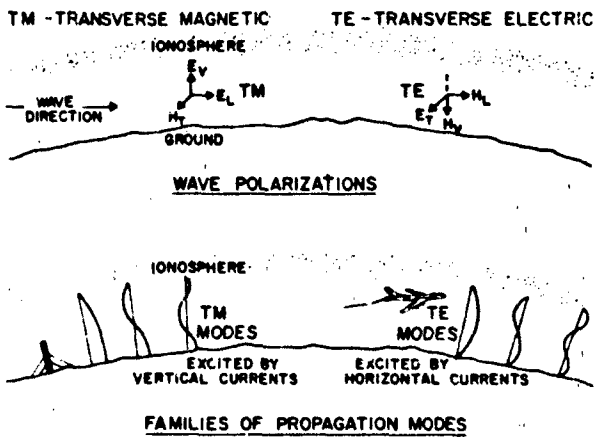


Figure 10-27. Characteristics of transverse magnetic (TM) and transverse electric (TE) waves [Kossey et al., 1982].

10.2.3.2 Waveguide Propagation Equations. The detailed equations for VLF/LF waveguide propagation are described, for example, by Wait [1970], Pappert and Bickel [1970], Galejs [1972], and Field et al., [1976]. To illustrate the key dependencies and to define the commonly used notations, they are given here for the case when geomagnetic anisotrophy (that is, polarization conversion) effects can be neglected. That approximation is fairly accurate for long-range VLF/LF propagation under normal daytime conditions.

Transverse Magnetic (TM) Modes

Conventional ground-based VLF/LF transmitters are vertical and their fields are composed of a superposition of TM waveguide modes. Following Field [1982c] the vertical electric field is given by

$$E_v = -120\pi i \exp\left(-\frac{i\pi}{4}\right)\frac{IL\cos\psi}{\sqrt{\lambda d}}\sqrt{\frac{d/a}{\sin d/a}}\sum_j S_j^{1/2} A_j \exp\left(-\frac{\alpha_j d}{8.7}\right) \exp\left(-\frac{2\pi i}{\lambda}\frac{c}{v_j}d\right) G_j(h_t) G_j(h_R)\ \text{V/m}, \quad (10.54)$$

where the subscript j denotes quantities associated with the jth TM mode, IL is the effective electric dipole moment of the transmitting antenna; λ is the free-space wavelength; d is the distance from the transmitter; a is the earth's radius; and c is the speed of light. Included is a factor cos ψ—where ψ is the angle between the dipole orientation and the vertical—to account for inclined transmitting antennas (cos ψ = 1 for a vertical electric dipole). Although most quantities are in mks units, all distances (L, λ, d, a) are expressed in megameters.

The quantity S_j is essentially the eigenvalue of the jth TM mode and must be computed numerically. At VLF, however, S has a magnitude close to unity, so the term $S_j^{1/2}$ in Equation (10.54) does not appreciably influence the field. The magnitude of the vertical electric field depends on the state of the ionosphere through three parameters: A_j, the excitation factor for the TM mode; α_j, the attenuation rate in decibels per megameter of propagation (dB/Mm); and G_j, the height gain function for transmitter and receiver heights h_T and h_R, respectively. The phase of the jth mode is governed by the relative phase velocity, c/v_j. These propagation parameters must all be computed numerically for model ionospheres having arbitrary height profiles.

Transverse Electric (TE) Modes

Airborne VLF/LF transmitters use long trailing-wire antennas that radiate a complicated superposition of TM and TE modes. Here much of that complexity is avoided by considering broadside propagation, where the great-circle path connecting transmitter and receiver is perpendicular to the plane containing the inclined electric-dipole transmitting antenna. The vertical electric field produced by the vertical component of the inclined transmitting antenna is given by Equation (10.55). The broadside horizontal electric field produced by the horizontal component is given by

$$E_H = -120\pi i \exp\left(-\frac{i\pi}{4}\right) \frac{IL \sin\psi}{\sqrt{\lambda d}} \sqrt{\frac{d/a}{\sin d/a}} \sum_m S_m^{1/2} A_m \exp\left(-\frac{\alpha_m d}{8.7}\right) \exp\left(-\frac{2\pi i}{\lambda}\frac{c}{v_m}d\right) G_m(h_T) G_m(h_R) \text{ V/m.} \quad (10.55)$$

The symbols are the same as in Equation (10.54), except that m denotes the mth TE mode.

***0.2.3.3 TE/TM Mode Structure.** Equations (10.54) and (10.55) show that each mode's contribution to the field is proportional to the product of four quantities: the excitation factor A, the transmitter height-gain function $G(h_T)$, the receiver height-gain function $G(h_R)$, and the propagation factor $\exp(-\alpha d/8.7)$. This section gives calculated values of these four quantities for a nominal ambient daytime ionosphere and an assumed ground conductivity of 10^{-4} S/m.

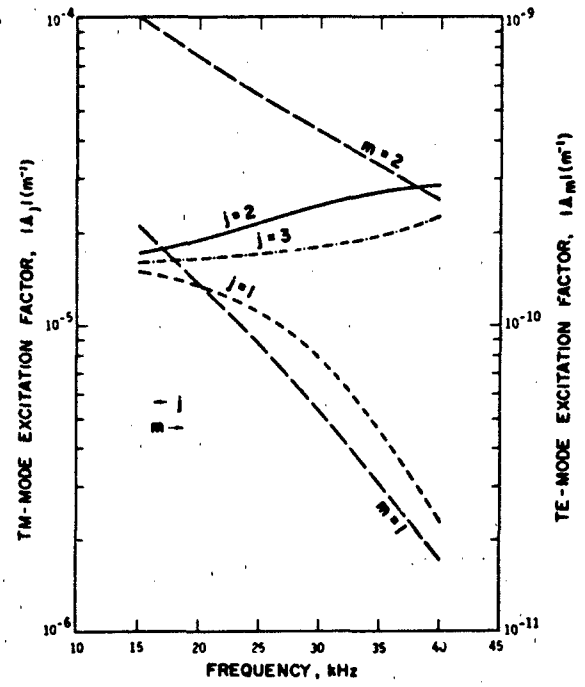


Figure 10-28. Excitation factors vs frequency for lowest three TM modes (j = 1,2,3) and lowest two TE modes (m = 1,2); ambient day, σ = 10^{-4} S/m [Field et al., 1976].

TE/TM Excitation Factors

Figure 10-28 shows the frequency dependence of the excitation-factor magnitudes of the first three TM modes and the first two TE modes. The first three TM modes are excited equally at the lower VLF frequencies, but above about 30 kHz the higher-order TM modes are much more effectively excited than the first one. The TE modes are excited much more poorly than the TM modes, by four or five orders of magnitude, as shown in Figure 10-28. The efficiency of TE mode excitation relative to TM mode excitation improves as the ground conductivity is reduced. For example, at 20 kHz the TE mode excitation factors are nearly two orders of magnitude greater for a 10^{-5} S/m conductivity than for 10^{-4} S/m. On the other hand, the excitation factor for the lowest TM mode is less by almost an order of magnitude if the conductivity is changed in the same fashion.

The excitation factors also depend on the state of the ionosphere. The excitation factors as defined here and shown in Figure 10-28 are inversely proportional to a quantity that becomes the "height of the ionosphere" in the limit of a sharply bounded ionosphere. For the diffuse ionospheres the excitation factors at the lower VLF frequencies are roughly proportional to the inverse of the average height at which important reflections occur. Thus, one would expect these factors to become somewhat larger under disturbed conditions.

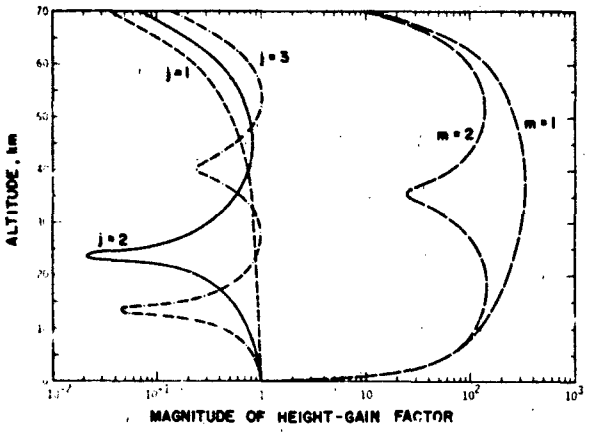


Figure 10-29. Height-gain factors for lowest three TM modes (j = 1,2,3) and lowest two TE modes (m = 1,2): ambient day, 20 kHz, $\sigma = 10^{-1}$ S/m [Field et al., 1976].

TE/TM Height-Gain Factors

The height-gain factor of the waveguide mode accounts for the effects of non-zero transmitter and receiver heights. The transmitter and receiver height-gain factors for a given mode are identical and, therefore, are equal when the transmitter and receiver are at the same altitude. Figure 10-29 shows computed height-gain factors for the first three TM modes and first two TE modes for a frequency of 20 kHz. These height-gain factors exhibit the classic height-dependences for antennas over a highly conductive ground; the TM mode height-gain factors are of the order of unity over most of the waveguide, except for some rather sharp nulls; and, above a few kilometers, the TE mode height-gain factors increase sharply to values well in excess of one-hundred. For elevated antennas, the large TE mode height-gain factor mitigates the effects of the small excitation factor, and these modes can be excited about as effectively as TM modes.

At frequencies above about 30 kHz the first TM mode develops a broad maximum in its height-gain factor in the 40–60 km altitude range. Such "whispering gallery" behavior is not important for ground-based or airborne terminals, but may be significant for very high, balloon-borne, terminals operating at higher frequencies [Videberg and Sales, 1973].

TE/TM Attenuation Rates

Figure 10-30 shows attenuation rates as a function of frequency for the first three TM modes and the first two TE modes. The higher order modes are more heavily attenuated than the lower, which often allows them to be neglected at VLF for long path-lengths. At the higher frequencies, the attenuation of the higher order modes can be mitigated by efficient excitation. Under normal conditions, therefore, it is usually necessary to retain many terms in the mode sums for frequencies throughout the LF band. It that case, geometric optics is often a more convenient approach to propagation analysis than is mode theory. In intense disturbances, however, higher order modes are much more severely attenuated than lower, and the mode sum can be used well into the LF regime. Figure 10-30 shows that the first TE mode is slightly less attenuated than the first TM mode, although that result depends on the specific normal daytime conditions and ground conductivity assumed.

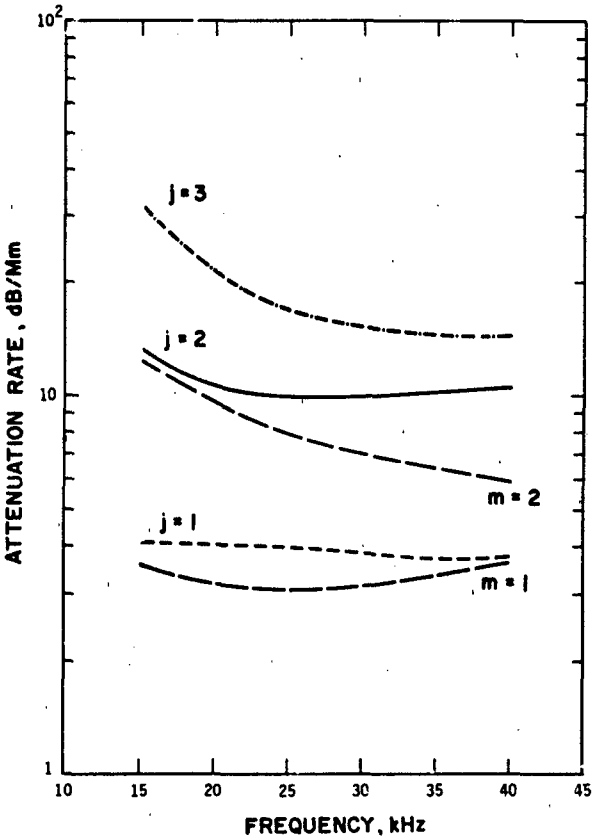


Figure 10-30. Attenuation rates vs frequency for lowest three TM modes (j = 1,2,3) and lowest two TE modes (m = 1,2): ambient day, $\sigma = 10^{-1}$ S/m [Field et al., 1976].

Figure 10-31 illustrates the ground conductivity dependence of the attenuation rates of the first TM and TE modes at 20 kHz. Results are given for normal daytime conditions and a moderate ionospheric disturbance, such as a solar proton event (SPE). The disturbance increases the attenuation rate of both polarizations over normal values. However, the TE attenuation rate is virtually independent of ground conductivity, whereas the TM rate exhibits a strong, broad maximum for conductivities between 10^{-5} and 10^{-4} S/m, where the TM eigenangle is near the Brewster's angle of the ground. The TM mode propagates somewhat better than the TE mode for most common ground conductivities, but propagates much worse over low-conductivity ground, such as occurs throughout Greenland and much of Canada.

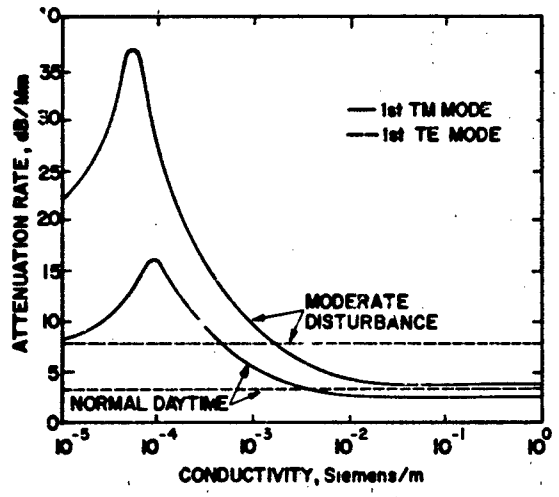


Figure 10-31. Attenuation rates of the lowest order TM and TE models vs ground conductivity, for 20 kHz under normal daytime and moderately disturbed ionospheric conditions [Field et al., 1976].

Although these results pertain to a frequency of 20 kHz, curves for other frequencies exhibit the same general behavior. The main difference is that the Brewster's angle peak in the TM attenuation rate occurs at higher values of conductivity for higher frequencies, and vice versa.

Except for propagation over very low-conductivity ground, TE modes are more vulnerable than TM mode to degradation in disturbed ionospheric environments. In the VLF band, for example, the TE mode attenuation becomes prohibitive for intense disturbances, as illustrated in Figure 10-32.

TE/TM Signal Strengths Versus Distance

Figure 10-33 shows 20 kHz TE and TM signal strengths, computed as a function of distance for ambient and intensely disturbed ionospheric conditions. The curves provide a direct comparison of TE and TM polarized signals radiated broadside from a trailing-wire antenna inclined 10 degrees with respect to the horizontal. Both the transmitter and the receiver are assumed to be at an altitude of 12.2 km, and the propagation path is over poorly conducting earth ($\sigma = 10^{-3}$ S/m).

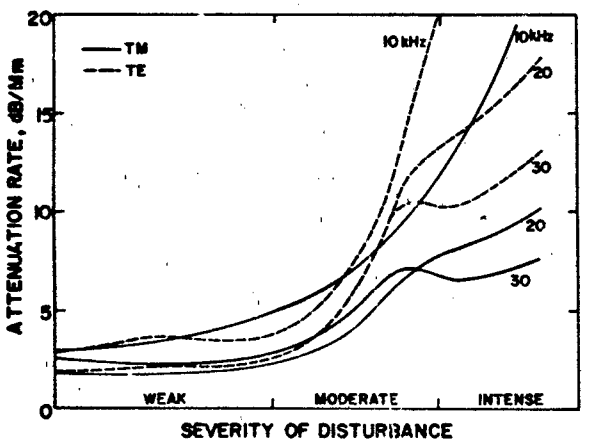


Figure 10-32. Attenuation rates of the lowest order TM and TE modes, for disturbed ionospheric conditions and a perfectly conducting earth [Field et al., 1976].

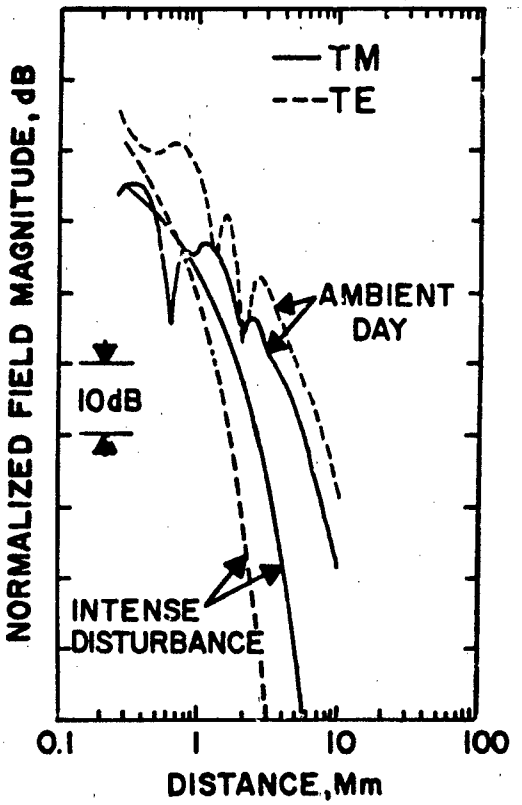


Figure 10-33. TE/TM signal strengths vs distance for ambient day and intensely disturbed ionospheric conditions: 20 kHz, $\sigma = 10^{-3}$ S/m, 10 degree antenna inclination and transmitter and receiver at 12.2 km [Field, 1982c].

In Figure 10-33 the ambient signals exhibit nulls and enhancements at ranges up to several megameters, caused by interference among several propagating waveguide modes. At greater distances, the higher order modes, which are more heavily attenuated than the lower order modes, diminish in importance and the signals fall off smoothly with distance. The curves also reveal several differences between signals in ambient and disturbed environments. First, and most important, if at least 2 or 3 Mm of the path are disturbed the disturbed signals fall well below the ambient signals. Such behavior is typical during strong solar proton events in the polar regions, and results when such widespread ionospheric disturbances depress the effective height of reflection in the ionosphere significantly. Second, mode interference patterns are nearly absent in disturbed environments, indicating that heavy attenuation of the higher order modes leaves only the lowest order modes to contribute

significantly to the signal strengths. Third, at ranges under approximately 1.5 Mm, the disturbed signals can be stronger than the ambient signals, because the disturbed environment increases the mode excitation factors but destroys interference nulls. However, if the disturbance covers most of the path, but not the transmitter and receiver, the attenuation rates increase but not the the excitations.

In Figure 10-33 the ambient TE signal is stronger than the TM signal at all distances, owing primarily to the relatively high transmitter and receiver altitudes and the nearly horizontal antenna orientation assumed in the calculations. Under the disturbed condition, the TE signal is more adversely affected than the TM signal and falls below it at most distances; however, as indicated by Figure 10-31, if the surface conductivity was reduced sufficiently, the TE signal would again become stronger than the TM signal.

10.2.3.4 Numerical Modeling of VLF/LF Waveguide Propagation. Numerous sophisticated computer programs have been developed for making VLF/LF field strength predictions. As described by Morfitt et al. [1982], the model developed at the United States Naval Ocean Systems Center (NOSC) is particularly attractive in that it incorporates (1) arbitrary electron and ion density distributions and collision frequency (with height), and (2) a lower boundary that is a smooth homogeneous earth characterized by an adjustable surface conductivity and dielectric constant. The model also allows for earth curvature, ionospheric inhomogeneity, and anisotropy resulting from the geomagnetic field. In addition, air-to-air, ground-to-air, and air-to-ground TE/TM propagation predictions can be made involving a horizontally inhomogeneous waveguide channel. The NOSC waveguide model can be used for computing long wave fields at fre-

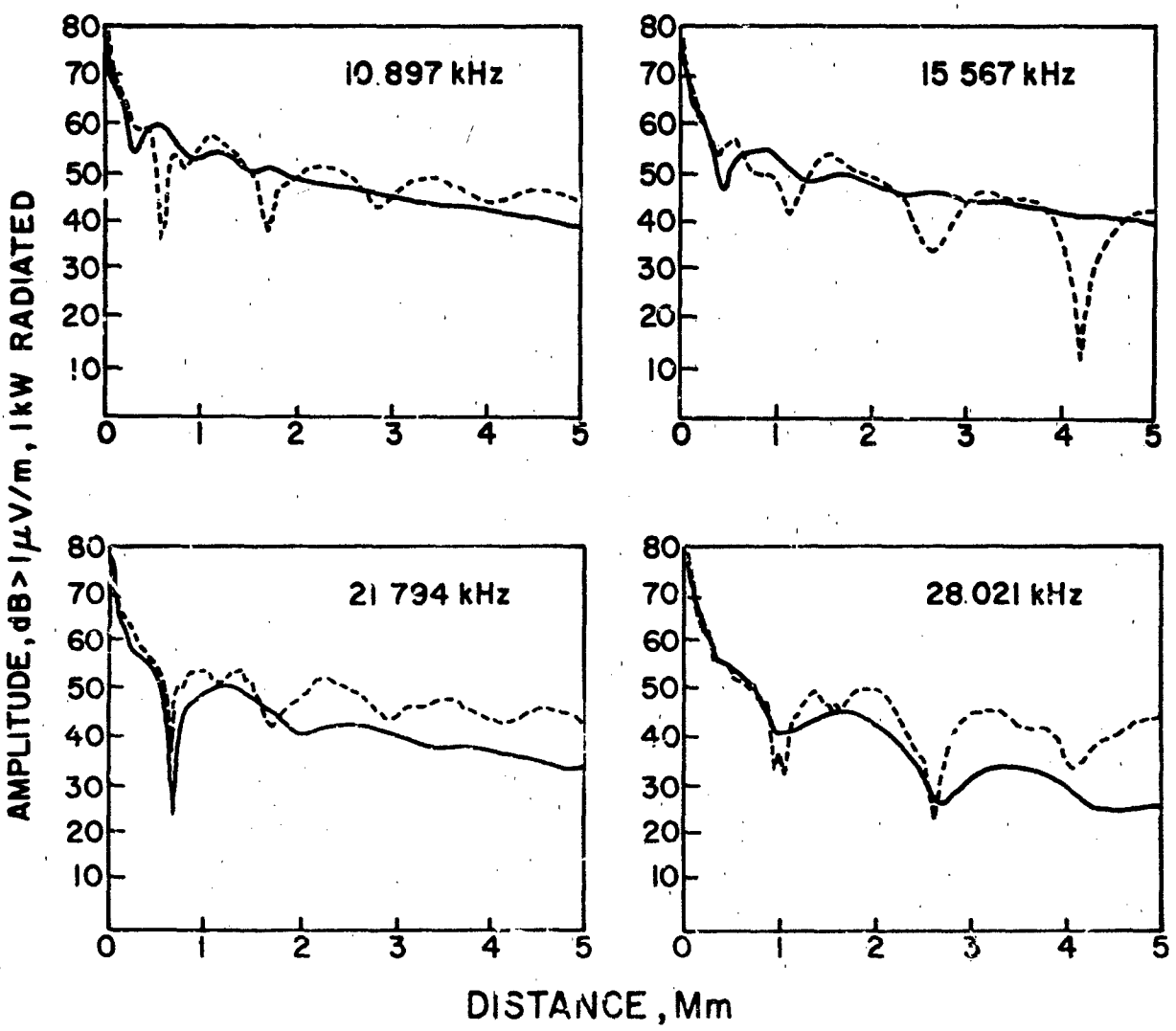


Figure 10-34. VLF signal calculations for a mostly seawater path from Hawaii to Sentinel, Arizona: —daytime, ----nighttime [Morfitt, 1977].

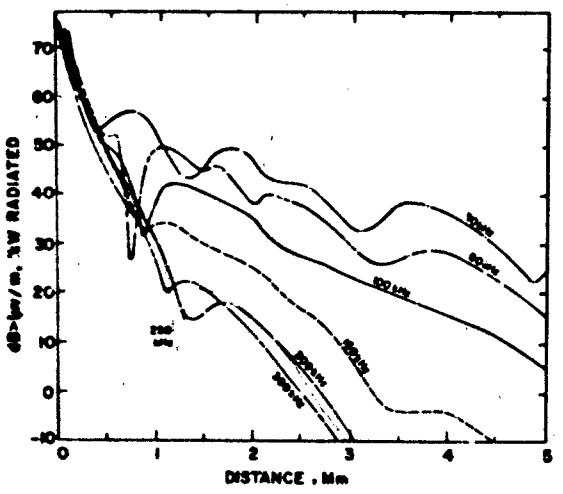


Figure 10-35. LF daytime signal calculations for midlatitude propagation [Pappert, 1981]

quencies as high as 300 kHz for daytime propagation [Pappert, 1981] and as high as 60 kHz for nighttime propagation [Morfitt et al., 1982]. Figures 10-34 and 10-35 show representative TM signal strengths computed by NOSC for nominal daytime and nighttime models of the ionosphere appropriate for midlatitudes, and for propagation over water. At VLF, many of the propagation predictions have been validated by NOSC airborne measurements, such as those described by Bickel et al. [1970].

10.2.3.5 Other VLF/LF Propagation Prediction Techniques. In addition to the waveguide mode formulations a number of other mathematical techniques have been developed for describing the propagation of VLF/LF waves. These include the zonal harmonic or spherical wave analysis method [Johler, 1964, 1966] and the wave-hop method [Berry, 1964, and Berry and Chrisman, 1965]. The spherical wave technique has the attractive feature that it can model variations in height in the earth-ionosphere cavity, but it requires the use of large-scale, very fast, digital computers for its implementation. For the higher VLF and LF frequencies the wave-hop method requires the least computing time, but it has not been formulated in a way that lends itself to modeling of discontinuities in the earth-ionosphere duct [Jones and Mowforth, 1982].

In the wave-hop approach the field strength at any point is the sum of the groundwave (see Section 10.2.1) and a series of "hops," which represent waves that have been reflected from the ionosphere and/or the ground. The hops are numbered according to the number of times they have been reflected from the ionosphere. Each reflection results in a reduction in hop amplitude so that usually a relatively small number of hops are needed to provide good field-strength estimates. For VLF propagation at distances less than about 1000 km and for LF propagation, in general, the number of hops required to describe the signal strengths are less than the number of waveguide modes that are required. The most general wave-hop formulations describe the propagation of TE/TM waves excited by an inclined dipole over a spherical earth with an anisotropic ionosphere [Lewis, 1970]. In addition to these sophisticated wave-hop techniques, other (simpler) approaches have been developed that provide good, quick, estimates of VLF/LF signal strengths, but without showing such propagation features as wave interference phenomena or polarization conversion effects [e.g., see Lewis and Kossey, 1975].

10.2.3.6 VLF/LF Probing of the Ionosphere. The propagation of long radio waves to great distances is controlled by the lowest regions of the ionosphere (usually the lower E region and the D region). As such, the variations in the amplitudes and phases of propagating long waves are very sensitive indicators of changes in the lower ionosphere. It is not surprising, therefore, that in addition to their uses for long-range communication and navigation long waves are used to assess the state of the lower ionosphere and as a tool for characterizing some of its properties.

Because of the extremely long wavelengths, ELF waves are affected by the electron and ion densities that are present over a very large range of altitudes. At night, for example, that altitude range can extend from below 50 km, to well up into the F region. Similarly, under disturbed conditions the electrons and ions at altitudes appreciably below 50 km can play an important (if not dominant) role in ELF propagation. However, owing to the difficulties involved in interpreting long-path ELF propagation data (which tend to represent an "average" of the state of the ionosphere), ELF has not been used extensively for ionospheric research.

The propagation of VLF/LF radio waves is controlled by the region of the ionosphere below about 90 km at night and below 75 km during the day. Unlike ELF, the observation of the signal characteristics of VLF/LF waves has provided a relatively simple ground-based technique for exploring the state and nature of the lower ionosphere. The technique has proven to be especially sensitive for monitoring ionospheric disturbances, such as those produced by solar x ray flares [for example, see Reder, 1969 and Kossey and Lewis, 1974], geomagnetic storms [Belrose and Thomas, 1968], electron precipitation events [Potemra and Rosenberg, 1973], ionospheric substorms [Svennesson, 1973], polar cap absorption events [Oelbermann, 1970], and high altitude nuclear bursts [Frisius et al., 1964 and Field and Engle, 1965]. The observation of continuous-wave transmissions over very long propagation paths has often been used for monitoring because such paths provide coverage over very large geographical areas. As with ELF, however, a disadvantage of such long path observations is that the effects of relatively localized disturbances are integrated, or smoothed-out, making it difficult to obtain information on the severity, extent, and structure of the disturbed region of the ionosphere.

Steep-incidence (that is, short-path) VLF/LF propagation techniques provide data on more localized regions of the ionosphere. However, with continuous-waves the direct and reflected components (groundwave and skywaves) overlap in space and time and can only be resolved indirectly by observing the interference pattern on the ground [Hollingworth, 1926] or by direct interpretation of diurnal phase and amplitude changes [Bracewell et al., 1951]. For example, the interference patterns produced by the groundwave and skywave from a 16 kHz transmitter were used by Bracewell and Bain [1952] to first suggest the presence of two ionized layers well below the ionospheric E-region.

Phase and amplitude observations can be used to characterize the steep-incidence VLF/LF reflection properties of the lower ionosphere. Of particular interest is the use of the data to determine effective heights of reflection and effective plane wave reflection coefficients of the ionosphere [Bracewell et al., 1951]. Such experimental data can be compared directly with that obtained theoretically, using full-wave computational techniques in conjunction with electron density and collision frequency models of the ionosphere [for example, see Budden, 1961, Pitteway, 1965 and Inoue and Horowitz, 1968]. Thus, the experimental data can be used to validate theoretical models of the ionosphere, such as those obtained from the chemistry of the upper atmosphere. In addition the data can be used to develop and validate phenomological models of the lower ionosphere, important for long wave propagation prediction [for example, see Bain, 1982].

More recently, with the advent of high-resolution VLF/LF pulse ionosounding [Lewis et al., 1973] it became possible to observe ionospheric reflections free of the ambiguities of the groundwave and skywave interference phenomena characteristic of continuous-wave measurements. The technique has been used to obtain a variety of steep-incidence reflection data at low-, mid-, and high geomagnetic latitudes [Lewis et al., 1973 and Kossey et al., 1974], and to investigate features of the C-layer of the daytime ionosphere [Rasmussen et al., 1980 and Rasmussen et al., 1982]. The technique provides a relatively direct means for determining VLF/LF ionospheric reflection heights and effective plane wave reflection coefficients, which then can be used to develop electron density models of the lower ionosphere [Kossey et al., 1983].

The inversion of steep-incidence VLF/LF reflection data to obtain electron density models of the lower ionosphere is not an easy task. Under quiet ionospheric conditions it is especially difficult, since usually the polarization conversion effects of the geomagnetic field cannot be ignored. Nevertheless, mathematical approaches that employ full-wave and iterative computational techniques have been developed and applied with some success [for example, see Shellman, 1970 and Field and Warber, 1984]. Under disturbed ionospheric conditions and certain daytime ambient conditions, when the VLF/LF reflections are controlled primarily by ionization below about 70 km, the effects of the geomagnetic field are greatly diminished and the ionosphere can be assumed isotropic. Under such conditions the mathematical inversion problem becomes somewhat simpler. Field et al. [1983] have developed an inversion technique, appropriate for isotropic propagation, which has been used in conjunction with VLF/LF pulse reflection data to derive conductivity profiles of the severely disturbed polar ionosphere.

A problem with profiles calculated by inversion is that of nonuniqueness, which can be caused by either incompleteness of data or the nonlinear dependence of the reflected signal on the propagation medium. In addition, the profiles characterize narrow regions of the ionosphere, since the propagation data contain information about only those altitudes where the ionosphere interacts appreciably with the reflected wave.

The altitude constraints are even more severe if long path propagation data are used, rather than steep-incidence reflection data. Nevertheless, some effort has been devoted to deducing the structure of the ionosphere from long path data. As described by Crain [1970], the data in this case are the attenuation rates and phase velocities of the propagating waveguide modes, and the analysis is a trial-and-error technique effectively to find an ionospheric conductivity profile which provides a waveguide mode or wave hop structure that agrees with the observed distribution of radio field strength.

In essence the long-path technique is similar to the steep-incidence approach insomuch as ionospheric reflection coefficients are calculated as an intermediate step in obtaining the mode constants. In order to synthesize the total field as measured, one has to take care to add in as many modes as contribute to the field. This can add a great deal of complexity to the application of the technique. Nevertheless, the technique has been applied with much success to develop phenomenological models of the lower ionosphere. Although such models may not be consistent in all respects with those derived from detailed analyses of the aeronomy of the upper atmosphere, they have found widespread application in long wave propagation prediction codes.

The results obtained by Bickel et al. [1970], Morfitt [1977] and Ferguson [1980] are especially noteworthy in that regard. They have performed detailed analyses of a large volume and a wide variety of VLF/LF propagation data and have derived analytic models of the lower ionosphere for propagation prediction. Those models are simply exponential height-profiles of conductivity, which can be specified by only two parameters, scale height and reference height. Following Wait and Spies [1964] the conductivity parameter, ω_r, depends on the ratio of electron density to electron-neutral collison frequency, and is taken to be of the form $\omega_r(z) = 2.5 \times 10^5 \exp(\beta(z-H'))$, where z (km) is altitude, β is the inverse scale-height (km^{-1}) and H' (km) is a reference height. The value of electron density $N(z)$, in electrons/cm^3, is calculated as a function of height by the equation $N(z) = 1.43 \times 10^7 \exp(\beta(z-H') - 0.15z)$. The

Table 10-5. Suggested exponential profiles for use in long wave propagation prediction codes. Frequencies, f, are in kHz [Morfitt et al., 1982].

Seasonal-Diurnal Propagation Condition	H' (km)	β (km^{-1})	Magnetic Dip (°)
Summer day	70	0.5	
Summer night	87	0.0077f + 0.31	
Winter day	74	0.3	
Winter night	80	0.035f − 0.025 (10 < f < 35)	90–75 (high latitudes)
	Linear change between high and middle latitudes		(transition latitudes)
		0.0077f + 0.31	<70 (middle latitudes)

collision frequency ν (collisions/s) is given by $\nu(z) = 1.82 \times 10^{11} \exp(-0.15z)$. Table 10-5 gives exponential profiles, based on VLF/LF propagation data, which are suggested for use in long wave propagation prediction codes [Morfitt et al., 1982].

10.2.4 MF Propagation

At night medium frequency skywaves can propagate to considerable distances with relatively little attenuation, but during the day the skywaves are severely attenuated in passing through the ionospheric D-region, so that only the groundwave provides usable signals. Thus, during the daytime MF signals are very stable, while at night they are much less so owing to the variability of the lower E-region of the ionosphere, and to interferences between the groundwave and skywaves.

10.2.4.1 MF Groundwave Propagation. The propagation of MF groundwaves can be described using the techniques discussed in Section 10.2.1. Because of the shorter wavelengths, however, such factors as the earth's atmosphere (and hence, the effective radius of curvature of the earth), terrain elevation, conductivity changes, and trees an l buildings along the propagation path usually influence MF groundwave characteristics to much larger extents than they do for VLF/LF waves. Most of these effects can produce strong local interference patterns in the amplitudes of MF groundwaves [Knight, 1982 and Hizal and Fer, 1982].

10.2.4.2 MF Skywave Propagation. Because there are so many factors that affect the characteristics of MF skywaves, it is difficult to draw a representative set of propagation curves. These factors include the losses associated with an imperfectly conducting earth at the transmitter and receiver, polarization coupling losses that depend on the geomagnetic field, ionospheric absorption losses, and losses associated with ground-reflections on multi-hop paths (see Figure 10-36). The problem is further compounded in that the ionospheric absorption losses alone show significant short-period and day-to-day variations, as well as diurnal, seasonal, latitudinal, and solar-cycle effects [Knight, 1982].

A number of techniques have been developed to estimate MF field strengths. A relatively simple technique, based on an empirical formula has been adopted by the CCIR [CCIR, 1978]. A more complex wave-hop method has also been developed by Knight [1975]. Figure 10-37 shows computed ionospheric losses over the 500-1500 kHz band for single-hop paths at mid- and low geomagnetic latitudes. In the auroral zones the ionospheric losses are somewhat greater than those shown in the figure, and the estimates do not apply if the waves penetrate the E-layer and are reflected by the F-layer. The latter is most likely to occur at frequencies above 1500 kHz [Knight, 1982].

10.2.4.3 Effect of MF Waves on the Ionosphere. Even relatively small electromagnetic fields impart appreciable energies to the electrons in the ionosphere causing their temperatures, and consequently their thermal velocities, increase. This increases the effective electron-neutral collision frequency; as a result, the complex dielectric constant of

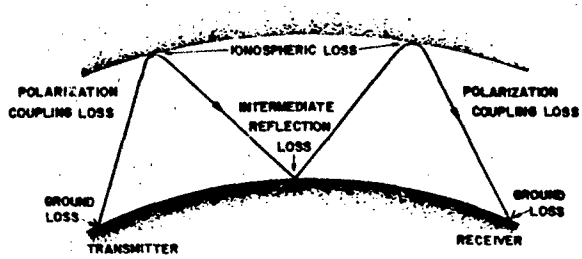


Figure 10-36. Losses associated with MF skywave propagation [Knight, 1982].

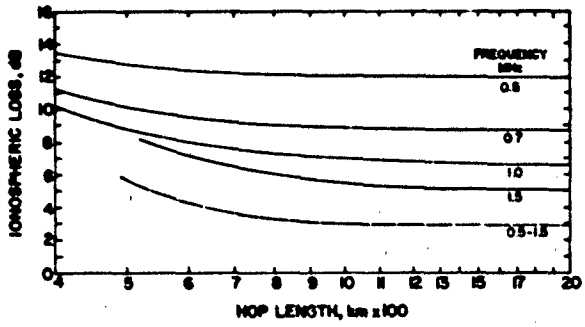


Figure 10-37 Computed MF ionospheric reflection losses:—east-west propagation at all latitudes,—north-south propagation at the magnetic equator [Knight, 1982].

the medium becomes appreciably dependent on the field. Thus, the associated physical processes, and the differential equations which describe the radio wave propagation in the ionosphere, become non-linear. This gives rise to various phenomena, including cross-modulation and de-modulation which, in principle, can be observed experimentally. The effects depend on wave frequency and the collision frequency and are such that they have been mainly observed in the MF band, particularly at night.

The main influence of the non-linear effects on the properties of radio waves reflected from the ionosphere manifest themselves throug', self-interference of the wave; that is, its influence on itself, and the interaction between a number of waves. These waves can be of the same frequency, or two independent modulated or unmodulated waves of different frequencies. An early observation of such an effect, the so-called Luxemburg effect, was reported by Tellegen [1933] who noted that the signal received in Holland from a 650 kHz Swiss station appeared to be modulated by the signals from a powerful station at Luxemburg (252 kHz). This phenomenon of cross-modulation has been found to be quite common when the unwanted, or disturbing, station is situated near the transmission path of the wanted wave [see Davies, 1969 and Al'pert, 1960].

Although the phenomenon of ionospheric cross-modulation or Luxemburg effect was discovered accidentally, it has been systematically investigated ever since its discovery. The motives of the earliest theoretical and experimental investigations were interest in the effect itself and the possible use of the effect for probing the ionospheric regions in which the cross-modulation occurs [Fejer, 1970]. In the latter case techniques have been developed to obtain D-region parameters such as electron density and collision frequency profiles.

Another MF probing technique, the "partial reflection" experiment, has proven to be a valuable method for obtaining quantitative measurements of electron distributions in the lower ionosphere. It has been one of the most extensively employed techniques for synoptic studies of the ionosphere below 100 km [Belrose, 1970, 1972].

10.2.5 Long Wave Propagation Through the Ionosphere

Electromagnetic waves cannot propagate in an ideal plasma unless the wave frequency is less than the plasma frequency, or approximately 8980 N, where N is the number of electrons per cubic centimeter. Thus, a density only slightly more than 1 el/cm^3 would suffice to completely reflect a 9 kHz wave. The well-known "whistler" phenomenon, however, demonstrates that under certain conditions long waves can penetrate even through the F-max region of the ionosphere, where the electron density is one hundred thousand times larger than would produce complete opacity if the ionosphere were a simple plasma. The long wave ionospheric transmission window is due to the geomagnetic field of the earth, which constrains the electron motion produced by electromagnetic waves incident on the ionosphere. As such the magnetic field provides a propagation mechanism.

The term "whistler" refers to an audio-frequency phenomenon associated with lightning discharges in the lower atmosphere. Electromagnetic energy at audio frequencies emitted by such discharges propagate in the ionosphere in a highly dispersive manner. The higher frequencies travel faster than the lower ones with the result that the signal, which was originally impulsive, is received over a relatively long time interval with the frequency generally descending with time (hence, a whistling sound).

Extensive studies have been made to determine the properties of whistlers [for example, see Storey, 1953, Ratcliffe, 1959, Pitteway, 1965 and Helliwell, 1965]. Their results [Watt, 1967] show that from the complete expression of the Appleton-Hartree equations for the refractive index, approximate expressions can be developed which give insight into the nature of the whistler mode. For example, the energy transmitted through the ionosphere is well-coupled to the whistler mode when the direction of propagation is in the same direction as the earth's magnetic field. For transverse propagation this coupling is very poor. Also, the coupling is increased for sharp gradients of refractive index at the ionosphere boundary. When collisions are included, there is a finite range of angles between the wave normal and the direction of the earth's magnetic field for coupling to the whistler mode. In an anisotropic ionosphere, the wave normal and the direction of energy flow along the field lines are different depending upon the magnitude of the refractive index and the static magnetic field. If ducts of ionization that are aligned with the earth's magnetic field exist, the waves can be guided in them.

VLF signals from terrestrial transmitters have been observed at satellite altitudes [Leiphart et al., 1962], and have been tracked from ground level to altitudes of 500 km or more [Orsak et al., 1965 and Harvey et al., 1973]. Such probes showed that the polarization changes from linear to circular as the wave penetrates the ionosphere, and that very significant delays in the signals occur.

Such phenomena can be calculated using full-wave tech-

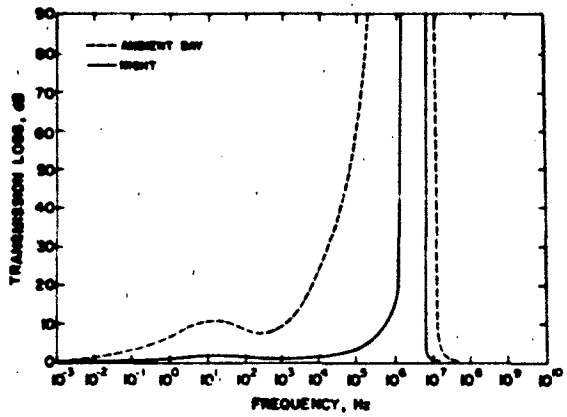


Figure 10-38. Computed long wave transmission losses through ambient day and night ionospheres [Booker et al., 1970].

niques [Pitteway, 1965] or WKB methods [Booker et al., 1970], in conjunction with appropriate models of the ionosphere. Figure 10-38 gives computed long wave transmission losses for plane waves incident on the ionosphere in a direction parallel to the geomagnetic field. The results are representative of those expected for ambient ionospheric conditions at mid- and high geomagnetic latitudes. Under disturbed conditions or at very low geomagnetic latitudes the penetration losses are much more severe, especially for frequencies above a few kilohertz [Booker et al., 1970 and Harvey et al., 1973].

10.3 IONOSPHERIC MODELING

For successful radio communication, it is essential to predict the behavior of the ionospheric region that will affect a given radio communication circuit. Such a prediction will identify the time periods, the path regions and the sections of high frequency bands that will allow or disrupt the use of the selected high frequency communication circuit. This need for prediction leads to modeling of the ionosphere.

A model is a numerical statistical description of the ionosphere in terms of location (geographic or geomagnetic latitudes and longitudes), time (solar zenith angle), seasons, and other factors such as the solar activity (10.7 cm flux, sunspot number). The empirical equations are derived from the dependence of the observed phenomena on variables mentioned above. These observed phenomena include: the behavior of critical frequencies* foE, foF1, foF2, and foEs for the E, F1, F2, and sporadic E layers; the altitudes (h_mE, h_mF1, and h_mF2) for peak (maximum electron densities for the layers, and the half-thickness widths** y_mE, y_mF1, and y_mF2. These models are called phenomenological models [Barghausen et al., 1969; Bent et al., 1972; Ching and Chiu, 1973; Chiu, 1975; and Kohnlein, 1978].

The ITS-78 model [Barghausen et al., 1969] based on the analysis by Jones et al. [1966] of world-wide, ground based ionosonde data, predicts only the bottomside of the ionosphere. The Bent model [Bent et al., 1972] predicts the total electron content of the ionosphere in the altitude range from 150 to 2000 km, without a direct consideration of the bottomside E and F1 layers. The Ching and Chiu [1973] model covers the altitude range from 110 to 1000 km. Instead of parabolic layers [Barghausen et al., 1969] they assume Chapman functions for the electron density distributions in the E, F1, and F2 layers. Later Chiu [1975] modified the Ching-Chiu model to incorporate the polar ionosphere. Their models are useful only for studying the large scale phenomena such as global thermospheric and ionospheric calculations.

Using ionospheric data from ESRO satellites, Kohnlein [1978] extended the altitude range up to 3500 km. He suggested a "differential approach" for the ionospheric modeling. He separated small scale spatial structures such as the equatorial trough, the midlatitude trough and the polar ionosphere, from the large scale global structure. He modeled these individual structures and added them into the global structure. His method reduces the number of coefficients otherwise needed to model the complicated ionospheric behavior.

The other approach for ionospheric predictions is to use theoretical models [Stubbe, 1970; Strobel and McElroy, 1970; Nisbet, 1971; Oran et al., 1974; and Oran and Young, 1977]. These are based on the physical processes responsible for the production, maintenance and decay of the ionosphere. A theoretical model would thus rely on the process of ionization of neutral atmospheric constituents by the incident solar extreme ultraviolet radiation, the transport processes such as diffusion and neutral winds, and also on the effect of electric and magnetic fields on the transport processes. Essentially the theoretical model tries to explain the experimental observations in terms of known physical processes. In addition, this approach seeks new physical processes to explain the differences between the observational results and the predictions based on the theoretical considerations.

Radio communication can be divided in two main categories. Ground-to-ground radio communication is based on the reflection and scattering characteristics of the ionospheric layers. On the other hand, ground to satellite, satellite to ground, or satellite to satellite radio communications depend on the transmission and refraction characteristics of

*The critical frequency is the limiting radio frequency below which a radio wave is reflected by, and above which it penetrates and passes through, the ionized medium (an ionospheric layer) at vertical incidence.

**The half-thickness width y_m of the ionospheric layer is determined with the assumption that the layer has a parabolic shape [Appleton and Beynon, 1940].

the ionosphere. The main goal of any modeling effort is to predict the periods of good or poor radio communications for the selected paths to enable a continuous undisrupted communication through the ionosphere or by some other means.

We will consider several ionospheric models that are routinely used (or are available) for the prediction and specification of the ionosphere. The emphasis here is on acquainting the user with the modeling programs and their limitations. We do not attempt to review the scientific literature for a determination of the state of the art of modeling efforts. Therefore only the essential references will be cited.

First, we will consider the numerical-phenomenological models. Then we will consider the theoretical models. This will be followed by the modifications to models to take into account high latitude phenomena such as the auroral E layer and the midlatitude F region trough. In the concluding section, we will look at the limitations of these models and a possible approach to overcome these limitations.

10.3.1 The Numerical-Phenomenological Models

At present the three most widely used numerical models for ionospheric predictions are (1) The ITS-78 model, (2) The Bent model, and (3) the *Ion*ospheric *C*ommunications *A*nalysis and *P*rediction Program (IONCAP). In addition, the 4-D model of the Air Force Global Weather Central and the Bradley model will be considered. We will also look at the International Reference Ionosphere-IRI 79.

10.3.1.1 The ITS-78 Model. The main purpose of this model [Barghausen et al., 1969] is to predict long term performance of communication systems operating in the 2–30 MHz frequency range. The ITS-78 model and its computer program was developed by the Institute of Telecommunication Sciences, ESSA, Boulder, Colorado. The model is based on the presentation of the ionospheric characteristics in a form of synoptic numerical coefficients developed by Jones and Gallet [1960] and improved by Jones et al. [1966]. The important features of the ITS-78 model are the parameters for the D, E, Es (sporadic), and F2 layers of the ionosphere.

The model provides (output) circuit operational parameters such as the maximum usable frequency (MUF), optimum traffic frequency (FOT), and the lowest usable frequency (LUF). In addition to the regular E-layer propagation mode, it takes into account propagation via the sporadic E layer. The program computes all the probable modes. It computes the system performance. For that purpose it calculates the antenna patterns and gains for 10 most commonly used antennas. It also has a program to determine MUF as a function of the magnetic activity index Kp.

The inputs for the ITS-78 model are the date, Universal Time, geographical location (latitude and longitude) of the transmitter and receiver, and sunspot number. To compute the system performance the model needs the antenna parameters, the radiation power of the transmitter, and the signal to noise ratio of the receiver.

For the D region the model considers only the absorption losses. The non-deviative absorption is in the form of a semi-empirical expression. It enables the user to compute the losses for the HF frequencies penetrating the D layer. The deviative absorption losses are included in the loss calculations as uncertainty factors.

For the E region the model computes the parameter foE. It assumes a constant height of 110 km for the maximum (peak) electron density of the E layer, with a constant semi-thickness of 20 km. The numerical coefficients for foE are based on the experimental ground ionosonde data during high solar activity phase in 1958, and the low solar activity phase in 1964.

For the F2 region the model computes the parameters foF2, the height of maximum electron density h_mF2 and the semi-thickness y_mF2 of the F layer. These are in the form of numerical coefficients for the high (1958) and low (1964) phases of the solar activity. Both the E and F2 layers are assumed to be parabolic in shape.

The sporadic E (Es) layer could be very helpful or harmful to radio communications depending on the nature of the Es layer. A blanketing, totally reflecting Es layer extends the frequency range of the E-mode communications. However a semi-transparent or partly reflecting E_s layer would cause serious multipath and mode interference and would be detrimental to communication systems. Using numerical coefficients, the ITS-78 model computes foE_s only for the ordinary ray. (The earth's magnetic field splits the incident ray into the ordinary and the extraordinary rays.) The numerical coefficients are for both the high (1958) and low (1964) phases of solar activity. As the model does not predict the occurrence of Es, the foEs maps are used only when propagation via regular E layer is not possible. To compute the system performance, the model incorporates three kinds of noise: galactic, atmospheric, and manmade.

To determine the operational parameters such as the maximum usable frequency (MUF), the model computes the path geometry (between the transmitter and the receiver). The parameters computed in the path geometry are the path length, path bearing (azimuth), and the solar zenith angle χ of the sun. The model computes paths for reflections from E, Es, and F2 layers. These are called the E, Es, and F2 modes. The paths could involve more layers (multiple modes) and more reflections (multihop).

To determine wave propagation the electron density distribution with altitude is needed. Both the E and F layers are assumed to have parabolic shapes. The maximum usable frequency (MUF) is obtained by multiplying the critical frequency of the layer by the MUF factor M(3000). The term in parentheses refers to the standard ground distance

of 3000 km between the (hypothetical) transmitter and the receiver. The experimental data for the numerical factor M(3000) (in terms of coefficients) come from 13 ionosonde stations covering the geomagnetic latitude range from 71°S to 88°N.

The stability and predictability of the E layer results in a 99% probability (highest) of supporting radio propagation and communication via the E layer. The next highest probability is via the regular F layer. When neither of the above modes is possible, the Es mode is considered for communication.

For computing the system performance the program allows a selection from 10 antenna patterns. The program takes into account the ground losses, ionospheric losses, free space losses, and the excess losses. The program computes the radio communication circuit reliability, service reliability, and the multipath evaluations.

The ITS-78 model has several limitations. The results from the model are useful only when the operating frequency is below the maximum usable frequency. The model assumes that transmission will be by reflection from the ionosphere. For this the transmitter and the receiver must be on the same side of the ionospheric layer (for example, ground-to-ground communication. The model does not take into account the daytime F1 layer which usually develops between the E and F2 layers. The model does not adequately account for the electron density above the altitude of h_mF2. Finally, the model does not take into account the dependence of absorption on the operating frequency in considering the D layer absorption.

10.3.1.2 The Bent Model. The Bent Model [Bent et al., 1972; Llewellyn and Bent, 1973] is basically for ground-to-satellite communication but can be adapted for ground-to-ground or satellite-to-satellite communication. The main purpose of the Bent model is to determine the total electron content (TEC) of the ionosphere as accurately as possible in order to obtain high precision values of the delay and directional changes of a wave due to refraction. Ground-to-satellite communication demands operating frequencies which are higher than the MUF. Thus the mode involves the transmision refraction characteristics of the F2 layer and the electron density distribution above the height of the F2 peak must be known.

The model provides (output) the vertical total electron content above the transmitter, the profile of vertical electron density with attitude, and the total electron content along the path between the satellite and the ground. It also provides the refraction corrections to the elevation angle, the range, and the range rate.

The input parameters to the model are the date, Universal Time, locations of the transmitter and receiver (ground and satellite), rate of change in elevation and altitude of the satellite, operating frequency, the solar flux (10.7 cm flux), and sunspot number.

The data base of the Bent model consists of 50 000 topside ionospheric soundings, 6000 satellite measurements of electron density and 400 000 bottomside soundings of the ionosphere. The data extend from 1962 to 1969 to cover the maximum and the minimum of the solar cycle.

The bottomside data are foF2 hourly values from 14 stations in the American sector covering geographic latitudes 76° N to 12° S or geomagnetic latitudes from 85° to 0°. The topside soundings cover the period 1962 to 1966, with geomagnetic latitude range 85° to −75°, and the electron density profiles are from 1000 km down to the altitudes of the foF2 peak (h_mF2). The satellite data are from the Ariel 3 satellite covering the period May 1967 to April 1968 and are linked with real time foF2 from 13 ground stations. Thus the data base of the Bent model refers to solar cycle 20, while the data base of the ITS-78 model is from solar cycle 19.

The Bent model uses foF2 from the ITS-78 model. Instead of the monthly median values, the Bent model computes average values for every 10-day interval of the month from the 10.7 cm solar flux input. For the height h_mF2, it uses M(3000) factors of NOAA (ITS-78) in terms of the sunspot number. It uses an empirical polynominal for M in computing h_mF2, in place of the Shimazaki equation [1955] used by ITS-78.

The distribution of electron density with altitude, assumed by the Bent model for the computation of the total electron density is shown schematically in Figure 10-39. Starting from the bottom, it divides the profile into five sections: a bottom bi-parabolic F2 layer; a parabolic F2 layer above the peak; and three exponential sections to cover the altitude above h_o (h_mF2 to 1000 km). The construction of the profile needs the parameters k_1, k_2, k_3, Y_1, Y_m, foF2, and h_mF2. The last two have already been explained. The dependence of the other parameters on geomagnetic latitude, solar flux, foF2, and the season is from work of Bent [Llewellyn and Bent, 1973]. The topside and the first adjoining exponential section are matched at a height d (above h_m) by the equation

$$d = h_o - h_m = \frac{1}{k_1}\left[\sqrt{1 + Y_1^2 k_1^2} - 1\right], \quad (10.56)$$

where Y_1 is the half thickness of the F_2 layer and k_1 is the exponential constant.

The remaining profile above the F2 peak [of altitude range (h_m + d to 1012) km] is divided in three equal intervals of altitude.

The model can predict with an accuracy of 75%–80%. If the model is updated with observed recent data within a range of 2000 km radius (from the transmitters), the predictability is improved to 90%.

Though the model predicts total electron content (TEC) with good accuracy, the model does not have separate E and F1 layers. As the model was constructed for the TEC,

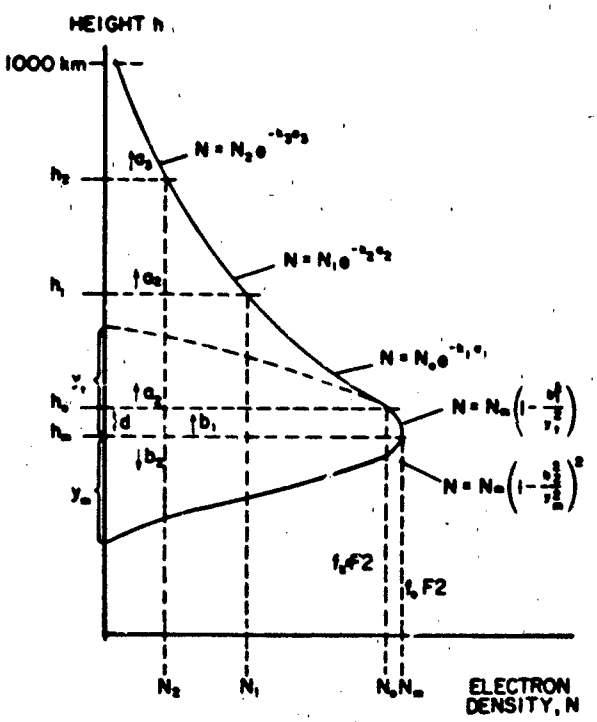


Figure 10-39 Schematic for exponential and bi-parabolic profiles for the electron density distribution with altitude for the Bent model.

the E and F1 layers are included as the bi-parabolic bottomside of the F2 layer. Also, the program does not take into account the non-deviative absorption in the underlying D layer.

10.3.1.3 The Ionospheric Communications Analysis and Prediction Program (IONCAP). The IONCAP [Lloyd et al., 1978] is essentially the latest, improved, and more versatile and flexible version of the ITS-78 model.

It provides 30 output options which can be divided into four categories, (1) ionospheric description, (2) antenna patterns, (3) MUF predictions, and (4) system performance predictions.

For ionospheric predictions it provides monthly median values for the parameters foE, foF1, foF2, h_mE, h_mF1, h_mF2, Y_mE, Y_mF1, and Y_mF2. It also provides the lower, median, and upper decile values of the minimum foE or foE_s. It can also provide a prediction in the form of a plot of operating frequency with virtual height and also with true height.

The MUF option of the output provides the minimum radiation angle and the M factors for all four modes, E, F1, F2, and Es. The plots for the diurnal variation of the MUFs are also available. The MUFs provide the description of the state of the ionosphere and do not include any system parameters. The operating frequency for a given radio communication circuit is the critical frequency of the layer multiplied by the MUF factor.

In the system performance options, 22 performance parameters are available. The program for the antenna output option computes the elevation angles and the operating frequencies for optimum antenna geometry and its gain.

Inputs for the program are the date, Universal Time, geographical locations of the transmitter and the receiver, and sunspot number. The program can accept external ionospheric parameters as input to the program. For antenna pattern, one can select the antenna from 17 antennas in the program (7 antennas from ITS-78 have been modified). For the system performance additional inputs such as radiation power of the transmitter, and the S/N ratio of the receiver are needed.

The schematic for the electron density distribution with altitude for the IONCAP program is shown in Figure 10-40. The model has 3 parabolic layers, E, F1, and F2. The altitudes for the peak electron densities are h_mE, h_mF1, and h_mF2. The half thickness widths for the layers are Y_mE, Y_mF1, and Y_mF2 respectively. For the F1 layer, h_mF1/Y_mF1 is assumed to be 4. The E layer has fixed altitudes h_mE = 110 km and Y_mE = 20 km. IONCAP improves on the ITS-78 model by incorporating D and F1 layers. The D layer contribution is considered indirectly by adding an exponential tail for the E layer down to the altitude of 70 km. In the transition region between the E and F1 (or F2 if F1 is absent)

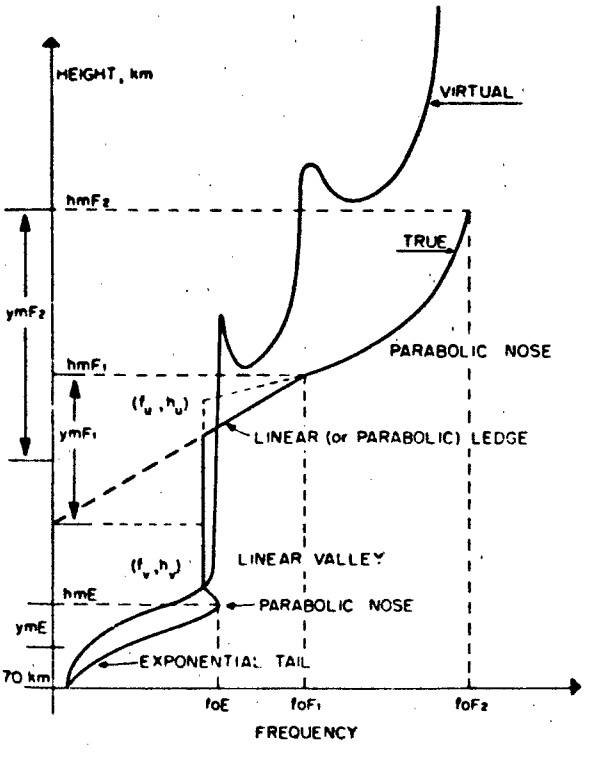


Figure 10-40 Schematic for the electron density profile and virtual height for the IONCAP model.

layers, the electron density is assumed to be linear for the frequency range f_l to f_u where $f_l = x_l \times foE$ and $f_u = x_u \times foE$. Typical values for x_l and x_u are 0.85 and 0.98 respectively. Thus the electron density decreases above the parabolic nose 0.85 foE (<foE) and continues upwards up to 0.98 foE (<foE) producing a linear valley in the transition region. When the $x_u = x_l = 1$ the valley is absent in the transition region and the curve is a vertical line starting at the tip of the parabolic nose of the E layer. The F1 layer forms a linear or parabolic ledge depending on the magnitudes of h_mF1, h_mF2, foF1, and foF2. In the ITS-78 model, the F1 layer is assumed to be absent. In the IONCAP model the numerical coefficients for foE are functions of geographic latitude for both solar maximum and minimum from the work of Leftin [1976]. The model uses foF1 maps of Rosich and Jones [1973]. It also takes into account the retardation below the F2 layer.

For the MUF computations the model uses the corrected form of Martyn's theorem. As the absorption equations using the secant law do not work for lower frequencies at altitudes below 90 km, these equations have been modified in the IONCAP program. The IONCAP provides two programs (1) the ITS-78 short path geometry and (2) the long path (>10 000 km) geometry. In addition to the ITS-78 model, the path computations now include the F1 mode, the over-the-MUF mode, D and E region absorption losses, and sporadic E losses. A correction to frequency dependence is added for low frequencies reflected from altitudes below <90 km.

The improvements over ITS-78 can be summarized (see IONCAP) as follows:

1. The description of the ionosphere is now more complete.
2. The loss equations have been supplemented. This includes E mode adjustments, sporadic E effects, over-the-MUF losses, and losses for low reflection heights.
3. The ray path geometry calculations have been revised. This was an empirical adjustment to Martyn's theorem.
4. The loss statistics were revised to include the effects of the sporadic E layer and of over-the-MUF modes.
5. A separate long path model was developed.
6. The antenna gain package was revised.

All models predict only the quiet ionosphere, which shows a large systematic dependence on latitude, longitude, season, time, and sunspot activity.

10.3.1.4 The Bradley Model. The Bradley model contains two modifications to the existing models: (1) the filling of the valley between the E and F layers (F1, or F2 if F1 is absent), by parabolic layers [Bradley and Dudeney, 1973] and (2) a simple formulation of the prediction of the probability of the high-frequency propagation [Bradley and Bedford, 1976].

The assumption that the electron density distribution in the E and F layers is parabolic in shape, results in a valley—a reduction in electron density between the two peak values, N_mE and N_mF—at altitudes h_mE and h_mF, respectively. The rocket observations have shown that in reality the electron density at any altitude between the altitudes h_mE and h_mF is rarely smaller than the peak electron density N_mE. Thus the assumption of the parabolic shapes for E and F layers underestimates the electron density in the altitude region between h_mE and h_mF. To correct such an underestimation Bradley and Dudeney [1973] suggested a linear distribution of electron density from foE to 1.7 foE. The lower end is at h_mE. At the upper end, the F layer is parabolic in shape down the altitude where the plasma frequency is 1.7 foE. This linear interpolation has not yet been incorporated in the IONCAP model (see Figure 10-40) of Lloyd et al. [1978].

In high frequency prediction it is essential to know the probability of communication at any particular operating frequency f. For convenience the observed data are expressed as follows: F_u is the upper decile, f_m is the median, and F_l is the lower decile of the ratios of f/f_m. The distribution functions of F_u and F_l are not simple Gaussian distributions ($F_u - MUF = MUF - F_l$ for Gaussian). The distributions are χ^2-distributions of F_u and F_l. For probability determinations these two χ^2-distributions (of F_u and F_l) have to be used. Two variables F_u and F_l with their associated degrees of freedom, and the need to integrate the χ^2-distribution curve makes the process of determining the probability distribution very cumbersome. Bradley and Bedford [1976] derived simple empirical equations for this probability distribution. The equations are

$$Q = 130 - \frac{80}{1 - \dfrac{1 - (f/f_m)}{1 - F_l}} \qquad (10.57)$$

or 100, whichever is smaller for $f \leq f_m$

and

$$Q = \frac{80}{1 - \dfrac{(f/f_m) - 1}{F_u - 1}} - 30 \qquad (10.58)$$

or 0, whichever is larger for $f > f_m$,

where

Q - is the cumulative probability,
f - is the operating frequency,
f_m - is the predicted median frequency,
F_l - is the lower decile (of f/f_m),
F_u - is the upper decile (of f/f_m).

They note that the probability distribution from their simple empirical equation is as good as, though not always better than, that from the χ^2 distribution. Therefore they highly recommend a replacement of χ^2 distribution proce-

dure by these equations for a determination of the present probability that signals will propagate at a given hour over a given sky-wave path.

The latest computer model like the IONCAP has not incorporated "Bradley Features" in its program.

10.3.1.5 The Air Force Global Weather Central 4-D Model. The input data to the 4-D numerical model [Flattery et al., personal communication] are the critical frequencies for the layers and M(3000), real time or near real time observations from 40 ground stations around the world, and total electron content (TEC) from eleven stations. The frequency of observations varies from hourly (best) at one end to weekly (worst) at the other end. The desired purpose of the 4-D model is to produce a consistent ionospheric specification anywhere in the northern hemisphere for a 24-h period. In that sense it is not a forecasting model like the other models mentioned above.

This model has three ionospheric layers, E, F1, and F2. Each layer is represented by a Chapman distribution function

$$N_e(h) = Ne_{max} \exp\{a[1 - z - \exp(-z)]\}, \quad (10.59)$$

where a refers to the loss mechanism and z is given by $z = \frac{h - h_{max}}{h_s}$ and h_s is the scale height for the layer. At a given altitude the total contribution to electron density is the sum of contributions by all three layers.

For any height the electron density is approximated by

$$N_e(\ell) = \sum_{k=1}^{N} a_k W_k(\ell),$$

where a is the weighting factor and $W(\ell)$ is an empirically derived set of discrete orthogonal functions for the altitude interval ℓ. The 95 to 2000 km range is divided into 127 intervals. The widths of the intervals range from 5 km at the lowest altitudes to 50 km at highest altitudes. The empirically derived function $W(\ell)$ is in two parts, spherical harmonic functions for spatial dependence and trigonometric functions (sine, cosine) for temporal dependence.

With the help of these variables a_k and W_k, the entire data base for the ionosphere is reduced to a limited number of coefficients. These can be used to construct the electron density profile for any location in the Northern Hemisphere valid for a 24-h specification period. The model is still being developed. The specification accuracy of the model will depend strongly on the frequency and reliability of the input data—real-time experimental observations from the 40 ground stations of the northern hemisphere. Also the quality of specifications interpolated for locations inside the network will be better than those extrapolated outside the observational network.

10.3.1.6 International Reference Ionosphere—IRI 79. The IRI 79 [Rawer, 1981] is the latest addition to the continued efforts of ionospheric modeling. The emphasis of IRI 79 is to summarize the experimental data from rockets and satellites to provide true height profiles of the ionosphere. The model serves as a standard reference for various purposes such as design of experiments, estimation of environmental and other effects, and testing theories. The model gives the altitude dependence of four parameters: electron density, electron and ion temperatures, and the composition of positive ions. It computes the density for atomic ions O^+, H^+, He^+ and for molecular ions O_2^+ and NO^+.

For the worldwide description of the peak electron density, the model uses foF2 from CCIR [1967] coefficients, with modified dip coordinates [Rawer, 1963]. As the foF2s are from the ground based ionosonde stations, the model really computes a relative distribution of electron density with true height, with respect to that of the foF2 peak. For a true peak height the model uses an empirical relation [Bilitza et al., 1971] with M(3000)foF2 coefficients from CCIR [1967]. This empirical relation is based on the incoherent scattering measurements which yield electron density with true height. The model has an alternate procedure based on the results of Chiu [1975] to replace the foF2 and M(3000)foF2 coefficients from CCIR. This procedure when used limits the ability of reproducing the complex ionospheric features available from the CCIR coefficients. The IRI 79 can also use direct data of the peak electron density and the peak altitude for computing the profiles.

The schematic for the altitude dependence of the electron density for the IRI 79 model is shown in Figure 10-41. The altitude range from 80 to 1000 km is divided into six sections: topside, F2 bottomside, F1, intermediate, valley, and E/D regions respectively. The topside region is modeled with the use of 'harmonized Bent' model [Ramakrishnan et al., 1979]. The bottomside F2 is expressed as the sum of Epstein Transition Functions [Rawer, 1981]. The F1 region is based on the work of Eyfrig [1955] and Ducharme et al. [1973]. The intermediate region fills the gap between the valley region and the F1 layer. The rocket measurements compiled by Maeda [1971] determine the shape and the depth of the valley region. The foE is from Kouris and Muggleton [1973a,b]. The model also takes into account the contribution from the D layer. The model does not account for the highly variable Es layer.

The IRI 79 is the only numerical model with information on additional parameters such as the electron and ion temperatures and the composition of positive ions. The composition is determined with the assumption that the plasma is electrically neutral above 84 km. The model also computes the distribution of cluster ions in the altitude range 80 to 90 km.

The inputs for the program are location (latitude and longitude), sunspot number and time. The optional inputs are the peak altitude and peak electron density. The output consists of 11 parameters: absolute electron density, relative

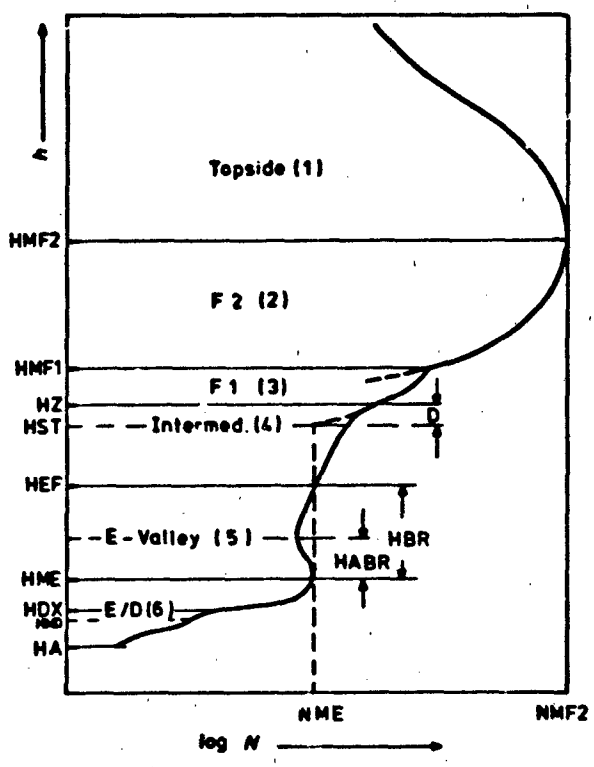


Figure 10-41 The IRI 79 model profile. For details refer to Rawer [1981].

electron density, neutral temperature, electron temperature, ion temperature, ratio of electron to ion temperature, percent concentrations of O^+ (and N^+), H^+, He^+, O_2^+ and NO^+ ions. The accuracy (σ) of predictions is as below:

	Height	Peak Density	Temperature
F region	± 15%	± 30%	± 30%
E region	± 5%	± 10%	± 10%

The programs are in FORTRAN-4 and ALGOL-60 computer codes. These programs are available from the World Data Center, Boulder, Colorado.

10.3.2 The Theoretical Models

The theoretical models for the ionosphere are based on the physical processes responsible for the observed ionospheric phenomena. The processes responsible for the ionosphere are production, maintenance, and decay of the ionosphere. As the approach deals directly with the physical processes, and not with the observed phenomena, the emerging model is called a physical model.

Four models are summarized in Table 10-6 to show several variations in the same processes considered by different workers. Strobel and McElroy [1970] considered only the F2 region (200 to 700 km), whereas others took into account the altitude range from 120 to 1200 km. Nisbet [1971] constructed the first computer-based simple physical model MK-1 for the ionosphere. He considered only three neutral constituents N_2, O_2 and O, whereas Stubbe [1970] and Oran and Young [1977] also considered the minor constituents He and H. For the dissociation and ionization of the neutral species, the incident solar EUV radiation in the range 30 to 1912 Å is used, along with the wavelength dependent absorption and ionization cross sections for the neutral species. Nisbet considered three basic predominant ionic species: O^+, NO^+, and O_2^+. Oran and Young [1977] took into account the additional ionic species H_2^+, Ne^+, N^+ and H^+. One has to consider the chemical reactions that produce ions by charge exchange processes. Nisbet [1971] used 5 reactions whereas Oran and Young [1977] used 24 chemical reactions [see Strobel and McElroy, 1970]. For maintenance of the ionosphere, the processes of diffusion and photoionization are assumed. The processes of dissociative recombination and radiative recombination are

Table 10-6 Variations in the physical processes used in the theoretical models

Processes	Nisbet [1971]	Stubbe [1970]	Strobel and McElroy [1970]	Oran and Young [1977]
In the Altitude Region (km)	120–1250	120–1500	200–700	120–1200
Neutral Constituents for Ionization	N_2, O_2, O	N_2, O_2, O, He, H	N_2, O_2, O, He	N_2, O_2, O, He, H
Chemical Reactions (Charge-Exchange)	5 Reactions	10 Reactions	4 Reactions	24 Reactions
Ionized Constituents	O^+, NO^+, O_2^+	O^+, NO^+, O_2^+, H^+	O^+	O^+, NO^+, O_2^+, H^+ N^+, He^+, N_2^+
Neutral Winds	—	Horizontal Winds	Horizontal Winds	Horizontal Winds
Electric Fields	—	Yes	—	—
Magnetic Fields	—	Yes	—	—
Additional Features	—	—	—	Solar Flare Effects

responsible for the decay of the ionosphere. For his simple model, Nisbet neglected the transport processes such as neutral winds, electric fields, and magnetic fields. The procedure is further complicated because coupled simultaneous equations must be solved for neutral winds, mass transport, and energy transport. For determining electron density in the ionosphere, the gas consisting of both ions and electrons, is considered electrically neutral. Thus, in every elementary volume, the number of electrons is equal to the number of ions. All the models reproduce many of the observed features such as the diurnal variation, seasonal variation, and solar cycle dependence of the midlatitude ionosphere under quiet conditions. The accuracy of the theoretical models depends upon the understanding of the physical processes considered in the models. For accurate predictions from the theoretical models, precise information on the large number of variables used in the models is necessary. Also, the models use several observed average boundary conditions which could have a large variability dependent on other geophysical parameters such as solar activity and magnetic activity. The results from the models are adequate for long term planning of science and engineering applications. Though these models reproduce main observed average features of the ionosphere, they are unable to specify the ionosphere within an accuracy of $\pm 20\%$ needed by the systems in operation. At present, the main input information of solar EUV radiation needed for the theoretical models is not routinely available for predicting the ionosphere.

10.3.3 Comparison of the Phenomenological Models, Their Limitations and Ability

In comparing the models one must note that IONCAP is the modified and more flexible version which replaces the ITS-78 model. As the ITS-78/IONCAP and the Bent models serve entirely different purposes, it is essential to understand the difference in their approaches and final output parameters computed by the models. These are summarized in Table 10-7. The left-hand column in Table 10-7 lists the parameter under consideration. The next four columns summarize the features in each of the models, ITS-78, IONCAP, the Bent, and the IRI 79 models, respectively. From the table it is seen that the selection of a model will depend more upon the information sought under the parameter headings, than on accuracy. The IONCAP model is basically useful for wave propagation using operating frequencies which would be reflected by the E, Es, F1, and F2 layers. On the other hand, the Bent model relies on the transmission, refraction, and absorption characteristics of the ionosphere, with the operating frequency much larger than the foF2 frequency. The IRI 79 model basically provides a distribution of electron density with altitude. All the models predict quiet ionospheric conditions only. The models do not hold for disturbed ionospheric conditions.

The additional limitations of these models are

1. All the models are poor in predicting the high latitude ionosphere.
2. None of the models take into account the effects of particle precipitation in the auroral region which enhance the E(Es) and F layers.
3. The mid-latitude trough which exhibits large horizontal gradients in electron density is not incorporated in these models.
4. These models are good for latitudes $\pm 20°$ to $\pm 60°$, and are poor predictors for the equatorial region and the high latitude region.

Nonetheless these models serve two useful functions: (1) to predict ionospheric parameters, and (2) to determine physical phenomena and/or to modify existing coefficients for explaining the deviations between the experimentally observed value and the predictions from these models.

The computer programs for the ITS-78 [Barghausen et al., 1969] and the IONCAP program [Lloyd et al., 1978] are available from the Institute for Telecommunications Sciences, Boulder, Colorado 80303. The computer programs for the Bent model [Llewellyn and Bent, 1973] are available from the Atlantic Science Corporation, P.O. Box 3201, Indialantic, Florida 32903. The computer programs for IRI 79 are available from World Data Center A, Boulder, Colorado 80303.

10.4 HIGH FREQUENCY RADIO PROPAGATION

The high frequency (HF) band of the electromagnetic spectrum extends from 3 to 30 MHz, corresponding to a wavelength range of 100 to 10 m. Many services have frequencies allocated in this band—Local/International Broadcast, Amateurs, Standard Frequencies, Maritime and Land Mobile, Point-to-point Communications, Industrial, Scientific, Medical Diathermy, Aero Fixed, Citizens' Band, and so on. The band is also used for ionospheric sounding and over-the-horizon surveillance. Its use in most applications depends on the fact that HF waves are reflected by the ionosphere.

HF is used for broadcasting because of its greater area coverage relative to the bands on either side, which are restricted to either ground wave or line-of-sight propagation. Its use for communications stems mainly from the fact that it is often the only means of communication. It is also very often the simplest and least expensive form of communication.

With the advent of satellite communications, which use signals of such high frequency that the normal ionosphere has little effect on them, and improvements to submarine cables, the proportion of traffic that goes by HF is significantly smaller than it used to be. However, the *total* use of HF radio circuits is actually greater now than ever before, and a substantial research effort is still being devoted to

Table 10-7. Intercomparison of the empirical-computer based ionospheric models.

Parameter	Ionospheric Models			
	ITS-78	IONCAP	Bent	IRI 79
D Region	Non-deviative and deviative absorptions only	Same as ITS-78 + E Layer exponential extension down to 70 km	Not modeled	Modeled
E Region				
foE	Modeled by Leftin et al. [1968]	Same as ITS-78 + exponential down to 70 km Leftin [1976] coefficients	Not modeled	Modeled Kouris and Muggleton [1973a,b]
h_mE	110 km fixed			
Y_mE	20 km parabolic shape			
F1 Region	Not modeled		Not modeled	Modeled
foF1		Rosich & Jones coefficients [1973]		Eyfrig [1955] Ducharme et al. [1973]
h_mF1				
Y_mF1		$h_mF1/_{ym}F1 = 4$ (fixed)		
F2 Region	Haydon-Lucas coefficients [1968]	Same as ITS-78	Bi-parabolic	Modeled Rawer [1981]
Bottomside	Shimazaki eq [1955] + E layer retardation Kelso [1964]			
foF2			Bent coefficients	
h_mF2				
Y_mF2				
F2 Region Topside	Not modeled	Not modeled	Up to 1000 km	Modeled Rawer [1981]
E-F Transition Region	Not modeled	Modeled	Not modeled	Maeda [1971]
Electron-Density Profile	Not computed	Available up to h_mF2	Available up to 3500 km	Available up to 1000 km
Electron, Ion Temperatures	Not modeled	Not modeled	Not modeled	Modeled
Ion Composition	Not modeled	Not modeled	Not modeled	Modeled
Total Electron Content (TEC)	Not computed	Not computed	Computed	Not computed
MUF	For short path only	Also for long path (>10,000 km)	Not modeled	Not modeled
Short-Term Prediction of MUF	Function of Kp	Not modeled	Not modeled	Not modeled
Input Parameters required	Sunspot number	Sunspot number	Sunspot number and 10.7 cm solar flux	Sunspot number

Table 10-7. (*Continued*)

Parameter	Ionospheric Models			
Noise Parameters	Galactic Atmospheric Manmade	Same Modified Same	Not modeled	Not modeled
MUF 50% FOT 90% HPF 10%	Modeled	Modeled	Not modeled	Not modeled
System Performance	Modeled for short path (<3000 km)	Also has a long path option ≥ 10,000 km	Not modeled	Not modeled
Antenna Patterns	Uses ITSA-1 Package with 10 antenna options	Modified ITS-78 package with 17 antenna options	Not modeled	Not modeled
Sporadic E	Modeled in terms of occurrence frequency	Same as ITS-78	Not modeled	Not modeled
Circuit Reliability Service Probability Multipath Evaluation	Modeled	Modeled as ITS-78	Not modeled	Not modeled

improving our knowledge of the ionosphere and HF propagation.

Some of the difficulties associated with using HF for communications, broadcasting, or surveillance stem from the ionosphere itself and success in any of these fields demands a good knowledge of the ionosphere and its vagaries. Irreducible difficulties associated with HF propagation can usually be traced to characteristics of the ionosphere or of radio waves propagating through any lightly ionized medium. Thus it is essential for the professional user of HF to have a good knowledge of both the ionosphere and radio wave propagation.

Much has already been written about the ionosphere and radio wave propagation and the reader should look elsewhere for details. See, for example, the books by Davies [1966, 1969], David and Voge [1969], Rishbeth and Garriot [1969], Hargreaves [1979], Ratcliffe [1970], Lied [1967], and Picquenard [1974]. The four volume report "Solar Terrestrial Predictions Proceedings" [Donnelly, 1979, 1980] is an excellent supplement to these books, providing more recent reports on the general problem of forecasting the solar-terrestrial environment. The reports of Study Group 6 of the Consultative Committee for International Radio (CCIR) are also a very useful source of information, and are particularly valuable because they are regularly updated. AGARD Lecture Series No. 127 on "Modern HF Communications" is also a valuable source of information [AGARD, 1983].

It is the intention of this section to provide a broad overview of HF propagation, its relationship to the ionosphere, its problems, and to indicate those areas of current interest to users of HF. A basic knowledge of the ionosphere itself is assumed (see Chapter 9 of the present volume). Emphasis will be placed on the use of HF for communications. The same concepts and problems also apply to the use of HF for broadcasting and surveillance (over-the-horizon radar). Section 10.6 covers the effects of over-the-horizon radars in the HF band.

10.4.1 Morphology of the Ionosphere

An understanding of the morphology of the ionosphere is an essential prerequisite for its successful use as a communications medium. The basic theory of the ionosphere and its variations has been outlined in Chapter 9 of this volume-here we are concerned mainly with how the ionosphere varies, rather than why it does.

There are five main variations of the electron density of the ionosphere that must be taken into account:

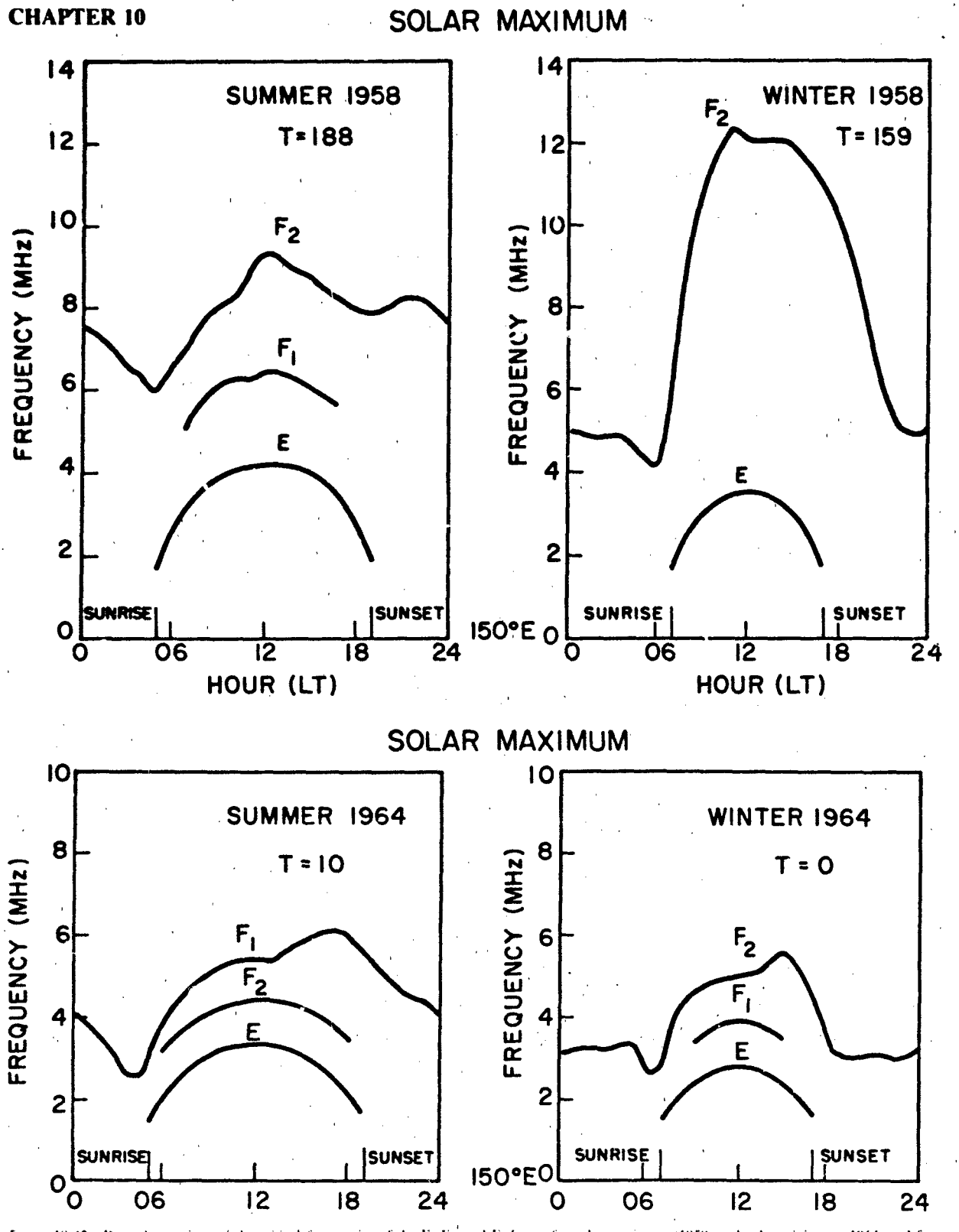


Figure 10-42 Diurnal variations of the critical frequencies of the E, F_1, and F_2 layers for solar maximum (1958) and solar minimum (1964) and for summer (January) and winter (June), at a typical midlatitude station (Canberra). The parameter T is an ionospheric index related to the level of solar activity.

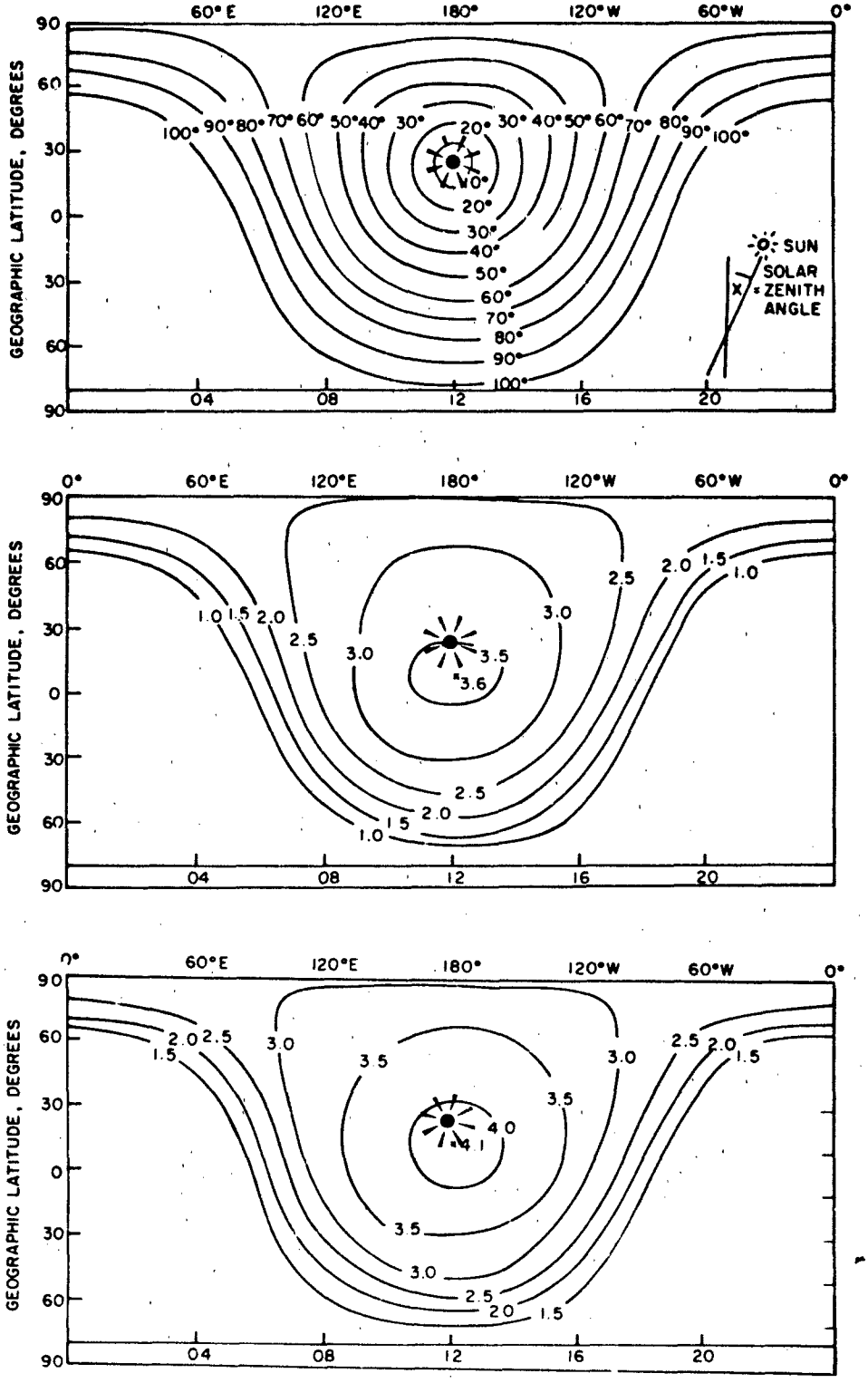


Figure 10-43 The geographical variation of the critical frequency of the E layer as a function of local time for June at solar minimum (center panel) and solar maximum (bottom panel). The top panel shows the variation of the solar zenith angle for the same month.

1. DIURNAL-variation throughout the day, which is largely due to the variation of the solar zenith angle.
2. SEASONAL-throughout the year.
3. LOCATION-both geographic and geomagnetic.
4. SOLAR ACTIVITY-both long term and disturbances
5. HEIGHT-the different layers.

These variations have all been deduced experimentally, by world-wide observations of the ionosphere over the past few decades. The reader may refer to Davies [1966, Chapter 3] and Hargreaves [1979, Chapter 5] for details. The diurnal, seasonal, solar cycle and height variations of the ionosphere may all be deduced by routine monitoring of the ionosphere at one location. Figure 10-42 shows these four variations for a typical midlatitude station. 1958 was a period of high solar activity, as indicated by the high values of the ionospheric index, T (see Section 10.4.4). The figure also illustrates the mid-latitude seasonal anomaly, the name given to the initially unexpected fact that foF2 is higher in the winter than in the summer, in spite of the larger solar zenith angle. This anomaly and others are described by Hargreaves [1979, Chapter 5].

Once the diurnal, seasonal, solar cycle and height variations of the ionosphere at a given location have been deduced, the next step is to measure and understand the variation with location. This has been achieved through an international effort of observations and data exchange, and we now have reliable maps of the world-wide distribution of the important ionospheric parameters. The accuracy of these maps over the ocean areas, where no observations are available, still remains somewhat limited [Rush et al., 1983].

The easiest part of the ionosphere to model on a world-wide basis is the E layer. Figure 10-43 shows the variation of foE, the critical frequency for the E layer, for June at solar minimum and solar maximum. The figure also shows the variation of the solar zenith angle for the month, and it can be seen that the variation of foE follows closely that of the solar zenith angle. In fact, the variations are so close, indicating that foE is very largely solar controlled, that it is possible to use a simple empirical representation to deduce foE for a given zenith angle, χ [for example, Hargreaves 1979; Muggleton, 1971]:

$$foE = 3.3\,[(1 + 0.008R)\cos\chi]^{1/4}\ \text{MHz}\ .$$

Note that foE also varies linearly with sunspot number, R, increasing by about 20% over a typical slow cycle for a zenith angle of zero.

Sporadic E (Es) layers also occur in the height range of the normal E region. These layers are patchy and only a few kilometers thick at mid latitudes. They tend to appear and disappear almost at random (hence the name), but have well-defined gross seasonal and latitudinal variations. See, for example, Hargreaves [1979, p.90].

The F1 layer is similar to the E layer (see Figure 10-44), except that the F1 layer tends to disappear in winter. Hargreaves [1979] gives the following formula for foF1:

$$foF1 = 4.25\,[(1 + 0.0015R)\cos\chi]^{1/4}\ \text{MHz}\ .$$

The variations of foE and foF1 with $(\cos\chi)^{1/4}$ identify both the E and F layers as well behaved Chapman layers. Other, more accurate, world-wide representations of foF1 have been given by Rosich and Jones [1973] and Ducharme et al. [1971].

Moving on to the F2 layer, due to its large height and electron density the most important layer as far as HF propagation is concerned, we find that the simple situation that holds for the E and F1 layers does not hold very well for

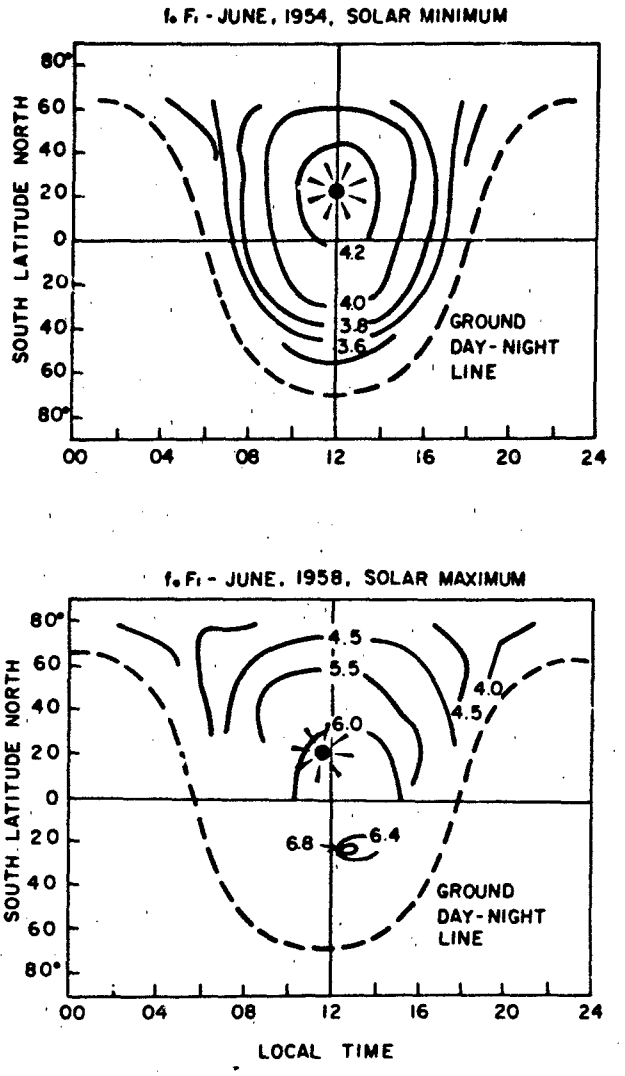


Figure 10-44. The geographical variation of the critical frequency of the F1 layer as a function of local time for June at solar minimum and maximum.

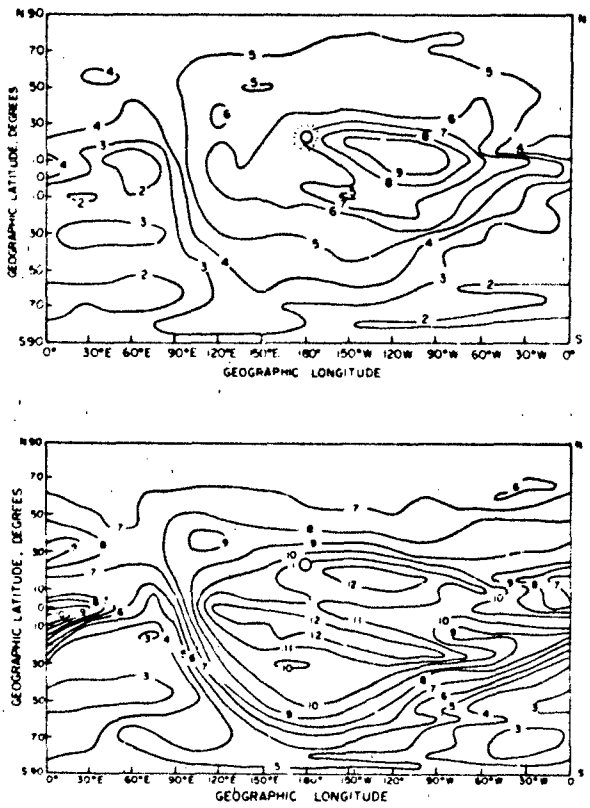


Figure 10-45. The geographical variation of the critical frequency of the F2 layer for June at solar minimum and maximum, for 00 UT. World maps such as these are made for each hour of each month (576 maps).

the F2 layer. Figure 10-45 shows, for example, how the F2 layer critical frequency, foF2, varies over the earth at 00 UT in June, for low and high solar activities. It can be seen that the simple structure obtained for the E and F1 layers, with the contours of foE and foF1 closely following the contours of the solar zenith angle, no longer applies although a clear zenith angle dependence can be seen around sunrise (~90 to ~120°E). In fact, the departures of foF2 from a simple $R_z \cos^{1/4}\chi$ dependence are so great that it is necessary to make world-wide observations to determine the actual variations of foF2. Detailed studies of foF2 have shown that as well as depending on R and cos χ, foF2 also depends on other factors such as electric fields, and neutral winds to name a few, and its large scale morphology is controlled by the geomagnetic field.

foF2 is also found to have variations with latitude which are not seen in foE and foF1. For example, Figure 10-45 shows that foF2 exhibits two afternoon peaks (~12 MHz in the solar maximum portion of the figure) situated on either side of the equator. This feature is known as the equatorial anomaly and is due to electrodynamic lifting of the layer at the equator under the combined influence of horizontal electric and magnetic fields. Theoretical modeling studies of the equatorial ionosphere have been performed by Anderson [1981], among others. Empirical maps of foF2 and other ionospheric parameters have been published [illegible] [1966].

The morphology of the high latitude ionosphere is even more complicated than that [illegible] and much remains unknown about [illegible]. An important feature of the high latitude [illegible] is the mid-latitude ionosphere trough, which lies equatorward of the auroral oval. The trough is a narrow feature that moves in step with geomagnetic activity and thus fails to appear in monthly median maps of foF2. However, it can have very serious effects on HF communications at high latitudes because of the strong horizontal gradients associated with it. The morphology of the high latitude ionosphere has been reviewed, for example, by Hunsucker [1979], Hunsucker et al. [1979] and CCIR [1981a].

Lastly, we must consider the D region. This region is of no direct concern for HF radio propagation since the electron densities are always too low to reflect HF waves. However, the D region is very important from the point of view of absorption of the energy of an HF wave, especially at the lower end of the HF band. A review of the D region and the prediction of its effects on radio propagation has been presented by Thrane [1979]. Synoptic models of the D region electron density are unreliable because of the complexities of the D region and the difficulties encountered in measurement of the electron density profiles.

Absorption of HF waves occurs mainly in the D and lower E regions of the ionosphere. The free electrons absorb energy from the incident wave and reradiate it in a continuous process. However, if an energetic electron collides with a neutral particle before it can reradiate its energy, this energy will be taken up by the neutral particle as kinetic energy and will be lost to the HF wave, that is, energy will be absorbed by the medium.

This type of absorption is known as non-deviative absorption and is roughly proportional to $1/f^2$ where f is the wave frequency. Extra absorption, known as deviative absorption, also occurs near the reflection level. The non-deviative term usually dominates for oblique propagation. If the operational frequency, f, becomes too low, the absorption will increase to the point where the signal disappears below the level of the noise at the receiver site and becomes unusable. This frequency is known as the lowest usable frequency, or LUF. A good treatment of absorption is given by Davies [1969, Chapter 6] while the variations of absorption are discussed by Davies [1966, Chapter 3].

10.4.2 Simple Ray Propagation

Many of the operational aspects of HF propagation may be studied using simple ray concepts. Figure 10-46 illustrates the basic geometry of a one-hop HF circuit. Note that

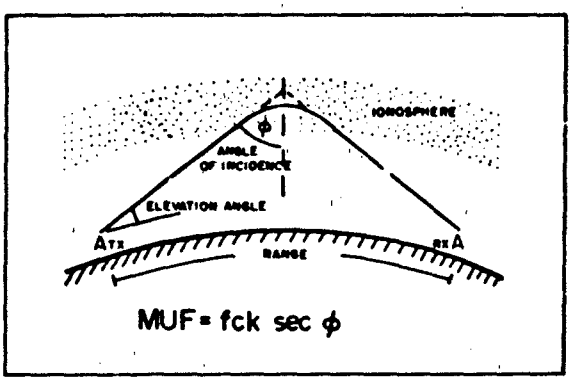


Figure 10-46. Simple geometry of an HF oblique circuit, illustrating the essential elements of the circuit. The formula for the maximum usable frequency (MUF) is all that is required in many calculations for HF propagation.

the diagram ignores the ground wave which is usable for ranges up to about 50–100 km, depending on the frequency, antenna, ground conductivity, etc. The reflected ray is continuously retracted as it passes through the ionosphere and if sufficient refraction occurs the ray will be bent down sufficiently to reach the receiver. Figure 10-46 also shows one of the most basic formulas of ray propagation:

$$MUF = f_c \times \text{obliquity factor} = f_c \times k \times \sec \phi,$$

that is, the maximum usable frequency (MUF) is equal to the product of the critical frequency, f_c, of the reflecting layer and an obliquity factor related to the geometry of the circuit. For a flat earth/ionosphere approximation, this factor is sec ϕ, where ϕ is the angle of incidence. For a curved earth and ionosphere, the factor k is introduced to allow for the different geometry. This factor is typically of the order 1.1.

In practice, the obliquity factor for a given circuit relying on reflection from the F2 layer for example, is obtained from

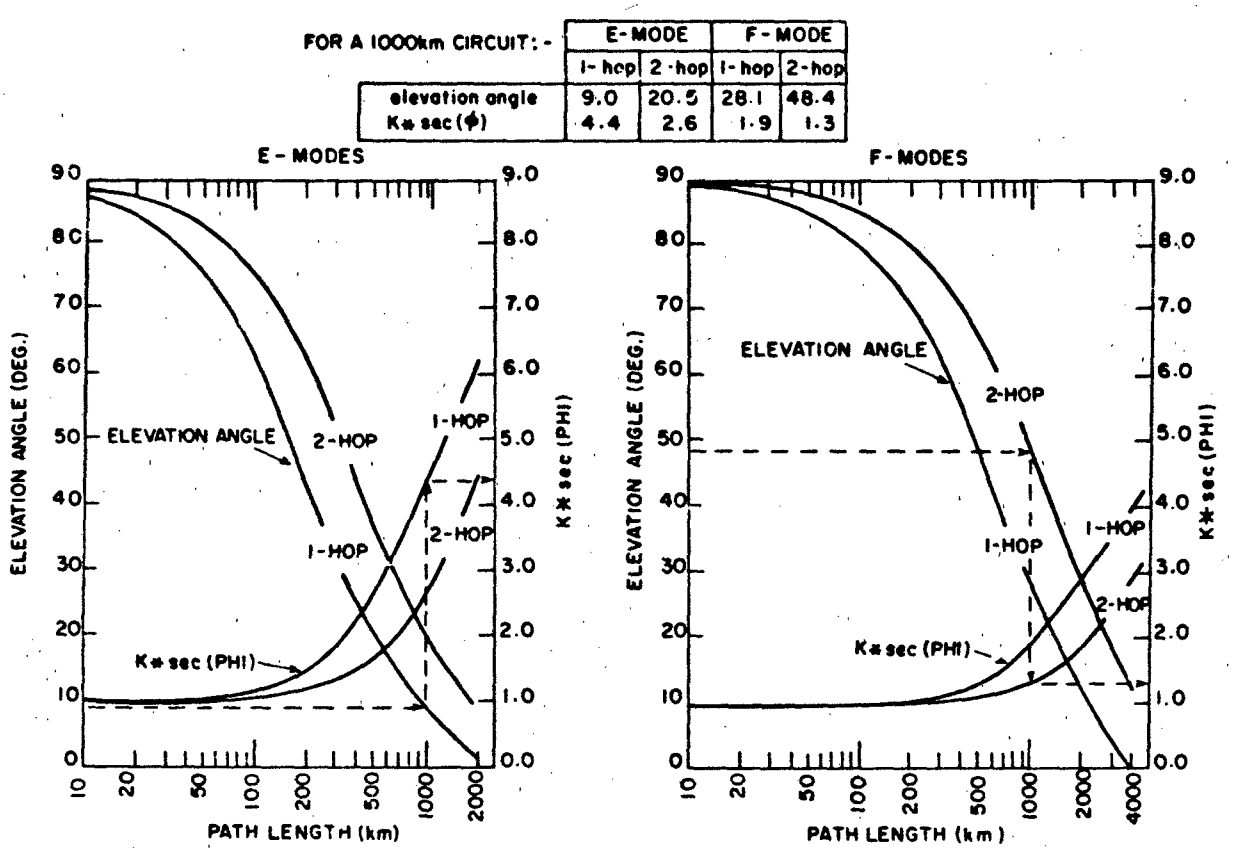


Figure 10-47. Plots showing the elevation angle and obliquity factor (k sec ϕ) for propagation via the E and F layers. For a given circuit length and number of hops, these plots show the required elevation angle (necessary for selecting an appropriate antenna) and obliquity factor for the two layers (useful, among other things, for consideration of possible E layer screening).

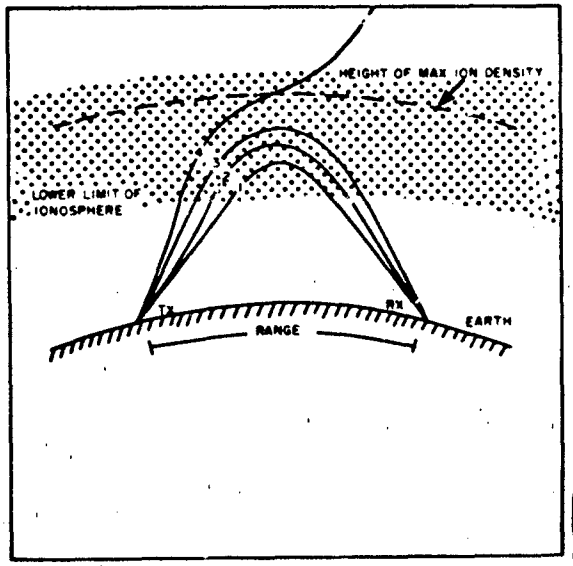


Figure 10-48. Sample ray paths for a fixed distance but different frequencies. As the frequency increases, the ray must penetrate further into the ionosphere. If the frequency is too high, that is, above the MUF for the circuit, the ray penetrates.

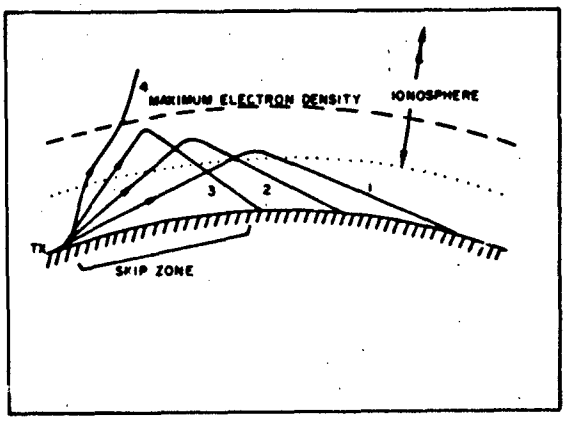


Figure 10-49. Sample ray path for a fixed frequency but varying elevation angle. If the operating frequency is above the local critical frequency, high elevation rays will penetrate and there will be a "skip zone" around the transmitter which cannot be reached by an ionospherically propagated ray.

values of the obliquity factor for a 3000 km hop, M(3000)F2, which is scaled routinely from vertical incidence ionograms [Piggott and Rawer, 1972]. The obliquity factor for a distance D is related in an empirical fashion to M(3000)F2. Figure 10-47 shows how the value of k sec ϕ varies with circuit length for one and two hop E and F layer modes. Typical values for the E and F one-hop modes are ~5 and ~3. Figure 10-47 also shows the elevation angles corresponding to the different propagation modes.

Figure 10-48 illustrates ray propagation for different frequencies on a fixed circuit. As the frequency increases, the ray must penetrate further into the ionosphere before it is refracted to the horizontal and thence back to the ground. The highest frequency that can be reflected back to the ground is the MUF for that circuit. Note that the ray corresponding to the MUF does not reach the altitude of the peak density of the layer, h_{max}. It is only for vertical incidence that the ray actually reaches h_{max}.

Figure 10-49 illustrates the concept of the skip zone, which is a zone into which an ionospherically reflected signal cannot propagate. The figure illustrates the effect of different elevation angles for a fixed frequency. As the elevation angle increases, corresponding to a shorter circuit length, the ray must penetrate deeper into the ionosphere in order to be reflected. However, as the elevation angle increases the obliquity factor k sec ϕ decreases. While the product foF2 · k sec ϕ remains greater than the operating frequency f, the signal will be reflected. When f exceeds foF2 · k sec ϕ, the signal will penetrate the ionosphere.

The area surrounding the transmitter, which is defined by k sec $\phi \leq$ f/foF2, is known as the skip zone for that frequency (and location and time). Signals at the frequency f cannot penetrate into the skip zone, although the ground wave would propagate out to about 50–80 km. This phenomenon can also be used to advantage by ensuring that an unwanted receiver lies in the skip zone of the transmitter. When the value of foF2 above the transmitter exceeds the operating frequency there is, of course, no skip zone.

HF-communication via ground wave is important in many areas, particularly over sea and flat land with high conductivities, where reliable circuits may be established up to distances of several hundred kilometers. The conductivity of the surface is strongly frequency dependent with rapid attenuation at the higher frequencies. In the past CCIR has published a set of curves of ground wave field strength versus distance. CCIR is in the course of implementing a computer program to estimate ground wave field strengths. Ground wave propagation may be quite complex, particularly over rough terrain and over mixed land-sea paths. There is a need for better charts of ground conductivity, and in some cases terrain modeling may be useful and important. Large topographical features such as mountain ranges and glaciers may cause reflections and strong attenuation, and vegetation, soil humidity and snow cover also influence the propagation characteristics.

We have seen that the MUF for a given circuit is set by the density of the ionosphere at the reflection point and the geometry of the circuit. For a multi-hop circuit, the MUF is set by the lowest of the MUFs for the individual hops. The lowest usable frequency (LUF), on the other hand, is set by absorption of the signal by the ionosphere and by the generally poor performance (low gain) of most HF antennas at low frequencies.

10.4.3 Requirements for Successful Communications

Under normal operating conditions, there are three factors that must be considered to achieve successful communications. These are

1. Choosing a suitable operating frequency.
2. Choosing a suitable antenna system.
3. Ensuring that the wanted signal is at a level above that of the local radio noise at the operating frequency.

The choice of a suitable operating frequency is the subject of main concern in the present context since this is where a knowledge of ionospheric physics and radio wave propagation is required. This choice will be discussed in Section 10.4.4. The choice of a suitable antenna system will require matching the antenna pattern to the propagation angles of the HF signals—these can be deduced using the ionospheric models developed for choosing a suitable operating frequency. However in many cases practical considerations will intervene and a far from ideal antenna will be used. The ideal antennas tend to be very expensive and mobile operators especially will often be forced to use a random wire hung over a tree. As a general rule of thumb, horizontal antennas are required for short circuits and vertical antennas are required for long circuits. The ubiquitous whip antenna is absolutely the worst (but probably the most common!) choice for short sky-wave circuits since its antenna pattern has a null in the vertical direction. Antenna patterns for the common antennas can be obtained from Lloyd et al. [1978]. Table 10-8 gives a brief summary of

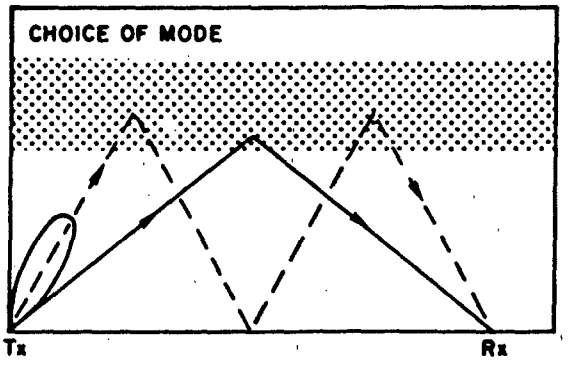


Figure 10-50. Sketch illustrating the fact that the antenna pattern should match the required propagation mode for a given circuit.

the situation. Note that the radiation pattern of an antenna is a function of frequency so that an antenna appropriate for a low HF frequency may have a very poor performance at a higher frequency. Selection of the correct antenna will not only ensure that the bulk of the transmitted power goes at the required elevation angle, but can also be used to select a particular propagation mode and thus avoid multipath interference. Multipath interference arises when the transmitted signal arrives at the receiver over two or more separate propagation paths with different time delays.

Figure 10-50 illustrates propagation via 1F and 2F propagation modes, with an antenna pattern whose lobe favors the higher angle 2F mode and almost completely prohibits

Table 10-8. Suitable simple antennas for use on paths from 0 to more than 3000 km.

Path Length (km)	Required Radiation (Elevation) Angles	Suitable Simple Antenna
0–200	60°–90°	Horizontal dipole: broadside is required azimuth. 0.25 wavelength (λ) above ground.
200–500	40°–70°	Horizontal dipole: broadside to required azimuth. 0.3 λ above ground.
500–1000	25°–50° and 10°–20°	0.25 λ vertical monopole, or horizontal dipole, broadside to required azimuth 0.5 λ above ground.
1000–2000	10°–30° and low angles	Vertical monopole up to 0.3 λ long with ground screen.
2000–3000	5°–15° and 20°–30°	Vertical monopole up to 0.3 λ long with ground screen.
> 3000	low angles	Vertical monopole up to 0.6 λ long with ground screen.

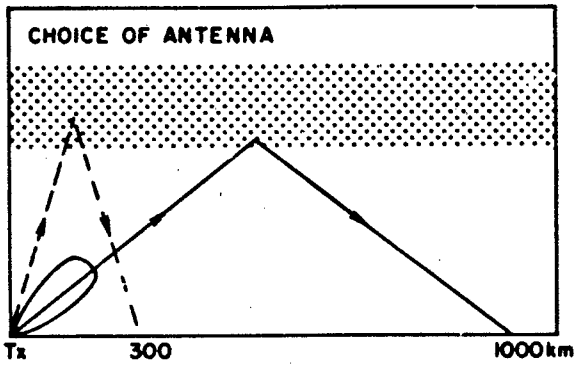


Figure 10-51. Sketch illustrating the fact that the antenna should be chosen to match the circuit length. The antenna pattern illustrated is not appropriate for short circuits (R ~ 300 km).

propagation via the 1F mode. The impact of the choice of an antenna for a given circuit is illustrated in Figure 10-51 for hop lengths of 300 and 1000 km. In this case, the antenna heavily favors the longer circuit and is quite inadequate for short-haul circuits.

Given a frequency that the ionosphere will support and an antenna which emits sufficient power in the direction taken by the signal that arrives at the receiving site, the third thing to ensure is that the signal strength is above the strength of the local radio noise. This noise can be natural or manmade. Below about 20 MHz, natural noise is caused by either distant thunderstorms, which cause a general increase in background noise level, or local thunderstorms which are usually much more obvious causes of poor signal-to-noise ratio. Galactic cosmic noise becomes the dominating natural noise above ~20 MHz when it penetrates the ionosphere from above at frequencies above foF2. The worldwide distribution and characteristics of atmospheric radio noise can be obtained from the CCIR Report 332 [CCIR, 1963]. Manmade noise includes such things as industrial noise due to welders, diathermy machines, car ignition, power lines and so on. Interference from other communicators using the same operating frequency can also be regarded as noise.

There are many techniques available to ensure an adequate signal-to-noise ratio. The more environmentally acceptable method is to aim at lowering the noise level. This can be achieved by choosing as a receiver site some location that is remote from the major sources of manmade noise, which usually entails being away from major cities. Some noise rejection is also possible using horizontally polarized antennas which de-select the local noise that tends to be vertically polarized. Careful attention to the azimuth of the main lobe of the receiving antenna can also result in the beam not being aimed at a nearby source of noise. In heavily urbanized areas (Europe, East and West Coast of the United States), the noise level is often set by the large number of other communicators.

If an adequate signal-to-noise ratio cannot be obtained by lowering the noise level, it is necessary to increase the signal level. Some increase in signal level can be achieved by choosing an antenna that has more gain at the given operating frequency and elevation angle. This is one approved solution. An alternate solution is to increase the transmitter power.

10.4.4 Predictions for HF Communications

The first step in predicting HF communication conditions is to set up an appropriate model of the ionosphere. To have practical application, this model must include all five main variations of the electron density distribution of the ionosphere (altitude, location, diurnal, seasonal, solar cycle), must include some measure of the remaining variability of the ionosphere after these main variations have been accounted for, must exist as a reasonably efficient and fast computer code, and must possess some method for projection forward in time. Ideally, it should also be capable of a modelled response to short time scale events such as shortwave fadeouts and ionospheric storms. (These are discussed in Section 10.4.6.)

Most of the ionospheric models used for communication predictions are empirical models based on world-wide observations of the ionosphere over the past four decades. Observations of the main parameters of the ionosphere, foE, foF1, foF2, M(3000)F2, and Es have been used to produce worldwide contour maps of monthly median values of these parameters for each hour of each month and for low and high levels of solar activity. An example of such maps is given in Figure 10-45. To calculate the maximum usable frequency (MUF) for a given circuit, hour and month, the values of foF2 and M(3000)F2, for example, are determined by interpolation in the appropriate world map for the expected point(s) of reflection of the signal. For a level of solar activity other than the low or high levels for which the maps are drawn, values of the ionospheric parameters are obtained by interpolating in each of the low and high activity level maps and then interpolating again in these results to find the values appropriate to the given level of activity.

In practice, the only prediction actually performed in making predictions of suitable HF operating frequencies is that of the general level of solar activity that can be expected to pertain at the epoch for which the predictions have been made. This prediction is usually made in the form of some ionospheric or solar index, which is a single parameter describing the gross behavior of the ionosphere, and which is used to drive the computer program. The most common index used for prediction programs such as the U.S. programs IONCAP and ITS-78, and by the CCIR, is the average sunspot number, R_{12}. However, Wilkinson [1982] has shown

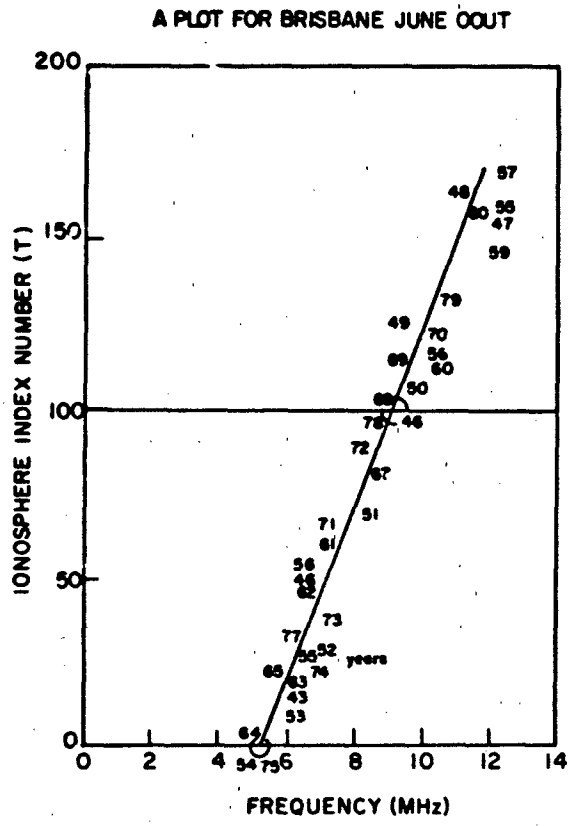


Figure 10-52 The variation of the monthly median value of foF2 for June, 00 UT, at Brisbane, Australia, as a function of the ionospheric index, T. Each data point represents the value for June for years 1943–1980.

that any index based on the ionosphere itself, such as IF2 [Minnis and Bazzard, 1960] is usually preferable to a purely solar index such as R_{12}. Figure 10-52 illustrates the relationship between the ionospheric index, T [Turner, 1968; Turner and Wilkinson, 1979] and the monthly median values of foF2 at Brisbane, Australia for 00 UT in June (10 LT). Each data point represents data for June in the years 1943 to 1980. The low dispersion of the data points about the regression line indicates the usefulness of the T index for describing ionospheric conditions. The selection of the correct index for some future epoch will become more uncertain as the lead time increases because of our general inability to predict the detailed behavior of the sun.

Predictions of the main critical frequencies and heights of the ionosphere, together with some simple geometry (Section 10.4.2), are adequate for the calculation of the appropriate frequencies to use for communication on a given circuit. It is also possible to determine elevation angles of the possible propagation modes (and thence choose a good antenna) and to determine the conditions under which some adverse propagation conditions will exist.

The calculation of the transmitter power required to yield an appropriate signal-to-noise level at a receiver site is a much more complicated procedure than the calculation of the MUF. The required transmitter power is of no small matter since the cost of the transmitter increases dramatically with the power. The noise level may be either measured at the site or estimated from empirical data bases such as provided by CCIR. See for example Davies [1966, Chapter 7] and CCIR [1963].

There are two other main concerns here, the antenna and the transmitter, but let us assume that an appropriate antenna with a known gain has already been selected. To calculate the required transmitter power, we need to calculate the losses. These include deviative and non-deviative absorption losses, basic free-space transmission loss, multihop ground reflection loss, polarization coupling loss, sporadic E obscuration loss and horizon focus gain (a negative loss). The largest of them is the free-space loss, and this increases with the length of the propagation path.

Because the circuit losses depend on the propagation path, all possible paths must be considered and compared. Reliable determination of the actual propagation paths requires a reliable model of how the electron density varies with altitude up to the peak of the F2 layer. Suitable models of this N(h) variation have been developed by Bradley and Dudeney [1973], among others. As well as being an accurate representation of the real N(h) distribution, such a model must also be computationally simple because it is used many times in the determination of the virtual reflection heights, h', and thence the possible propagation modes. The International Reference Ionosphere [Lincoln and Conkright, 1981] may also be used to deduce an N(h) profile, but this model has not been designed for the speed that is required in routine field strength calculations.

The ionospheric models currently available for use in studies of HF propagation have been reviewed by Dandekar [1982] and Goodman [1982] (see also Section 10.3). Full descriptions of the techniques used to calculate field strengths at a receiving site are given by Davies [1966, Chapter 7], Lloyd et al. [1978] and CCIR [1970, 1978]. A straight forward approach to obtaining MUF's for various modes and their respective path losses which also permits the inclusion of realistic antenna patterns for transmit and receive sites is available in the IONCAP program [Lloyd et al., 1978]. The model computations are based on a three layer (F2, F1, E) representation of the electron density profile with ITS78 coefficients being used to determine the (R_z-dependent) foF2 value. The ionospheric parameters are evaluated at the reflection points and two dimensional ray tracing is applied, assuming local horizontal stratification. For the most reliable results, simplistic approaches to propagation mode determination must be replaced by ray tracing techniques such as described by Davies [1969, Chapter 7]. These techniques are essential when large horizontal gradients exist in the ionosphere, but are rarely used for routine calculations because they are very time consuming.

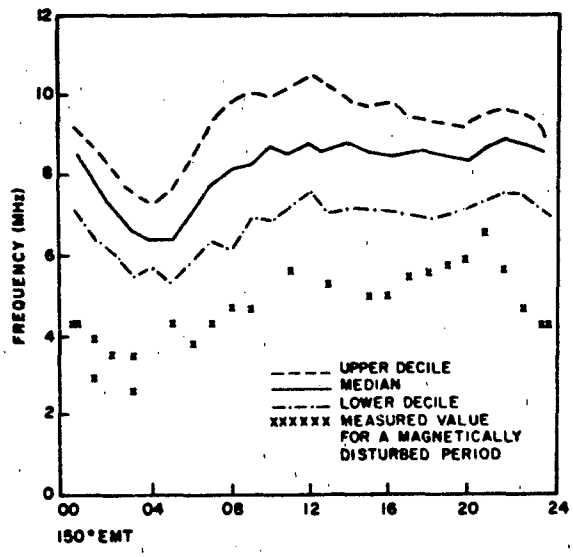


Figure 10-53. Plot of the observed values of foF2 a midlatitude station, December 1980, as a function of local time. At each hour there are 31 data points. 80% of observations lie within the dashed upper and lower decile curves.

10.4.5 Problems with HF Communications

The use of the ionosphere for communication, broadcasting, and surveillance is fraught with difficulties, some of which have already been mentioned. Predictions of the MUF, LUF, and field strengths rely on median models of the ionosphere and can therefore specify only mean propagation conditions. A knowledge of the spread of values about the median is required for successful communication for more than 50% of the time, and such statistics are in fact usually provided by the better prediction programs. For quiet, undisturbed conditions, this spread about the median is typically 15%–20%. See for example Figure 10-53.

Even when conditions are undisturbed and the communicator has chosen an appropriate operating frequency, antenna, and transmitter power, propagation problems are still encountered. Some of these are associated with events on the sun and will be discussed in Section 10.4.6. Here we wish to consider some problems that can occur even when the ionosphere is being cooperative.

10.4.5.1 Multipath Propagation. We have already seen in Section 10.4.3 that propagation will normally occur by several paths, for example, the 1F and 2F modes (see Figure 10-50). The received signal will be the vector sum of all waves arriving at the receiver. If the different signal paths change with time in different ways, deep and rapid fading may occur, sometimes causing the signal level to drop below the local noise level. Different modulation techniques are affected in different ways by this multipath fading.

Multipath propagation may sometimes be eliminated from a circuit by choosing a frequency that is supported by only one propagation mode (for example, use a frequency above the 2F MUF but below the 1F MUF) or by choosing an antenna that heavily favors one propagation mode. Diversity techniques, in which the same information is sent in two or more ways and recombined at the receiver, can also be used to overcome the effects of fading. Example are space (spaced antennas), frequency, angle of arrival, and polarization diversity.

10.4.5.2 Sporadic E. Sporadic E is a two-faced phenomenon having both advantages and disadvantages. There is no doubt that when a strong sporadic E layer is present, it presents an ideal propagation mode suitable even for medium speed digital data transfer. However, a dense Es layer can screen the F layer, preventing signals from ever reaching that layer and causing Es-F combination propagation modes that may have none of the desired characteristics—the signals may miss the target or may arrive at an angle not specified in the original receiving antenna design. Many prediction schemes for HF propagation include statistical occurrence rates for Es propagation modes, but the Es layer remains essentially unpredictable.

10.4.5.3 Problems in High Latitudes. After major solar flares with particular characteristics, high energy protons penetrate into the lower ionosphere at the poles and causes wide-spread and long-lasting disruptions to HF communications. The polar cap absorption event is discussed briefly in Section 10.4.6. The penetration of electrons with energies in excess of 10 keV into the D region leads to increased absorption in the auroral zone. This auroral absorption may have severe consequences for circuits crossing the auroral zone, but strong absorption is usually limited geographically to patches a few hundred kilometers in extent and the duration is typically half an hour to a few hours.

An important feature of the undisturbed high latitude ionosphere is the great variability in space and time. This variability severely limits the usefulness of a median model of the ionosphere. The F region trough, which marks a transition between the midlatitude and high-latitude ionospheres, can have severe and detrimental effects on signals propagating through it. The sharp gradients at the walls can cause reflections and result in off-great-circle propagation.

Propagation in the auroral region may introduce rapid and severe fading of HF signals. Diversity techniques or some sort of real time channel evaluation technique (see Section 10.4.9) therefore become almost mandatory.

The polar cap ionosphere in darkness, of importance to long distance HF communication between higher latitude stations such as in Canada and northern Europe, is essentially unpredictable for HF purposes. Enhanced F ionization regions resulting from particle precipitation are randomly distributed within a low density background ionization. Especially during slightly disturbed conditions polar cap plasma convection moves high density particles anti-sunward at

high velocities (Buchau et al, 1983]. These changes will result in rapid MUF and mode variations and will in general lead to poor channel performance. MUF variations of typically several MHz from hour to hour and day to day were observed on an arctic link [Petrie and Warren, 1968]. In summertime the F1 mode tends to dominate the mid-morning MUF, and during winter/solar minimum conditions Es modes over transpolar circuits (for example, Andoya, Norway to College, Alaska) have a greater than 50% occurrence [Hunsucker and Bates, 1969].

10.4.5.4 Problems in Low Latitudes. HF communication problems at low latitudes due to steep spatial and temporal gradients have been discussed by Lakshmi et al. [1980]. The very steep gradients in foF2 during sunrise hours give rise to several difficulties [Lakshmi et al., 1980]:

1. HF link operators are expected to get their frequencies cleared from the appropriate governmental authority well in advance and it is usual practice to fix one frequency for the daytime and another for the nighttime. The use of the night frequency during sunrise will require much more power than is normally permitted while the frequency allocated for the daytime will be higher than the MUF during the transition period.

2. Point-to-point links normally use inexpensive tuned directional antennas, and frequent change of operational frequency is deleterious from the point of view of antenna efficiency.

3. In the case of long distance circuits in the east-west direction involving multi-hop F region propagation, the problem of the sunrise period will extend to a large number of hours, because the different F region reflection points will fall in the sunrise transition location at different periods.

The steep gradients associated with the equatorial anomaly cause problems with north-south circuits. For example, if we consider the anomaly peak in the northern hemisphere to be at 15°N geomagnetic latitude, if a north-south HF circuit is operating such that the reflection point is on either of the sides of the peak, and if the frequency of the link is very close to the MUF, a peculiar situation arises. If the point of reflection is equatorward of this anomaly peak, the radio waves incident on the ionosphere for the northern circuit will continuously encounter increasing levels of electron density due both to the vertical gradient as the radiowave penetrates higher into the ionosphere and the horizontal gradient as the wave progresses in the direction of increasing electron density. On the other hand for the same link in the return direction, the horizontal gradient is reversed. Thus the real MUF values for the two opposite directions in the same circuit can vary by a large margin depending on the angle of incidence and on the magnitude of the horizontal gradient. In fact, rather frequently, especially when the operating frequency is close to the MUF (calculated ignoring horizontal gradients), only one way communication would be possible. This has been one of the unusual complaints in the Indian Subcontinent [Lakshmi et al., 1980].

10.4.5.5 Spread F and Irregularities. Small scale irregularities of ionization seem to exist at every level of the ionosphere superimposed on the background of ionization discussed in previous sections. They affect the propagation of radio waves and their characteristics may, therefore, be determined by radio techniques (see for example, Rishbeth and Garriott, [1969], Section 6.4). When seen on ionograms, the presence of irregularities is described as "spread F", the spreading being either in the range or frequency domain. The corresponding descriptions are "range spreading" and "frequency spreading", and parameters describing this spreading are scaled routinely from ionograms [Piggott and Rawer, 1972]. Many morphological studies of spread F, showing its diurnal, seasonal, solar cycle and latitudinal variations, have been published (see for example Herman, [1966] or Singleton, [1980], and references cited therein). An earlier review of spread F and some of the effects on radio propagation has been given by Newman [1966]. Most morphological studies of spread F have been concerned with frequency spreading. The characteristics of range spreading are, however, quite different and only a few studies of this type of variation have been made (see for example Cole and McNamara, [1974]). At equatorial regions, range spreading is often associated with the large scale electron density depletions known as bubbles or plumes [Basu et al., 1980c; Tsunoda, 1980; and Tsunoda et al., 1982].

The effects of irregularities on radio propagation are most important on paths that cross the equator when propagation actually relies on the presence of the irregularities [Nielson and Crochet, 1974]. Most recent efforts on describing the effects of irregularities on communications have concentrated on the effects of irregularities on transionospheric propagation, where the problem of scintillation is encountered. This subject is covered in Section 10.7, this Chapter.

10.4.6 Disturbances to Normal Communications

We have previously seen that HF communications are subject to sudden and often large disturbances due to events that originate on the sun. Solar flares can cause immediate and complete absorption of HF radio waves by greatly enhancing the absorption suffered as the waves pass through the D region. These events, called short-wave fadeouts (SWF), are due to increased ultraviolet and x-ray fluxes, and can therefore occur on only the sunlit side of the earth. Major flares can also eject a stream of protons which can penetrate the ionosphere near the magnetic poles and give rise to complete blackout of HF communications in polar regions. Such events are called polar cap absorption events (PCA) and can last for days depending on the size of the flare and

how well it is connected magnetically to the earth. The SWFs, on the other hand, usually last from a few minutes to a few hours at the greatest and are most severe at low latitudes (see Section 10.5).

Many large flares also eject a cloud of plasma, which if geometrical and interplanetary magnetic field conditions are favorable, can intersect the earth's magnetosphere and cause both geomagnetic and ionospheric storms. The geomagnetic storm, in which the earth's magnetic field is usually depressed below its normal quiet day value, is a result of a strong enhancement of the ring current and of no intrinsic interest to the HF communicator, but is a useful and readily available indicator of an accompanying disturbance to the ionosphere and thus to communicators. The ionospheric storms are of concern, however, especially when the effect of the storm is to decrease the MUF on a circuit to well below the predicted levels.

The physics of ionospheric storms is not yet completely understood, but it is well established that both electric fields and thermospheric winds (see Chapter 17) play a role. See for example Hargreaves [1979, Chapter 11].

The effect of a storm at a given location can be either to increase the value of foF2 (a so-called positive phase or enhancement) or to decrease the value of foF2 (a negative phase or depression), or to do both, a long depression following short enhancement. What actually happens depends on such things as local time of onset of the storm, station latitude, and season. During winter, most storm effects tend to be enhancements, whereas in summer and equinox depressions often follow short-lived enhancements. It is the depression that causes the major communications problems—the enhancements often go unnoticed by communicators. Major depressions last typically for a day and can decrease the MUF by a factor of two. Storm effects are much more marked at higher latitudes (see for example Rishbeth and Garriott, [1969], Chapter 8).

The lower regions of the ionosphere are not usually affected during ionospheric storms and the lower frequency limit for HF communication remains unchanged.

Ionospheric storms due to solar flares are a high solar activity phenomenon. However, ionospheric storms also occur away from the peak of the solar cycle. These are attributed to enhanced solar wind streams emanating from magnetically open features in the corona known as coronal holes. As the stream sweeps over the earth, the electric currents flowing in the magnetosphere and ionosphere are modified, yielding both geomagnetic and ionospheric storms. These storms are called recurrent storms because they tend to recur every 27 days (solar rotation period as seen from the earth) as the solar wind stream passes over the earth again. The effects on the ionosphere are usually less marked than those of a flare-induced storm but can last longer (a few days) because of the time taken for the stream to pass over the earth.

The scientific community keeps a 24-h watch on the sun through an international effort coordinated through the IUWDS (International Ursigram and World Days Service) and many national agencies issue forecasts of solar and geophysical conditions. This effort has been described, for example, by several authors in Volume 1 of Donnelly [1979], including Heckman [1979]. CCIR Report 727 [CCIR, 1981c] gives a review of the subject and refers to many of the latest available papers.

The USAF Air Weather Service (AWS), through its operational centers of the Air Force Global Weather Central (AFGWC), provides space environmental support to the entire Department of Defense. The overall driving requirement is to minimize system effects caused by impulsive solar/geophysical activity and ionospheric variations. The techniques for geophysical forecasting used at AFGWC have been described by Thompson and Secan [1979]. A major advance in observational equipment during the last decade was the deployment of the new solar observing network (SOON) and radio solar telescope network (RSTN). With the data from SOON, RSTN and x-ray data from satellites, AFGWC can provide a real-time comprehensive analysis of a flare and its effects on the space environment.

The action taken by an HF user to overcome the effects of these disturbances depends on the nature of the disturbance. During a SWF, a move to higher operating frequencies is appropriate since the absorption decreases with increasing frequency. During the negative phase of an ionospheric storm, lower frequencies must be used. The effect of a PCA, on the other hand, is normally so severe that it becomes necessary to reroute the traffic around the disturbed area. This usually requires avoiding the whole high latitude ionosphere.

10.4.7 Unusual Propagation Modes

The usual monthly-median HF predictions normally assume simple propagation modes such as the 1F and 2F modes and in general these predictions are quite successful. Some prediction systems also include Es and such propagation modes as 1Es and Es-F. However, it is found in practice that other unusual propagation modes can also exist. In general these unusual modes have one feature in common—they rely for their support on some particular feature of the ionosphere that is restricted in latitude. We shall consider here a few examples of such modes.

Possibly the most useful propagation mode not normally predicted is the F^2 super-mode encountered on transequatorial circuits during periods when the equatorial anomaly is well developed, that is, during the afternoon and early evenings, during equinoxes, at high levels of solar activity. This mode involves two F region reflections without an intervening ground reflection and is characterized by high signal strength, low fading rates and an MUF 10-15 MHz higher than for the normal (2F) mode. It is often described as the afternoon-type TEP (transequatorial propagation) mode.

A second mode associated with the equatorial ionosphere, and which can coexist with the afternoon-type TEP mode in the early evening, is called the evening-type TEP mode. It is characterized by strong signals, rapid flutter fading and frequencies well above the normal 2F MUF. A propagation frequency of 102 MHz has been regularly observed on a Japan-Australia circuit. The propagation mode in this case is probably a ducted mode, the signals traveling within the walls of equatorial "bubbles" [Heron and McNamara, 1979 and Winkler, 1981]. This mode is also most likely to occur during the equinoxes at solar maximum. A review of TEP has been published by Nielson and Crochet [1974].

Enhancements of the MUF on transequatorial circuits have also been attributed to combination modes in which one hop is via a reflection from the equatorial Es belt which stretches a few degrees either side of the magnetic equator [McNamara, 1974a]. In an Es-F combination mode, the MUF is enhanced because the F layer hop is longer than the usual hop length on a 2F mode. Similar MUF enhancements on nighttime circuits have been attributed to a combination mode in which one hop relies on scatter from F layer irregularities [McNamara, 1974b]. Propagation above the MUF due to scatter by small scale irregularities can routinely be observed [Rawer, 1975].

The high latitude also produces its share of unusual propagation modes, mostly because of the presence of the midlatitude trough. Strictly speaking, the modes are not different from those expected—they are just heavily affected by the presence of the trough.

Another interesting propagation mode is the round-the-world (RTW) propagation mode. This mode relies on aspects of the vertical distribution of electron density rather than on latitudinally localized features of the ionosphere. Consideration of observed elevation angles, signal losses and time delays have led to the conclusion that RTW signals propagate via a chordal mode of propagation, rather than by a uniform multihop mode.

10.4.8 Short-Term Forecasting of HF Conditions

Long-term predictions of monthly median parameters of the ionosphere such as described in Section 10.4.4 are the traditional approach to frequency management. Frequencies are selected which should ensure at least a 90% success rate for communications at all times of day, during any season, and at all parts of the solar cycle. Engineering decisions such as transmitter power and antenna configuration are also made at this stage. Optimum working frequencies (OWF) are derived for each hour of the day, usually for a month at a time, using empirically derived frequencies and signal-to-noise ratios. Use of predictions in this manner continues to provide acceptable, though not high quality, communications for many purposes such as voice and telegraph conditions. For a survey of the historical development of forecasting methods for HF propagation see Rawer [1975].

The effects on the ionosphere and communications of solar disturbances which cause ionospheric storms have been discussed in Section 10.4.6. However daily values of foF2 and MUF are known to vary by about 15% to 20% from the monthly median values during quiet times as well as during ionospheric/magnetic storms. See, for example, Figure 10-53. These variations may be superimposed on slower upward or downward drifts in values over several days. It is desirable to predict all of these variations for the purpose of efficient radio communications.

There are several possible approaches to the short-term forecasting of ionospheric parameters.

1. Using associated geophysical parameters.
2. Using the ionospheric parameters themselves.
3. Using ionospheric indices.

The day-to-day variability of foE and foF1 is so small that monthly median values of these parameters can be used to represent the daily variation in the E and F1 regions [Rush and Gibbs, 1973] and therefore offer little difficulty.

Short-term forecasts of foF2 have been made by relating changes in foF2 to corresponding changes in selected solar and geophysical variables such as the 10.7 cm solar flux and the geomagnetic index Kp [Bennett and Friedland, 1970; Ichinose et al, 1980]. The disadvantage of this type of approach is that the independent variables upon which changes in foF2 are assumed to depend must themselves first be predicted. Even if this is done successfully, only limited success is possible because the correlations of these parameters with parameters of the ionosphere are not very high.

A more successful approach to the short-term forecasting of foF2, or alternatively of an operational MUF on a given circuit, is a prediction scheme based directly upon immediate past observations of foF2 or an operational MUF. Such prediction schemes for foF2 are described, for example, by Rush and Gibbs [1973], Lyakhova and Kostina [1973], McNamara [1976], and Wilkinson [1979].

Rush and Gibbs [1973] used a five-day weighted mean value of foF2 to predict daily and hourly values of that parameter. The method of Lyakhova and Kostina [1973] is based on the observation that correlation coefficients between the deviations of foF2 from median values remain greater than 0.5 for up to four hours. The high correlation between hour-to-hour variations of foF2 has been discussed by Lyakhova [1960], Radinov [1963], Gautier and Zacharisen [1965] and Rush [1972].

McNamara [1976] made predictions of foF2 at a particular location up to 3 hours ahead by projecting forward the trend in the departures of the last few hours' observations from a 15-day running median. Wilkinson [1979] on the other hand, simply projected forward in time the deviation of an observed foF2 value from the predicted monthly median value of foF2. He found this technique to be effective for lead times of up to about 3 hours.

Similar techniques have been applied to oblique circuits

by Ames and Egan [1967], Ames et al. [1970], Krause et al. [1970, 1973a,b], D'Accardi [1978], and Uffelman and Harnish [1982].

The success of any of these forecasting schemes will depend on the particular circumstances of its intended use, especially as regards the requried accuracy of the forecast and the lead time required. Most schemes are reasonably successful in forecasting an operational MUF that is closer to the actual value than is the predicted monthly median value, but only for lead times of the order of an hour or less.

Forecasts can be made with different lead times, depending on how closely the variations in the ionosphere need to be tracked, and the sampling interval for actual observations must match that lead time. In general, the error in the forecast values will increase with the lead time and if this error becomes too great, the forecasting scheme would offer no advantages over the use of a monthly median value.

To make short-term forecasts for circuits for which no real-time observations exist, the behavior of the ionosphere must be inferred from such data as are available, using these observations to infer the values along the required circuit. Numerous studies have reported correlation coefficients that illustrate the degree to which hourly deviations from median values of ionospheric parameters at two or more locations are related [Gautier and Zacharisen, 1965; Zacharisen, 1965; Zevakina et al., 1967; Rush, 1972]. These correlations may be used to infer values of foF2 at locations remote from the observing site, with an error depending on the separation of the two locations in latitude and longitude, and on the local time and season. Rush [1976] has used the observed correlations to establish the usefulness of, and requirements for, a network of ground-based and satellite-borne ionospheric observations whose measurements are to be used for short-term forecasting of radio propagation conditions. However, it should be noted that the relatively high correlations used by Rush originate from the very disturbed days (that is, ionospheric storm days) and these same high correlations are not always obtained for days when the deviations from the median values are relatively small.

A third approach to short-term forecasting is to use observations at all available ionospheric stations to determine an effective ionospheric index, which can then be used in conjunction with synoptic monthly median maps of foF2 and other parameters (see Section 10.4.4) to predict the values of these parameters at the reflection points of the given circuit. The limitations of this technique are set by the accuracy with which an appropriate index can be determined from a restricted subset of observations and the accuracy with which the ionospheric model, driven by this index, can reproduce the actual ionosphere at the reflection points. An example of the use of ionospheric indices to update models of the ionosphere has been given by McNamara [1979]. Other, more complicated, methods have been described by Thompson and Secan [1979] and Tascione et al. [1979]. Actual forecasts of propagation conditions requires the projection forward in time of the effective ionospheric index. In the limit of zero lead time, the "forecast" in fact becomes a real-time assessment of the ionosphere.

A brief review of short-term forecasting for HF propagation has been given by CCIR [1981b]. The special case of high latitude propagation is considered in CCIR [1983].

10.4.9 Real Time Channel Evaluation

To take full advantage of the HF communications potential of the ionosphere and to overcome its inherent variability, frequency management should be implemented in three stages, namely long-term predictions (Section 10.4.4), short-term forecasting (Section 10.4.8) and real-time channel evaluation (RTCE). It is at the first stage that engineering decisions such as site location, antenna configuration, and transmitter power are made, and suitable operating frequencies applied for from the appropriate regulating body. The second stage, that of short-term forecasting, determines which of the allocated frequencies will actually propagate now or a short time ahead. These two stages are generally all that would be required for voice and low-speed telegraphy circuits.

This two-stage approach has several limitations:

1. The signal-to-noise data are not always reliable and this ratio is not necessarily a useful criterion for choosing frequencies for some forms of communications.

2. Reliable long-term predictions of sporadic E are not available. (Es modes are often the best to use.)

3. No account is taken of interference from other users.

4. The forecast available frequency range will be uncertain to some extent because of inevitable errors in the forecasting models.

5. The approach does not indicate which of the assigned frequencies propagating at a given time would be the best to use for a given form of communication.

RTCE is the third stage of a frequency management system required to maintain reliable high quality HF communications even under the most adverse conditions. It becomes especially important for medium and high speed digital data transfer. With an RTCE system, all channels are sounded in turn to determine which is actually the best to use at a given time for the particular type of communication/modulation system.

Darnell [1978] has given the following definition of RTCE: "Real-time channel evaluation is the term used to describe the processes of measuring appropriate parameters of a set of communication channels in real time and of employing the data thus obtained to describe quantitatively the states of those channels and hence the relative capabilities for passing a given class, or classes, of communication traffic."

A review of RTCE has been prepared by Study Group 6 of CCIR [CCIR, 1981d], and the interested reader is

referred to that report. One of the more recently developed techniques of RTCE, which can serve as an example of the technique, is that of pilot tone sounding.

In pilot tone sounding [Betts and Darnell, 1975], low level CW tones are either inserted into the data spectrum or transmitted in potentially available alternative channels. At the receiver, simple measurements on the tones enable the relative states of the channels to be specified in terms of predicted (digital) data error rates. In a situation where a multicomponent broadcast is being used to radiate identical data simultaneously on each of several frequencies, this type of RTCE system allows the best component to be selected automatically at the receiver.

One of the main advantages of the pilot-tone RTCE technique is the extreme simplicity of the concept and the implementation when compared with other techniques for RTCE, for example, oblique incidence ionosondes. A further advantage of the technique is that it permits RTCE and data signals to be combined in a simple format, rather than requiring a separate stimulus. The technique also lends itself readily to automation. With an automated system, there will be little difficulty in principle in including into the algorithm for selecting the optimum working frequency any measurements of other pertinent properties of the pilot tone, for example, amplitude, Doppler shift, and spectral width. The noise level in the channel could also be included.

The pilot-tone sounder technique does not permit the determination of the operational MUF or identification of propagation modes. The former may cause difficulties when the MUF is decreasing (for example, due to normal diurnal variation) if the working frequency is just below the MUF. Difficulties could be severe if the MUF changes significantly during the RTCE time (typically 3 min per channel).

The signal optimization problem at HF is potentially very complex and to date RTCE has been used almost exclusively as a frequency selection aid. Darnell [1975b] lists the various parameters on an HF link which could be under the control of the communicator and hence be made adaptive in response to the RTCE data:

Transmission frequency
Transmitter power level
Bandwidth
Data rate
Modulation type
Time at which the transmission is made
Signal processing algorithm at the receiver
Elevation angles for antenna array beams
Diversity type

To this list may be added antenna polarization for near vertical incidence propagation.

The potential advantages accruing from the use of RTCE techniques have been summarized by Darnell [1975a]:

1. The effect of man made noise and interference can be measured and specified quantitively [Darnell, 1978].

2. The facility for real-time, on line measurement of propagation and interference allows the use of relatively transient propagation modes, for example, sporadic E layer propagation.

3. RTCE evaluation allows a more efficient use of the frequency spectrum by tending to select frequency channels higher than those which would be chosen via prediction techniques. Thus spectrum congestion is reduced.

4. RTCE will provide a means of automatically selecting the best frequency and simultaneously indicating preferred stand-by channels.

5. Transmitter power can be minimized, consistent with providing an acceptable quality of received traffic.

6. In the longer term, RTCE data can be used to adapt other parameters of a communication system, apart from frequency, optimally for the prevailing path conditions.

10.4.10 Conclusion

It has not been possible within the present limitations of space to present more than a cursory overview of the subject of HF propagation. However, it is hoped that most readers will be able to follow up on topics of special interest by going to the references cited. In this regard, the value of the proceedings of the "Boulder Workshop" (Donnelly, 1979; 1980) cannot be stressed too highly. The CCIR documents and the AGARD Lecture Series No. 127 should also prove very useful, but possibly somewhat harder to obtain.

10.5 IONOSPHERIC DISTURBANCES

10.5.1 Sudden Ionospheric Disturbance (SID)

The SID is caused by x-ray and ultraviolet emissions from solar flares. These emissions produce increased ionization of the sunlit ionosphere. This excess ionization typically lasts for a period of 10 to 60 minutes, depending on the intensity and duration of the responsible flare. SIDs are observed by monitoring manmade and natural High Frequency (HF) or Very Low Frequency (VLF) radio signals that propagate via the ionosphere. The various types of SIDs are named according to the method of their observation:

1. SWF—short wave fadeout: a decrease (either sudden or gradual) in the signal received from a distant HF (2–32 MHz) transmitter.
2. SCNA—sudden cosmic noise absorption: a decrease in the intensity of the constant galactic radio noise as measured by riometers (*r*elative *i*onospheric *o*pacity *m*eters) operating between 15–60 MHz.
3. SPA—sudden phase anomaly: a change in phase of a received VLF (10–150 kHz) signal relative to a frequency standard at the receiving site.
4. SES—sudden enhancement of signal: an increase in the strength of an incident VLF signal occurring at the same time as an SPA.

5. SEA—sudden enhancement of atmospherics: an increase in the background VLF noise from distant thunderstorms,
6. SFD—sudden frequency deviation: a short-lived increase in the frequency of the signal from a distant HF transmitter,
7. SFE—solar flare effect or geomagnetic crochet: sudden variation in the H component of the earth's magnetic field.

The first five of these effects are all attributable to increased ionization in the D layer, primarily by soft (1–8 Å) flare x-rays. The absorptive effects (SWF and SCNA) are caused by a thickening of the absorbing layer through which the HF waves pass, while the effects on VLF signals are due to the lowering (SPA) and strengthening (SES and SEA) of the D layer from which these lower frequency waves are reflected. In contrast, SFDs arise from the flare associated increase in ionization in the E and F regions produced by extreme ultraviolet (EUV) radiation in the 10–1030 Å range. Finally, the geomagnetic crochet (SFE) appears to be a hybrid effect, having an impulsive component associated with the flare "flash phase" EUV emission, and a more gradual component associated with the flare soft (1–8 Å) x-rays [Richmond and Venkateswaran, 1971]. These ionizing emissions temporarily increase the D and E layer conductivity and alter the normal ionospheric currents to give rise to the SFE. For a comprehensive review of the SID phenomenon, see Mitra [1974].

10.5.2 Polar Cap Absorption (PCA)

While SIDs are caused by flare electromagnetic emission, PCAs result from bombardment of the ionosphere by flare-accelerated protons with energies $\lesssim$ 10 MeV. These particles stream into the earth's polar regions along geomagnetic field lines and produce increased ionization in the D layer (as well as aurora at E layer heights). PCAs are normally observed by means of riometers. In contrast to the SID, the PCA is a long-lived effect, with durations ranging from tens of hours to several days. In general, PCAs follow only the most intense solar flares and begin within a few hours after flare maximum, dependent on the flare longitude. A review of the PCA phenomenon has been published by Hultqvist [1969].

10.6 HF RADAR IONOSPHERIC CLUTTER

Backscatter reflections from E layer auroral ionization and F layer electron density irregularities can be an important source of clutter for a radar operating in the high latitudes. The ionospheric clutter can degrade the performance of an over-the-horizon (OTH) radar designed for the surveillance of aircraft at ranges of approximately 1000 to 3500 km.

The radar reflections are the result of scattering from electron density irregularities aligned along the lines of force of the earth's magnetic field. The characteristics of the field-aligned scatterers are such that the radar echoes originate in a small range of angles about perpendicular incidence to the magnetic field lines and that the amplitude of the auroral echo is aspect-angle sensitive. In addition to the orthogonality condition, it is necessary that this geometric configuration take place at ionospheric altitudes, that is, 80 km and above.

The probability of observing radar reflections from ionization irregularities is also dependent upon the frequency of occurrence of E layer auroral or spread-F irregularity activity in the region of interest. Reflections can be expected to occur in regions where both the conditions of near-perpendicularity at ionospheric heights and high auroral or spread-F irregularity activity are satisfied. The fulfillment of only one requirement is not sufficient to warrant a radar reflection.

An appropriate model to use to define the condition of auroral activity is the Feldstein-Starkov [1967] auroral belt (oval) model. Since the location and extent of the auroral oval are a function of time and geomagnetic activity, E layer auroral echoes can be expected to appear over a wide area during magnetically disturbed periods for an HF backscatter radar located at high latitudes.

The F layer irregularities can be described in terms of the probability of occurrence of spread-F derived by Penndorf [1962] or in terms of Aarons' [1973] irregularity scintillation region.

In this section, an estimate is made of the characteristics of ionospheric clutter that could be observed by an HF backscatter radar operating in the midlatitude with the antenna beam oriented towards the polar region. The topics to be discussed are the amplitude, the backscatter cross-sectional area, the angular extent, the Doppler frequency variation, the frequency of occurrence, the diurnal and seasonal variation and the correlation with solar-geophysical conditions.

10.6.1 Signal Amplitude

The amplitude of E layer auroral clutter that could be encountered by an HF backscatter radar is deduced by extrapolation of radar-auroral data recorded by SRI International at Fraserburgh, Scotland, during 1959 and 1960 [Leadabrand et al., 1965]. The radar measurements were made simultaneously at frequencies of 30, 401 and 800 MHz.

To predict the SRI International auroral data effects on an HF backscatter radar in terms of the corresponding radar, it is necessary to determine the relative sensitivities of the

CHAPTER 10

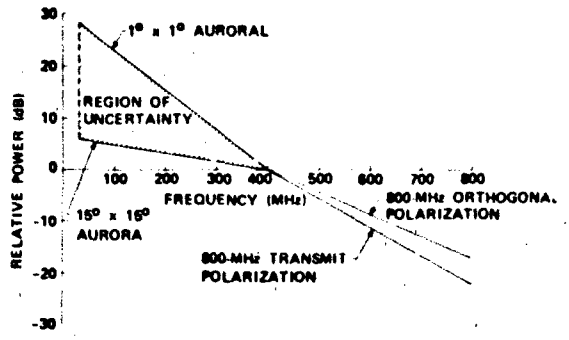


Figure 10-54. Relative signal strength of auroral echoes at 30, 401, and 800 MHz obtained by the SRI International Scotland radars [Leadabrand et al., 1965].

radars to the aurora. This is accomplished by comparing the various parameters that enter into the radar equation.

According to the radar equation, the signal-to-noise ratio of the received signal is given by

$$\frac{P_r}{N} = \frac{P_t G_t G_r \lambda^2 \sigma}{(4\pi)^3 R^4 \, KTB \, \overline{NF} \, L_s L_p} \qquad (10.60)$$

where P_t is the transmitted power, G_t and G_r are the gains of the transmitting and receiving antennas, respectively, λ is the wavelength, σ is the radar-target cross section, R is the radar range, K is Boltzmann's constant, T is the ambient temperature, B is the receiver noise bandwidth, $\overline{NF}$ is the receiver noise figure, L_s is the system loss and L_p is the two-way loss due to the propagation medium. This relationship is applicable to a point target, that is, the antenna beamwidth is much larger than the dimensions of the target.

For aurora which fills the antenna beam, σ is replaced by the volume scattering coefficient, σ_v, that is, radar cross section per unit volume, such that, for a rectangular antenna,

$$\sigma = \sigma_v V = \sigma_v R^2 \beta_H \beta_V \frac{c\tau}{2}, \qquad (10.61)$$

where V is the volume filled by the scatterers, c is the free space velocity, τ is the pulse length, and β_H and β_V are the horizontal and vertical antenna beamwidth, respectively.

Figure 10-54 is a plot of the relative signal strengths of auroral echoes recorded by SRI International at the three frequencies in April 1960. The region of uncertainty between 30 and 401 MHz results from the fact that there was no way to specify to what degree the auroral scatterers filled the 30-MHz antenna beam. At 401 and 80 MHz, the antenna beamwidth was 1.2°, while at 30 MHz it was 15°. Thus, it was assumed that the scatterers completely filled the beam at the two higher frequencies. However, it was doubtful that the beam was completely filled at 30 MHz. The 1° × 1° aurora is assumed to be a point target for the 30-MHz antenna.

The data in Figure 10-54 are normalized with respect to 401 MHz. For a filled beam antenna, the relative signal strength of the 30-MHz echoes was 6-dB greater than those at 401 MHz, while, for the point target case, there was a 28-dB difference between the 30- and 401-MHz data.

The absolute magnitude of the 30-MHz auroral echoes can be deduced by comparing the system sensitivities of the 30- and 401-MHz radars. According to Table 10-9, the 401-MHz radar was 48.9-dB more sensitive than the 30-MHz radar assuming that the aurora was a point target and 26.9-dB more sensitive, assuming that the 30-MHz antenna beam was completely filled with scatterers. Thus, it follows that

Table 10-9. Comparison of SRI International-Scotland 30-MHz and 401-MHz radar sensitivities utilized in the April 1960 radar-auroral observations.

Parameter	30-MHz Radar	401-MHz Radar	Point Target (401 MHz/30 MHz) (dB)	Scatterers Fill Antenna Beam (401 MHz/30 MHz) (dB)
λ^2 (m^2)	$(10.0)^2$	$(0.748)^2$	−22.5	−22.5
P_t (Mw)	0.0015	0.12	+19	+19
$G_t G_r$ (dB)	40	84	+44	+44
β_H (DEG)	15	1.2	−	−11
β_V (DEG)	15	1.2	−	−11
τ(μs)	300	300	0	0
B (kHz)	6	6	0	0
$\overline{NF}$ (dB)	9	4.5	+4.5	+4.5
L_s (dB)	7	3.1	+3.9	+3.9
			+48.9	+26.9

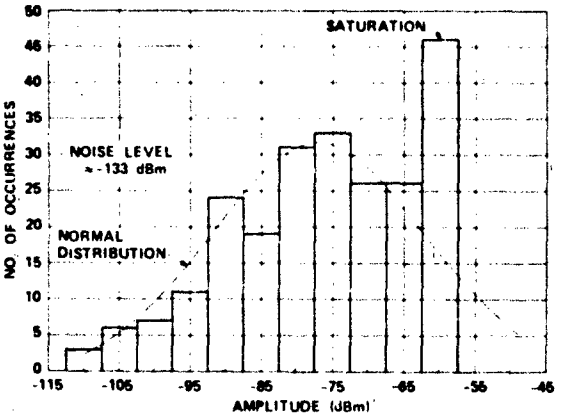


Figure 10-55 Amplitude histogram of SRI International 401-MHz Scotland radar auroral data [Leadabrand et al., 1965].

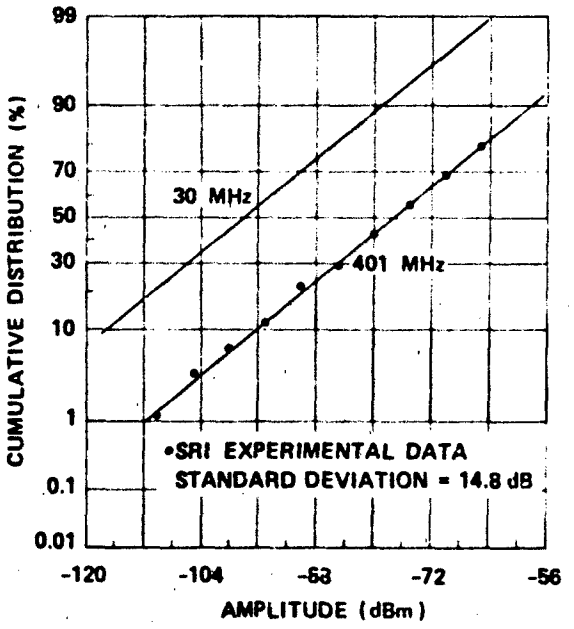


Figure 10-56. Cumulative distribution function of SRI International-Scotland radar auroral data.

the magnitude of the 30-MHz radar-auroral data was 20.9-dB less than that of the 401-MHz data.

The amplitude distribution of the 401-MHz radar-auroral echoes is presented as a histogram form in Figure 10-55. The system noise level was approximately -133 dBm and that the system saturated at -60 dBm, which accounts for the large number of echoes at that amplitude.

The histogram data, when replotted as a cumulative distribution on probability paper, shown in Figure 10-56, closely fit a straight line. This characteristic is indicative that the 401-MHz amplitudes are log-normally distributed. The theoretical normal distribution function plotted in Figure 10-55 was computed utilizing the value of -77.2 dBm for the mean as obtained from Figure 10-56. The 30-MHz amplitude data, also shown in Figure 10-56, was derived on the basis that the 30-MHz auroral echoes followed a Gaussian distribution and were 20.9-dB less than the 401-MHz results.

Table 10-10 Comparison of the sensitivity of an HF backscatter radar to that of the SRI International-Scotland 30-MHz radar

Parameter	HF Backscatter Radar	SRI-Scotland Radar	Point Target (HF Radar/SRI) (dB)	Scatterers Fill Antenna Beam (HF Radar/SRI) (dB)
λ^2 (m^2)	$(10.0)^2$	$(10.0)^2$	0	0
P_t (MW)	0.8	0.0015	+27.3	+27.3
$G_t G_r$ (dB)	45	40	+5	+5
β_H (DEG)	5	15	–	−4.8
β_V (DEG)	15	15	–	0
τ (μs)	400	300	–	+1.2
	10			−14.8
B (kHz)	2.5	6	+3.5	+3.8
	100		−12.2	−12.2
$\overline{NF}$ (dB)	12	9	−3	−3
L_s (dB)	3	7	+4	+4
			(τ = 400 μs) +37.1 dB	+33.5 dB
			(τ = 10 μs) +21.1 dB	+1.5 dB

Table 10-11. Estimate of radar-auroral clutter levels of a hypothetical HF backscatter radar at 30 MHz based on SRI International data.

Statistical Parameter	Point Target		Scatterers Fill Antenna Beam	
	τ = 400 μs (dBm)	τ = 10 μs (dBm)	τ = 400 μs (dBm)	τ = 10 μs (dBm)
Upper Decile	−42.0	−58.0	−45.6	−77.6
Upper Quartile	−51.0	−67.0	−54.6	−86.6
Median	−61.0	−77.0	−64.6	−96.6
Lower Quartile	−71.0	−87.0	−74.6	−106.6
Lower Decile	−80.0	−96.0	−83.6	−115.6
Standard Deviation	14.8	14.8	14.8	14.8

To estimate the magnitude of the auroral echoes that could be observed by an HF backscatter radar, it is necessary to compare the sensitivity of the HF backscatter radar to that of the SRI International 30-MHz radar. The parameters of an hypothetical HF radar, considered in this analysis, are given in Table 10-10. For an assumed pulse length of 400 μs, the HF radar is 37.1-dB more sensitive than the SRI International 30-MHz radar for the point target case and 33.5-dB more sensitive for the filled beam case. When the pulse length is decreased to 10 μs, the difference in sensitivities decreases to + 21.1 dB and + 1.5 dB for the point target and filled beam case, respectively.

The predicted HF backscatter radar-auroral clutter levels at 30 MHz are presented in Table 10-11. The SRI International 30-MHz radar-auroral data given in Figure 10-56 and radar sensitivities in Table 10-10 are used in the calculations.

In deducing the radar-auroral signal levels at other frequencies in the HF band, it is assumed that the auroral echo power is frequency dependent according to the law

$$P = k f^{-n}, \qquad (10.62)$$

where k and n are constants. Utilizing the data in Figure 10-54, it can be shown that for this power law n = 2.5 for a point target and n = 0.5 for the filled beam case. However, an examination of other radar-auroral data taken at Scotland revealed a value of n = 6.8 for a point target and n = 2.7 for a filled beam [Leadabrand et al., 1965]. An analysis of auroral echoes from simultaneous multiple frequency observations in Alaska by Leadabrand et al., [1967] revealed than n = 2 for frequencies between 50 and 850 MHz and n = 5 between 850 and 3000 MHz.

Radar backscatter measurements of artificial electron clouds in the E region of the ionosphere by Gallagher and Barnes [1963] yielded a constant, n, of 4 for frequencies between approximately 20 and 50 MHz and = 4 between 5 and 20 MHz. At times, however, it was found that the amplitude returns were insensitive to frequency in approximately the 10- to 20-MHz range. From simultaneous auroral echo measurements at frequencies of 49.7, 143.5 and 226 MHz, Flood [1960] has deduced a value of n = 3.5 between 49.7 and 143.5 MHz and n = 6.5 between 143.5 and 226 MHz.

Measurements of E region radar aurora at frequencies in the HF band and at VHF (143.8 MHz) from locations in Finland and Germany indicated a frequency dependency of n = 3 [Oksman et al., 1979].

Since there is a wide discrepancy in the experimental measurements of the frequency dependence of auroral backscatter, which could be due to the characteristics of the auroral ionization, that is, inhomogeneous distribution of auroral electrons, varying scale sizes of ionization irregularities and different scattering altitudes, a value of 4 for the exponent in Equation (10.62) is assumed in this analysis.

The estimated median auroral clutter amplitudes that could confront the hypothetical HF backscatter radar in the 5- to 30-MHz band are shown in Figure 10-57. The external

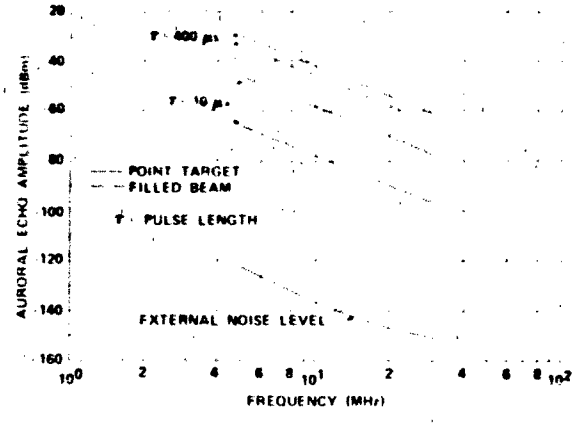


Figure 10-57. Estimated median auroral clutter amplitude based on SRI International-Scotland radar data, f^{-4} frequency dependency.

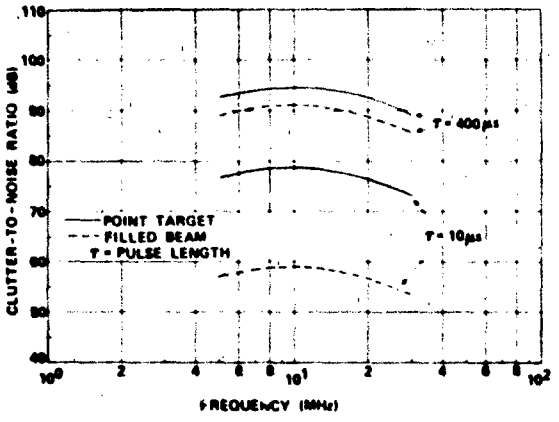


Figure 10-58. Estimated median radar-auroral clutter-to-noise ratio based on SRI International-Scotland radar data, f^{-4} frequency dependency

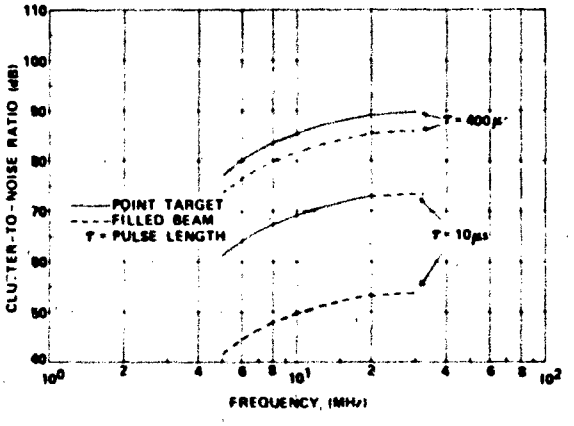


Figure 10-59. Estimated median radar-auroral clutter-to-noise ratio based on SRI International-Scotland radar data, f^{-2} frequency dependency.

noise level of a rural environment also shown in Figure 10-57 is the average median value as predicted for a location in Maine at an azimuth of 60° for all seasons, all times of day and a sunspot number of 70. It is seen that the auroral clutter amplitude increases with decreasing frequency and bandwidth and that the amplitude of the point target case is greater than that of the filled beam case.

Figure 10-58 is a replot of the data in Figure 10-57 in terms of the auroral clutter-to-noise ratio. It is of interest to note that, for a f^{-4} frequency dependency, the clutter-to-noise ratio maximizes in the vicinity of 10 MHz. This is due to the fact that the external noise level increases with decreasing frequency.

When a frequency dependency of f^{-2}, that is, $n = 2$ in Equation (10.62), is assumed, the ionospheric clutter levels over the HF band are decreased, with respect to the f^{-4} estimates, by an amount of 15.6, 9.5, 6.0, 3.5 and 1.6 dB at 5, 10, 15, 20 and 25 MHz, respectively. As shown in Figure 10-59, the clutter-to-noise ratio for the f^{-2} case, monotonically increases with increasing frequency. The clutter-to-noise ratio at 30 MHz is 12.2 dB greater than that at 5 MHz. The upper and lower decile values of the data presented in Figures 10-58 and 10-59 are + 19 dB and − 19 dB with respect to the median level.

The height distribution of the auroral echoes observed by SRI International in Scotland was peaked at about 100 to 120 km, that is, in the E region, although heights as great as 200 km were observed. Thus, the data presented in Figures 10-58 and 10-59 can be considered to apply only to E region reflections.

It should also be mentioned that the amplitude distribution of the 401-MHz auroral echoes shown in Figure 10-55, which was used as a basis of extrapolation to the HF backscatter radar, was applicable to data obtained when the direction of propagation at the reflection point was perpendicular to the magnetic field lines. When directions of propagation and magnetic field lines are not orthogonal, the auroral clutter amplitude should decrease.

Bates and Albee [1969] report that, assuming a simple model of auroral backscatter from the E region that includes ionospheric refraction effects, a lower limit of 6 dB/deg in the 15- to 50-MHz range was observed for the aspect-sensitive decrease in backscatter cross section with off-perpendicular angle from the geomagnetic field. The aspect-sensitive decrease of the cross section of F layer backscatter echoes was found by Bates and Albee [1970] to be on the order of 5 dB/deg of off-perpendicularity from the magnetic field in the 6- to 15-MHz region. An aspect angle decay of 2 dB/deg at 30 MHz was deduced from the Scotland-auroral measurements by Leadabrand et al. [1965].

10.6.2 Cross Section

HF radar observations conducted at Caribou, Maine, under the Polar Fox II program, at frequencies of 8.125 and 14.875 MHz, during the period between December 1971 and November 1972 are used to estimate the cross section of ionospheric clutter.

The radar cross section given by Equation (10.61) can also be expressed by

$$\sigma = \sigma_A R \beta_{II} \frac{c\tau}{2} \Delta h \qquad (10.63)$$

where Δh is the thickness of the aurora in the vertical direction. Since Δh was an unknown factor in the Polar Fox II experiment, it was assumed to be unity (1m). Thus, Equation (10-63) can be written as

$$\sigma = \sigma_A A = \sigma_A R \beta_{II} \frac{c\tau}{2} \qquad (10.64)$$

Table 10-12. Average area scattering coefficient of E layer and F layer irregularities based on Polar Fox II data, January 1972.

Frequency (MHz)	Statistical Parameter	Area Scattering Coefficient (dBsm/m²)	
		E Layer	F Layer
8.125	Upper Quartile	−34.2	−27.2
	Median	−47.6	−33.6
	Lower Quartile	−58.6	−42.0
14.875	Upper Quartile	−56.5	−37.1
	Median	−63.9	−47.2
	Lower Quartile	−74.0	−54.3

where σ_A is the area scattering coefficient and A is the area containing the scatterers.

The area scattering coefficients of E layer and F layer irregularities as deduced from the Polar Fox II data are given in Table 10-12. The data show the averages over all times in the month of January 1972, and over all azimuths in the 90° sector between −30° and +60° with respect to true North. It is seen that, for all the statistical parameters, the 8.125-MHz data are from 12- to 21-dB greater than the 14.875-MHz results. The most interesting feature of the basic data, from which Table 10-12 was derived, was the lack of evidence of the influence of the geographic locations of the ionospheric scatterers on the magnitude of the area scattering coefficient. It was hypothesized that the scattering coefficient would be greater in the northerly directions since this is the region encompassed by the auroral oval [Feldstein and Starkov, 1967].

The cross section of ionospheric clutter that could be observed by the HF backscatter radar is presented in Table 10-13. The calculations are based on an assumed radar range of 700 km for the E layer irregularities and 1500 km for the F layer irregularities. These ranges correspond to a ray path oriented at an elevation angle of 5° and intersecting the ionosphere at altitudes of approximately 100 and 300 km, respectively. For this configuration, the E layer scattering area evaluates to 95.6 dBsm and 79.6 dBsm for a pulse length of 400 μs and 10 μs, respectively. For the corresponding pulse lengths, the F layer scattering area is 99.0 dBsm and 82.9 dBsm. According to Table 10-13, it is evident that the F layer clutter cross section is about 10 to 20 dB greater than the E layer cross section at 8.125 MHz and 20 to 23 dB at 14.875 MHz. The differences between the F layer and E layer cross sections are most likely due to the fact that the F layer ionization level is many times that of the E layer.

It should be noted that the Polar Fox II estimates of the E layer and F layer area scattering cross section contain all the propagation loss terms, that is, the ionospheric propagation losses had not been removed from the calculations. Because of this, it is not possible to use the Polar Fox II data for extrapolation to other frequencies in the HF band.

10.6.3 Angular Extent

The angular extent of HF ionospheric clutter was determined from Polar Fox II radar data recorded in January 1972 at six frequencies ranging between 8 and 23 MHz. Figure 10-60 contains plots of the cumulative distribution of the angular extent of E layer, F layer, and the combined E and F layer clutter. According to Table 10-14 which

Table 10-13. HF backscatter radar estimated average cross-section of E layer and F layer irregularities based on Polar Fox II data, January 1972.

Frequency (MHz)	Pulse Length (μs)	Statistical Parameter	Average Cross Section (dBsm)		Cross Section Ratio (dB)	
			E Layer	F Layer	F Layer	E Layer
8.125	400	Upper Quartile	61.2	71.8	10.6	
		Median	48.0	65.4	17.4	
		Lower Quartile	37.0	57.0	20.0	
	10	Upper Quartile	45.4	55.7	10.3	
		Median	32.0	49.3	17.3	
		Lower Quartile	21.0	40.9	19.9	
14.875	400	Upper Quartile	39.1	61.9	22.8	
		Median	31.7	51.8	20.1	
		Lower Quartile	21.6	44.7	23.1	
	10	Upper Quartile	23.1	45.8	22.7	
		Median	15.7	35.7	20.0	
		Lower Quartile	5.6	28.6	23.0	

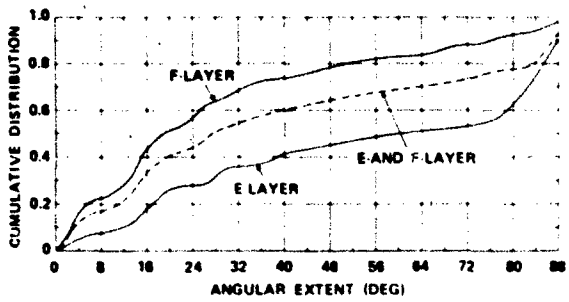


Figure 10-60. Cumulative distribution of the angular extent of ionospheric clutter based on Polar Fox II data

summarizes the statistical characteristics of the three curves in Figure 10-60, the median values evaluate to 60.2°, 18.8° and 27.6°, respectively. The angular dimensions characterized in Table 10-14 were deduced from ionospheric clutter present only in the 1-hop propagation mode.

10.6.4 Doppler Frequency Spectrum

The Doppler spectrum of radar pulses reflected from the aurora differs from that of the original transmitted pulse. The changes that can take place are (1) the center frequency can be shifted, and (2) the spectral width can be increased.

The shift in the center frequency corresponds to a drift motion of the auroral ionization. Radar-auroral data indicate an east-west drift before magnetic midnight and a west-east drift after magnetic midnight [Leadabrand et al., 1965]. Drift velocities of the order of 500 m/s are typical. According to the analysis of E layer radar-auroral echoes recorded in Scotland at a frequency of 401 MHz, the maximum Doppler shift is normally ±2.15 kHz [Larson and Hodges, 1967].

The spread in the Doppler spectrum is due to the random, turbulent motion of the irregularities of electron density in the auroral ionization. According to Larson and Hodges

Table 10-14. Statistical distribution of the angular extent of ionospheric clutter based on Polar Fox II data

Statistical Parameter	Angular Extent (Deg)		
	E Layer Clutter	F Layer Clutter	E and F Layer Clutter
Upper Decile	87.7	76.0	87.0
Upper Quartile	83.8	42.4	74.4
Median	60.2	18.8	27.6
Lower Quartile	19.0	10.0	13.6
Lower Decile	11.6	2.8	3.4

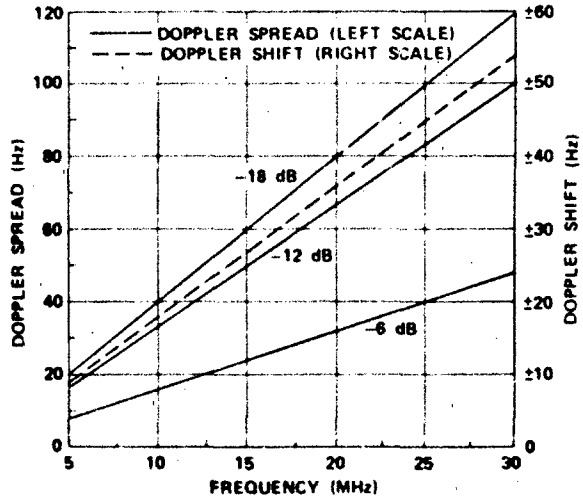


Figure 10-61. Estimate of E layer HF radar-auroral doppler frequency spread and shift based on SRI International-Scotland measurements at 401 MHz

[1967], the maximum Doppler frequency spread at 401 MHz is on the order of 1.90 kHz at −6 dB below the peak, 4.00 kHz at −12 dB, and 4.75 kHz at −18 dB.

The expected (one standard) deviation of Doppler frequency shift and spread at HF is presented in Figure 10-61. The estimates are based on the fact that the Doppler frequency variations are directly proportional to frequency and on the assumption that shifts and spreads are normally distributed. Thus, the maximum shift and spread are equivalent to the 3-sigma (standard deviation) value of the distribution

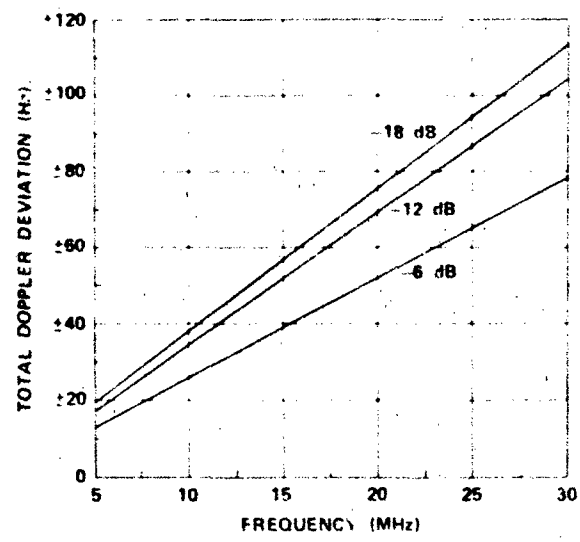


Figure 10-62. Estimate of E layer HF radar-auroral total doppler frequency deviation based on SRI International-Scotland measurements at 401 MHz.

At 10 MHz, the Doppler frequency shift evaluates to ± 18 Hz and the Doppler frequency spread becomes 16 Hz at the −6 dB level, 33 Hz at −12 dB and 40 Hz at −18 dB.

The estimates of E layer HF radar-auroral total Doppler frequency deviation are plotted in Figure 10-62. At 10 MHz, it is possible that the 1-sigma value of the total Doppler deviation could be on the order of ± 26 Hz at the −6 dB level, ± 35 Hz at −12 dB, and ± 38 Hz at −18 dB.

E layer radar-auroral echoes, having radial velocities as high as 1450 m/s, have been observed at 17 MHz by Brooks [1966]. This corresponds to a Doppler frequency shift of 164.3 Hz. The maximum radial velocity of F region field-aligned irregularities detected at 17.3 MHz [Baggaley, 1970] was found to be on the order of 165 m/s with a mean of 65 m/s. In terms of the Doppler frequency shift, these velocities evaluate to 19.0 and 7.5 Hz, respectively.

The extrapolation to the HF band of the Doppler measurements of E region irregularities conducted by Hofstee and Forsyth [1969] at a frequency of approximately 40 MHz, Balsley and Ecklund [1972] at 50 MHz, Balsley et al. [1973] at 50 MHz, Greenwald et al. [1975] at 50 MHz and Haldoupis and Sofko [1976] at 42 MHz are in agreement with the results presented in Figures 10-61 and 10-62.

10.6.5 Frequency of Occurrence and Correlation with Solar-Geophysical Conditions

The experimental observations of field-aligned echoes at 19.4 MHz conducted over a three-year period (1961–1963) at a site located in the vicinity of Boston disclosed that the echoes were present for as long as 11 h/day [Malik and Aarons, 1964]. Although a seasonal pattern was clearly defined, there appeared to be a tendency for the reflections to occur on a greater number of days during the summer than during the winter. The correlation of echo activity with sunspot number was found to exist. For example, in 1961 when the sunspot number was 53, echoes appeared on 73% of the days at an average of 3.0 h/day. In 1963 when the sunspot number was 28, the echoes appeared on 55% of the days at an average of 1.5 h/day. Maximum echo activity occurred between 1800 and 2000 h local time.

A more thorough analysis of the 19.4 MHz field-aligned echoes recorded from 1961 through 1965 had been performed by the Air Force Cambridge Research Laboratories [Basu et al., 1973, 1974]. It was found that under quiet magnetic conditions field-aligned E layer echoes showed a summer evening maximum and appeared to be associated with the ground-backscatter echoes from sporadic E [Basu et al., 1973]. A weak secondary maximum existed in the winter with no detectable field-aligned echoes during the daytime. During disturbed magnetic conditions, the field-aligned E layer echoes increased with magnetic activity and appeared during the daytime with a corresponding decrease in sporadic E. The data did not reveal a seasonal dependence.

Under quiet magnetic conditions, the field-aligned F layer echoes were found to be a sunset phenomenon and correlated with the sunspot cycle [Basu et al., 1974]. In general, the echoes occurred in the range from 1050 to 1500 km. However, during the daytime, long range echoes, at times, were observed at ranges of 3000 to 3300 km. An interesting disclosure was the fact that the occurrence of the F layer field-aligned echoes increased directly with the increase in magnetic activity until the level K_{Fr} (Fredericksburg, Maryland, three-hour K index) attained a value of four. Beyond this level, there was a decrease in the occurrence of the F layer echoes with no echoes being observed when $K_{Fr} \geq 7$. Radar auroral measurements made at Stanford University at a frequency of 17.3 MHz reveal that maximum auroral activity occurred between 1900 and 0400 hours local time [Peterson et al., 1955].

According to the HF radar echo observations conducted at 12 and 18 MHz in the state of Washington, there was a pronounced peak in the frequency of occurrence of the echoes from field-aligned irregularities near the time of local sunset [Hower et al., 1966]. The characteristics of the echoes detected at sunset were different from those occurring at night. It was found that the local sunset echoes appeared at ranges of about 2000 km and were generally spread in range (diffuse). The occurrence of the sunset echoes was relatively independent of magnetic activity. The nighttime echoes, on the other hand, appeared at ranges of about 1000 km and tended to display little spread in range (discrete). The occurrence of nighttime echoes was highly dependent on magnetic activity.

Hower et al. [1966] noted a decreasing echo activity with solar cycle, the maximum percentage occurrences during the nighttime being approximately 80% in December 1958, 37% in December 1963, and 35% in March 1964. The (Zurich) relative sunspot number during the corresponding period was 185.2, 11.8, and 14.5, respectively. Hower and Makhtiani [1969] have concluded that HF F layer backscatter echoes arise from the same general irregularity regions in which spread F is detected.

Sprenger and Glode [1964] have reported that the diurnal variation of the frequency of occurrence of radar-auroral echoes recorded from late 1958 to mid 1962 at 33 MHz at Kuhlungsborn (geographic coordinates: 54°N, 12°E; geomagnetic latitude: 54°N) showed a double maximum at approximately 0100 and 1700 hours local time. The probability of occurrence of E layer auroral echoes was found to be a function of magnetic activity. No auroral echoes were observed at a magnetic index, $Kp < 5$, and 100% occurrence at $Kp = 9$. With regard to the correlation of auroral echo activity with sunspot cycle, Sprenger and Glode [1964] have concluded that the maximum of auroral activity occurred in 1960 about two years after the sunspot maximum and that the activity decreased to zero within a period of only one and a half years.

Brooks' [1965, 1966] investigation of radar-auroral echoes at 17 MHz conducted in 1959 and 1960 indicated that, for the discrete echoes, the maximum number appeared between 0000–0300 hours local time, and for the diffuse echoes, between 1900–2100 hours local time.

A maximum number of backscatter echoes at 13.866 MHz from the F region on 75% of the nights for March 1958 was reported by Weaver [1965]. A minimum number of echoes was observed in June, the data being collected during 1957 and 1958 at Ithaca, New York, with the antenna beam oriented in a northward direction.

Backscatter observations of F region field-aligned irregularities made at 17 MHz near Sheffield, England (geographic coordinates: 53.43°N, 1.58°W; geomagnetic latitude: 65.4°N) between mid-October, 1964, and mid-January, 1966, [Baggaley, 1970] which was a period of low sunspot activity, showed that, on the average, the echoes were present on 23.3% of the days while only 1.5% of the total observing time was occupied by the echoes. A summer maximum and a winter minimum were found to exist. This seasonal variation correlates with Malik and Aarons' [1964] results.

Baggaley [1970] noted that there was a correlation between the onset of F layer echoes and the solar zenith angle at the reflection point in the F layer (assuming a height of 300 km). That is, a maximum number of echoes occurred at a solar zenith angle of 90°–95°, and a minimum at angles less than 85° and greater than 125°. In addition, no correlation was found to exist between F layer echo occurrence and magnetic activity. Of the total time for which F layer echo activity was present, only 11% was associated with magnetic index $Kp \geq 4$.

Field-aligned echoes from the F region observed at Brisbane, Australia, at 16 MHz occurred preferentially during geomagnetic disturbances and correlated strongly with spread-F and radio star scintillations occurring in the same region of the ionosphere [Swenson, 1972].

An analysis of 29-MHz backscatter measurements carried out in Northern (West) Germany and Scandanavia (geomagnetic latitudes of 55°–77°N) by Czechowsky et al., [1974] confirmed the close correlation between radio-auroral occurrence and geomagnetic activity. Most of the radio auroras appeared in the afternoon and evening hours.

10.6.6 Conclusions

The radar-auroral clutter-to-noise ratio in the 5–30MHz frequency range, as predicted from the SRI International radar data acquired in Scotland, is dependent on the frequency dependence law of the auroral backscatter echo power. Assuming a frequency variation of f^{-4}, the clutter level maximizes at approximately 10 MHz. When a f^{-2} law is assumed, the maximum clutter level is shifted to a frequency of 30 MHz.

For an assumed f^{-2} dependence, the radar-auroral clutter-to-noise ratio at 5 MHz is 15.6-dB less than that derived on the basis of an f^{-4} law. Assuming that the radar aurora is a point target, the clutter-to-noise ratio for a 400 μs pulse is estimated to be 3.6-dB greater than that deduced for auroral scatterers that completely fill the antenna beam. For a 10-μs pulse, the corresponding difference is 19.6 dB.

An analysis of the HF ionospheric backscatter data obtained in Maine under the Polar Fox II program reveals that the F layer cross section could be 10 to 20 dB greater than the E layer cross section at 8.125 MHz and 20 to 23 dB at 14.375 MHz.

The median value of the angular extent of E layer, F layer and combined E and F layer ionospheric clutter appears to be on the order of 60.2°, 18.8° and 27.6°, respectively.

Extrapolating from the SRI International-Scotland UHF radar-auroral data, it is estimated that for 5-MHz radar transmissions reflected from E layer auroral ionization, the 1-sigma value of the total Doppler frequency deviation, that is, Doppler frequency shift and spread, could be on the order of ±13 Hz at the −6 dB normalized signal level, ±17 Hz at −12 dB, and ±19 Hz at −18 dB. At 30 MHz, the 1-sigma Doppler deviations should increase to approximately ±78 Hz, ±105 Hz, and ±114 Hz, respectively.

HF field-aligned backscatter echoes usually occurs most often near local sunset. However, experimental observations have shown, at times, peak activity near the midnight hours.

There is a strong correlation of backscatter echoes with the solar cycle. That is, the percentage of days displaying backscatter reflections decreases with decreasing sunspot number.

The characteristics of sunset echoes are found to be different from those occurring during the night. The sunset echoes are generally of the diffuse type, and are independent of magnetic activity while the nighttime echoes are discrete and correlate with magnetic activity.

10.7 SCINTILLATION ON TRANS-IONOSPHERIC RADIO SIGNALS

A radio wave traversing the upper and lower atmosphere of the earth suffers a distortion of phase and amplitude. When it traverses ionospheric irregularities, the radio wave experiences fading, phase deviations, and angle of arrival variations. These signal fluctuations, known as ionospheric scintillation, vary widely with frequency, magnetic and solar activity, time of day, season, and latitude.

The irregularities producing scintillations are predominantly in the F layer at altitudes ranging from 200 to 1000 km with the primary disturbance region for high and equatorial latitude irregularities between 250 and 400 km. There are times when E layer irregularities in the 90 to 100 km region produce scintillation, particularly sporadic E and auroral E.

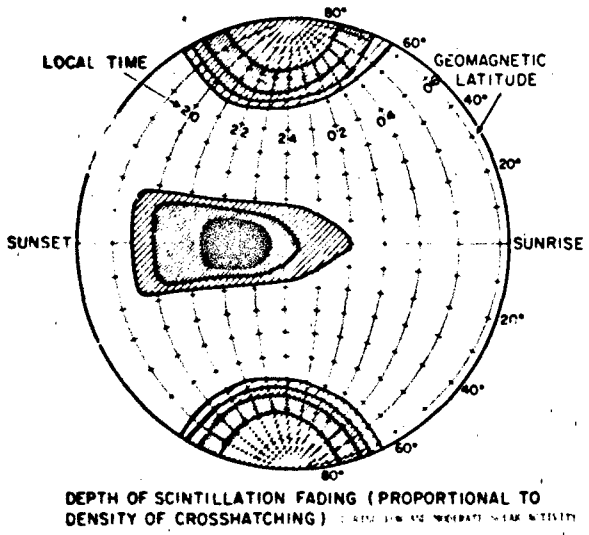


Figure 10-63. Global depth of scintillation fading (proportional to density of crosshatching) during low and moderate solar activity.

10.7.1 Global Morphology

From the global point of view there are three major sectors of scintillation activity (Figure 10-63). The equatorial region comprises an area within ± 20° of the magnetic equator. The high latitude region, for the purposes of the scintillation description, comprises the area from the high latitude edge of the trapped charged particle boundary into the polar region. We shall term all other regions "middle latitudes".

In all sectors, there is a pronounced nighttime maximum. At the equator, activity begins only after sunset. Even in the polar region, there appears to be greater scintillation occurrence during the winter months than during the months of continuous solar visibility.

10.7.2 Scintillation Examples

The intensity fading can best be characterized by an idealized example such as in Figure 10-64. The mean signal level is modulated by the passage through the irregularities so that the signal level very rapidly increases and decreases. In Figure 10-64 the mean signal level at times fades below the 3 dB level and below the 6 dB level. The number of fades and the fade duration for a typical 15 minute length of signal from a synchronous satellite is shown in Figure 10-64 along with the cumulative probability distribution function (pdf). In this example the signal was above the 6 dB fade level 91.7% of time.

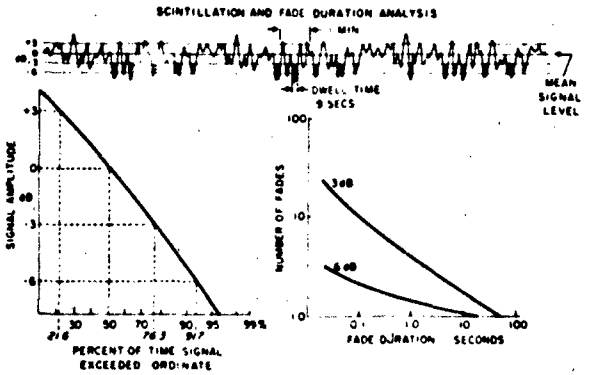


Figure 10-64. Sample of intensity fading produced by signal passing through irregularities. Fade duration and cumulative probability distribution are also shown.

10.7.3 Signal Characteristics

The amplitude, phase and angle of arrival of a signal will fluctuate during periods of scintillation. The intensity of the scintillation is characterized by the variance in received power with the S_4 index commonly used for intensity scintillation and defined as the square root of the variance of received power divided by the mean value of the received power that is, $S_4^2 = \dfrac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}$ [Briggs and Parkin, 1963]. An alternative, less rigorous but simple measure of scintillation index has been adopted by many workers in the field [Whitney et al., 1969] for scaling long-term chart records.

The definition is

$$SI = \frac{P_{max} - P_{min}}{P_{max} + P_{min}} \qquad (10.65)$$

where P_{max} is the power level of the 3rd peak down from the maximum excursion of the scintillations and P_{min} is the level of the 3rd peak up from the minimum excursion, measured in dB [Whitney et al., 1969].

The equivalence of selected values of these indices is indicated below.

S_4	dB
0.075	1
0.17	3
0.3	6
0.45	10

Scaling of the chart records is facilitated by simply measuring the decibel change between the P_{max} and P_{min} levels.

The phase variations are characterized by the standard deviation of phase, σ_ϕ, over a given interval of fluctuation frequency.

Attempts have been made to model the observed amplitude pdf. Whitney et al. [1972] and Crane [1977] have constructed model distribution functions based upon the use of the Nakagami-m distribution ($m = (S_4)^{-2}$) and have shown that the empirical models provide a reasonable approximation to the calculated distribution functions. Fremouw et al. [1980] showed that the Nakagami distribution for intensity and the normal distribution for phase may be used to characterize the statistics of the scintillation signal. In addition, the Rayleigh pdf provides a good fit to the data under conditions of very strong scintillation ($S_4 \sim 1.0$).

10.7.4 Frequency Dependence

Observations [Fremouw et al., 1978] employing ten frequencies between 138 MHz and 2.9 GHz transmitted from the same satellite, show a consistent wavelength (λ) dependence of the form $\lambda^{1.5}$ of S_4 for S_4 less than about 0.6. The frequency dependence becomes less steep for stronger scintillation, as S_4 approaches a maximum value near unity with a few rare exceptions. When S_4 exceeds 0.6 (peak to peak values > 10 dB) the frequency dependence exponent decreases. (If two frequencies are being compared and both experience strong scattering to the extent that each displays Rayleigh fading ($S_4 \sim 1.0$), then there is effectively no wavelength dependence over the frequency interval.) When strong scattering occurs but is not constant over the frequency interval, the wavelength dependence is difficult to determine. The observations [Fremouw et al., 1978] also show that the phase scintillation index, σ_ϕ, varies as λ for both weak and strong scattering. However, in extremely strong scattering environment, the frequency dependence of phase scintillation is also weakened.

10.7.5 Fading Spectra

During their passage through the ionospheric irregularities of electron density, radio waves from satellites undergo spatial phase fluctuations. As the wave emerges from the irregularity layer and propagates towards the ground, these phase fluctuations cause interference to occur and a diffraction pattern in both intensity and phase develops on the ground. In the presence of a relative motion between the diffraction pattern and the observer, a temporal variation of intensity and phase results. By transforming the temporal pattern to the frequency domain, the frequency spectra of intensity and phase fluctuations are obtained.

If a thin irregularity layer lies in the plane $z = 0$ of a cartesian coordinate system (x,y,z), and a plane wave of wavelength λ traveling in the z-direction is incident on the layer, then for weak scattering (rms phase $\phi_o \ll 1$) the spatial spectrum, $\Phi(\kappa_x,\kappa_y)$, of phase (s) and log amplitude (χ) (Yeh and Liu, 1982) on the ground ($z = z$) is given by

$$\Phi_{\chi,s}(\kappa_x,\kappa_y) = \pi(r_e\lambda)^2 L \left\{1 \mp \frac{2k}{\kappa_\perp^2 L}\sin\left(\frac{\kappa_\perp^2 L}{2k}\right) \times \cos\left[\frac{\kappa_\perp^2}{k}\left(z - \frac{L}{2}\right)\right]\right\}\Phi_{\Delta N}(\kappa_x,\kappa_y,0) \quad (10.66)$$

the upper and lower signs within the bracket referring to χ and s respectively. In the above equation,

$\kappa_x,\kappa_y,\kappa_z$ - spatial wave numbers in x, y and z directions

$\kappa_\perp^2 = \kappa_x^2 + \kappa_y^2$

$\Phi_{\Delta N}(\kappa_x,\kappa_y,\kappa_z)$ - power spectrum of irregularity electron density deviation

k - the wavenumber of the radio wave $\left(= \frac{2\pi}{\lambda}\right)$

r_e - the classical electron radius ($= 2.818 \times 10^{-15}$ m)

L - irregularity layer thickness.

In the presence of a relative motion between the propagation path and the irregularities the spatial spectrum $\Phi(\kappa_x,\kappa_y)$ is convected past the observer and a temporal variation of signal phase and amplitude is observed. If the irregularity structure does not change during the convection ("frozen-in" hypothesis) and the irregularities have a uniform velocity u in the x-direction then the power spectra of log amplitude and phase in the frequency domain are given by

$$\Phi_{\chi,s}(\omega) = \frac{1}{u}\int_{-\infty}^{+\infty}\Phi_{\chi,s}\left(\kappa_x = \frac{\omega}{u}, \kappa_y\right)d\kappa_y = \pi\frac{(r_e\lambda)^2 L}{u}\int_{-\infty}^{+\infty}\left\{1 \mp \frac{2k}{q^2 L}\sin\left(\frac{q^2 L}{2k}\right) \times \cos\left[\frac{q^2}{k}\left(z - \frac{L}{2}\right)\right]\right\}\cdot\Phi_{\Delta N}\left(\frac{\omega}{u},\kappa_y,0\right)d\kappa_y \quad (10.67)$$

where

ω—angular frequency of phase and amplitude fluctuation

$$q^2 = \frac{\omega^2}{u^2} + \kappa_z^2.$$

The term within the bracket with the upper sign is known

as the amplitude or Fresnel filter function and that with the lower sign as the phase filter function. The Fresnel filter function oscillates with the variation of frequency f and attains its first maximum at the frequency $f_F = u/\sqrt{2\lambda z}$. The behavior of the phase filter function is very different from the amplitude filter function as it fails to attenuate the low frequency regime. Equation (10.67) provides a relationship between the irregularity power spectrum in the ionosphere and the amplitude or phase scintillation spectrum obtained on the ground. Since the irregularity power spectrum has a power law variation [Dyson et al., 1974; Phelps and Sagalyn, 1976] of the form κ^{-p}, the power spectrum of amplitude scintillation shows a maximum at the frequency f_F due to the Fresnel filter function. On the other hand, the phase fluctuations are dominated by the low frequency regime. At $f \gg f_F$, both amplitude and phase scintillation spectra show an asymptotic variation f^{1-p} when the three-dimensional irregularity spectrum is of the form κ^{-p}. Thus from a study of weak scintillation spectra the spectral form of the irregularities causing scintillations may be deduced.

Figures 10-65a and 10-65b show two samples of weak phase and amplitude spectra obtained from 244 MHz scintillation observations made at Goose Bay, Labrador by the use of the Fleetsat geostationary satellite [Basu et al., 1982]. The phase and amplitude scintillation data were detrended by a high pass detrend filter with a cut-off frequency of 0.0067 Hz and the data sample yields an rms phase deviation of 2.4 radians and amplitude scintillation index $S_4 = 0.51$ conforming to weak scintillation criterion. Both spectra show an asymptotic variation in the high frequency region, the amplitude spectrum showing $f^{-2.35}$ variation, and the phase spectrum showing $f^{-2.6}$ variation. If we consider that the scintillation spectra have an average $f^{-2.5}$ variation, the corresponding three-dimensional irregularity power spectrum is expected to have a power law wavenumber spectrum of the form $\kappa^{-3.5}$. The decrease of power spectral density at the low frequency end of the phase spectrum is caused by the detrend filter. It may be noted that in the amplitude scintillation spectrum the high frequency roll-off starts at the Fresnel frequency of about 40 mHz. For an irregularity layer height of 350 km, the observed Fresnel frequency yields the irregularity drift velocity as 37 m/sec.

When scintillations become intense, the theory outlined above does not hold, and strong single scattering as well as multiple scatter theories appropriate to such cases have been developed [Yeh and Liu, 1982 and references therein]. In such cases it becomes difficult to relate in a straightforward manner, the scintillation spectra to the irregularity structures in the ionosphere. Figures 10-65c and 10-65d show a sample of intense amplitude scintillation data and its spectrum ($S_4 = 0.88$) analyzed from 257 Hz scintillation data obtained at Ascension Island near the crest of the equatorial anomaly by the use of transmission from the geostationary satellite, Marisat [Basu and Whitney, 1983]. In contrast to the weak amplitude scintillation spectrum shown earlier, these spectra show a flat low frequency portion and increased bandwidth of the frequency spectra, implying the development of shorter scales in the diffraction pattern. It is found that the correlation lengths get progressively smaller with increased strength of scattering [Rino and Owen, 1981; Basu and Whitney, 1983]. Under conditions of strong scattering, the phase scintillation spectra are also believed to suffer from the refractive scattering from very large scale irregularities [Booker and MajidiAhi, 1981]. Since wave propagation through a strong irregularity environment has considerable systems applications, intensive work in this area is in progress.

10.7.6 Geometrical Considerations

The intensity at which scintillations are observed depends upon the position of the observer relative to the irregularities in the ionosphere that cause the scintillation. Keeping both the thickness of the irregularity region and ΔN, the electron density deviation of the irregularity, constant, geometrical factors have to be considered to evaluate data and to predict scintillation effects at a particular location. Among these are:

(a) Zenith distance of the irregularity at the ionospheric layer. One study [Wand and Evans, 1975] found the intensity of scintillation related approximately to the secant of the zenith angle up to 70°; at angles >70° the dependence ranges between $^1/_2$ and the first power of the secant of the zenith angle.

(b) Propagation angle relative to the earth's magnetic field. Performing this calculation demands the use of an irregularity configuration and the consideration of Gaussian or a power law model for the irregularities. At high latitudes, irregularities in one study were elongated along the earth's magnetic field with a cylindrical form of axial ratio of 5 along the lines of force. Sheet-like irregularities with forms of 10 : 10 : 1 have also been found in recent auroral studies [Rino et al., 1978]. For equatorial latitudes, this elongation along the lines of force may be of the order of 50 to 100 [Koster, 1963].

(c) The distance from the irregularity region to the source and to the observer (near the irregularities, only phase fluctuations are developed). As noted in Mikkelsen et al., [1978] and Crane, [1977] the theoretical scintillation index can be expressed in terms of the above factors when dealing with ionospheric irregularities represented by a Gaussian power spectrum.

Mikkelsen et al. [1978] have attempted to determine the theoretical scintillation index, S_4, when the irregularities are described by a power-law power spectrum.

Mikkelsen assumed the approximate dividing line between weak and strong scintillation is ~9 dB, with SI < 9 dB denoting the weak case.

The geometric variation of S_4 is provided in Mikkelsen et al. [1978].

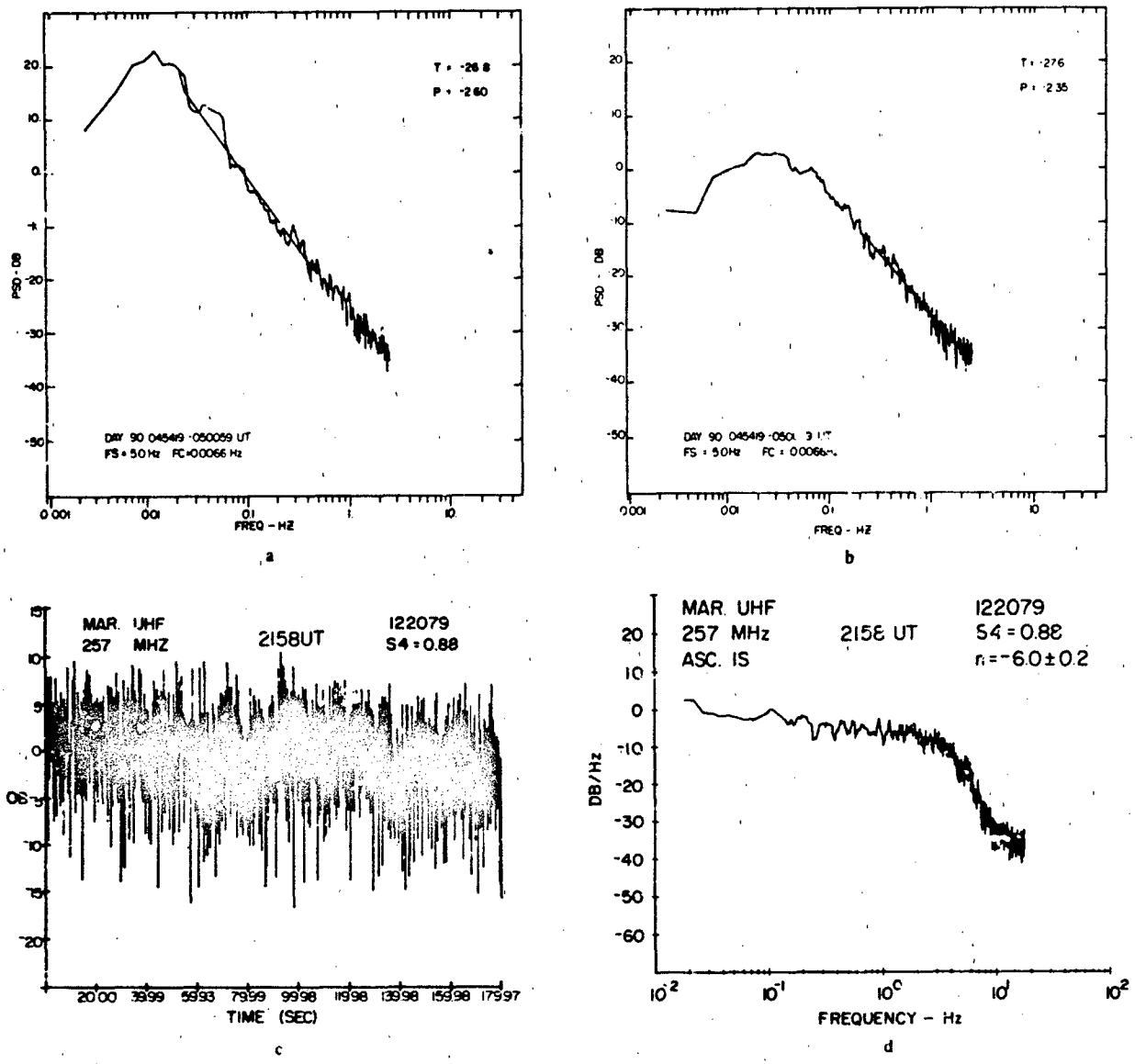


Figure 10-65. Spectra of (a) phase scintillation and (b) intensity scintillation under weak scatter conditions at 244 MHz observed at Goose Bay, Labrador on 31 March, 1979. (c) Data of 257 MHz intensity scintillation and (d) its spectrum under strong scatter conditions observed at Ascension Island on 20 December, 1979.

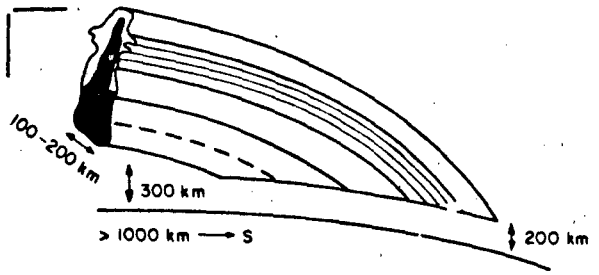
10.7.7 Spread F and Scintillation

The evidence from the corelation of scintillation occurrence and spread F [Rastogi, 1980] is that equatorial range spread is associated with scintillation activity and frequency spread is not. Thus the available spread F maps cannot be used for scintillation observations in these regions; they are dramatically misleading in many cases. In the high latitude region no statistical study has been made to correlate types of spread F with scintillation activity. It might be noted that even range spread occurrence and scintillation have important differences.

10.7.8 Equatorial Scintillations

10.7.8.1 Patch Characteristics. It has been established that nighttime ionospheric equatorial irregularity regions emerging after sunset develop from bottomside instabilities, probably of the Rayleigh-Taylor type. The depleted density bubble rises into the region above the peak of the F2 layer. Steep gradients on the edges of the hole help to generate the smaller-scale irregularities within the patch which produces intense scintillation effects [Basu and Kelley, 1979].

A plume-like irregularity region develops, finally forming a patch of irregularities that has been likened to a banana



THREE DIMENSIONAL
PATCH MODEL

Figure 10-66. Three dimensional model of an equatorial irregularity patch in the form of a banana or orange segment.

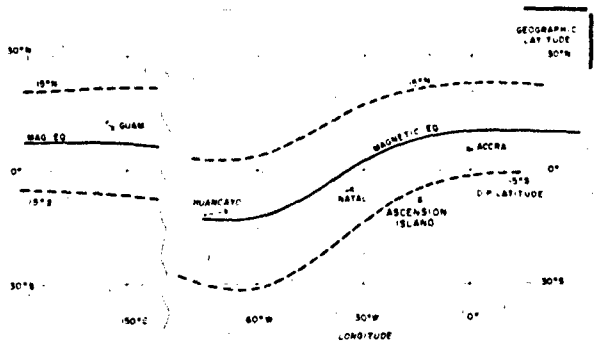


Figure 10-67. Map of equatorial regions using the 1975 epoch of the DMA magnetic inclination map. X marks sub-ionospheric intersection.

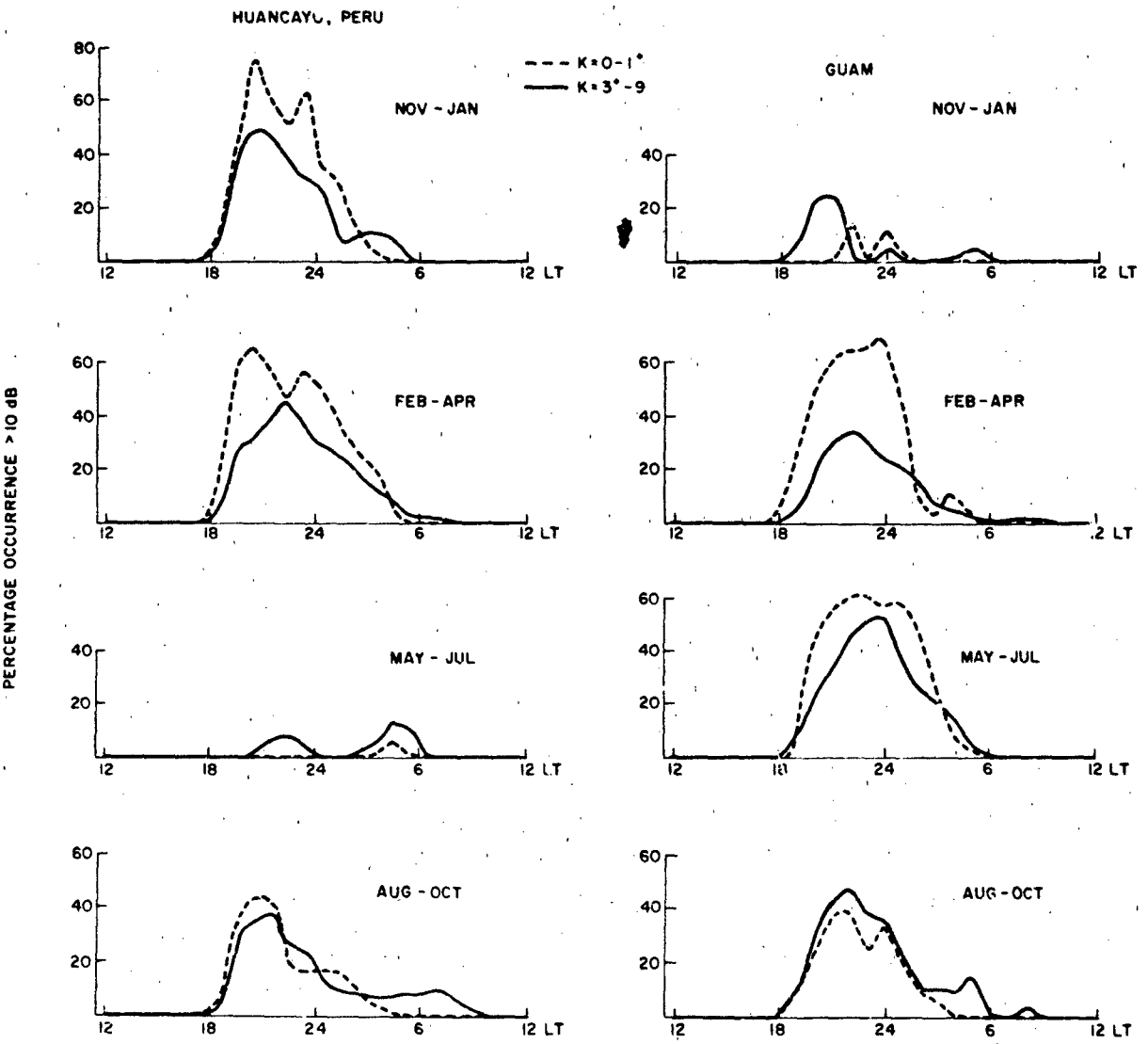


Figure 10-68. Comparison of seasonal patterns of occurrence of scintillation activity > 10 dB at 250 MHz at Guam and Huancayo under very quiet (Kp = 0 - 1⁻) and disturbed (Kp = 3⁻ - 9) magnetic conditions.

or an orange segment. A cut through the center of the "banana" is shown in Figure 10-66.

The characteristics of the patch development, motion and decay can be summarized as follows:

1. A new patch forms after sunset by expanding westward in the direction of the solar terminator with velocities probably similar to that of the terminator. It comes to an abrupt halt after typically expanding to an east-west dimension of 100 to several hundred kilometers. It appears to have a minimum size of ~100 km.

2. It is composed of field aligned elongated rod or sheet-like irregularities. The vertical thickness of the patch is 50 to several hundred kilometers. The patch has maximum intensity irregularities in a height region from 225 to 450 km, with irregularities extending to over 1000 km.

3. Its north-south dimensions are of the order of 2000 km or greater.

4. Once formed, the patch drifts eastward with velocities ranging from 100 to 200 m/s.

5. The patch duration as measured by scintillation techniques is known to be greater than 2½ h; individual patches have been tracked by airglow techniques up to 3 hours where they have maintained their integrity [Weber et al., 1978]. Effects have been seen over 8 h.

Studies of the variation in electron content in the patches have been made by measuring the change in relative phase between the two characteristic waves (ordinary and extraordinary modes) with polarimeters. It is found that the patches are regions of depletion in electron content. While the electron content depletions are found to be only of the order of 20%, the satellite *in situ* data may indicate density depletions as large as two or three orders of magnitude at one fixed altitude. In a strong irregularity environment, however, fast fluctuations in polarization are often obtained. Lee et al. [1982] have shown that scattering suffered by each characteristic wave may induce fast polarization fluctuations and obtained expressions for the variance of these fluctuations for irregularities with Gaussian and power law type spectra.

10.7.8.2 Variation of Scintillation Activity. A variety of observatories used data taken over the same time period to compare scintillation activity at 250 MHz [Aarons et al., 1980b]. One set of data was taken at Huancayo, Peru, Natal, Brazil, and Accra, Ghana with all observations made at elevation angles greater than 20° and with distance between the most separated stations about 70° of longitude; a map of both geographical and magnetic coordinates is shown on the right side of Figure 10-67.

A second comparison of data at 250 MHz was made between observations from Huancayo and from Guam. The data are shown in Figure 10-68; activity minima occur from May-July in Huancayo and from November-January in Guam. The conclusion is that the occurrence patterns are longitudinally controlled.

It should be noted that in general maximum intensity occurs in the equinoctial months. This can best be illustrated

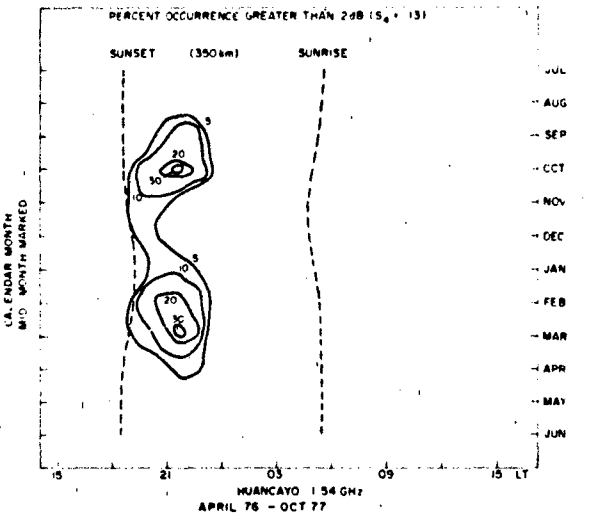


Figure 10-69. Percentage occurrence of 1.5 GHz scintillation ≥ 2 dB during Apr 1976–Oct 1977.

by the occurrence of L band, 1500 MHz activity, at Huancayo, Peru. That evidence is shown in Figure 10-69 [Basu et al., 1980b] L-band activity at Huancayo does not suffer from strong scattering or from saturation (as does 136 MHz and 250 MHz data on occasion); the data show clear equinoctial maxima.

From available data it appears as if geomagnetic control of the occurrence of scintillation varies with longitude. The generalization can be made that increased magnetic activity inhibits scintillation activity before midnight—except during those months with very low scintillation activity (May–July for the region (0°–70°W) and November–January in the Pacific longitudes (135°–180°E). After midnight the scintillation activity in general increases slightly with the presence of magnetic storms. The data shown in Figures 10-70 and 10-71 are for a year's observation in each case. The complexities of the magnetic control of scintillation occurrence are illustrated by the variations in the curves of occurrence at each station in each season. For further details see Mullen [1973].

10.7.8.3 In-Situ Data. Basu and Basu [1980] have developed a model from *in-situ*, theoretical, and scintillation studies. In their morphological model of scintillations, *in situ* measurements of irregularity amplitude, $\Delta N/N$, as computed from T secs of data are utilized in conjunction with simultaneous measurement of electron density N. A combination of $\Delta N/N$ and N data provides the required ΔN parameter as a function of position and time.

10.7.8.4 Sunspot Cycle Dependence. From the viewpoint of electron density variations, the equatorial region around the magnetic equator displays a complex pattern. During the day an increase in maximum electron density occurs away from the equator. The electron density contours

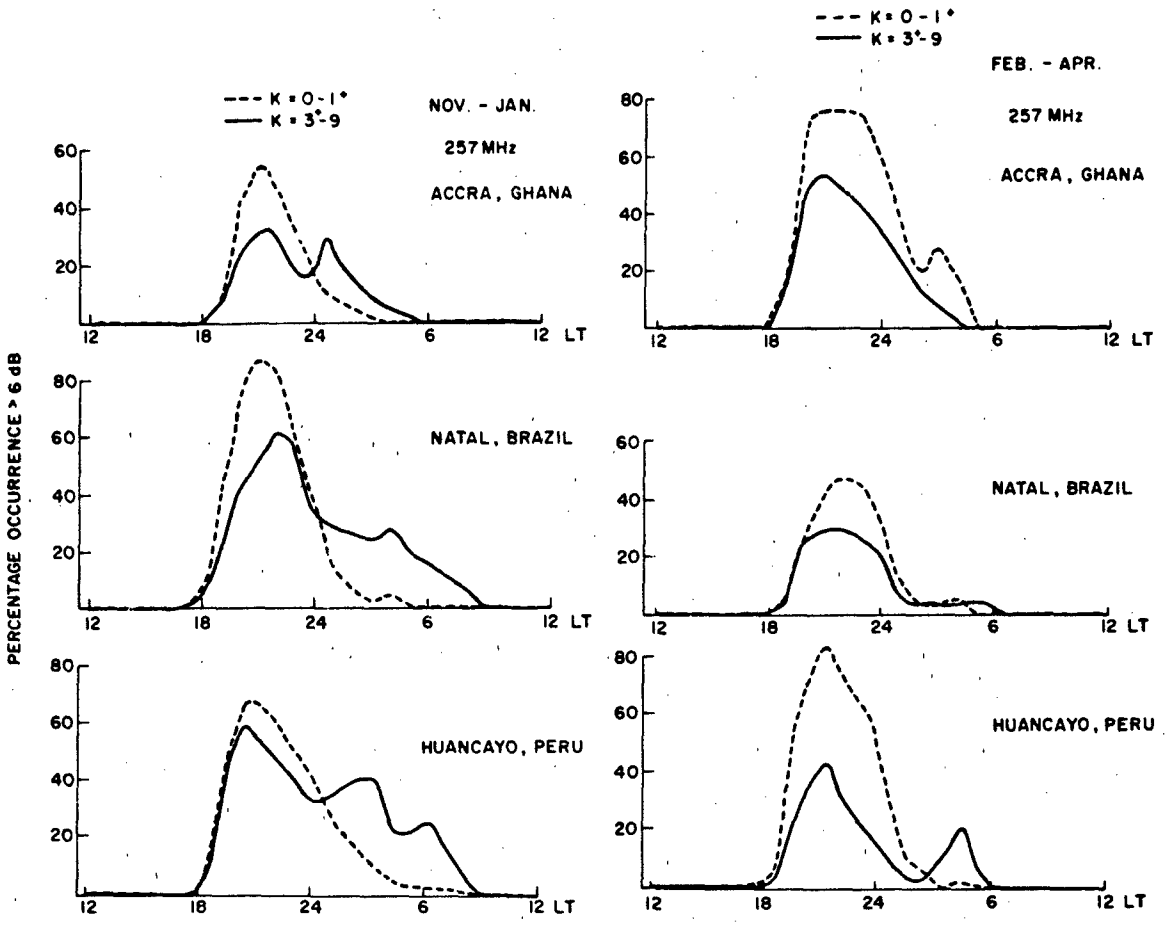


Figure 10-70. Seasonal patterns of occurrence of scintillation activity > 6 dB (S_4 = 0.3) at 250 MHz for very quiet (Kp = 0 – 1°) and for disturbed (Kp = 3° – 9) magnetic conditions for Nov–Apr.

display a distinct trough of electron density in the bottomside and topside ionosphere at the magnetic dip equator with crests of ionization at ± 15°–20° north and south dip latitudes; this is the Appleton anomaly with the region within ± 5° dip latitude of the magnetic equator termed the electrojet region.

From the solar cycle minimum in 1974 and maximum in 1969–1970, Aarons [1977] found that there was a higher occurrence of deep scintillations during a year of high solar flux than during a year with low solar flux for observations at both Accra, Ghana and Huancayo, Peru.

Recent observations of L band scintillations from both MARISAT and GPS during the period of maximum solar flux (1979–1981) [DasGupta et al., 1981] have revealed that scintillation intensities maximize in the Appleton anomaly region rather than near the magnetic equator.

The conclusion in the study [Aarons et al., 1981a] was that the intense scintillation activity during years of high solar flux are due to two factors:

1. The equatorial anomaly has considerably higher electron density values in high sunspot number years than in years of low solar activity.

2. The occurrence of maximum electron density for anomaly latitudes is near sunset in the years of high sunspot number and in the afternoon in years of low solar activity. Thus the post sunset irregularity patches attain high ΔN levels in the years of high solar flux.

10.7.9 Middle Latitude Scintillation

The middle latitude scintillation activity is not as intense as that encountered at equatorial, auroral or polar latitudes. However, activity may reach levels, primarily at VHF and UHF, that will increase error rates of systems with low fade margins. The reader is referred to Aarons [1982] and Bramley [1974].

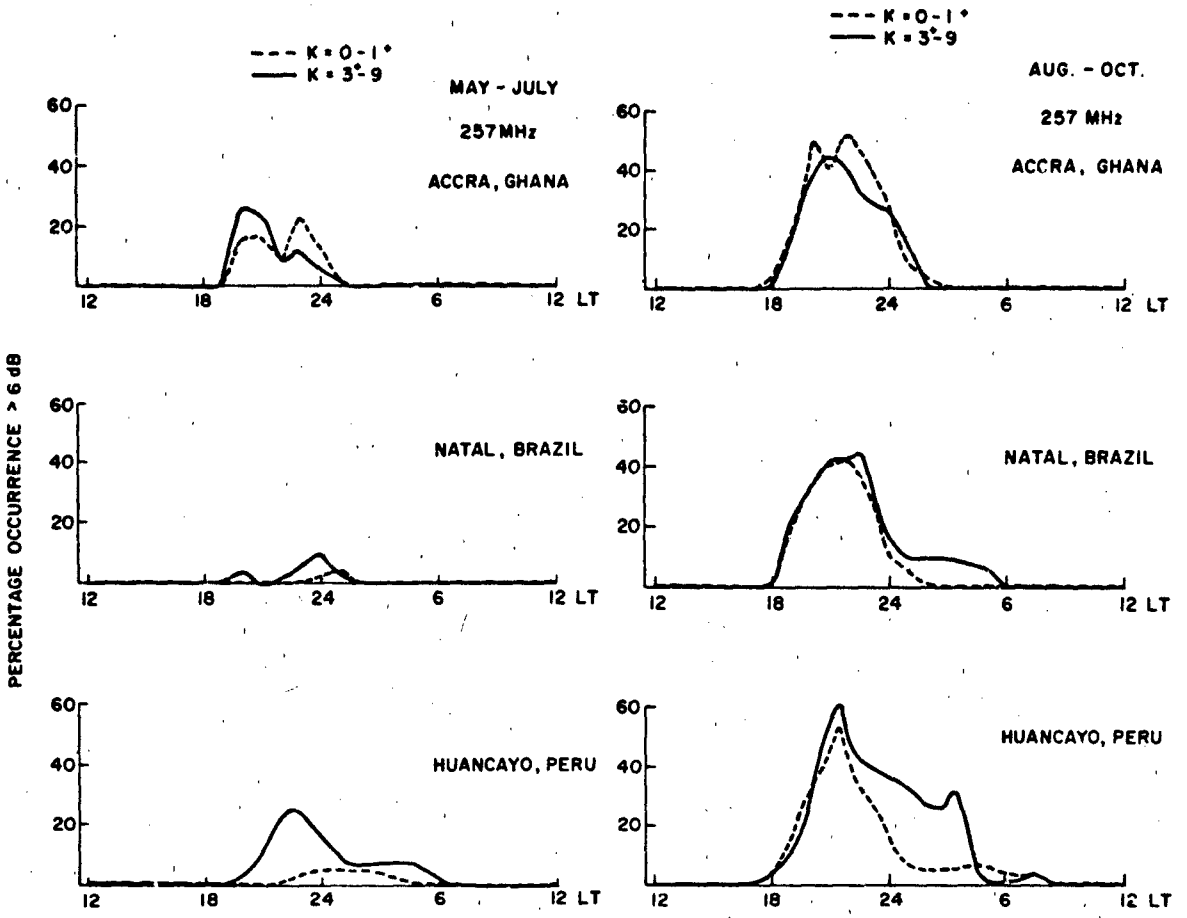


Figure 10-71. Seasonal patterns of occurrence of scintillation activity > 6 dB (S_4 = 0.3) for very quiet (Kp = 0 − 1+) and for disturbed (Kp = 3− − 9) magnetic conditions for May–Oct.

10.7.9.1 Effect of Magnetic Index on Midlatitude Scintillation. At latitudes below the auroral oval, various sets of data have yielded behavior indicating little correlation with magnetic conditions. Evans [1973] found no correlation of their 400 MHz radar scintillations with magnetic index south of their station at 56° invariant latitude. Aarons and Martin [1975] found that during the August 4–10, 1972 magnetic storms there was a negative correlation of scintillation and magnetic index for Athens, Greece and Camp Parks, California and little correlation at the 45° invariant latitude intersection for Aberystwyth, Wales. Bramley [1974] found that except for the December 1971 magnetic storm (when the irregularity region probably encompassed the intersection point of ~45°), there was no correlation between magnetic activity and scintillations.

This type of data essentially corroborates the early radio star obvservations in the U.K. which found little correlation with magnetic index except in paths to the north (with the exception of some intense magnetic storms).

10.7.10 The High Latitude Region

Figure 10-72 depicts the intensity of scintillation at high latitudes in a very broad manner for the period of time around midnight.

10.7.10.1 The Plasmapause and the Trough. The present evidence is that there is a boundary at the high latitudes where weak irregularities commence. It is probably a few degrees equatorwards of the plasmapause, between 45°–55° Corrected Geomagnetic Latitude (CGL), the system used at high latitudes in this review.

At night a trough or region of low electron density and total electron content exists between the end of normal ionospheric plasma behavior and the auroral region where energetic electron precipitation and current systems are dominant factors in producing both the normal ionospheric layers and the irregularities.

All observers of irregularities see a dramatic change in

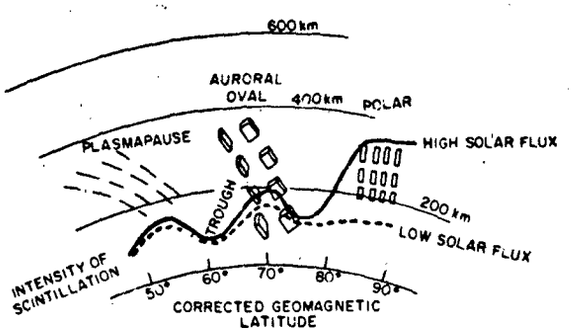


Figure 10-72. Depiction of high latitude irregularities ~22-02 LT. Sheet-like irregularities are seen in the auroral oval, rod irregularities at higher and lower latitudes.

irregularities in the auroral oval at the poleward edge of the trough. In the auroral oval, the intensity of scintillations is a function of local magnetic activity. Poleward of the aurora there may again be a lowering of scintillation activity until the observing path transits the polar region [Aarons et al., 1981b]

10.7.10.2 Auroral Scintillations. From studies of radio star and low altitude satellite scintillations, a series of height measurements have pointed to F layer heights as the primary seat of the irregularities producing the signal fading.

Maximum irregularity intensity appears above the region showing maximum intensity aurora [Martin and Aarons, 1977]. Vickrey et al. [1980] have shown that there is a collocation of scintillation patches in the auroral oval and F region ionization enhancements.

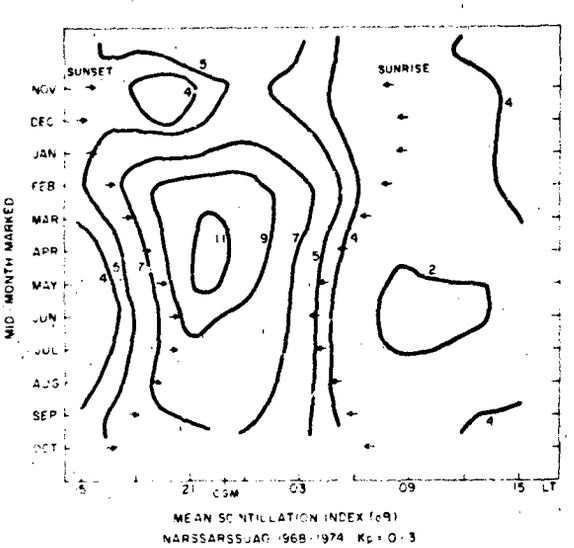


Figure 10-73. Contours of monthly mean scintillation index in dB at 137 MHz as a function of local time for quiet conditions (Kp = 0 – 3) obtained at Narssarssuaq during 1968-1974.

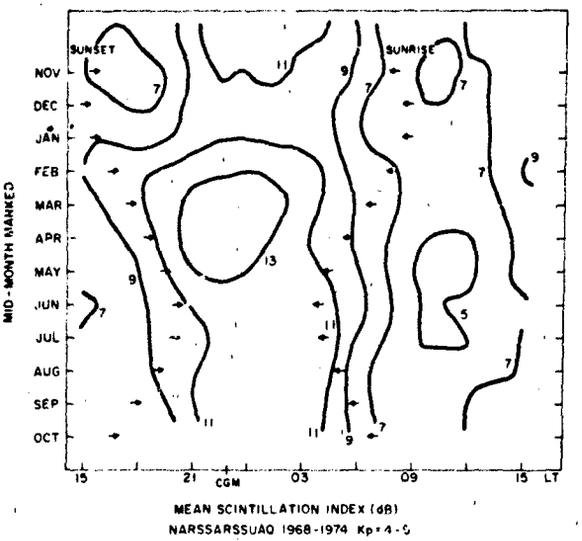


Figure 10-74. Contours of monthly mean scintillation index in dB at 137 MHz as a function of local time for disturbed (Kp = 4 – 9) magnetic conditions obtained at Narssarssuaq during 1968–1974.

Perhaps the most consistent studies of long term behavior of scintillations have been made in the auroral zone, at Alaskan longitudes and along the 70°W meridian.

Both the diurnal pattern of scintillation activity and the seasonal behavior as observed from one site can be noted in Figures 10-73 and 10-74. The data used for this long term study [Basu and Aarons, 1980] were taken over a period of 6 years from Narssarssuaq by observing 137 MHz scintillations of the ATS-3 beacon, the propagation path traversed the ionosphere at ~63° CGL.

The long term study used for Figure 10-75 incorporated data from three observatories (Narssarssuaq, Greenland; Goose Bay, Labrador; and Sagamore Hill, Massachusetts). The contours are of reduced data for one season (May–July) for magnetically active periods of time [Basu and Aarons, 1980]. The boundary of active scintillation is pushed equatorwards extending into what was the quiet trough and plasmapause latitudes. Thus during magnetic storms scintillations and

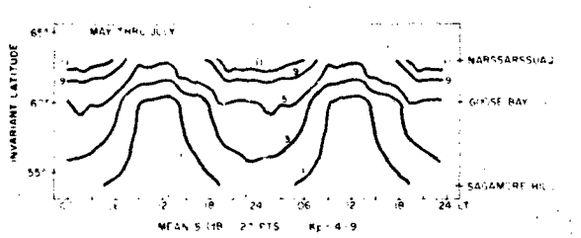


Figure 10-75. Variation of mean scintillation index during the northern solstice in dB at 137 MHz with local time and invariant latitude derived from hourly data at the 3 stations under disturbed magnetic conditions (Kp = 4 – 9).

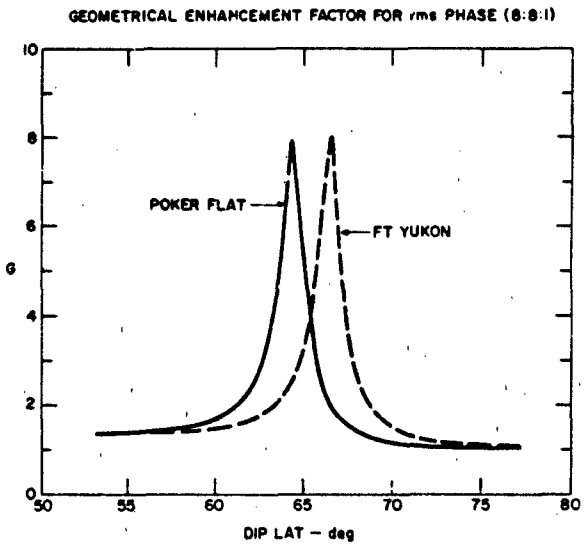


Figure 10-76. Model computations of phase geometrical enhancement factor for sheet-like structures with an 8 : 8 : 1 anisotropy. Because of the meridional pass traiectory, the location of the enhancement is independent of the pass elevation. [Rino and Owen 1980]

optical aurora can be noted farther south than 55°. In the 70°W longitude region this extends below the latitude of Boston.

Geometry and Enhancement. Sheet-like irregularities produce strong enhancements when observations are made in specific directions. For two sites in Alaska, Rino and Owen [1980] have constructed the theoretical geometrical enhancement factor for rms phase fluctuations for an 8:8:1 irregularity (Figure 10-76) [Rino and Matthews, 1980]. They found this enhancement in phase fluctuations as can be seen by the data in Figure 10-77. The amplitude enhancement, less dramatic but present, is also shown in Figure 10-77. Daytime scintillation does not show the sheet-like struc-

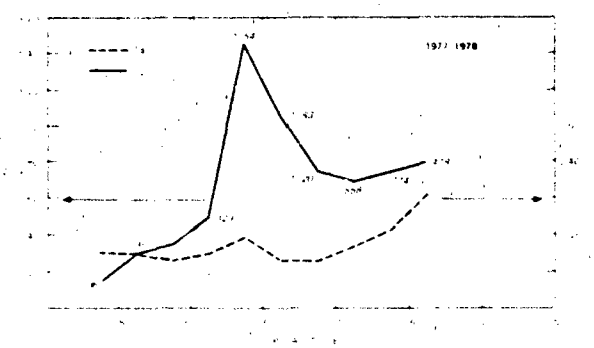


Figure 10-77. RMS phase and S_4 at 50% exceedance level vs magnetic latitude for nighttime data during 1977–1978.

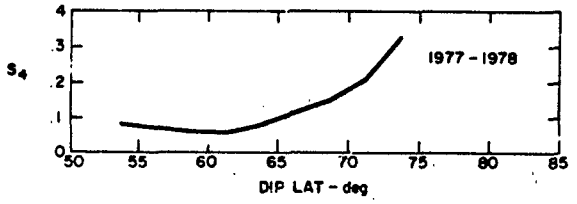


Figure 10-78. S_4 at 50% exceedance level vs magnetic latitude for daytime data during 1977–1978.

ture—at least as observed from Alaska. Figure 10-78 illustrates the daytime increase with increasing latitude.

10.7.10.3 Polar Scintillations. A long term consistent series of measurements has been taken at Thule, Greenland with observations at 250 MHz [Aarons et al., 1981b]. The scintillations for this study ranged from very low values of 3–6 dB peak to peak on occasion during a period of low sunspot number to saturation fading of 28 dB peak to peak for hours during winter months of years of high sunspot number.

One set of measurements was taken between April and October 1975. During this period of low solar activity, there was an absence of strong scintillation activity to such an extent that only the occurrence of scintillation greater than 6 dB could be plotted. Figure 10-79 shows the contrast between the 1975 period when solar flux was low (10.7 cm flux was ~75) and the same months in 1979 when the solar flux was high (~150–225).

A contour plot of the percent occurrence of scintillation index greater than 10 dB is shown in Figure 10-80. The plot was developed from hourly average values of the 15 minute SI for each month for low magnetic activity (Kp = 0-3). Two patterns emerge: (1) Maximum occurrence of activity takes place in the months of little or no sunlight at F

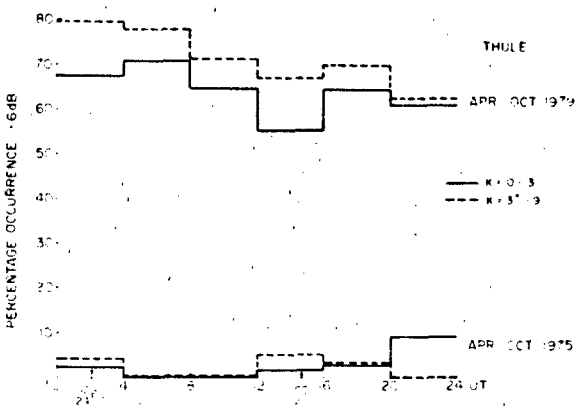


Figure 10-79. Percentage occurrence of scintillation greater than 6 dB for low solar flux period April–October 1975 is contrasted with that for high solar flux period April–October 1979 for both quiet and disturbed magnetic conditions.

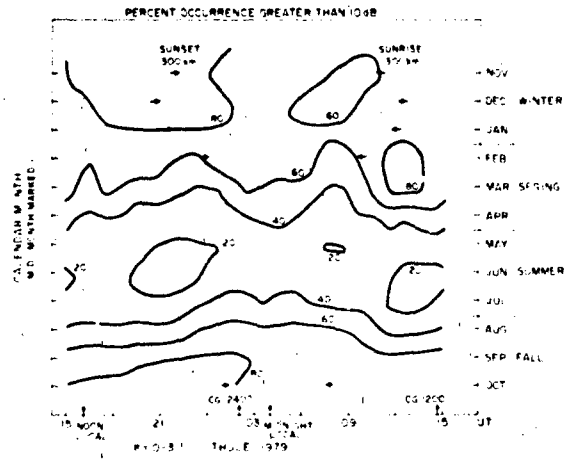


Figure 10-80. Contour plot of diurnal pattern of monthly percent occurrence of scintillation greater than 10 dB for low magnetic activity (Kp = 0 – 3). Observations were taken during Mar 1979–Feb 1980.

region heights. Much lower scintillation occurrence takes place in the sunlit months. (2) The diurnal variation is weak, and apparent only during the winter months.

Auroral arcs in the polar cap are approximately aligned with the noon-midnight magnetic meridian [Davis, 1962]. These arcs generally drift in the dawn to dusk direction [Danielson, 1969]; however, reversals have been noted [Akasofu, 1972; Weber and Buchau, 1981]. Recently Weber and Buchau [1981] described the orientation and motion

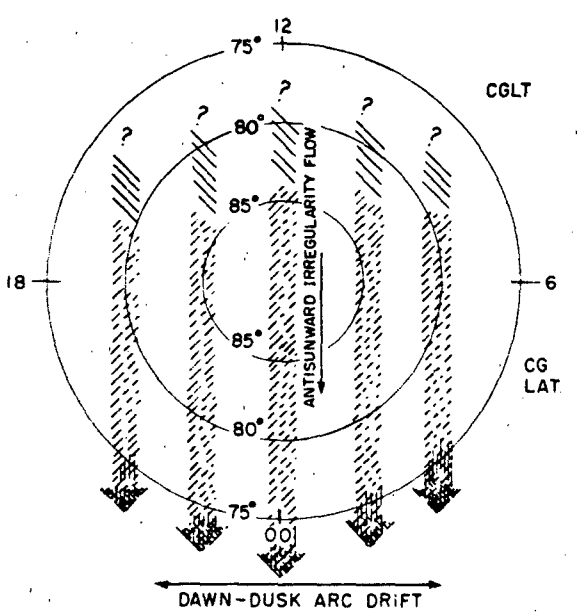


Figure 10-81. Schematic of small scale anti-sunward irregularity drift and the patch motion.

of subvisual F layer (λ = 6300 Å OI) polar cap arcs. Kilometer-size irregularities within the arcs produced intense (saturated) amplitude scintillation at 250 MHz as the arcs drifted through a satellite to ground ray path. Outside the arcs, scintillation frequently persisted at a lower level (SI ~ 6 dB).

A pictorial representation of both the small scale antisunward irregularity drift and the patch motion (predominantly dawn to dusk) is shown in Figure 10-81 (E. Weber, private communication). Results point to two irregularity components in the polar cap; antisunward drifting irregularities which produce a background level of weak to moderate scintillation and intense irregularities within F layer polar cap arcs that produce more discrete (~1 h duration) intense scintillation events as the arcs drift through the ray path.

10.7.11 Empirical Model Of Global Scintillation Behavior

10.7.11.1 WBMOD. Over a period of years, starting from available data and from weak scintillation theory, a model of scintillation termed WBMOD has been developed [Fremouw and Bates, 1971; Fremouw and Lansinger, 1981; Fremouw and Rino, 1978, 1976; Fremouw, 1980; Fremouw et al., 1977]. The program provides for phase and amplitude information. Input user parameters include frequency, location, local time, sunspot number, and planetary magnetic index, Kp. The user also must specify the longest time the system needs phase stability. Scintillation indices are the output. A model of the irregularity drift velocity is contained in the program.

Program WBMOD permits a user to specify his operating scenario. The code returns the spectral index p for power-law phase scintillation, the spectral strength parameter T, the standard deviation σ_ϕ of phase, and the intensity scintillation index, S_4, as functions of a changing independent variable chosen by the user.

The descriptive irregularity model is based on numerous observations [Fremouw and Bates, 1971; Fremouw and Rino, 1978], but most particularly on observations of phase scintillation performed in the DNA Wideband Satellite Experiment [Sagalyn et al., 1974]. The most significant caveat about use of WBMOD, however, is that it has been calibrated quantitatively against Wideband data from only a single station in the northern auroral zone (Poker Flat, Alaska). The descriptive model was developed by iterative comparison with most of the Wideband data population from Poker Flat, with a portion of the population reserved for final comparative tests.

The basic calculations are made of two central quantities T and p. T is the spectral strength of phase at a fluctuation frequency of 1 Hz, p is the power-law spectral index of

phase; T is highly variable, unlike p. The program calculates T and p and the two commonly used indices of scintillation activity based on them, one for phase, σ_ϕ, and one for intensity, S_4.

In order to calculate T, p, σ_ϕ, and S_4, one must have values for eight parameters describing ionospheric irregularities. They are (1) the height, h; (2) vector drift velocity, V_d, of the irregularities; (3) an outer scale, α; (4,5,6,7) four "shape" parameters describing the irregularities' three-dimensional configuration and spatial "sharpness", a, b, δ, and ν; and (8) the height integrated strength of turbulence, C,L. Program WBMOD contains models for the foregoing eight parameters, but the degree of detail is very much less for some than for others.

The most variable and the most important of the eight is the height-integrated strength of turbulence, C,L. The irregularity strength is modeled by

$$\sqrt{C,L} = E(\lambda_m,\lambda_g,T,D,\bar{R}) + M(\lambda_m,T) + H(\lambda_m,T_{in},K_p,\bar{R}) \quad (10.68)$$

where λ_m = geomagnetic invariant latitude,

λ_g = geographic latitude,

T = local meridian time,

D = day of the year,

$\bar{R}$ = smoothed Zurich sunspot number,

T_{in} = geomagnetic time,

Kp = planetary geomagnetic activity index.

The three terms in Equation (10.68) respectively describe the strength of equatorial, midlatitude, and high-latitude irregularities. The first two have not been tested extensively against Wideband data but the high latitude term H has been.

The high-latitude term is based on the observation that there often is a more-or-less abrupt boundary [Aarons et al., 1969] between the midlatitude region of relatively smooth ionosphere and the highlatitude scintillation region. It is located, typically, equatorward of discrete-arc auroras in the general vicinity of the diffuse auroral boundary.

10.7.11.2 Formulas In Atlantic Sector. Since WBMOD has been developed and calibrated against data from only one longitude sector (Alaska), it is appropriate to note empirical formulas that, though not as complex, have been developed for another longitude sector, along the 70°W meridian. These formulations have been made [Aarons et al., 1980a] for Narssarssuaq, Greenland, Goose Bay, Labrador and Sagamore Hill, Massachusetts based on 3–7 years data base of 15-min scintillation indices. The forcing functions are time of day, day of the year, magnetic index and solar flux. However, these individual models are much more limited than WBMOD as (1) they are applicable only for the frequency of the data base, 137 MHz, (2) there is an equipment-biased limited excursion of the scintillations and, (3) these data have an implicit dependence on the geometry of the observations, namely, observing ATS-3 from the stations detailed above. This does not permit other viewing geometries or taking into consideration the configuration of the irregularities unless correcting factors are included.

With these caveats, the equations for each station are

Narssarssuaq (63° CGL intersection)

$$\begin{aligned} SI\ (dB) = {} & -6.4 + 9.2(1 - 0.2FD)\{1 + 0.23(1 - 0.3FD) \\ & \times \cos(HL + 2.0 + 0.34Kp) + 0.03 \\ & \times \cos[2(HL - 0.6)] + 0.02\cos[3(HL \\ & + 3.0)]\}2^{[0.14Kp(1 + 0.12FD) + 0.09As(1 + 1.76FD)]} \end{aligned}$$

$$FD = \cos(DA + 15.6) + 0.56\cos[2(DA - 22.4)] \quad (10.69a)$$

Goose Bay (60° CGL intersection)

$$\begin{aligned} SI(dB) = {} & -1.3 + 1.1(1 - 0.77FD)\{1 + 0.5(1 - 0.2FD) \\ & \times \cos(HL + 2.1 - 0.6Kp) + 0.06 \\ & \times \cos[2(HL - 2.1)] + 0.02\cos[3(HL \\ & + 5.2)]\}2^{[0.3Kp(1 + 0.1FD) + 0.8As(1 + 1.2FD)]} \end{aligned}$$

$$FD = \cos(DA + 0.5) + 0.2\cos[2(DA - 99)] \quad (10.69b)$$

Sagamore Hill (53° CGL intersection)

$$\begin{aligned} SI(dB) = {} & 0.33 + 0.02(1 + 0.2FD)\{1 + 1.2(1 - 0.01FD) \\ & \times \cos(HL - 0.4 - 0.15\,Kp) + 0.3 \\ & \times \cos[2(HL - 0.8)] - 0.1\cos[3(HL \\ & + 6.1)]\}2^{[0.38Kp(1 + 0.3FD) + 3.1As(1 - 0.2FD)]} \end{aligned}$$

$$FD = \cos(DA + 56) + 0.7\cos[2(DA - 143)]. \quad (10.69c)$$

DA is day number, As = S_f/100, HL is local time (hours) at subionospheric point (350 km), and S_f is solar flux at 2695 MHz in solar flux units; all angles are in radians. Arguments of the cosines with diurnal and yearly terms should be converted by factors of $2\pi/24$ and $2\pi/365$, respectively.

In Aarons et al. [1980a] corrections for frequency dependence are given thus allowing higher frequency scintillations to be estimated. In addition, corrections for geometry are also given similar to those cited in Section 10.7.3.1.

CHAPTER 10

10.8 IONOSPHERIC TIME DELAY EFFECTS ON EARTH-SPACE PROPAGATION

One of the most important effects of the ionosphere on radio waves that traverse it is a retardation, or group delay, on the modulation or information carried on the radio wave, due to its encounter with the free, thermal electrons in the earth's ionosphere. Other effects the ionosphere has on radio waves include (1) RF carrier phase advance (2) Doppler shift of the RF carrier of the radio wave (3) Faraday rotation of the plane of polarization of linearly polarized waves (4) angular refraction or bending of the radio wave path as it travels through the ionosphere (5) distortion of the waveform of transmitted pulses, and (6) amplitude and phase scintillation. With the exception of scintillation effects (see Section 10.7), all the other effects listed here are proportional, at least to first order, to the total number of electrons encountered by the wave on its passage through the ionosphere or to their time rate of change. In fact, phase scintillation also is merely the short term, time rate of change of total electron content (TEC) after the longer term variations have been removed.

In this section a short description is given of each ionospheric TEC effect upon radio waves, along with a representative value of the magnitude of each of the these effects under normal ionospheric conditions. This is followed by a discussion of the important characteristics of average ionospheric TEC behavior and the temporal and spatial variability of TEC.

10.8.1 Group Path Delay

The additional time delay, over the free space transit time, of a signal transmitted from above the ionosphere to a user on, or near, the earth's surface is given by

$$\Delta t = \frac{40.3}{cf^2} \times \text{TEC (s)}, \qquad (10.70)$$

where the TEC is the total number of electrons along the path from the transmitter to the receiver, c is the velocity of light in m/sec., and f is the system operating frequency in hertz. The TEC is generally expressed as the number of electrons in a unit cross section column of one square meter area along this path.

A plot of time delay versus system operating frequency for TEC values from 10^{16} to 10^{19} el/m^2 is given in Figure 10-82. These two values represent the extremes of observed TEC in the earth's ionosphere. Note that, at a system operating frequency of 1 GHZ, for example, a TEC of 10^{18}, a value frequently exceeded in many parts of the world, would produce a time delay of 134 ns or 40.2 m of range error. At a system operating frequency of 100 MHz this same TEC values would produce a range error of over 4

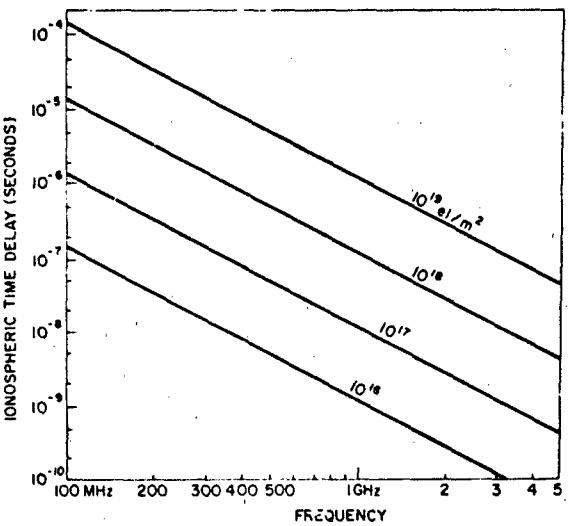


Figure 10-82. Time delay vs frequency for various values of TEC.

km! Obviously, the TEC parameter is of potentially great importance to precision satellite ranging systems.

10.8.1.1 Two-Frequency Ionospheric Time Delay Corrections. If the navigation or ranging system bandwidth is large enough so that two, fairly widely spaced bands can be used for ranging, the ionospheric time delay error can be reduced to an acceptable level automatically and can be made transparent to the system user. Because the ionospheric time delay is a function of frequency we can write:

$$\Delta t_1 = \frac{k}{cf_1^2} \times \text{TEC}\ ,\ \Delta t_2 \frac{k}{cf_2^2} = \times \text{TEC}, \qquad (10.71)$$

where Δt_1 is the ionospheric error on frequency f_1, and Δt_2 is the ionospheric error on the frequency f_2. If the normal system operational frequency is f_1 and we choose f_2 at a lower frequency for ionospheric correction purposes, we obtain:

$$\delta(\Delta t) = \frac{k}{c} \times \text{TEC}\ (1/f_2^2 - 1/f_1^2)$$

$$= \Delta t_1\ (f_1^2 - f_2^2)/\ f_2^2 \qquad (10.72)$$

or:

$$\Delta t_1 = f_2^2/(f_1^2 - f_2^2) \times \delta(\Delta t)$$

The value $\delta(\Delta t)$ is obtained from the difference of the simultaneous measurements of the total range, including ionospheric time delay, at the two frequencies, f_1 and f_2, since the geometric distance is, of course, the same at all frequencies. The quantity $f_2^2/(f_1^2 - f_2^2)$ is called the ionospheric scaling factor. For ratios of f_2/f_1 near unity, the required precision of the differential measurement may be unreasonably large. A plot of this quantity, normalized by f_1, is given in Figure 10-83. In this derivation the contribution of re-

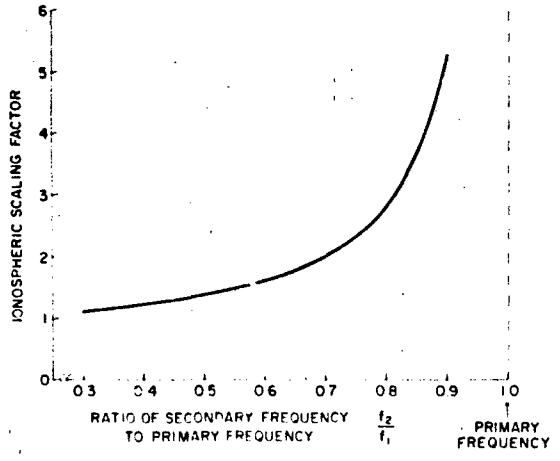


Figure 10-83. Ionospheric scaling factor vs ratio of primary (higher) to secondary (lower) frequency.

ceiver noise to the differential measurement accuracy has not been considered.

10.8.1.2 An Example of a Two-Frequency Ionospheric Time Delay System. The Department of Defense is currently testing an advanced navigation system, called the NAVSTAR-Global Positioning System (GPS), [Demaro, 1981; Milliken and Zaller, 1978] which uses coherently derived, identical modulation on two carrier frequencies, called L1 and L2, to measure the ionospheric group path delay directly and thereby correct for ionospheric time delay. The ratio of frequencies used in the GPS system is exactly 154/120, with the higher frequency (L1) at 1575 MHz. The two carrier frequencies transmitted by the GPS system are the 154th and 120th harmonics of 10.23 MHz. This 10 MHz frequency is bi-phase modulated on both carriers with a psuedo random code resulting in a $[(\sin x)/x]^2$ shaped spectrum of width 20 MHz to the first nulls. A user with knowledge of the transmitted code collapses the received spectrum to equivalent carriers with 10 MHz modulation. The 10 MHz modulation is transmitted with a known phase difference on the two carriers, and the received modulation phase difference is a direct measure of the ionospheric group path delay.

For the GPS carrier and modulation frequencies the ionospheric group path delay at frequency L1, as obtained from Equation 10.73 is

$$\Delta t_1 = -1.5457\ \delta(\Delta t), \qquad (10.73)$$

where $\delta(\Delta t)$ is the difference between the ionospheric time delay measured at the two frequencies. This difference in range is directly related to absolute ionospheric time delay as, of course, the satellite is at the same range at both frequencies. The only frequency dependent parameter in range measurements is the ionospheric time delay effect, assuming the satellite transmitted modulation phase at L1 and L2 is known and the receiving system frequency dispersive characteristics can be independently measured and corrected for.

For a typical daytime high solar activity TEC value of 10^{18} el/m^2 column the $\delta(\Delta t)$ measured by a GPS receiver would be 35 ns or 10.5 m of ionospheric error. For a direct measure of absolute TEC from the modulation phase delay at L2 minus L1 we have

$$\text{TEC} = 2.852 \times 10^{16} \times \delta(\Delta t), \qquad (10.74)$$

where $\delta(\Delta t)$ is measured in nanoseconds (ns). Since, at 10.23 MHz, one complete cycle of modulation phase of 360° is 97.75 ns, we obtain TEC = 9.7745×10^{16} el/m^2 per degree of 10.23 MHz modulation phase difference, or: TEC = 278.8×10^{16} el/m^2 per cycle of modulation phase difference. Thus, the cycle ambiguity in absolute values of TEC is trivial to resolve using the GPS system as a means of determining ionospheric time delay.

Absolute ionospheric time delay measurements can be made with an accuracy approaching 1 to 2 ns, depending upon the received signal to noise ratio on both frequencies. For the power levels transmitted by the GPS satellites, an omnidirectional receiving antenna, and a receiver with a modulation tracking bandwidth of approximately 15 Hz, the differential modulation phase has been measured to within approximately ±2 ns. The contribution of receiver noise for the two-frequency ionospheric time delay corrections on the GPS system has been considered by Cretcher [1975].

10.8.2 RF Carrier Phase Advance

In addition to group path delay, or modulation time delay, over the free space delay, the phase of the carrier of radio frequency transmissions is changed by the ionosphere. The RF phase is advanced with respect to its phase in the absence of an ionosphere. This effect is extremely important in determining space object velocities by means of range rate measurements. The amount of phase increase or phase path decrease can be expressed as

$$\Delta\Phi = \frac{1.34 \times 10^{-7}}{f}\ \text{TEC (cycles)}, \qquad (10.75)$$

where f is the system operating frequency in hertz, and TEC is in el/m^2 column. In practice, the amount of this phase advance cannot readily be measured on a single frequency and two, coherently derived, frequencies are required for this measurement.

10.8.2.1 Differential Carrier Phase. In addition to the dual frequency identical modulation transmitted from the GPS satellites for ionospheric group path correction, these satellites also transmit two, coherently-derived carrier fre-

quencies for ionospheric differential carrier phase measurements. For the pair of frequencies used by GPS, approximately 1.2 and 1.6 GHz, the differential carrier phase shift, referenced to the lower frequency, is

$$\Delta\phi = \frac{1.34}{f_2} \times 10^{-7} \times \frac{(m^2 - 1)}{m^2} \times \text{TEC (cycles)}, \quad (10.76)$$

where $m = f_1/f_2 = 1.2833$, $\Delta\phi = 4.31 \times 10^{-17} \times$ TEC or 2.32×10^{16} el/m^2 per complete 2π cycle of differential carrier phase between L1 and L2, measured at L2. The differential carrier phase [Equation (10.76)] is related to the differential modulation phase, [Equation (10.72)] simply by the ratio of carrier to modulation frequencies. With a reasonable carrier signal to noise ratio, this differential carrier phase can be measured to within a few degrees, or less than approximately 0.04×10^{16} el/m^2. Since the TEC is generally much greater than 2.32×10^{16}, corresponding to 2π of differential carrier phase, there is a $2n\pi$ ambiguity in the differential phase measurement.

The differential carrier method of measuring TEC cannot, in practice, be used to measure absolute values of TEC by itself due to the large $2n\pi$ phase ambiguity in the measurement, but this is not important for navigation systems which require a correction only for range rate errors due to the ionosphere between two measurement times.

The US Navy Navigation Satellite System, NNSS [Black, 1980; Kouba, 1983], determines position for stationary and slowly moving vehicles by measuring satellite transmitted RF carrier phase changes as a function of low-orbit satellite motion across the sky. This method of positioning requires only range rate information. The primary NNSS frequency is 400 MHz. A second RF carrier at 150 MHz is used only for ionospheric range rate corrections. While various techniques have been proposed for determining the absolute TEC from the differential carrier phase information received from the NNSS satellites they all involve assumptions concerning some *a priori* knowledge of the ionosphere, and they cannot be used in the general case.

As an ionospheric monitoring tool the combination of differential carrier phase and differential modulation phase provides an excellent means of determining ionospheric electron content along the ray path to the satellite. The absolute value of TEC can be determined by the group delay technique and relative TEC changes can be measured with great accuracy by the differential carrier phase technique.

10.8.2.2 Second Difference of Carrier Phase. The second difference in phase between an RF carrier and that of its upper and lower sidebands can be used to measure absolute values of TEC, as described by Burns and Fremouw [1970]. If three coherently derived frequencies, $f - f_m$, f, and $f + f_m$ are transmitted the second difference of phase is given by

$$\Delta_2\phi = (\phi_u - \phi_c) - (\phi_c - \phi_L) = \phi_u + \phi_L - 2 \times \phi_c$$

from equation (10.75)

$$\Delta\phi = \frac{1.34}{f} \times 10^{-7} \times \text{TEC (cycles)}$$

thus

$$\Delta_2\phi = \frac{2.68 \times 10^{-7} f_m^2}{f(f^2 - f_m^2)} \text{TEC (cycles)}.$$

When

$$f^2 >> f_m^2 \quad (10.77)$$

$$\Delta_2\phi = \frac{2.68 \times 10^{-7} f_m^2}{f^3} \text{TEC (cycles)}.$$

For a carrier frequency of 100 MHz a modulation frequency of 1.93 MHz would be required to give 2π of second differential phase for a TEC value of 10^{18} el/m.^{2} A value of $\Delta_2\phi$ of 2π for 10^{18} el/m^2 is a reasonable compromise between the requirement for minimizing chances of an ambiguity in absolute TEC and accuracy in measuring TEC relative changes. The second difference of carrier phase has been used with the DNA-002 satellite to make estimates of the absolute value of TEC [Freemouw et al., 1978].

10.8.3 Doppler Shift

Since frequency is simply the time derivative of phase, an additional contribution to geometric Doppler shift results due to changing TEC. This additional frequency shift is generally small compared to the normal geometric Doppler shift, but can be computed by

$$\Delta f = \frac{d\phi}{dt} = \frac{1.34 \times 10^{-7}}{f} \frac{d}{dt} \text{TEC (Hz)}. \quad (10.78)$$

For high orbit satellites where the diurnal changes in TEC are greater than geometric ones, an upper limit to the rate of change of TEC is approximately 0.1×10^{16} el m^{-2}s^{-1}. This value yields an additional frequency shift of less than 0.1 Hz at 1.6 GHz which would not be significant compared with a typical required receiver loop bandwidth of at least a few hertz. At 400 MHz a similiar rate of change of TEC would produce a frequency shift of approximately 0.3 Hz, probably still not significant.

During times of severe phase scintillation, which can occur even at GHz frequencies, the TEC likely does not change in a consistent, rapid manner to yield greater ion-

ospheric Doppler shifts, but the phase of the incoming RF signal can have a large random fluctuation superimposed upon the changes associated with the normal rate of change in TEC. This large, random component may actually spread out the spectrum of the received signal sufficiently to cause the receiver to lose phase lock, as the receiver signal phase may have little energy remaining in the carrier, and instead may be spread over several Hz, with little recognizable carrier remaining. A knowledge of phase scintillation rates is required to determine the spread of received signal phase.

10.8.4 Faraday Polarization Rotation

When a linearly polarized radio wave traverses the ionosphere the wave undergoes rotation of the plane of polarization. At frequencies of approximately 100 MHz and higher the amount of this polarization rotation can be described by:

$$\Omega = \frac{2.36 \times 10^{-5}}{f^2} \int B \cos \theta N dl. \qquad (10.79a)$$

where the quantity inside the integral is the product of electron density times the longitudinal component of the earth's magnetic field, integrated along the radio wave path. Many ionospheric workers have used this effect, named for Michael Faraday who first observed polarization changes in an optical experiment, to make measurements of the TEC of the ionosphere. Since the longitudinal magnetic field intensity changes much slower with height than the electron density of the ionosphere, the equation can be rewritten as

$$\Omega = \frac{K}{f^2} B_L \times TEC, \qquad (10.79b)$$

where $B_L = \overline{B \cos \theta}$ is taken at a mean ionospheric height, usually near 400 km. $K = 2.36 \times 10^{-5}$ and TEC is $\int Ndl$. Typical values of polarization rotation for northern midlatitude stations viewing a geostationary satellite near their station meridian are given in Figure 10-84 as a function of system frequency and total electron content. In fact, the largest portion of TEC data available today from stations throughout the world have come from Faraday rotation measurements from geostationary satellite VHF signals of opportunity.

For satellite navigation and communication designers, however, the Faraday polarization rotation effect is a nuisance. If a linearly polarized wave is transmitted from a satellite to an observer on or near the surface of the earth, the amount of polarization rotation may be nearly an odd integral multiple of 90 degrees, thereby giving no signal on the receiver's linearly polarized antenna, unless the user is careful to realign his antenna polarization for maximum received signal.

As shown in Figure 10.84 at 4 GHz, a commercial

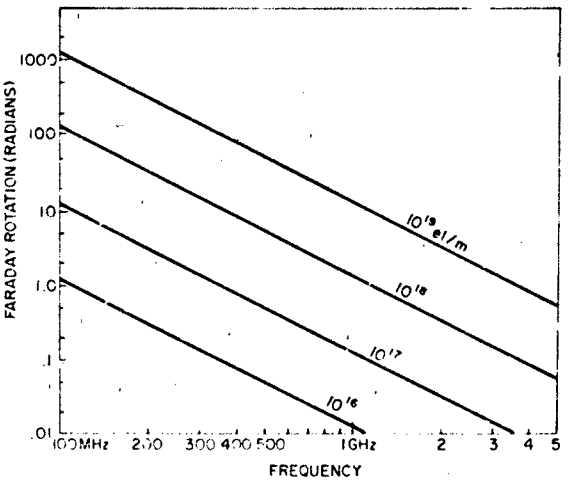


Figure 10-84. Faraday polarization rotation vs frequency for various values of TEC.

satellite transponder frequency band, the amount of Faraday rotation can be a tenth of a radian, well in excess of that required for dual, linear orthogonal channel separation.

The Faraday rotation problem is overcome by the use of circular polarization of the correct sense at both the satellite and at the user's receiver. Generally the mobile user finds it difficult to utilize circular polarization due to the continual vehicle directional changes; thus he settles for a received linear polarization. The 3 dB loss between transmitted circular polarization and received linear polarization is a necessary price to pay for user antenna maneuverability and simplicity.

10.8.5 Angular Refraction

The refractive index of the earth's ionosphere is responsible for the bending of radio waves from a straight line geometric path between satellite and ground. This angular refraction or bending produces an apparent higher elevation angle than the geometric elevation. Millman and Reinsmith [1974] have derived expressions relating the refraction to the resultant angular bending. Perhaps the easiest expressions to use, as given by Millman and Reinsmith [1974] relate the ionospheric range error to angular refraction:

$$\Delta E = \frac{R + r_0 \sin E_0 \, (r_0 \cos E_0)}{h_i \, (2r_0 + h_i) + r_0^2 \sin E_0} \times \frac{\Delta R}{R}, \qquad (10.80)$$

where E_0 is the apparent elevation angle, R is the apparent range, ΔR is computed from $\Delta R = (40.3/f^2) \times TEC$, r_0 is the earth's radius, and h_i is the height of the centroid of the TEC distribution, generally between 300 and 400 km.

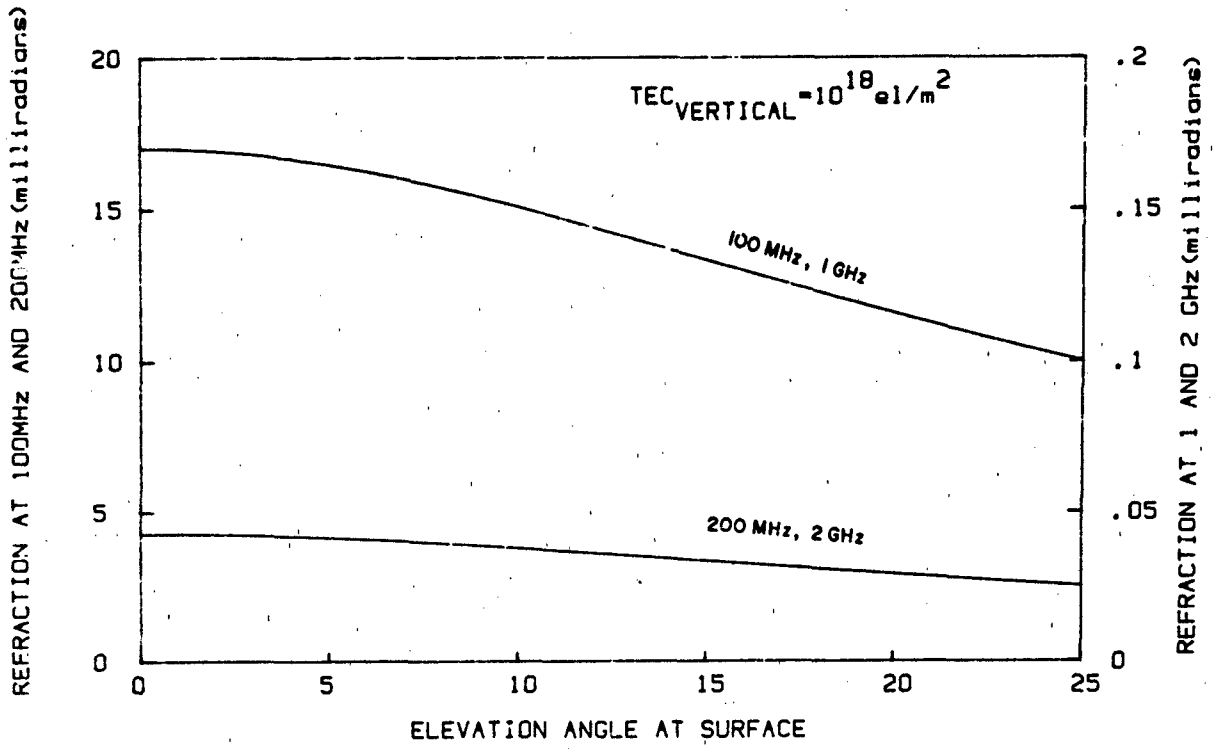


Figure 10-85. Refraction in elevation angle vs elevation angle for indicated frequencies and values of TEC.

For low elevation angles and satellites well above most of the ionization, $R \gg r_o \sin^2 \varepsilon_o$, and the angular refraction can be expressed as:

$$\Delta E = \frac{\cos E_o}{2h_i} \Delta R. \qquad (10.81)$$

Typical values of elevation refraction error for a TEC of 10^{18} el/m^2 column are shown in Figure 10-85 for several frequencies. Note that, at the lowest frequency, 100 MHz, near the horizon the refraction is well over 1.5 degrees! The curves shown in Figure 10.85 have been constructed using the approximation derived by Millman and Reinsmith [1974] for low elevation angles given in Equation (10.81).

Generally, the range error itself is the main ionospheric problem for advanced navigation systems, and elevation angle errors are insignificant. Satellite detection radar systems, on the other hand, do have the requirement to know accurate pointing elevation angles for their large aperture arrays, though generally the accurate tracking is done by using range rate information, and elevation angle is of secondary importance as long as the beamwidth of the antenna is large enough to see the target.

Errors in the azimuth of radio waves transmitted through the ionosphere can also occur; they depend upon azimuthal gradients in TEC which are generally small and which can usually be neglected in practical cases.

10.8.6 Distortion of Pulse Waveforms

Two characteristics of the ionosphere can produce distortion of pulses of RF energy propagated through it. The dispersion, or differential time delay due to the normal ionosphere, as derived by Millman [1965] is proportional to $1/f^2$, and produces a difference in pulse arrival time across a bandwidth Δf of

$$\Delta t = \frac{80.6 \times 10^6}{c\, f^3} \Delta f \times TEC, \qquad (10.82)$$

where c is the velocity of light in m/s, f and Δf are expressed in Hertz, and TEC is in el/m^2 column. The dispersive term for pulse distortion is thus proportional to TEC. When the difference in group delay time across the bandwidth of the pulse is the same magnitude as the width of the pulse it will be significantly disturbed by the ionosphere. Millman and Bell [1971] also derived mathematical relationships for ionospheric dispersive effects on an FM Gaussian shaped pulse.

In addition to pulse distortion by the dispersive effects due to the TEC of the normal background ionosphere, radio pulses are also modified by scattering from ionospheric irregularities. Yeh and Liu [1979] have computed pulse mean arrival time and mean pulsewidth due to both dispersion and scattering.

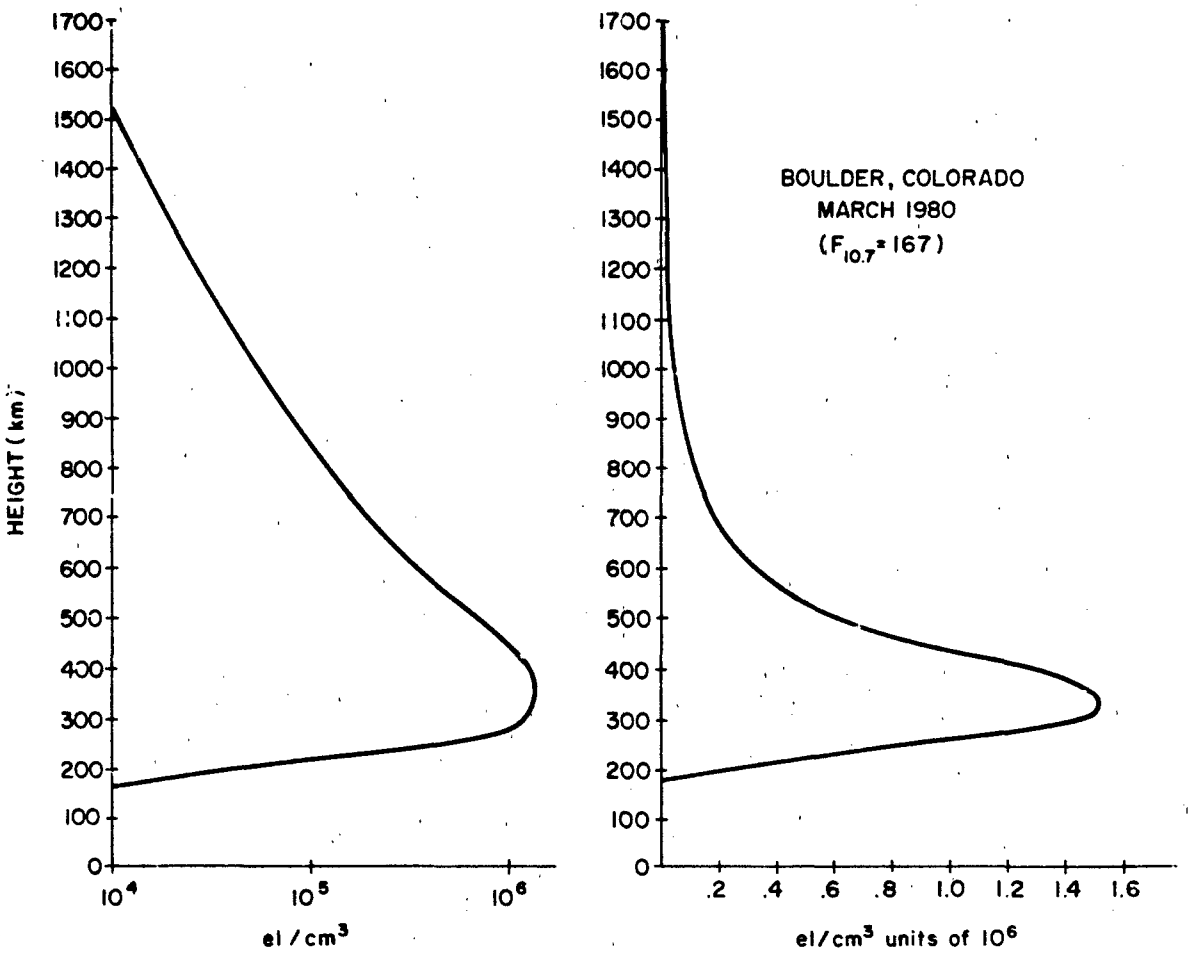


Figure 10-86. Typical profile of electron density vs height. In (a) log N_e is plotted; in (b) N_e is plotted on a linear scale.

10.9 IONOSPHERIC TOTAL ELECTRON CONTENT (TEC)

10.9.1 Average TEC Behavior

The ionospheric parameter responsible for the effects described in section 10.8 is the total number of free electrons, TEC, or its rate of change, along the path from a satellite to a ground station. The greatest contribution to TEC comes from the F2 region of the ionosphere. A typical daytime midlatitude, high solar maximum electron density profile is illustrated in Figure 10-86. The curve on the left side of Figure 10-86 is the log of N_e plotted versus height as normally shown by ionospheric workers. Since the TEC is represented by the area under the curve of a linear plot of N_e versus height, the right hand plot of Figure 10-86 illustrates the actual linear plot. Note that most of the contribution to TEC occurs near the peak of the F2 region. The reason for making this point is as follows: ground-based ionosondes have been used since the 1930s to make continuous, routine measurements of the density at the peak of the F2 region, measured by ionosondes as foF2, and equated to N_{max} by

$$(foF2)^2 = 80.6\ N_e \qquad (10.83)$$

where foF2 is in MHz, and N is in units of 10^6 el/cc.

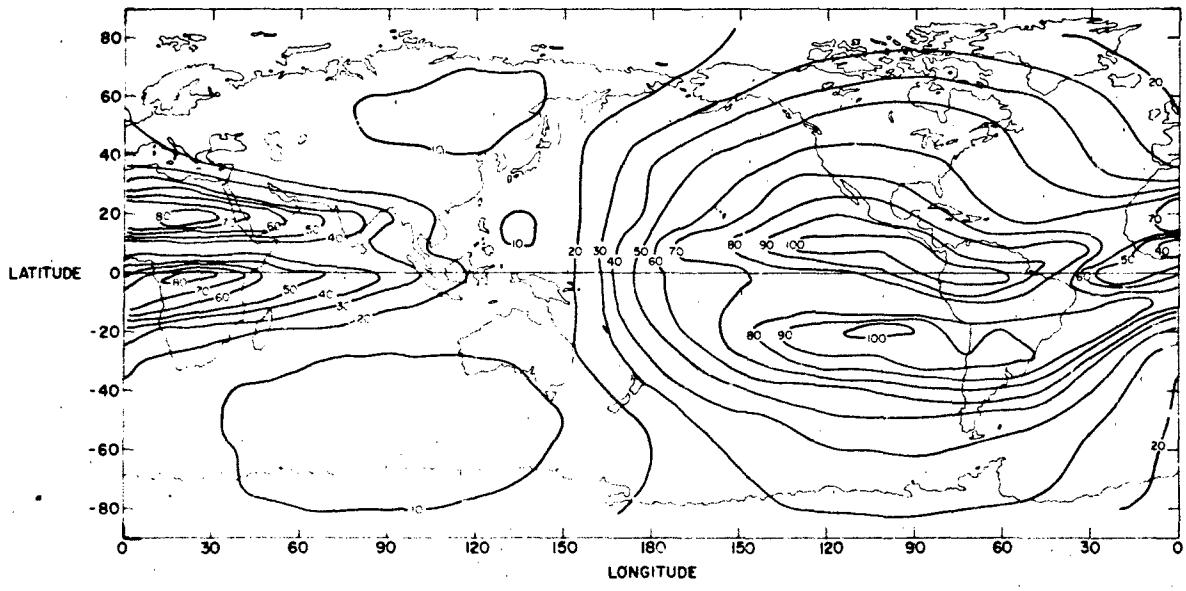
In the 1950s and 1960s, continuing to a more limited extent even today, upwards of 150 ionosondes were operated to provide improved prediction capability for long distance high frequency propagation by means of ionospheric refraction. Various models of foF2 were developed for this purpose, one of the more popular ones being commonly known as ITS-78 [Barghausen et al., 1969] after the report number which described the model. This model, among other things, characterized the 10 day average worldwide behavior of foF2 by Fourier temporal components and Legendre polynomial geographic coefficients ordered by magnetic, rather than geographic, latitude. The success of this

experimental, data based, or empirical model, in representing the actual worldwide foF2 is due to the large amount of data available from ionosondes in many regions of the world. Other characteristics of this model are discussed by Dandekar [1982] and in Section 10.3 of this chapter.

For the TEC parameter, data availability have been, and will likely continue to be, much more sparse. First, TEC measurements have generally been calculated from measurements of Faraday polarization rotation using VHF signals of opportunity transmitted from geostationary satellite telemetry transmitters. A few lunar reflected Faraday rotation measurements in the late 1950s and early 1960s and the TEC obtained from a few low orbit satellites did not contribute significantly to our knowledge of world-wide TEC behavior, at least not for modeling average ionospheric conditions. Only since the early to mid-1960s have TEC values been obtained on a more-or-less regular basis. Even today fewer than one dozen stations regularly contribute TEC data, which can be used in TEC modeling purposes, to a world data center.

Fortunately most of the contribution to TEC comes from near the F2 region density peak where models of foF2 are available. These foF2 models can be combined with some limited knowledge of topside ionospheric thickness obtained from topside sounders, along with topside *in situ* density measurements, to produce a complete ionospheric height profile model. The most well known of these models is the one by Bent [Llewellyn and Bent, 1973] which uses ITS-78 coefficients for foF2, and topside exponentials for computing TEC. A representation of world-wide average behavior of TEC is illustrated in Figure 10-87 for 2000 hours UT. To first order the TEC contours shown in Figure 10-87 move westward along magnetic, rather than geographic, latitude lines, at the earth's rotation rate. The Bent model was constructed using solar maximum data from the 1968-1969 period and had to be adjusted upward somewhat to account for the much higher 1979-1980 solar maximum than that of 1968-1969. This adjustment was necessary to adequately represent the actual TEC values from stations making observations in March 1980, which was near the maximum of the second highest solar cycle ever recorded in the more than 200 year history of solar cycle observations. The Bent model, appropriately adjusted for high solar cycle values, does however, represent fairly well the average behavior of TEC for many locations tested. Other worldwide ionospheric electron density profile models from which average TEC can be obtained include ones by Ching and Chiu [1973], and Chiu [1975], Kohnlein [1978], the 4-D model [VonFlotow, 1978], and the International Reference Ionosphere (IRI)[Rawer, 1981]. The characteristics of some of these models are described in Dandekar [1982] and in Section 10.3

Other empirical models of TEC have been developed directly from TEC data alone, though these have necessarily been limited in temporal and geographic extent by the available data base. These include models of TEC over Europe and the Mediterranean [Klobuchar, 1973] for low and medium activity portions of the 11 year solar cycle, and a



CONTOURS OF TOTAL ELECTRON CONTENT UNITS OF 10^{16} el / m^2 COLUMN
MARCH 1980 ($F_{10.7}$ = 167)
2000 UT

Figure 10-87. Contours of vertical TEC in units of 10^{16} el/m^2 column for 2000 UT, March 1980

model of TEC over the Indian subcontinent for both solar minimum and for an average solar maximum [Klobuchar, et al.1977]. Models of the slab thickness parameter, the ratio of TEC/N_{max} have been developed for specific regions such as the one for northern Europe by Kersley [1980], and one for the eastern USA by Klobuchar and Allen [1970], from which TEC can be obtained from a model of foF2. An algorithm designed for an approximate 50% correction to world-wide TEC, for use in an advanced navigation system, has been developed by Klobuchar [1975].

All of the models listed here, and the list is by no means complete, are empirical models that attempt to correct for average TEC behavior only. However, the variability from average TEC behavior can be large and may be important to some radio wave systems that must propagate through the ionosphere.

10.9.2 Temporal Variability of TEC

10.9.2.1 Variability from Monthly Mean TEC Values. The ionosphere is a weakly ionized plasma and the resultant TEC is a function of many variables including solar ionizing radiation, neutral wind and electric field effects, neutral composition, and temperature changes. A monthly overplot of curves of diurnal changes in TEC for a northern mid-latitude station for twelve months during a solar maximum period is shown in Figure 10-88. The standard deviation from monthly mean diurnal behavior is approximately 20%-25%, during the daytime hours when the absolute TEC values are greatest. Figure 10-89 shows the standard deviation from monthly average TEC behavior for the mid-day hours for a number of stations during the solar maximum period 1968-1969. Again 20%-25% is a good value for the standard deviation from the monthly average behavior. The standard deviation is somewhat higher during the nighttime hours, but the absolute TEC values are much lower during these periods.

If a satellite ranging system has error requirements such that it must correct for monthly average ionospheric time delay, but still can tolerate the approximate 20%-25% variability of TEC from monthly average conditions, approximately 70%-80% of the ionospheric effect on the system can be eliminated by the use of an average TEC model such as the one constructed by Bent [Llewellyn and Bent, 1973]. If the system only requires an approximate 50% rms correction of the ionospheric time delay, the algorithm developed by Klobuchar [1975] can be used. On the other hand, if corrections for some portion of the remainder of the ionospheric time delay are required, after a state of the art TEC

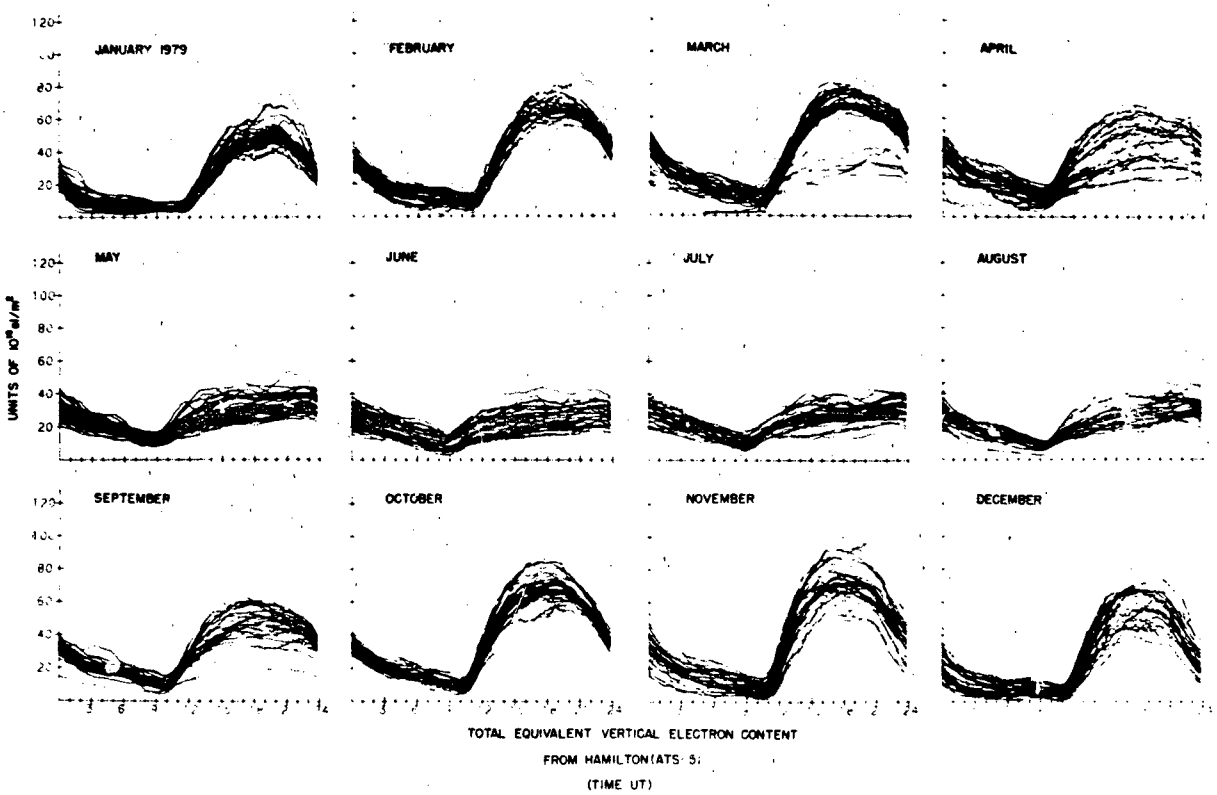


Figure 10-88 Monthly overplots of TEC diurnal curves for Hamilton, Mass. for 1979.

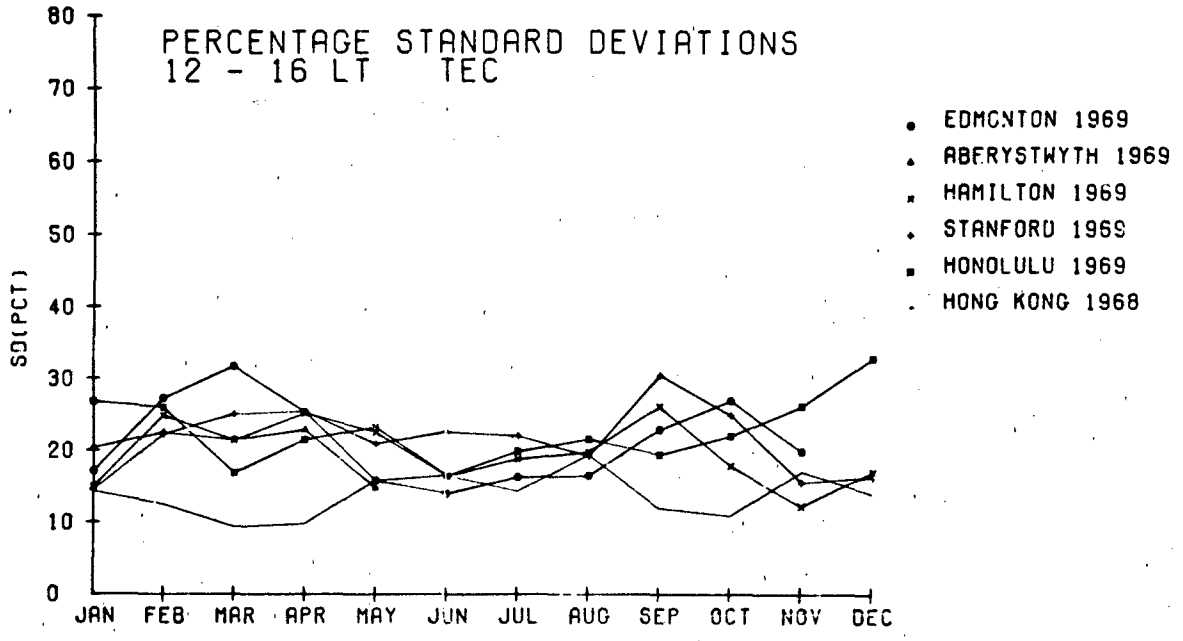


Figure 10-89. Percentage standard deviations for daytime TEC from the stations indicated.

model, such as the Bent one, has been used to take out the monthly mean TEC, then the short term (a few hours) temporal variability as well as the geographic variability, of TEC must be considered.

10.9.2.2 Short Term Temporal Variability of TEC. The correlation time of departures of TEC from monthly average curves has been studied by Donatelli and Allen [1981]. They concluded, for the midlatitude station they studied, that the useful prediction time was a function of local time, season, and long term sunspot activity. However, in most cases they gained no significant improvement over the use of monthly mean predictions when they used actual data more than 3 hours old. The longest useful prediction interval occurred, fortunately, during solar maximum daytime hours when absolute TEC values are highest. During solar minimum periods their useful prediction time interval was often as short as one hour.

In their study, Donatelli and Allen [1981] predicted TEC data for the same geographic location and direction in the sky as their measurements. If the prediction is for a different location, the temporal correlation will be lower.

10.9.2.3 Geographic Variability of TEC. The variability of TEC at the same local time, but as a function of distance has been studied by Klobuchar and Johanson [1977]. They utilized TEC data from two sets of stations, one aligned approximately along an east-west direction, with the other set of stations aligned along an approximate north-south direction. Their results are shown in Figures 10-90a and b for the east-west and the north-south station alignments, respectively. No significant difference in correlation distance was found with season.

The percent improvement, P.I., in TEC from the average value is related to the correlation coefficient r by

$$\text{P.I.} = 100 \times [1-(1-r^2)^{0.5}]$$

(see Gautier and Zacharisen, [1965]) (10.84)

Note that a correlation coefficient of 0.7 explains only 29% of the variance between the data at station pairs; hence a measurement at one station, with a correlation coefficient of 0.7 between data sets with a second station would result in an improvement at the second station over the average predicted value of only 29%.

10.9.3 TEC in the Near-Equatorial Region

All of the preceding sections have concentrated on the behavior of TEC in the midlatitude regions of the world, mainly because most of the available data are from that region. The near-equatorial region deserves special mention due to the fact that the highest TEC values in the world occur in this region, as shown in Figure 10-87. This region extends to approximately $\pm 20°$-$25°$ either side of the magnetic equator, with the highest TEC values not at the equator, but rather at the so called "equatorial anomaly" regions located at approximately $\pm 15°$ from the magnetic equator.

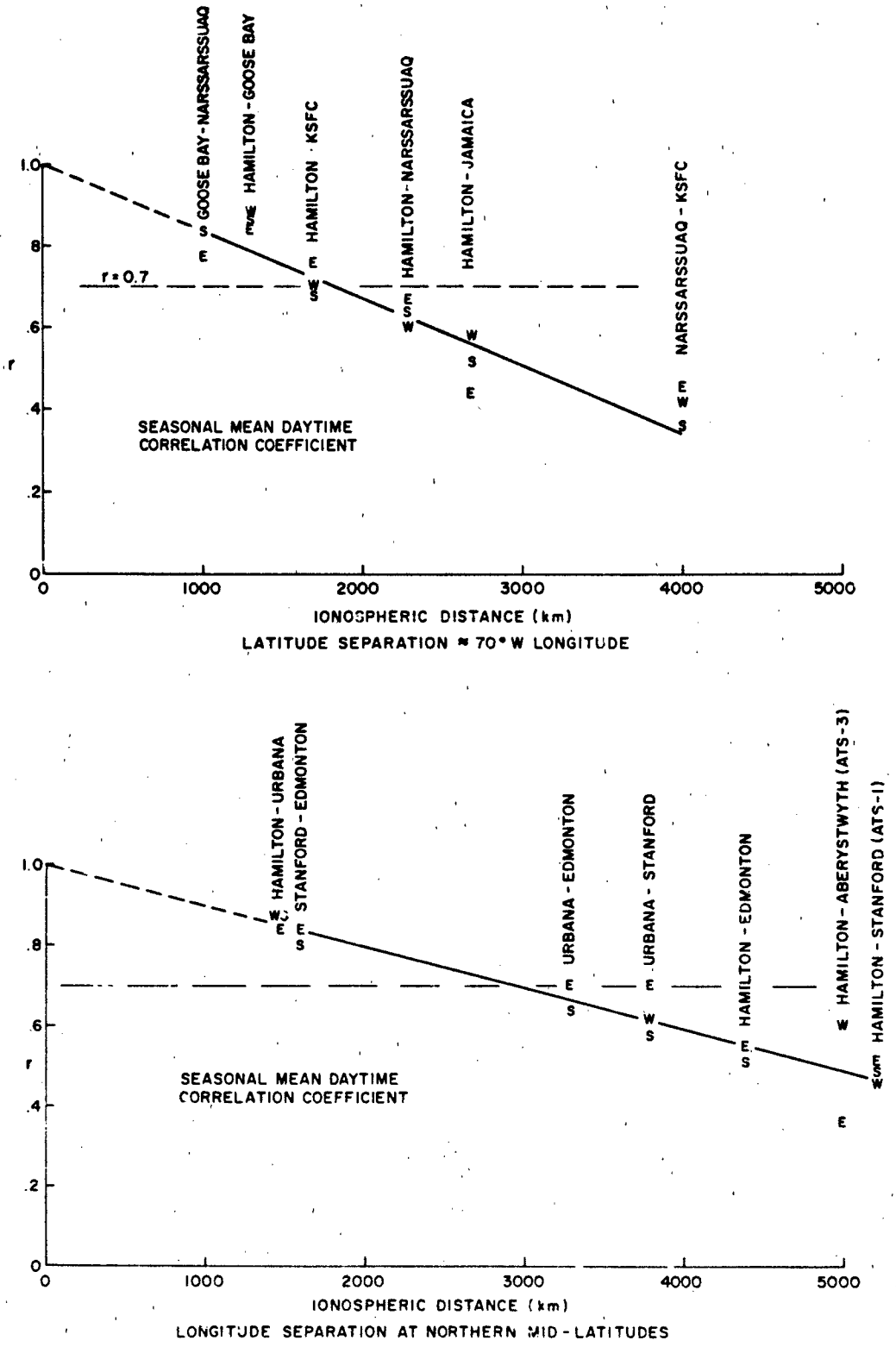


Figure 10-90. Correlation coefficient vs station separation in (a) latitude and (b) longitude.

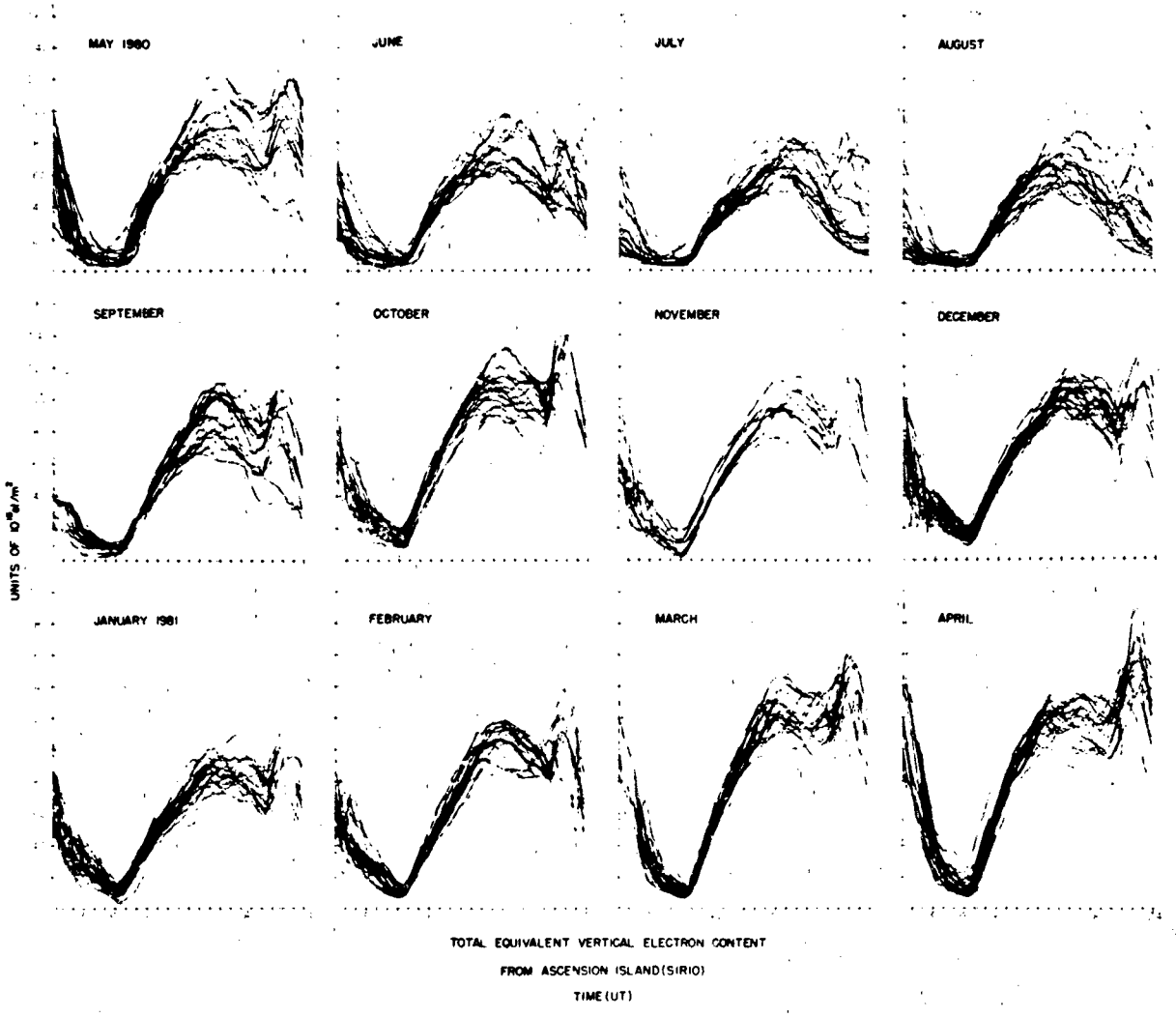


Figure 10-91. Monthly overplots of TEC diurnal curves for Ascension Island, May 1980–April 1981.

The regions of highest TEC values at 2000 hours UT are clearly seen in Figure 10-87 near 100° west longitude.

Most of the day-to-day geographic variability of TEC in the equatorial anomaly region during solar minimum conditions can be explained by the variability of equatorial electrojet strength. Unfortunately, no such similar TEC data are available for solar maximum.

An example of the high temporal variability of TEC for solar maximum conditions for Ascension Island, a station located near the peak of the southern TEC equatorial crest region, is shown in Figure 10-91. Note the extremely large day-to-day TEC variability in the afternoon and evening hours in some months. Any satellite ranging system requiring ionospheric TEC corrections in the near-equatorial region should not use the midlatitude standard deviation values of approximately 20%-25% to represent the variability of the near-equatorial region.

10.9.4 TEC in the Auroral and Polar Cap Regions

Since most available TEC values have been measured using radio signals transmitted from geostationary satellites, which can be viewed only at low elevation angles from high latitudes, knowledge of the variability of TEC in the auroral and polar cap regions is sparse. In the American longitude sector, where the magnetic latitudes are lowest for a given geographic latitude, there is considerable TEC data from Goose Bay, Labrador, over which the aurora passes southward, even during moderately magnetically disturbed periods. The behavior of TEC during those periods can be highly irregular, especially during the nighttime hours. TEC values often exhibit rapid changes and occasionally even exceed the daytime maximum values briefly. While the occurrence of general auroral activity may be predictable, the

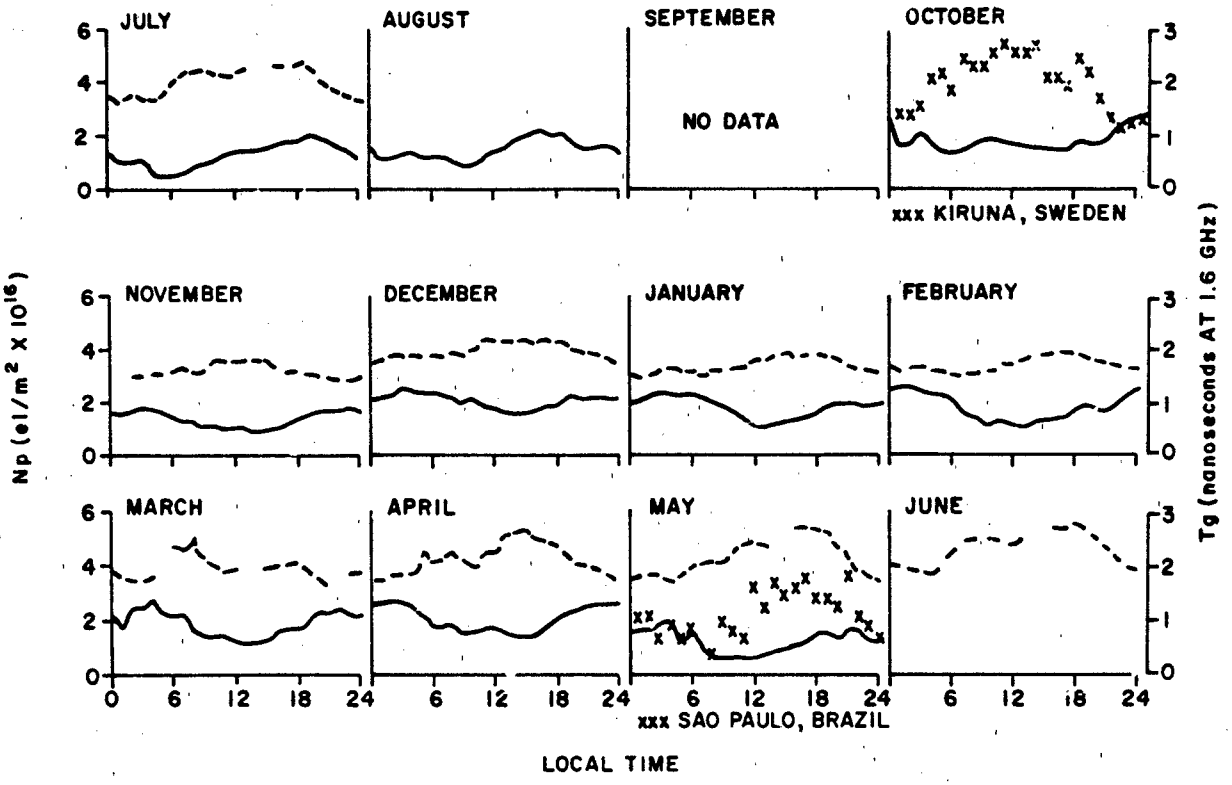


Figure 10-92. Monthly average plasmaspheric electron content vs. local time for Aberystwyth, Wales (dashed line) and for Hamilton, Mass. (solid line). Also plotted are values from Kiruna, Sweden for October 1975 and from Sao Paulo, Brazil for May 1975.

specific large increases in TEC, likely due to auroral precipitation, are not individually predictable, but may be statistically characterized as a function of magnetic activity.

In the polar cap region a negligible amount of TEC data exists. The absolute TEC values are probably lower in this region than in the midlatitudes, and the variability of the polar cap TEC is probably very high.

10.9.5 Protonospheric Electron Content

Most of the available TEC data has been taken by measurements of Faraday rotation of single frequency radio waves transmitted from geostationary satellites to ground observers. The electron content obtained from Faraday rotation observations, while made from radio waves transmitted from satellites at geostationary satellite height above the earth's surface, only includes the contribution of electrons up to heights of approximately 2000 km. This is because the integrated product of the longitudinal component of the earth's magnetic field times the electron density, above approximately 2000 km, is negligible. The only measurements of the additional contribution of electrons above the Faraday maximum height have been made using signals from an ionospheric beacon on the geostationary satellite ATS-6. Davies [1980] has reviewed the overall results of the ATS-6 experiment. A summary of typical protonospheric electron content data is shown in Figure 10-92 taken from Klobuchar et. al. [1978]. Note that the protonospheric values are fairly low in absolute value.

During the nighttime hours when the ionospheric TEC is low, the protonospheric contribution may become a fairly large percentage of the total number of electrons between a satellite at geostationary height and an observer on, or near the earth's surface. Unfortunately, no protonospheric electron content data are available during solar maximum conditions.

10.9.6 Short Term Variations in TEC

The time rate of change of TEC, in addition to the normal diurnal variations, also has periodic variations due to perturbations of the ionospheric F region from various potential sources as geomagnetic substorms, meteoriogical sources such as weather fronts, shock waves from supersonic aircraft, volcanic explosions, rocket launches, and other miscellaneous sources. While these short term variations in TEC cover a large range of periods and amplitudes, common periods range from 20 to over 100 minutes with amplitudes

of a few percent of the background TEC. A 10% ionospheric disturbance with respect to the background TEC is uncommon, while a 1% TEC perturbation is common. Titheridge [1968] and Yeh [1972] have made studies of the statistics of traveling ionospheric disturbances (TIDs), in TEC for midlatitude regions.

A system that requires correction for the rate of change of TEC cannot rely on models of TEC to provide reliable information on short term rate of change of TEC information, and can use available TID information only in a statistical manner. The only recourse for a system significantly affected by rate of change of TEC is to use a dual frequency measurement technique to directly measure the ionospheric contribution to range rate.

10.9.7 Conclusions

There are at least three categories of systems potentially affected by ionospheric time delay. For the first category of user the potential systems effects may be small, at least under any naturally occurring worst case ionospheric conditions. In the second category, a user may require a nominal correction for average ionospheric time delay, but is able to tolerate the 20%-25% standard deviation from average conditions. He should expect at least a 50% correction for ionospheric time delay effects using a relatively simple time delay algorithm, and up to 70%-80% for a state of the art, fairly complex model. These model corrections can be improved by the use of actual ionospheric measurements within a reasonable temporal and spatial frame. For the third category of user ionospheric model corrections, even updated with near-real-time measurements, may not be sufficient to correct for ionospheric time delay, and the system must then make its own ionospheric correction. Fortunately, the ionosphere is a dispersive medium and the use of identical modulation on two, widely-spaced frequencies will allow a direct measurement to be made of ionospheric range delay. Two coherently-derived carrier frequencies may be used to obtain accurate time rate of change information for TEC. Details of measuring ionospheric effects directly by a system's use of multiple frequencies are available in Burns and Fremouw [1970].

10.10 ARTIFICIAL MODIFICATION

The field of ionospheric modifications is a subset of a more general class of research today called "Active Experiments" in space plasmas. This field was initiated early in the space program by using rocketborne chemical releases as tracers and/or modifiers of upper atmospheric processes. The physical basis for such experiments was reviewed in some detail by Haerendel [1976] during the first international meeting devoted entirely to artificial modification studies [Albrecht, 1976]. A second major symposium on Active Experiments [Russell and Rycroft, 1980] dealt with a broad spectrum of experiments: energetic particle injections, plasma wave (VLF) injections, mass (neutral gas) injections, as well as with laboratory and computer simulation experiments. The most recent summary of Active Experiments in space treated particle beams, neutral gas injections, wave injections and high power heating experiments [Burke, 1983]. The common thread that binds all of these methods is the use of well-defined input/output experiments to probe the system response functions for specific atmospheric and space plasma systems.

In terms of purely ionospheric phenomena, the modification of ambient electrons and ions are most often achieved by chemical injections or by radiowave heating experiments. Each of these areas is treated in the following sections.

10.10.1 Chemical Releases

The history of chemical release experiments dates from the the earliest days of space exploration when, shortly after Sputnik-1 in 1957, rocket-borne payloads of highly reactive chemicals were injected into the upper atmosphere in attempts to use artificial perturbation techniques as a way of investigating the structure and dynamics of the neutral and ionized components of the upper atmosphere. The Air Force Cambridge Research Laboratories (ARCRL) carried out the initial work with *plasma cloud injections* [Marmo et al., 1959], and later a pioneering and comprehensive series of chemical injection experiments under PROJECT FIREFLY [Rosenberg, 1964]. Experiments using barium releases (or similar, easily ionized species) have formed the major activity in this field, tracing and/or modifying ionospheric processes from auroral locations [Holmgren et al., 1980] to the equator [Kelly et al. 1979]. The symposium proceedings referenced above [Albrecht, 1976; Russell and Rycroft, 1980; Burke, 1983] offer comprehensive summaries of these experiments.

The field of *neutral mass injections* was first concerned with the environmental impacts that might result from the larger and more powerful rockets being developed for space exploration [Kellogg, 1964]. In 1973, when the last Saturn V rocket to be used in the U.S. Space Program launched NASA's Skylab Workshop, the resultant deposition of approximately 1000 kg/s of H_2 and H_2O exhaust molecules into the 200-440 km altitude region initiated a rapid and large-scale depletion of the ionosphere to an extent never seen before (see Figure 10-93). The artificially-created "ionospheric hole" amounted to nearly a 50% decrease in the total electron content (TEC) of the ionosphere over an area of approximately a million square kilometers. Mendillo et al. [1975 a,b] attributed the effect to the explosive expansion of an exhaust cloud of highly reactive molecules that initiated a rapid recombination of the ionospheric plasma.

The introduction of such typical rocket exhaust products as H_2, H_2O and CO_2 into the upper atmosphere causes the

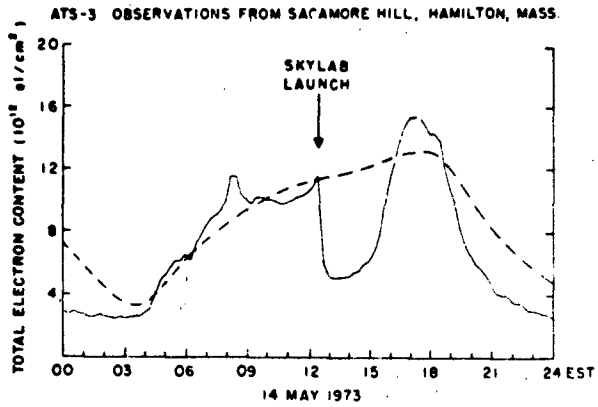


Figure 10-93. Total Electron Content (TEC) data used to detect the "SKYLAB effect" on 14 May 1973. The dashed curve gives the anticipated diurnal TEC behavior based upon a monthly median prediction updated for geomagnetic storm effects [Mendillo et al., 1975b].

atomic ion F region plasma to be transformed to a molecular ion plasma at rates 100 to 1000 times faster than occur with the naturally present molecules of nitrogen (N_2) and oxygen (O_2).

These important reactions are

$$O^+ + H_2O \xrightarrow{\kappa_1} H_2O^+ + O \qquad \kappa_1 = 2.4 \times 10^{-9}\ cm^3/s \tag{10.85}$$

$$O^+ + H_2 \xrightarrow{\kappa_2} OH^+ + H \qquad \kappa_2 = 2.0 \times 10^{-9}\ cm^3/s \tag{10.86}$$

$$O^+ + CO_2 \xrightarrow{\kappa_3} O_2^+ + CO \qquad \kappa_3 = 1.2 \times 10^{-9}\ cm^3/s \tag{10.87}$$

Once a molecular ion is formed, its dissociative recombination with an ambient electron occurs rapidly.

$$H_2O^+ + e \xrightarrow{\alpha_1} OH + H \qquad \alpha_1 = 3.0 \times 10^{-7}\ cm^3/s \tag{10.88}$$

$$OH^+ + e \xrightarrow{\alpha_1} H + O \qquad \alpha_1 = 1.0 \times 10^{-7}\ cm^3/s \tag{10.89}$$

$$O_2^+ + e \xrightarrow{\alpha_1} O + O \qquad \alpha_1 = 2.0 \times 10^{-7}\ cm^3/s \tag{10.90}$$

and hence an "ionospheric hole" is formed. A review of rocket induced ionospheric disturbances has been given by Mendillo [1981].

Figure 10-94 contains a schematic showing the many physical and chemical processes associated with artificially-induced depletions in the F region. It should be noted that the lack of large-scale/long-lived modification effects upon the lower regions of the ionosphere is due primarily to the high neutral densities and molecular ion chemistry already dominant at D and E region heights, as discussed in detail by Forbes [1980].

Computer simulation models for the F region effects have been constructed by Bernhardt et al. [1975], Mendillo and Forbes [1978], Anderson and Bernhardt [1978] and Zinn and Sutherland [1980]. The emphasis in these studies has ranged from environmental impacts of proposed in-space construction scenarios [Rote, 1979], to laboratory-in-space experiments using "dedicated engine-burns" of the space shuttle as part of the Spacelab-2 mission in 1985, to a series of chemical modification experiments planned for the Combined Release and Radiation Effects Satellite (CRRES) scheduled for the late 1980's. Some of these concepts have been tested using rocketborne chemical payloads during projects LAGOPEDO [Pongratz and Smith, 1978], WATERHOLE [Whalen et al., 1981], BIME [Narcisi, 1983] and COLOURED BUBBLES [Haerendel et al., 1983]. In the AFGL Ionospheric Modification Study [Narcisi, 1983], attempts were made to study effects associated with SF_6 induced negative ion plasmas [Mendillo and Forbes, 1978]. During so-called "experiments of opportunity," where scheduled rocket launches are monitored by a variety of techniques, satellite radio beacon observations have been reported by Mendillo, et al. [1980] incoherent scatter measurements by Wand and Mendillo [1984], and optical diagnostics by Kofsky [1981] and Mendillo and Baumgardner [1982]. Figure 10-95 offers an example of the artificial airglow clouds associated with F region hole-making experiments.

10.10.2 High Power HF Transmissions

Ground based high power high frequency transmitters operating below the critical frequency of the ionosphere have been used to artificially modify the ionospheric electron thermal budget and plasma characteristics [Utlaut, 1970; Gordon et al., 1971; Silyger, 1974; for comprehensive reviews, see Carlson and Duncan, 1977 and Gurevich and Fejer, 1979]. The power aperture product of these high power transmitters have been typically of the order of 10^4 Mwm^2 providing power densities of about 10-100 μwm^{-2} at ionospheric heights. The ionospheric "modification" or so-called "heating" experiments have been observed to cause not only the initially intended enhancements of electron gas temperature with associated plasma redistribution but give rise to a variety of nonlinear plasma phenomena. Figure 10-96 [after Carlson and Duncan, 1977] summarizes, in a schematic form, the striking variety of observed effects of ionospheric heating. The enhancements of electron gas temperature have been observed to be a few hundred degrees K [Gordon et al., 1971] caused by the deviative

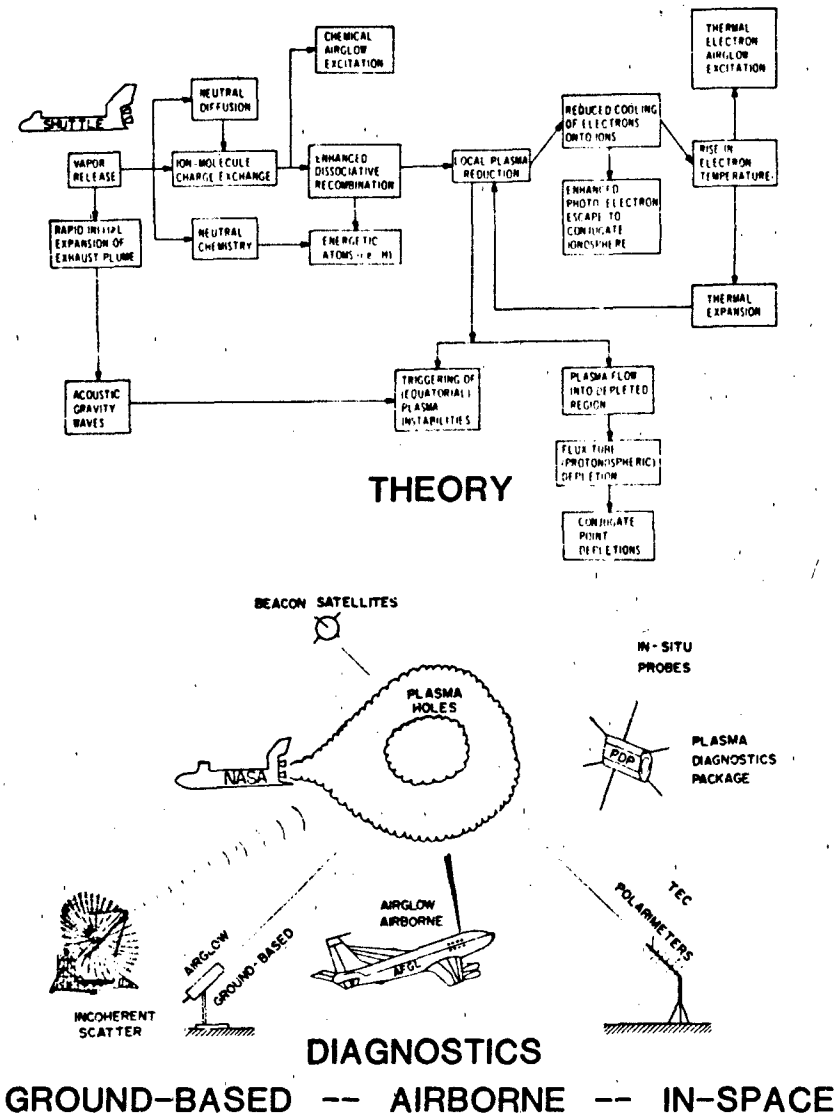


Figure 10-94. Schematic summary of possible rocket effluent effects upon the upper atmosphere (h > 200 km), associated with the NASA Spacelab-2 mission scheduled for Spring-Summer 1985.

absorption of the heater wave near the altitude of HF reflection. The following manifestations of plasma instabilities have been observed: (1) artificially created spread F; (2) strongly enhanced radio wave absorption; (3) the creation of field-aligned density irregularities which scatter (this phenomenon is sometimes called field-aligned scatter of FAS) an incident HF, VHF, or UHF wave with virtually no frequency change and make certain types of scatter communication circuits possible; (4) scattering process in which the frequency of the scattered wave differs from the frequency of the incident wave by roughly the frequency of the high-power HF transmissions; and (5) strongly enhanced airglow at 6300 Å; some enhancement at 5577 Å is also observed (see the Special Issue of Radio Science, [1974]).

The short-wavelength (1 cm - 10 m) field-aligned irregularities produced by ionospheric heating are a result of parametric decay instability or wave interaction between the high power radio wave (pump) and the ion-acoustic and Langmuir waves. This was predicted from theory [Perkins and Kaw, 1971; Perkins et al., 1974; DuBois and Goldman, 1972] and experimentally confirmed at Arecibo [Carlson et al., 1972].

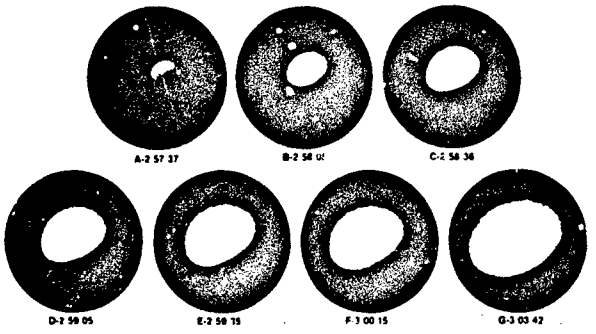


Figure 10-95. The growth of an ionospheric hole is shown in this sequence of image-intensified, wide-angle photographs of the expanding 6300 Å airglow cloud produced by excited oxygen atoms created from the recombination of free electrons and molecular ions (O_2^+, OH^+, H_2O^+) produced by exhaust molecules (CO_2, H_2, H_2O) and ambient atomic ions (O^+). Times are a.m., PST [Mendillo and Baumgardner, 1982].

The long-wavelength (~1 km) field-aligned irregularities giving rise to artificial spread F [Utlaut et al., 1970; Utlaut and Violette, 1972; Wright, 1973] could not, however, be explained in terms of the above instability process. The causative mechanism for the generation of long wavelength irregularities remained obscure for quite a while and is now attributed to either a thermal self-focusing mechanism [Perkins and Valeo, 1974; Thome and Perkins, 1974] or the alternative mechanisms of stimulated Brillouin scattering [Cragin and Fejer, 1974] and stimulated diffusion scattering [Goldman, 1974].

The above range of irregularity scale sizes has sufficient power spectral intensity to cause scintillation of radio signals received from radio stars and artificial satellites. This was demonstrated when VHF/UHF signals transmitted through the artificially heated ionospheric F region were found to exhibit scintillations [Rufenach, 1973; Pope and Fritz, 1974; Bowhill, 1974]. Radio star scintillation measurements at 26 MHz during ionospheric modification indicated the presence of either rapid and random or deep long-period (~5 mins) fluctuations. In order to avoid some of the difficulties of radio star observations, Bowhill [1974] performed scintillation measurements with both geostationary and orbiting satellites and established the field-aligned nature of the irregularities causing VHF and UHF scintillations, their transverse scale and drift speed. One feature common to all the above studies was the fact that the heater frequency was below the plasma frequency of the F region. The magnitude of scintillations observed on transionospheric communication channels is found to be of the order of 5 dB at 250 MHz when the nighttime F region is heated by an incident power density of about 50 μwm^{-2} [Basu et al., 1980a].

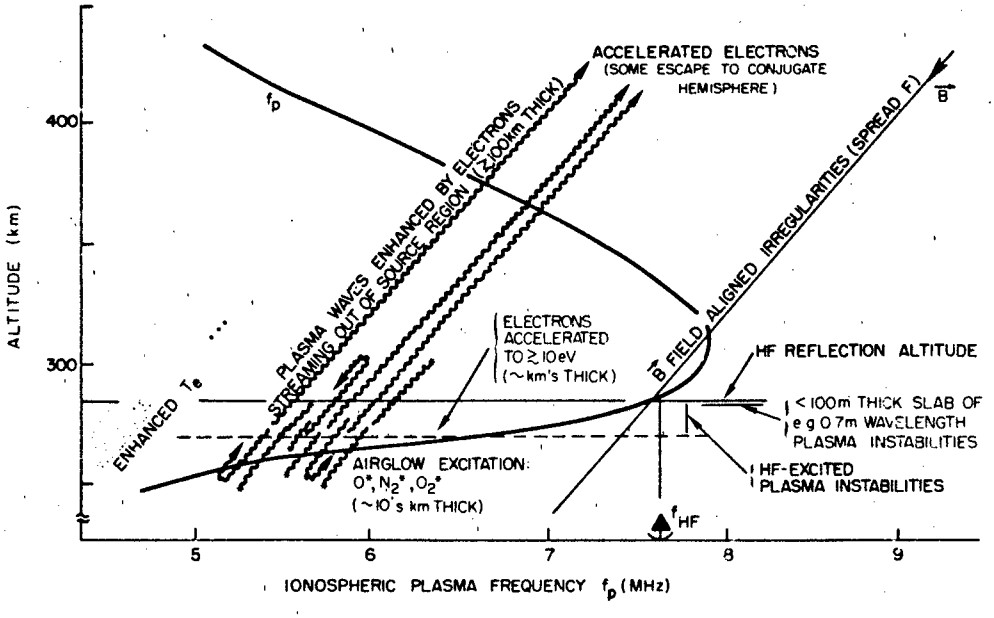


Figure 10-96. Effects produced by ground-based transmitter of power aperture of the order of 10^4 Mwm^2 in the 4–12 MHz frequency range. Energy deposited in the ionospheric plasma alters both the thermal and nonthermal properties of its charged particle population. Controlled experiments have applications to aeronomy, chemical rates, atomic cross sections, communications, and a number of areas of plasma physics [Carlson and Duncan, 1977].

CHAPTER 10

REFERENCES

Aarons, J., "A Descriptive Model of F-Layer High-Latitude Irregularities as Shown by Scintillation Observations," *J. Geophys. Res.*, **78:** 7441–7450, 1973.

Aarons, J., "Equatorial Scintillations. A Review," *IEEE Trans. Antennas Propagat.*, **AP-25,** 729, 1977.

Aarons, J., "Global Morphology of Ionospheric Scintillations," *Proc. IEEE*, **70:** 360–378, April 1982.

Aarons, J. and R.S. Allen, "Scintillation Boundaries During Quiet and Disturbed Magnetic Conditions," *J. Geophys. Res.*, **76:** 170, 1971.

Aarons, J. and E. Martin, "The Effects of the August 1972 Magnetic Storms on Ionospheric Scintillations," *Radio Sci.*, **10:** 547–554, 1975.

Aarons, J., J.P. Mullen, and H.E. Whitney, "The Scintillation Boundary," *J. Geophys. Res.*, **74:** 884–889, 1969.

Aarons, J., E. MacKenzie, and K. Bhavnani, "High-Latitude Analytical Formulas for Scintillation Levels," *Radio Sci.*, **15:** 115–127, 1980a.

Aarons, J., J.P. Mullen, J.P. Koster, R.F. daSilva, J.R. Medeiros, R.T. Medeiros, A. Bushby, J. Pantoja, J. Lanat, and M.R. Paulson, "Seasonal and Geomagnetic Control of Equatorial Scintillations in Two Longitudinal Sectors," *J. Atmos. Terr. Phys.*, **42:** 861–866, 1980b.

Aarons, J., E. MacKenzie, and S. Basu, "Microwave Equatorial Scintillation Intensity During Solar Maximum," *Radio Sci.*, **16:** 939–945, 1981a.

Aarons, J., J.P. Mullen, H. Whitney, A. Johnson, and E. Weber, "VHF Scintillation Activity Over Polar Latitudes," *Geophys. Res., Lett.*, **8:** 277–280, 1981b.

AGARD, "Modern HF Communications," *AGARD Lecture Series* **127,** (avail. through NTIS, Springfield, Va) 1983.

Akasofu, S.-I., "Midday Auroras and Polar Cap Auroras," *Geofysiske Publikasjoner* (Norske Videnskaps-Akademi i Oslo), **29:** 73–85, 1972.

Albrecht, H.J., (ed.), Artificial Modification of Propagation Media. *AGARD Conf. Proc.* **192,** Brussels, 26–29 April 1976.

Al'pert, Ia. L., "Ionospheric Propagation of Radio Waves," Acad. of Sci. of USSR, Moscow, Section on Long Waves (avail) as Translation T5-60, National Bureau of Standards, Boulder, 1960.

Ames, J.W. and R.D. Egan, "Digital Recording and Short-Term Prediction of Oblique Ionospheric Propagation," *IEEE Trans. Antennas Propag.* **AP-15:** 382, 1967.

Ames, J.W., R.D. Egan and G.F. McGintie, "Short-Term Prediction of HF Communication Circuit Performance" in *Ionospheric Forecasting, AGARD Conf. Proc.*, **49,** AD700896, 1970.

Anderson, D.N., "Modelling the Ambient Low Latitude F-Region Ionosphere—A Review," *J. Atmos. Terr. Phys.*, **43:** 753, 1981.

Anderson, D.N. and P.A. Bernhardt, "Modeling the Effects of an H Gas Release the Equatorial Ionosphere," *J. Geophys. Res.*, **83:** 4777, 1978.

Appleton, E.V. and W.J.G. Beynon, "The Application of Ionospheric Data to Radio-Communication Problems: Part I," *Proc. Phys. Soc.*, **52:** 518, 1940.

Baggaley, W.J., "Backscatter Observations of F-Region Field-Aligned Irregularities During the I.Q.S.Y.," *J. Geophys. Res.*, **75:** 152–158, 1970.

Bain, W.C., "Models of the Mid Latitude D Region at Noon" in *Medium, Long and Very Long Wave Propagation, AGARD Conf. Proc.*, **305:** 34-1–34-8 ADA113960, 1982.

Balsley, B.B. and W.L. Ecklund, "VHF Power Spectra of the Radar Aurora," *J. Geophys. Res.*, **77:** 4746–4760, 1972.

Balsley, B.B., W.L. Ecklund, and R.A. Greenwald, "VHF Doppler Spectra of Radar Echoes Associated with a Visual Auroral Form: Observations and Implications," *J. Geophys. Res.*, **78:** 1681–1687, 1973.

Bannister, P.R., "Variations in Extremely Low Frequency Propagation Parameters," *J. Atmos. Terr. Phys.*, **37:** 1203–1210, 1975.

Bannister, P.R., "Extremely Low Frequency (ELF) Propagation," NUSC Scientific and Engineering Studies, Naval Underwater Systems Center, New London, Conn., February 1980.

Bannister, P.R., "Overview of ELF Propagation," in *Medium, Long, and Very Long Wave Propagation, AGARD Conf. Proc.*, **305:** 10-1–10-13 ADA113969, 1982.

Barghausen, A.L., J.W. Finney, L.L. Proctor, and L.D. Schultz, "Predicting Long-term Operational Parameters of High-Frequency Sky-Wave Telecommunication Systems," ESSA Tech. Rept. ERL 110-ITS 78, 1969.

Barr, R., "The Effect of Sporadic E on the Nocturnal Propagation of ELF Radio Waves," *J. Atmos. Terr. Phys.*, **39:** 1379–1387, 1977.

Barry, G.H., "A Low Power Vertical Incidence Ionosonde" *IEEE Trans. Geosci. Elect.*, **GE-9(2):** 86–89, 1971.

Basu, S. and Su. Basu, "Modeling of Equatorial Phase and Amplitude Scintillations from OGO-6 and AE Irregularity Data, Low Latitude Aeronomical Processes," *COSPAR Symposium Series*, **8:** 187–199, Pergamon Press 1980.

Basu, S. and H.E. Whitney, "The Temporal Structure of Intensity Scintillation Near the Magnetic Equator," *Radio Sci.*, **18:** 263–271, 1983.

Basu, S., Su. Basu, A.L. Johnson, J.A. Klobuchar, and C.M. Rush, "Preliminary Results of Scintillation Measurements Associated with Ionosphere Heating and Possible Implications for the Solar Power Satellite," *Geophys. Res. Lett.*, **7:** 609, 1980a

Basu, S., Su. Basu, J.P. Mullen, and A. Bushby, "Long-term 1.5 GHz Amplitude Scintillation Measurements at the Magnetic Equator," *Geophys. Res. Lett.*, **7:** 259–262, 1980b.

Basu, S., J.P. McClure, Su. Basu, W.B. Hanson, and J. Aarons, "Coordinated Study of Equatorial Scintillation and in-situ and Radar Observations of Nighttime F Region Irregularities," *J. Geophys. Res.*, **85:** 5119–5130, 1980c.

Basu, Su. and J. Aarons, "The Morphology of High-Latitude VHF Scintillation Near 70°W," *Radio Sci.*, **15:** 59–70, 1980.

Basu, Su. and M.C. Kelley, "A Review of Recent Observations of Equatorial Scintillations and their Relationship to Current Theories of F-Region Irregularity Generation," *Radio Sci.*, **14:** 471, 1979.

Basu, Su., R.L. Vesprini, and J. Aarons, "Field-Aligned Ionospheric E-Region Irregularities and Sporadic E," *Radio Sci.*, **8:** 235–246, 1973.

Basu, Su., R.L. Vesprini, and J. Aarons, "F-Layer Irregularities as Determined by Backscatter Studies at 19 MHz over Half of a Solar Cycle," *Radio Sci.*, **9:** 355–371, 1974.

Basu, Su., S. Basu, R.C. Livingston, E. MacKenzie, and H.E. Whitney, "Phase and Amplitude Scintillation Statistics at 244 MHz from Goose Bay using a Geostationary Satellite," AFGL TR-82-0222, ADA124291, 1982.

Bates, H.F. and P.R. Albee, "Aspect Sensitivity of HF Auroral Echoes," *J. Geophys. Res.*, **74:** 1164–1168, 1969.

Bates, H.F. and P.R. Albee, "Aspect Sensitivity of F-Layer HF Backscatter Echoes," *J. Geophys. Res.*, **75:** 165–170, 1970.

Belrose, J.S., "Radio Wave Probing of the Ionosphere by the Partial Reflection of Radio Waves (From Heights below 100 km)," *J. Atmos. Terr. Phys.*, **32:** 567–596, 1970.

Belrose, J.S., "COSPAR Symposium Record on D- and E-Region Ion Chemistry," *Aeron. Rept. No. 48*, Univ. of Illinois, Urbana, Ill., 1972.

Belrose, J.S., "LF Propagation: An Overview" in *Medium, Long, & Very Long Wave Propagation, AGARD Conf. Proc.*, **305:** 22-1–22-10 ADA113969, 1982.

Belrose, J.S. and L. Thomas, "Ionization Changes in the Middle Latitude D-Region Associated with Geomagnetic Storms," *J. Geophys. Res.*, **30:** 1397–1413, 1968.

Bennett, G.M. and A.B. Friedland, "Prediction of Daily Fluctuations of the F-Region Plasma Frequency," *Ionospheric Forecasting, AGARD Conf. Proc.*, **49** AD700896, 1970.

Bent, R.B., S.K. Llewellyn, and M.K. Walloch, "Description and Evaluation of the Bent Ionospheric Model," **I,** SAMSO-TR-72-239, 1972.

Bernhardt, P.A., C.G. Park, and P.M. Banks, "Depletion of the F Region Ionosphere and Protonosphere by the Release of Molecular Hydrogen," *Geophys. Res. Let.*, **2:** 341, 1975.

Bernstein, S.L., M.L. Burrows, J.E. Evans, A.S. Griffiths, D.A. McNeill, C.W. Neissen, I. Richer, D.P. White, and D.K. Wilson, "Long Range Communication at Extremely Low Frequencies," *Proc. IEEE*, **62(3),** 292–312, 1974.

Berry, L.A., "Wave Hop Theory of Long Distance Propagation of LF Radio Waves," *Radio Sci.*, **68D:** 1275–1282, 1964.

Berry, L.A. and M.E. Chrisman, "The Path Integrals of LF/VLF Wave Hop Theory," *Radio Sci.*, **64D:** 1469–1480, 1965.

Betts, J.A. and M. Darnell, "Real Time HF Channel Estimation by Phase Measurements on Low-Level Pilot Tones" in "Radar Systems and the Ionosphere," *AGARD Conf. Proc.*, **173** ADA022439, 1975.

Bibl, K., "Fixed and Variable Frequency Pulse Transmission at Oblique Ionospheric Incidence," AFCRL 64392 AD603045, 1962.

Bibl, K. and B.W. Reinisch, "Digisonde 128 PS," University of Lowell, Center for Atmospheric Research, Lowell, Mass., 1978a.

Bibl, K. and B.W. Reinisch, "The Universal Digital Ionosondes," *Radio Sci.*, **13:** 519–530, 1978b.

Bickel, J.E., J.A. Ferguson, and G.V. Stanley, "Experimental Observations of Magnetic Field Effects on VLF Propagation at Night," *Radio Sci.*, **5:** 19–25, 1970.

Bilitza, D., M. Noor Sheikh, and R. Eyfrig, "A Global Model for the Height of the F_2 Peak using M3000 Values from the CCIR Numerical Map," *Telecomm. J.*, **46:** 549–553, 1971.

Black, H.D. "The Transit System, 1977; Performance, Plans and Potential," *Philos. Trans. R. Soc. Lond.* **A294:** 217–236, 1980.

Booker, H.G., C.M. Crain, and E.C. Field, "Transmission of Electromagnetic Waves Through Normal and Disturbed Ionospheres," RAND Rept. R-558-PR, RAND Corp., Santa Monica, Calif., 1970.

Booker, H.G. and G. MajidiAhi, "Theory of Refractive Scattering in Scintillation Phenomena," *J. Atmos. Terr. Phys.*, **43:** 1199, 1981.

Bowhill, S.A., "Satellite Transmission Studies of Spread-F Produced by Artificial Heating of the Ionosphere," *Radio Sci.*, **9:** 975, 1974.

Bowles, K.L., "Observations of Vertical Incidence Scatter from the Ionosphere at 41 Mc/sec," *Phys. Rev. Lett.*, **1:**454, 1958.

Bowles, K.L. and Staff of Jicamarca Radio Observatory Institute, "Profiles of Electron Density Over the Magnetic Equator Obtained Using the Incoherent Scatter Technique," *Tech. Rept. Note 169*, National Bureau of Standards, Boulder, Colo., 1963.

Bracewell, R.N. and W.C. Bain, "An Explanation of Radio Propagation at 16 kc/sec in Terms of Two Layers Below E Layer," *J. Atmos. Terr. Phys.*, **2:** 216–225, 1952.

Bracewell, R.N., K.G. Budden, J.A. Ratcliff, T.W. Straker, and K. Weekes, "The Ionospheric Propagation of Low and Very Low Frequency Radio Waves Over Distances Less than 1000 km," *Proc. IEE*, **98:** 221, 1951.

Bradley, P.A. and C. Bedford, "Prediction of HF Circuit Availability," *Elec. Lett.*, **12:** 32, 1976.

Bradley, P.A. and J.R. Dudeney, "A Simple Model of the Vertical Distribution of Electron Concentration in the Ionosphere," *J. Atmos. Terr. Phys.*, **35:** 2131, 1973.

Bramley, E.N., "Fluctuations in Direction and Amplitude of 136 MHz Signals from a Geostationary Satellite," *J. Atmos. Terr. Phys.*, **36:** 1503–1513, 1974.

Breit, G. and M.A. Tuve, "A Test of the Existence of the Conducting Layer," *Phys. Rev.*, **28:** 554–575, 1926.

Briggs, B.H. and I.A. Parkin, "On the Variation of Radio Star and Satellite Scintillation with Zenith Angle," *J. Atmos. Terr. Phys.*, **25:** 339–365, 1963.

Brooks, D., "Observations of Radio Auroral Echoes and

Simultaneous Magnetic Disturbances," *J. Atmos. Terr. Phys.*, **27:** 1151–1158, 1965.

Brooks, D., "Observations of Auroral Echoes Following Worldwide Storm Sudden Commencements," *J. Atmos. Terr. Phys.*, **28:** 103–109, 1966.

Buchau, J., W.N. Hall, B.W. Reinisch, and S. Smith, "Remote ionospheric Monitoring from Effects of the Ionosphere on Space and Terrestrial Systems," *Proc. IES* (Ionospheric Effects Symposium), edited by John H. Goodman, U.S. Government Printing Office, Washington, D.C., 1978.

Buchau, J., B.W. Reinisch, E.J. Weber, and J.G. Moore, "Structure and Dynamics of the Winter Polar Cap F Region," *Radio Sci.*, **19:** 995–1010, 1983.

Budden, K.B., *Radio Waves in the Ionosphere*, Cambridge University Press, Cambridge, UK, 1961.

Buneman, O., "Scattering of Radiation by the Fluctuations in a Non-Equilibrium Plasma," *J. Geophys. Res.*, **67:** 2050, 1962.

Burke, W. (ed.), *Proceedings of International Symposium on Active Experiments in Space*, 23–29 May 1983, Alpbach, Austria, European Space Agency Report ESA-SP-195, Paris, 1983.

Burns, A.A. and E.J. Fremouw, "A Real-Time Correction Technique for Transionospheric Ranging Error," *IEEE Trans. Antennas Propag.*, **AP-18:**(6), Nov. 1970.

Burrows, M.L., *ELF Communications Antennas*, Peter Peregrinus Ltd., Stevenage, England, 1978.

Carlson, H.C. and L.M. Duncan, "HF Excited Instabilities in Space Plasmas," *Radio Sci.*, **12:** 1061, 1977.

Carlson, H.C., W.E. Gordon, and R.L. Showen, "High Frequency Induced Enhancements of the Incoherent Scatter Spectrum at Arecibo," *J. Geophys. Res.*, **77:** 1242, 1972.

CCIR (Comite Cosultatif International des Radiocommunications = International Radio Consultative Committee), "World Distribution and Characteristics of Atmospheric Radio Noise," Report 332, Doc. Xth Plenary Ass., International Telecommunications Union (ITU), Geneva, 1963.

CCIR, "CCIR Atlas of Ionospheric Characteristics," Report 340 and 340-2, Doc. XIth Plenary Ass., ITU, Geneva, 1966.

CCIR, "Interim Method for Investigating Sky-Wave Field Strength and Transmission Loss at Frequencies Between the Approximate Limits of 2 and 30 MHz," Report 252-2, Doc. XIIth Plenary Ass., ITU, Geneva, 1970.

CCIR, "Second CCIR Computer-Based Interim Method for Estimating Sky-Wave Field Strength and Transmission Loss Between 2 and 30 MHz," Suppl. Rept. 252-2, Doc. XIVth Plenary Ass., ITU, Geneva, 1978.

CCIR, "Prediction of Sky-Wave Field Strength Between 150 and 1500 kHz," CCIR Recommendation 453-3, Doc. XIV Plenary Ass., ITU, Geneva, 1978.

CCIR, "Special Properties of the High Latitude Ionosphere Affecting Radio-communications," Study Program 25B/6, Doc. 6/1005-E, 14 Oct 81, XVth Plenary Ass., ITU, Geneva, 1981a.

CCIR, "Short-Term Forecasting of Operational Maximum Usable Frequencies and Total Electron Content," Study Program 27C/6, Doc. 6/1017-E, 14 Oct 81, XVth Plenary Ass., ITU, Geneva, 1981b.

CCIR, "Short-Term Prediction of Solar-Induced Variations of Operational Parameters for Ionospheric Propagation," Study Program 27A/6, Doc. 6/1016-E, 14 Oct 81, XV Plenary Ass., ITU, Geneva, 1981c.

CCIR, "Real-Time Channel Evaluation of Ionospheric Radio Circuits," Study Program 27B/6, Doc. 6/1018-E of XVth Plenary Ass., Geneva, 1981d.

CCIR, "Prediction and Short-Term Forecasting of HF Radio Propagation Conditions at High Latitudes," Draft Report submitted ITU, Geneva, August 1983.

Chan, K.L. and L. Colin, "Global Electron Density Distribution from Topside Soundings," *Proc. IEEE*, **57:** 6, 1969.

Chapman, F.W., D.L. Jones, J.D.W. Todd, and R.A. Challinor, "Observations on the Propagation Constant of the Earth-Ionosphere Waveguide in the Frequency Band 8 c/s to 16 kc/s," *Radio Sci.*, **1(11)**, 1273–1282, 1966.

Ching, B.K. and Chiu, Y.T., "A Phenomenological Model of Global Ionospheric Electron Density in the E, F_1, and F_2 Regions," *J. Atmos. Terr. Phys.*, **35:** 1615, 1973.

Chiu, Y.T., "An Improved Phenomenological Model of Ionospheric Density," *J. Atmos. Terr. Phys.*, **37:** 1563, 1975.

Cole, D.G. and L.F. McNamara, "Variations of Spread-F Occurrence Rates at Near-Equatorial Stations in the Australian Zone," *Aust. J. Phys.*, **27:** 249, 1974.

Cragin, B.L. and J.A. Fejer, "Generation of Large Scale Field-Aligned Irregularities in Ionospheric Modification Experiments," *Radio Sci.*, **9:** 1071, 1974.

Crain, C.M., "Ionospheric Probing with Long Wavelength Radio Waves," *J. Atmos. Terr. Phys.*, **32:** 551–566, 1970.

Crane, R.K., "Spectra of Ionospheric Scintillation," *J. Geophys. Res.*, **81:** 2041–2050, 1976.

Crane, R.K., "Ionospheric Scintillation," *Proc. IEEE*, **65:** 180, 1977.

Cretcher, C.K., "Ionospheric Effects in NAVSTAR-GPS," in *Effects of the Ionosphere on Space Systems and Communications*, edited by J.M. Goodman, Naval Research Lab., Washington, D.C., U.S. Government Printing Office, Stock No. 08-051-00064-0, 1975.

Czechowsky, P., W. Dieminger, and H. Kochan, "Backscatter Results from Lindau-I. Observations of Radio Auroras," *J. Atmos. Terr. Phys.*, **36:** 955–966, 1974.

D'Accardi, R.J., "Time Series Modelling and Analysis of HF Vertical and Short-Path Oblique Incidence Ionospheric Soundings," U.S. Army Communications R & D Command Tech. Rept. CORADCOM 78-7, NTIS ADA058630, 1978.

Dandekar, B.S., "Ionospheric Modeling," AFGL TR-82-0024, ADA 115243, 1982.

Danielson, C., "Auroral Observations at Thule 1961-65," *Danish Met. Inst., Geophys. Papers R-9*, 1969.

Darnell, M., "Channel Evaluation Techniques for HF Communication" in *Radio Systems and the Ionosphere, AGARD Conf. Proc.*, **173,** 1975a.

Darnell, M., "Adaptive Signal Selection for Dispersive Channels and its Practical Implications in Communications System Design" in *Radio Systems and the Ionosphere, AGARD Conf. Proc.*, **173,** 1975b.

Darnell, M., "Channel Evaluation Techniques for Dispersive Communications Paths," *Communication System*

and Random Process Theory, edited by J.K. Skwirzynski, Sijthoff and Noordhoff Alphen aan den Rijn, Holland, 1978.

DasGupta, A., A. Maitra, and S. Basu, "Occurrence of Nighttime VHF Scintillations near the Equatorial Anomaly Crest in the Indian Sector," *Radio Sci.*, **16:** 1455–1458, 1981.

David, P. and J. Voge, *Propagation of Waves*, Pergamon Press, London, U.K., 1969.

Davies, K., *Ionospheric Radio Propagation*, National Bureau of Standards Monograph 80, U.S. Government Printing Office, Washington, D.C., 1966.

Davies, K., *Ionospheric Radio Waves*, Blaisdell, Waltham, Mass., 1969.

Davies, K., "Recent Progress in Satellite Radio Beacon Studies with Particular Emphasis on the ATS-6 Radio Beacon Experiment," *Space Sci. Rev.*, **25:** 357–430, 1980.

Davis, T.N., "The Morphology of the Auroral Displays of 1957–1958. (2) Detailed Analysis of Alaska Data and Analysis of High Latitude Data," *J. Geophys. Res* **67:** 75–110, 1962.

Demaro, R.P., "NAVSTAR: The All-Purpose Satellite," *IEEE Spectrum*, **18:** 35–40, May 1981.

Donatelli, D.C. and R.S. Allen, "Time Cells for Adaptive Predictions of Total Electron Content," *Radio Sci.*, **16:** 261–269, 1981.

Donnelly, R.F., ed., *Solar Terrestrial Prediction Proceedings*, U.S. Dept. of Commerce, U.S. Government Printing Office, Washington, D.C. 20402

Vol. 1 *Prediction Group Reports*, No. 003-023-0041-9, 1979.

Vol. 2 *Working Group Reports and Reviews*, No. 003-017-00471-6, 1979.

Vol. 3 *Solar Activity Predictions*, No. 003-017-00473-2, 1979.

Vol. 4 *Predictions of Terrestrial Effects of Solar Activity*, No. 003-017-00479-1, 1980.

Dougherty, J.P. and D.T. Farley, "A Theory of Incoherent Scattering of Radio Waves by a Plasma," *Proc. Roy. Soc. (London)*, **A259:** 79, 1960.

Dubois, D.F. and M.V. Goldman, "Spectrum and anomalous Resistivity for the Saturated Parametric Instability," *Phys. Rev. Lett.*, **28:** 218, 1972.

Ducharme, E.D., L.F. Petrie, and R. Eyfrig, "A Method for Predicting the F_1 Critical Frequency," *Radio Sci.*, **6:** 369, 1971.

Ducharme, E.D., L.E. Petrie, and R. Eyfrig, "A Method of Predicting the F_1 Layer Critical Frequency Based on Zurich Smoothed Sunspot Number," *Radio Sci.*, **8:** 837–839, 1973.

Dyson, P.L., J.P. McClure, and W.B. Hanson, "In situ Measurements of the Spectral Characteristics of Ionospheric Irregularities," *J. Geophys. Res.*, **79:** 1497, 1974.

Eccles, D. and J.W. King, "A Review of Topside Sounder Studies of the Equatorial Ionosphere," *Proc. IEEE* **57**(6): 1012–1018, 1969.

Evans, J.V., "Theory and Practice of Ionosphere Study by Thomson Scatter Radar," *Proc. IEEE*, **57:** 496, 1969.

Evans, J.V., "Millstone Hill Radar Propagation Study" Scientific Results," Tech. Rept. 509; Joint Radar Propagation Study, Lincoln Laboratory, MIT, Lexington, Mass., 1973.

Evans, J.V., "High-Power Radar Studies of the Ionosphere," *Proc. IEEE*, **63:** 12, 1636–1650, 1975.

Eyfrig, R., "Contribution a l'Etude de l'Effect du Geomagnetisme sur le Couche F_1," *Comptes Rendus Academie des Sciences*, **241:** 759–761, 1955.

Fejer, J.A., "Scattering of Radiowaves by an Ionized Gas in Thermal Equilibrium," *Can. J. Phys.*, 38: 1114, 1960; also *J. Geophys. Res.*, **65:** 2635, 1960.

Fejer, J.A., "Radio Wave Probing of the Lower Ionosphere by Cross-Modulation Techniques," *J. Atmos. Terr. Phys.*, **32:** 597–607, 1970.

Fejer, J.A., "Ionospheric Modification and Parametric Instabilities," *Rev. Geophys. Space Phys.*, **17:** 135, 1979.

Feldstein, Y.I. and G.V. Starkov, "Dynamics of Auroral Belt and Polar Geomagnetic Disturbances," *Plan. Space Sci.*, **15:** 209–229, 1967.

Ferguson, J.A., "Ionospheric Profiles for Predicting Nighttime VLF/LF Propagation," Tech. Rep. 530, Naval Ocean Systems Center, San Diego, Calif., 1980.

Field, E.C., "Propagation of ELF Waves Under Normal and Naturally Disturbed Conditions," *J. Geophys. Res.*, **74:** 3639–3650, 1969.

Field, E.C., "Low-Frequency Ground Wave Propagation Over Narrow Terrain Features," *IEEE Trans. Antennas Propag.*, **30:** 831–836, 1982a.

Field, E.C., "ELF Propagation in Disturbed Environments" in *Medium, Long and Very Long Wave Propagation AGARD Conf. Proc.*, **305:** 11-1–11-10, ADA113969 1982b.

Field, E.C., "VLF Propagation in Disturbed Environments" in *Medium, Long and Very Long Wave Propagation AGARD Conf. Proc.*, **305:** 17-1–17-10, ADA113969, 1982c.

Field, E.C. and R.D. Engel, "The Detection of Daytime Nuclear Bursts Below 150 km by Prompt VLF Phase Anomalies," *Proc. IEEE*, **53:** 2009–2017, 1965.

Field, E.C. and R.G. Joiner, "An Integral Equation Approach to Long Wave Propagation in a Non-Stratified Earth-Ionosphere Waveguide," *Radio Sci.*, **14:** 1057–1068, 1979.

Field, E.C. and R.G. Joiner, "Effects of Lateral Ionospheric Gradients on ELF Propagation," *Radio Sci.*, **17:** 693–700, 1982.

Field, E.C. and C.R. Warber, "Calculation of Ionospheric Electron Density Profiles by Inverting VLF/LF Reflection Data, 2. Anisotropic Propagation" RADC-TR-84-242, Rome Air Development Center, Griffiss AFB, N.Y., 1984.

Field, E.C., M. Lewinstein, and M.A. Dore, "Effects of Antenna Elevation and Inclination on VLF/LF Signal Structure RADC-TR-76-375 ADA035510, 1976.

Field, E.C., R.E. Warren, and C.R. Warber, "Calculation of Ionospheric Conductivity Profiles by Inverting VLF/LF Reflection Data, 1. Isotropic Propagation," *Radio Sci.*, **18:** 452–460, 1983.

Flattery, T.W., T.F. Tascione, J.A. Secan, and J.W. Taylor, Jr., *A Four-Dimensional Ionospheric Model* (private communication).

Flood, W.A. "Simultaneous VHF Auroral Backscatter Measurements," *J. Geophys. Res.*, **65:** 2261–2268, 1960.

Forbes, J.M., "Upper Atmosphere Modifications Due to Chronic Discharges of Water Vapor from Space Vehicle Exhausts," in *Space Systems and their Interactions with Earth's Space Environment* edited by H.B. Garrett and C.P. Pike, American Institute of Aeronautics and Astronautics, New York, 1980.

Franklin, C.A. and M.A. MacLean, "The Design of Swept Frequency Topside Sounders," *Proc. IEEE*, **57**(6): 897–929, 1969.

Fremouw, E.J., "Geometrical Control of the Ratio of Intensity and Phase Scintillation Indices," *J. Atmos. Terr. Phys.*, **42:** 775–782, 1980.

Fremouw, E.J. and J.F. Bates, "Worldwide Behavior of Average VHF-UHF Scintillation," *Radio Sci.*, **6:** 863–869, 1971.

Fremouw, E.J. and C.L. Rino, "Modelling of Transionospheric Radio Propagation," RADC-TR-76-35, Rome Air Development Center, Griffiss AFB, NY, 1976.

Fremouw, E.J., C.L. Rino, A.R. Hessing, and V.E. Hatfield, "A Transionospheric Communication Channel Model," *Quart. Tech. Rep. 7*, SRI International (Stanford Research Institute), Menlo Park, Calif., 1977.

Fremouw, E.J. and C.L. Rino, "A Signal-Statistical and Morphological Model of Ionospheric Scintillation," *Proc. AGARD Conf. Operational Modelling of the Aerospace Propagation Environment*, Ottawa, Canada, 1978.

Fremouw, E.J., R.L. Leadabrand, R.C. Livingston, M.D. Cousins, C.L. Rino, B.C. Fair, and R.A. Long, "Early Results from the DNA Wideband Satellite Experiment-Complex Signal Scintillation," *Radio Sci.*, **13:** 167–187, 1978.

Fremouw, E.J., R.C. Livingston, and D.A. Miller, "On the Statistics of Scintillating Signals," *J. Atmos. Terr. Phys.*, **43:** 717–731, 1980.

Fremouw, E.J. and J.M. Lansinger, "Recent High Latitude Improvement in a Computer Based Scintillation Model," *Proc. Ionospheric Effects Symposium*, 3rd, 1981, Alexandria, Va., edited by J.M. Goodman, Naval Research Lab., Washington, D.C., 1982.

Frisius, J., A. Ehmert, and D. Stratmann, "Effects of High Altitude Nuclear Tests on VLF-Propagation," *J. Atmos. Terr. Phys.*, **26:** 251–262, 1964.

Galejs, J., *Terrestrial Propagation of Long Electromagnetic Waves*, Pergamon Press, Oxford, UK, 1972.

Gallagher, P.B. and R.A. Barnes, "Radio-Frequency Backscatter of Artificial Electron Clouds," *J. Geophys. Res.*, **68:** 2987–3010, 1963.

Gautier, T.N. and D.H. Zacharisen, "Use of Space and Time Correlations in Short-Term Ionospheric Predictions," Conference Record, 1st Annual IEEE Communications Convention, 671–676, 1965.

Goidman, M.V., "Field Aligned Instability Due to Stimulated Scattering of Intense Radio Waves from Diffusion Quasimodes," *Radio Sci.*, **9:** 1077, 1974.

Goodman, J.M., "A Survey of Ionospheric Models. A Preliminary Report on the Development of an Ionospheric Model Thesaurus and Users Guide," NRL Memorandum Report 4830, Naval Research Lab., Washington, D.C., 1982.

Gordon, W.E., "Incoherent Scattering of Radio Waves by Free Electrons with Applications to Space Exploration by Radar," *Proc. IRE*, **46:** 1824, 1958.

Gordon, W.E., H.C. Carlson, and R.L. Showen, "Ionospheric Heating at Arecibo: First Tests," *J. Geophys. Res.*, **76:** 7808, 1971.

Gowell, R.W. and R.W. Whidden, "Ionospheric Sounders in Aircraft," AFCRL-68-0369, AD678047, 1968.

Greenwald, R.A., W.L. Ecklund, and B.B. Balsley, "Diffuse Radar Aurora: Spectral Observations of Non-Two-Stream Irregularities," *J. Geophys. Res.*, **80:** 131–139, 1975.

Griefinger, C. and P. Greifinger, "Approximate Method for Determinging ELF Eigenvalues in the Earth-Ionosphere Waveguide," *Radio Sci.*, **13:** 831–837, 1978.

Griefinger, "C. and P. Griefinger, "On the Ionospheric Parameters which Govern High Latitude ELF Propagation in the Earth-Ionosphere Waveguide," *Radio Sci.*, **14:** 889–895, 1979.

Grubb, R.N., "The NOAA SEL HF Radar Systems (Ionospheric Sounder)," NOAA Tech. Memo ERL SEL-55 NOAA Environmental Research Laboratories, Boulder, CO, 1979.

Haerendel, G., "Modification of Ionized Media by Chemical Substances—A Review of Physical Processes," in *Artificial Modification of Propagation Media*, edited by H.J. Albrecht, *AGARD Conf. Proc.*, **192,** Brussels, 1976.

Haerendel, G., O.H. Bauer, S. Cakir, H. Foppl, E. Reiger, and A. Valenzuela, "Coloured Bubbles—An Experiment for Triggering Equatorial Spread F," in *Active Experiments in Space* edited by W. Burke, European Space Agency Report **SP-195:** 295, 1983.

Hagfors, T., "Density Fluctuations in a Plasma in a Magnetic Field with Applications to the Ionosphere," *J. Geophys. Res.*, **66:** 1699, 1961

Hagg, E.L., E.J. Hewens, and G.L. Nelms, "The Interpretation of Topside Sounder Ionograms," *Proc. IEEE*, **57:** 949–959, 1969.

Haldoupis, C. and G. Sofko, "Doppler Spectrum of 42 MHz CW Auroral Backscatter," *Can. J. Phys.*, **54:** 1571, 1976.

Hargreaves, J.K., *The Upper Atmosphere and Solar-Terrestrial Relations*, Van Nostrand Reinhold, New York, 1979.

Harrison, R.P., J.L. Heckscher, and E.A. Lewis, "Helicopter Observations of Very Low Frequency Radio Waves Over Certain Mountains and Shorelines," *J. Atmos. Terr. Phys.*, **33:** 101–110, 1971.

Harvey, R.P., R.B. Harrison, V.C. Fields, G.C. Hirst, P.A. Kossey, and E.A. Lewis, "Rocket Investigations of the VLF Ionospheric Transmission Window," AFCRL-TR-73-0293, AD767185, 1973.

Haydon, G.W. and D.L. Lucas, "Predicting Ionosphere Electron Density Profiles," *Radio Sci.*, **3:** 111, 1968.

Headrick, J.M. and M.I. Skolnik, "Over-the-Horizon Radar in the HF Band," *Proc. IEEE*, **62:** 664, 1974.

Heckman, G., "Predictions of the Space Environment Services Center," *STPP*, (Solar Terrestrial Predictions Workshop) **1:** 322 (see R.F. Donnelly, ed.) 1979.

Heckscher, J.L. and E.J. Tichovolsky, "Height-Gain Atlas for an Elemental Vertical Dipole Above a Flat Earth,"

RADC-TR-81-39, Rome Air Development Center, Griffiss AFB, N.Y., 1981.

Helliwell, R.A., *Whistlers and Related Ionospheric Phenomena*, Stanford University Press, Stanford, Calif., 1965

Herman, J.R., "Spread F and Ionospheric F-Region Irregularities," *Rev. Geophys.*, **4:** 244–299, 1966. (Includes important bibliography.)

Heron, M.L. and L.F. McNamara, "Transequatorial VHF Propagation through Equatorial Plasma Bubbles," *Radio Sci.*, **14,** 897, 1979.

Hill, D.A., and J.R. Wait, "Excitation of the Zenneck Surface Wave by a Vertical Aperture," *Radio Sci.*, **13:** 969–977, 1978.

Hizal, A. and A.F. Fer, "Medium Wave Multi-Section Groundwave Propagation in Marmara and West Black-Sea Region—Numerical Results" in *Medium, Long & Very Long Wave Propagation AGARD Conf. Proc.*, **305:** 31-1–31-7, ADA113969 1982.

Hofstee, J. and P.A. Forsyth, "Ion Acoustic Waves in the Auroral Plasma," *Can. J. Phys.*, **47:** 2797, 1969.

Hollingworth, J., "The Propagation of Radio Waves," *J. IEE*, **64:** 579, 1926.

Holmgren, G., R. Bostrom, M.C. Kelley, P.M. Kintner, R. Lundin, U.V. Fahleson, E.A. Bering, and W.R. Sheldon, "Trigger and Active Release Experiment that Stimulated Auroral Particle Precipitation and Wave Emissions," *J. Geophys. Res.*, **85:** 5043, 1980.

Hower, G.L. and A.B. Makhijani, "Further Comparison of Spread-F and Backscatter Sounder Measurements," *J. Geophys. Res.*, **74:** 3723–3725, 1969.

Hower, G.L., D.M. Ranz, and C.L. Allison, "Comparison of HF Radar Echoes and High-Lattitude Spread-F Measurements," *J. Geophys. Res.*, **71:** 3215–3221, 1966.

Hufford, G.A., "An Integral Equation Approach to the Problem of Wave Propagation Over an Irregular Surface," *Quart. Appl. Math.*, **9:** 391–403, 1952.

Hultqvist, B., "Polar Cap Absorption and Ground Level Effects," in *Solar Flare and Research*, edited by G. de Jager and Z. Svestka, p. 216, North Holland, Amsterdam, 1969.

Hunsucker, R.D., "Morphology and Phenomonology of the High Latitude E and F Regions," *STPP*, **2:**543 (see R.F. Donnelly, ed.), 1979.

Hunsucker, R.D. (Chairman), "High Latitude E- and F-Region Ionospheric Predictions," *STPP*, **2:** 513 (see R.F. Donnelly, ed.), 1979.

Hunsucker, R.D. and R.F. Bates, "Survey of Polar and Auroral Region Effects on HF Propagation," *Radio Sci.*, **4:** 347–365, 1969.

Ichinose, M., R. Maeda and S. Ito, "MUF Short-Term Prediction for HF Radio Propagation Circuits over Japan Area," *J. Rad. Res. Labs. (Japan)*, **27:** 179, 1980.

Inoue, Y. and S. Horowitz, "Numerical Solution of Full-Wave Equation with Mode Coupling," *Radio Sci.*, **1:** 957, 1968.

Jelly, D.H. and L.E. Petrie, "The High-Latitude Ionosphere," *Proc. IEEE*, **57:** 1005–1018, 1969.

Johler, J.R., "Concerning Limitations and Further Corrections to Geometric Optic Theory for LF, VLF Propagation Between the Ionosphere and the Ground," *Radio Sci.*, **68D:** 67–78, 1964.

Johler, J.R., "Zonal Harmonics in Low Frequency Terrestrial Radio Wave Propagation," NBS Tech. Note TN335, U.S. Government Printing Office, Washington, D.C., 1966.

Jones, T.B. and K. Mowforth, "A Review of the Analytical Techniques for Determining the Phase and Amplitude of a VLF Radio Wave Propagating in the Earth-Ionosphere Waveguide," in *Medium, Long, and Very Long Wave Propagation, AGARD Conf. Proc.*, **305:** 16-1–16-16 ADA113969, 1982.

Jones, W.B. and R.M. Gallet, "Ionospheric Mapping by Numerical Methods," *ITU Telecomm. J.*, **12:** 260, 1960.

Jones, W.B., R.P. Graham, and M. Leftin, *Advances in Ionospheric Mapping by Numerical Methods*, NBS Tech. Note 337, U.S. Government Printing Office, Washington, D.C., 1966.

Kelley, M.C., K.D. Baker, and J.C. Ulwick, "Late Time Barium Cloud Striations and Their Possible Relationship to Equatorial Spread F," *J. Geophys. Res.*, **84:** 1898, 1979.

Kellog, W.W. "Pollution of the Upper Atmosphere by Rockets," *Space Sci. Rev.*, **3:** 275, 1964.

Kelso, J.M., *Radio Propagation in the Ionosphere*, McGraw Hill, New York, 1964.

Kersley, L., "An Empirical Model of Ionospheric Slab Thickness," AGARD-CPP-284, May 1980.

King, R.J. and J.R. Wait, "Electromagnetic Groundwave Theory and Experiment," *Symposia Mathematica* (Instituto Nazionale di Alta Mathematica, Bologna, Italy), **18:** 107–208, 1976.

Klobuchar, J.A., "Numerical Models of Total Electron Content over Europe and the Mediterranean and Multi-Station Scintillation Comparisons," AGARDOGRAPH: AG-166A, Nov. 1973.

Klobuchar, J.A., "A First-Order, Worldwide, Ionospheric Time Delay Algorithm," AFCRL-TR-75-0502, ADA018862, Sep 1975.

Klobuchar, J.A. and R.S. Allen, "A First-Order Prediction Model of Total-Electron-Content Group Path Delay for a Midlatitude Ionosphere," AFCRL-70-0403, AD711365 1970.

Klobuchar, J.A., M.J. Buonsanto, M.J. Mendillo, and J.M. Johanson, "The Contribution of the Plasmasphere to Total Time Delay" in *Effects of the Ionosphere on Space and Terrestrial Systems*, edited by J.M. Goodman, U.S. Government Printing Office stock no. 008-051-00069-1, 1978.

Klobuchar, J.A. and J.M. Johanson, "Correlation Distance for Mean Daytime Electron Content," AFGL TR-77-0185, ADA048117, 22 Aug 1977.

Klobuchar, J.A., K.N. Iyer, H.O. Vats, and R.G. Rastogi, "A Numerical Model of Equatorial and Low Latitude Total Electron Content for use by Satellite Tracking Systems for Ionospheric Corrections," *Indian J. Radio Space Phys.*, **6,** 1977.

Knight, P., "MF Propagation: A Wave-Hop Method for Ionospheric Field-Strength Prediction," *BBC Engineering*, **100:** 22–34, 1975.

Knight, P., "Medium Frequency Propagation: A Survey," in *Medium, Long, and Very Long Wave Propagation,*

AGARD Conf. Proc., **305:** 28-1–28-17, ADA113969 1982.

Kofsky, I., "Further Evaluations of Infrared Simulation Data," Tech. Rept. DNA Contract 001-81-C-003, Photo Metrics, Inc., Woburn, Mass., 1981.

Kohnlein, W., "Electron Density Models of the Ionosphere," *Rev. Geophys. Space Phys.*, **16:** 341, 1978.

Kossey, P.A. and E.A. Lewis, "Detection and Classification of Solar X-Ray Flares Using VLF Phase and Amplitude Information," AFCRL-TR-74-0468, ADA003383 1974.

Kossey, P.A., E.A. Lewis, and E.C. Field, "Relative Characteristics of TE/TM Waves Excited by Airborne VLF/LF Transmitters," in *Medium, Long, and Very Long Wave Propagation, AGARD Conf. Proc.*, **305:** 19-1–19-10 ADA113969 1982.

Kossey, P.A., J.E. Rasmussen, and E.A. Lewis, "VLF Pulse Ionosounder Measurements of the Reflection Properties of the Lower Ionosphere," COSPAR, *Lower Ionosphere Structure*, 133–138, Akademie Verlag, Berlin, 1974.

Kossey, P.A., J.P. Turtle, R.P. Pagliarulo, W.I. Klemetti, and J.E. Rasmussen, "VLF Reflection Properties of the Normal and Disturbed Polar Ionosphere in Northern Greenland," *Radio Sci.*, **18:** 907–916, 1983.

Koster, J.R., "Some Measurements of the Irregularities Giving Rise to Radio-Star Scintillations at the Equator," *J. Geophys. Res.*, **68:** 2579–2590, 1963.

Kouba, J. "A Review of Geodetic and Geodynamic Satellite Doppler Positioning, *Rev. Geophys. and Space Phys.* **21:**27-40, 1983.

Kouris, S.S. and L.M. Muggleton, "Diurnal Variation in the E Layer Ionization," *J. Atmos. Terr. Phys.*, **35:** 133–139, 1973a.

Kouris, S.S. and L.M. Muggleton, "A Proposed Prediction Method for Monthly Median foE," CCIR Report 252-2, Contrib. No. 6/3/07 to Interim Working Party 6/3, 1973b.

Krause, G.E., R.J. D'Accardi, and E.L. Roswell III, "Field Test of a Near-Real-Time Ionospheric Forecasting Scheme (60 km)," U.S. Army ECOM Rept. 3345, 1970.

Krause, G.E., R.J. D'Accardi, and E.L. Roswell III, "Field Test of a Near-Real-Time Ionospheric Forecasting Scheme (200 km)," U.S. Army ECOM Rept. 4144, 1973a.

Krause, G.E., R.J. D'Accardi, and E.L. Roswell III, "Field Test of a Near-Real-Time Ionospheric Forecasting Scheme (500 km)," U.S. Army ECOM Rept. 4145, 1973b.

Lakshmi, D.R., S. Aggarwal, P.K. Pasricha, and B.M. Reddy, "HF Communication Problems at Low Latitudes due to Steep Spatial and Temporal Gradients," *STPP*, **4:** D2-58 (see R.F. Donnelly, ed.), 1980.

Larson, A.C. and J.C. Hodges, "VHF/UHF Auroral Radar Measurements," Stanford Research Institute International, Calif., June 1967.

Leadabrand, R.L., A.G. Larson, and J.C. Hodges, "Preliminary Results on the Wavelength Dependence and Aspect Sensitivity of Radar Auroral Echoes Between 50 and 3000 MHz," *J. Geophys. Res.*, **72:** 3877–3887, 1967.

Leadabrand, R.L., J.C. Schlobohm, and M.J. Baron, "Simultaneous Very High Frequency and Ultra High Frequency Observations of the Aurora at Fraserburgh, Scotland," *J. Geophys. Res.*, **70:** 4235–4284, 1965.

Lee, M.C., A. DasGupta, J.A. Klobuchar, S. Basu, and S. Basu, "Depolarization of VHF Geostationary Satellite Signals near the Equatorial Anomaly Crest," *Radio Sci.*, **17:** 399, 1982.

Leftin, M., "Numerical Representation of Monthly Median Critical Frequencies of the Regular E Region (F_0E)," OT Rept. 76-88k, U.S. Office of Telecommunications, Boulder, Colo., 1976.

Leftin, M., S.M. Ostrow, and C. Preston, "Numerical Maps of foE, for Solar Cycle Minimum and Maximum," ESSA Tech Rept. ERL 73-ITS-63, Boulder, Colo., 1968.

Leiphart, J.P., R.W. Zeek, L.S. Bearce, and E. Toth, "Penetration of the Ionosphere by Very-Low-Frequency Radio Signals-Interim Results of the LOFTI I Experiment," *Proc. IRE*, **50:** 6-17, 1962.

Lewis, E.A. and P.A. Kossey, "POWERFLUX I: A Method of Estimating Wave Intensities at Large Distances from Ground-Based Low Frequency Transmitters," AFCRL-TR-75-0338, ADA014787, 1975.

Lewis, E.A., J.E. Rasmussen, and P.A. Kossey, "Measurements of Ionospheric Reflectivity from 6 to 35 kHz," *J. Geophys. Res.*, **78:** 3903–3911, 1973.

Lewis, R.L., "The Wave-Hop fields for an Inclined Dipole Over a Spherical Earth with an Anisotropic Ionosphere," Tech Rept. OT/ITS RR 5, Institute for Telecommunication Sciences, Boulder, Colo., 1970.

Lied, F., (ed.), *H.F. Radio Communications*, Technivision, Maidenhead, UK, 1967.

Lincoln, J.V. and R.O. Conkright (eds.), "International Reference Ionosphere-IRI 79," Report UAG-82, World Data Center A for Solar-Terrestrial Physics, Boulder, Colo., 1981.

Llewellyn, S.K. and R.B. Bent, "Documentation and Description of the Bent Ionospheric Model " AFCRL-TR-73-0657, AD 772733, 1973.

Lloyd, J.S., G.W. Haydon, D.L. Lucas, and L.R. Teters, *Estimating the Performance of Telecommunication Systems Using the Ionospheric Transmission Channel*, National Telecommunications and Information Administration, Boulder, Colo., 1978.

Lyakhova, L.N., "Short-Range Forecasting of the State of the Ionosphere," *Trudy IZMIRAN*, **17:** 240, 1960.

Lyakhova, L.N. and L.I. Kostina, "On Quantitative Ionospheric Forecasting," *Geomag. Aeron.*, **13:** 50, 1973.

Maeda, K.I., "Study on Electron Density in the Lower Ionosphere," *J. Geomag. Geoelect.*, **23:** 133–159, 1971.

Malik, C. and J. Aarons, "A Study of Auroral Echoes at 19.4 Megacycles per Second," *J. Geophys. Res.*, **69:** 2731–2736, 1964.

Marmo, F.F., L.M. Ashcenbrand, and J. Pressman, "Artificial Electron Clouds—I. Planet," *Space Sci.*, **1:** 277, 1959.

Martin, E. and J. Aarons, "F-Layer Scintillation and the Aurora," *J. Geophys. Res.*, **82:** 2717, 1977.

McNamara, L.F., "Ionospheric Predictions on Transequatorial Circuits," *Proc. IREE* (Aust.), **35:** 17, 1974a.

McNamara, L.F., "Unusual Nighttime Propagation Modes

on the Guam and Manila—N.W. Cape Circuits," Ionospheric Prediction Service Series R Reports, IPS-R27, Sydney, Australia, 1974b.

McNamara, L.F., "Short-Term Forecasting of foF2," Ionospheric Prediction Service Series R Reports, IPS-R33, Sydney, Australia, 1976.

McNamara, L.F., "The Use of Ionospheric Indices to Make Real and Near Real Time Predictions of foF_2 Around Australia," *STPP*, **1:** 249 (see R.F. Donnelly, ed.), 1979.

Mendillo, M., "The Effect of Rocket Launches on the Ionosphere," *Adv. Space Res.*, **1:** 275, 1981.

Mendillo, M. and J. Baumgardner, "Optical Signature of an Ionospheric Hole," *Geophys. Res. Lett.*, **9:** 215, 1982.

Mendillo, M. and J.M. Forbes, "Artificially Created Holes in the Ionosphere," *J. Geophys. Res.*, **83:** 151, 1978.

Mendillo, M. and J. Forbes, "Theory and Observation of a Dynamically Evolving Negative Ion Plasma," *J. Geophys. Res.* **87:** 8273, 1982.

Mendillo, M., G.S. Hawkins, and J.A. Klobuchar, "A Large-Scale Hole in the Ionosphere Caused by the Launch of Skylab," *Science*, **187:** 343, 1975a.

Mendillo, M., G.S. Hawkins, and J.A. Klobuchar, "A Sudden Vanishing of the Ionospheric F-Region Due to the Launch of Skylab," *J. Geophys. Res.*, **80:** 2217, 1975b.

Mendillo, M. D. Rote, and P. Bernhardt, "Preliminary Report on the HEAO-Hole in the Ionosphere," *Eos*, Trans. AGU, **61:** 529, 1980.

Mikkelsen, I.S., J. Aarons, and E. Martin, "Geometrical Considerations of 136 MHz Amplitude Scintillation in the Auroral Oval," *J. Atmos. Terr. Phys.*, **40:** 479–483, 1978.

Milliken, R.J. and C.J. Zaller, "Principle of Operation of NAVSTAR and System Characteristics," *Navigation: J. Institute of Navigation*, **25**(2): 95–106 Summer 1978.

Millington, G., "Ground-Wave Propagation Over an Inhomogeneous Smooth Earth," *Proc. IEE*, **96,** Part III: 53–64, 1949.

Millman, G.H. *Modern Radar*, ed. R.S. Berkowitz, Wiley, New York, 1965.

Millman, G.H., "An Evaluation of HF Ionospheric Backscatter Echoes," G.E. Tech. Information Series Rept. No. R75EMH19, Nov 1975.

Millman, G.H. and C.D. Bell, "Ionospheric Dispersion of an FM Electromagnetic Pulse," *IEEE Trans. Antennas Propag.*, **AP-19**(1), Jan 1971.

Millman, G.H. and G.M. Reinsmith, "An Analysis of the Incoherent Scatter Faraday Rotation Technique for Ionospheric Propagation Error Correction," Rept. No. R74EMH2, G.E., New York, Feb 1974.

Minnis, C.M. and G.H. Bazzard, "A Monthly Ionospheric Index of Solar Activity Based on F2-Layer Ionization at Eleven Stations," *J. Atmos. Terr. Phys.*, **18:** 297, 1960.

Mitra, A.P., *Ionospheric Effects of Solar Flares*, D. Reidel, Dordrecht, Holland, 1974.

Morfitt, D.G., "Effective Electron Density Distributions Describing VLF/LF Propagation Data," Tech. Rept. 141, Naval Ocean Systems Center, San Diego, Calif., 1977.

Morfitt, D.G., J.A. Ferguson, and F.P. Snyder, "Numerical Modeling of the Propagation Medium at ELF/VLF/LF," in *Medium, Long and Very Long Wave Propagation, AGARD Conf. Proc.*, **305:** ADA113969 32-1–32-14, 1982.

Muggleton, L.M., "Solar Cycle Control of $N_m(E)$," *J. Atmos. Terr. Phys.*, **33:** 1307, 1971.

Mullen, J.P., "Sensitivity of Equatorial Scintillation to Magnetic Activity," *J. Atmos. Terr. Phys.*, **35:** 1187, 1973.

Muldrew, D.B., "F-Layer Ionization Troughs Deduced from Alouette Data," *J. Geophys. Res.*, **70**(11): 2635–2650, 1965.

Narcisi, R.S., "Overview of Project BIMF (Brazil Ionospheric Modification Experiments)," in *Active Experiments in Space* edited by W. Burke, European Space Agency Report SP-195, p. 295, 1983.

Newman, P. (ed.), "Spread F and its Effects upon Radiowave Propagation and Communication," Technivision, Maidenhead, UK, 1966. (Also AGARDOGRAPH 90).

Nielsen, E. and J. Aarons, "Satellite Scintillation Observations over the Northern High Latitude Regions," *J. Atmos. Terr. Phys.*, **36:** 159–165, 1974.

Nielson, D.L. and M. Crochet, "Ionospheric Propagation of HF and VHF Radio Waves Across the Geomagnetic Equator," *Rev. Geophys. Space Phys.*, **12:** 688, 1974.

Nisbet, J.S., "On the Construction and use of a Simple Ionospheric Model," *Radio Sci.*, **6:** 437, 1971.

Norton, K.A., "The Calculation of Ground-Wave Field Intensity Over a Finitely Conducting Spherical Earth," *Proc. IRE*, **29:** 623–639, 1941.

Noxon, J.F. and A.F. Johanson, "Effects of Magnetically Conjugate Photoelectrons on OI (6300 A)," *Planet. Space Sci.*, **18:** 1367, 1970.

Oelbermann, E., "Solar Particle Effects on Polar Cap VLF Propagation," *J. Franklin Inst., Special Issue on the Polar Cap*, **290:** 281–296, 1970.

Oksman, J., H.G. Moller, and R. Greenwald, "Comparisons between Strong HF Backscatter and VHF Radar Aurora," *Radio Sci.*, **14:** 1121–1133, 1979.

Oran, E.S. and T.R. Young, "Numerical Modeling of Ionospheric Chemistry and Transport Processes," *J. Phys. Chem.*, **81:** 2463, 1977.

Oran, E.S., T.R. Young, D.V. Anderson, T.P. Coffey, P.C. Kepple, A.W. Ali., and D.F. Strobel, *A Numerical Model of the Mid-Latitude Ionosphere*, Naval Research Laboratory Memorandum Rept. No. 2839, 1974.

Orsak, L.E., L.H. Rorden, G.B. Carpenter, and B.P. Ficklin, "VLF Propagation and Noise in the Ionosphere Observed by Sounding Rockets," Final Report on Contract NASr-49(01), Stanford Research Institute, Menlo Park, Calif., 1965.

Pappert, R.A., "A Numerical Study of VLF Mode Structure and Polarization Below an Anisotropic Ionosphere," *Radio Sci.*, **3:** 219–233, 1968.

Pappert, R.A., "Effects of a Large Patch of Sporadic E on Nighttime Propagation at Lower ELF," *J. Atmos. Terr. Phys.*, **42:** 417–425, 1980.

Pappert, R.A., "LF Daytime Earth Ionosphere Waveguide Calculations," Tech. Rept. 647, Naval Ocean Systems Center, San Diego, Calif., 1981.

Pappert, R.A. and J.E. Bickel, "Vertical and Horizontal

VLF Fields Excited by Dipoles of Arbitrary Orientation and Elevation," *Radio Sci.*, **35:** 1445–1452, 1970.

Pappert, R.A. and W.F. Moler, "A Theoretical Study of ELF Normal Mode Reflection and Absorption Produced by Nighttime Ionospheres," *J. Atmos. Terr. Phys.*, **40:** 1031–1045, 1978.

Patenaude, J., K. Bibl, and B.W. Reinisch, "Direct Digital Graphics," *American Lab.*, **9:**5–101, 1973.

Paul, A.K., "Processing of Digital Ionograms," NOSC-TD-529, Naval Ocean Systems Center, San Diego, 1982.

Penndorf, R., "Geographic Distribution of Spread-F in the Arctic," *J. Geophys. Res.*, **67:** 2274–2288, 1962.

Perkins, F.W. and P.K. Kaw, "On the Role of Plasma Instabilities in Ionospheric Heating by Radio Waves," *J. Geophys. Res.*, 76: 282, 1971.

Perkins, F.W. and E.J. Valeo, "Thermal Self-Focusing of Electromagnetic Waves in Plasma," *Phys. Res. Lett.*, **32:** 1234, 1974.

Perkins, F.W., C. Oberman, and E.J. Valeo, "Parametric Instabilities and Ionospheric Modification," *J. Geophys. Res.*, **79:** 1478, 1974.

Peterson, A.M., O.G. Villard, R.L. Leadabrand, and P.B. Gallagher, "Regularly Observable Aspect-Sensitive Radio Reflections from Ionization Aligned with the Earth's Magnetic Field and Located within the Ionosphere Layer at Middle Latitudes," *J. Geophys. Res.*, **60:** 497–512, 1955.

Petrie, L.E. and E.S. Warren, "The Propagation of High Frequency Waves on the Winnipeg-Resolute Bay Oblique Sounder Circuit," in *Ionospheric Radio Communications*, Plenum, New York, 1968.

Picquenard, A., *Radio Wave Propagation*, MacMillan, 1974.

Piggot, R.W. and K. Rawer, "URSI Handbook of Ionogram Interpretation and Reduction," Report UAG-23 and UAG-23A, World Data Center A for Solar-Terrestrial Physics, NOAA, Boulder, Colo., 1972.

Pike, C.P., "An Analytical Model of the Main F-Layer Trough," AFGL-TR-76-0098, ADA026031, 1976.

Pitteway, M.L.V., "The Numerical Calculation of Wave Fields, Reflection Coefficients and Polarizations for Long Radio Waves in the Lower Ionosphere, I," *Phil. Trans Roy. Soc.*, **A275:** 219–241, 1965.

Phelps, A.D.R. and Sagalyn R.C. "Plasma Density Irregularities in the High Latitude Topside Ionosphere," *J. Geophys. Res.*, **81:** 515 1976.

Pongratz, M.B. and G.M. Smith, "The LAGOPEDO Experiments –An Overview," *Eos*, Trans. AGU, **59:** 334, 1978.

Pope, J.H. and R.B. Fritz, "Observations of Artificially Produced Scintillation using Satellite Transmissions," *J. Geophys. Res.*, **79:** 1074, 1974.

Potemra, T.A. and T.J. Rosenberg, "VLF Propagation Disturbances and Electron Precipitation at Mid-Latitudes," *J. Geophys. Res.*, **78:** 1572–1580, 1973.

Proceedings IEEE, **57**(6), 1969.

Radinov, Ya. S., "Autocorrelative Characteristics of Critical Frequency and Ionization Density Fluctuations," *Geomag. Aeron.*, **3:** 985, 1963.

Radio Science, **2**(10), 1967.

Radio Science, **9**(11), 1974.

Ramakrishnan, S., D. Bilitza, and H. Thiemann, *Limitations of the IRI-78 Model*, presented at COSPAR General Assembly, Bangalore, India, 1979.

Rasmussen, J.E., P.A. Kossey, and E.A. Lewis, "Evidence of an Ionospheric Reflecting Layer Below the Classical D region," *J. Geophys. Res.*, **85:** 3037–3044, 1980.

Rasmussen, J.E., P.A. Kossey, and J.P. Turtle, "VLF/LF Pulse Reflections from Layers Below the Ionospheric D-Region," *AGARD Conf. Proc.*, **305:** 5-1–5-10 ADA 113969, 1982.

Rastogi, R.G., "Seasonal and Solar Cycle Variations of Equatorial Spread-F in the American Zone," *J. Atmos. Terr. Phys.*, **42:** 593–597, 1980.

Ratcliffe, J.A., "The Magneto-Ionic Theory and its Application to the Ionosphere," Cambridge University Press, London, 1959.

Ratcliffe, J.A., *Sun, Earth and Radio*, World University Library, London, 1970.

Rawer, K., "Propagation of Decameter Waves (HF-Bands)" in *Dynamical and Chemical Coupling between the Neutral and Ionized Atmosphere*, edited by B. Grandal and J.A. Holtlet, Reidel, Dordrecht, Holland, 1963.

Rawer, K., "The Historical Development of Forecasting Weather for Ionospheric Propagation of HF Waves," *Radio Sci.*, **10**(7): 669–679, 1975.

Rawer, K., *International Reference Ionosphere—IRI-79*, edited by J.V. Lincoln and R.O. Conkright, World Data Center A, for Solar Terrestrial Physics (STP), NOAA, Boulder, Colo., 1981.

Rawer, K. and K. Suchy, "Radio Observations of the Ionosphere," in *Encyclopedia of Physics*, edited by S. Flugge, Springer-Verlag, New York, 1967.

Reddy, C.A. and M.M. Rao, "On the Physical Significance of the Es Parameters fbEs, fEs and foEs," *J. Geophys. Res.*, **73:** 215, 1968.

Reder, F.H., "VLF Propagation Phenomena Observed During Low and High Solar Activity," XVI General Assembly of URSI, Ottawa, Canada, 1969.

Reilly, M.H., L.O. Harnish, and J.M. Goodman, "Polarimetry Studies of Ionospheric Modifications by Rocket Boosters," NRL Mem. Rpt. No. 4517, Naval Research Laboratory, Washington, D.C., 1981.

Reinisch, B.W. and H. Xuequin, "Automatic Calculations of Electron Density Profiles from Digital Ionograms, 1. Automatic O and X Trace Identification for Topside Ionograms," *Radio Sci.*, **17:** 421–434, 1982.

Reinisch, B.W. and H. Xuequin, "Automatic Calculation of Electron Density Profiles from Digital Ionograms, 3. Processing of Bottomside Ionograms," *Radio Sci.*, **18:** 477–492, 1983.

Reinisch, B.S., J.S. Tang, W.N. Hall, and J. Buchau, "Automatic Ionogram Processing Shows Seasonal Variations in the Auroral Ionosphere," AFGL-TR-82-0122, ADA 119911, 1982a.

Reinisch, B.W., J.S. Tang, and R.R. Gamache, "Automatic Scaling of Digisonde Ionograms," Test and Evaluation Report, AFGL TR-82-0324, ADA121666, 1982b.

Richmond, A.D. and S.V. Venkateswaran, "Geomagnetic Crochets and Associated Ionospheric Current Systems," *Radio Sci.*, **6:** 139, 1971.

Rino, C.L. and S.J. Matthews, "On the Morphology of

Auroral Zone Radio Wave Scintillation," *J. Geophys. Res.*, **85:** 4139, 1980.

Rino, C.L. and J. Owen, "The Structure of Localized Nighttime Auroral Zone Enhancements," *J. Geophys. Res.*, **85:** 2941–2948, 1980.

Rino, C.L. and J. Owen, "On the Temporal Coherence Loss of Strongly Scintillating Signals," *Radio Sci.*, **16:** 31, 1981.

Rino, C.L., R.C. Livingston, and S.J. Matthews, "Evidence for Sheet-Like Auroral Ionospheric Irregularities," *Geophys. Res. Lett.*, **5:** 1039, 1978.

Rishbeth, H. and O.K. Garriott, *Introduction to Ionospheric Physics*, Academic Press, 1969.

Rosenberg, N.W. (ed.) *PROJECT FIREFLY 1962–1963*, AFCRL-64-364, 1964.

Rosich, R.K. and W.B. Jones, "The Numerical Representation of the Critical Frequency of the F_1 Region of the Ionosphere," OT Rept. 73-22, Boulder, Colo., 1973.

Rote, D.M., "Proceedings of the Workshop on Ionospheric and Magnetospheric Effects of Satellite Power Systems (SPS)," Conf. Rept. No. 7807115, U.S. Department of Energy, December 1979.

Rufenach, C.L., "Radio Scintillation of Stellar Signals during Artificial Ionospheric Modification," *J. Geophys. Res.*, **78:** 5611, 1973.

Rush, C.M., "Improvements in Ionospheric Forecasting Capability," AFCRL-72-0138, AD742258 1972.

Rush, C.M., "An Ionospheric Observation Network for use in Short-Term Propagation Predictions," *Telecomm. J.*, **43:** 544, 1976.

Rush, C.M. and J. Gibbs, "Predicting the Day-to-Day Variability of the Mid-Latitude Ionosphere for Application to HF Propagation Predictions," AFCRL-TR-73-0335, AD764711 1973.

Rush, C.M., M. Pokempner, D.N. Anderson, F.G. Steward, and J. Perry, "Improving Ionospheric Maps using Theoretically Derived Values of foF2," *Radio Sci.*, **18:** 95, 1983.

Russell, C.T. and M.J. Rycroft (eds.), *Active Experiments in Space Plasmas*, Proc. Sym. 9, COSPAR XXIII, Budapest, 2–14 June 1980, Pergamon Press, 1980.

Sagalyn, R.C., M. Smiddy, and J. Wisnia, "Measurements and Interpretation of Ion Density Distribution in the Daytime F-Region," *J. Geophys. Res.*, **68:** 199, 1963.

Sagalyn, R.C., M. Smiddy, and M. Ahmed, "High Latitude Irregularities in the Topside ionosphere based on ISIS—Thermal Probe Data," *J. Geophys. Res.*, **79:** 4252, 1974.

Saltpeter, E.E., "Electron Density Fluctuations in a Plasma," *Phys. Rev.*, **120:** 1528, 1960; also *J. Geophys. Res.*, **65:** 1851, 1960.

Shea, M.A., S.A. Militello, and H.E. Coffey, "Directory of Solar-Terrestrial physics Monitoring Stations," MONSEE Special Publication No. 2, AFGL TR 84-0237,1984.

Shellman, C.H., "Electron Density Distributions in the Lower Ionosphere with Associated Error Limits Derived from VLF and LF Sounder Data," *Radio Sci.*, **5:** 1127–1135, 1970.

Shimazaki, T., "World Wide Variations in the Height of the Maximum Electron Density of the Ionospheric F_2 Layer," *J. Radio Res. Labs. (Japan)*, **2:** 86, 1955.

Shlyger, I.S., "Self-Modulation of a Powerful Electromagnetic Pulse Reflection from the Upper Layers of the Ionosphere," *JETP* (Engl. Transl.), **19:** 162, 1974.

Singh, M., E.P. Szuszczewicz, and J.C. Holmes, "The STP/S3-4 Satellite Experiment," Equatorial F-Region Irregularities," Proc—Ionospheric Effects Symposium 1981, Alexandria, Va. edited by J.M. Goodman, Naval Research Lab., Washington D.C., 1982.

Singleton, D.G., "The Morphology of Spread F Occurrence over Half a Sunspot Cycle," *J. Geophys. Res.*, **73:** 295–308, 1968.

Singleton, D.G., "Dependence of Satellite Scintillations on Zenith Angle and Azimuth," *J. Atmos. Terr. Phys.*, **32:** 789–803, 1970.

Singleton, D.G., "Transionsopheric Propagation Prediction. An Improved Ionospheric Irregularity Model," *STPP*, **4:** D1-1 (see R.F. Donnelly, ed.), 1980.

Sinno, K. and M. Kan, "Ionospheric Scintillation and Fluctuations of Faraday Rotation Caused by Spread-F and Sporadic-E over Kokobunji, Japan," *J. Radio Res. Lab.*, **27:** 53–77, 1980.

Sommerfeld, A., "The Propagation of Waves in Wireless Telegraphy," *Ann. Physik*, **28:** 665, 1909.

Sprenger, K. and K. Glöde, "Some Properties of Radio Aurorae in Medium Latitudes," *J. Atmos. Terr. Phys.*, **26:** 193–198, 1964.

Storey, L.R.O., "An Investigation of Whistling Atmospherics," *Phil. Trans. Roy. Soc.*, London, **A246:** 113–141, 1953.

Stratton, J.A., *Electromagnetic Theory*, McGraw-Hill, New York, 1941.

Strobel, D.F. and M.B. McElroy, "The F_2 Layer at Middle Latitudes," *Planet. Space Sci.*, **18:** 1181, 1970.

Stubbe, P., "Simultaneous Solution of the Time Dependent Coupled Continuity Equations, and Equations of Motion for a System Consisting of a Neutral Gas, and Electron Gas, and a four Component Ion Gas," *J. Atmos. Terr. Phys.*, **32:** 865, 1970.

Svennesson, J., "Effects on VLF Propagation During Ionospheric Substorms," *J. Atmos. Terr. Phys.*, **35:** 761–773, 1973.

Swenson, E.M., "Aspect-Sensitive Reflections from Ionization Irregularities in the F-Region," *J. Atmos. Terr. Phys.*, **34:** 1469–1476, 1972.

Tascione, T.F., T.W. Flattery, V.G. Patterson, J.A. Secan, and J.W. Taylor, Jr., "Ionospheric Modelling at Air Force Global Weather Central," *STPP*, **1:** 367 (see R.F. Donnelly, Ed.) 1979.

Tellegen, B.D.H., "Interaction Between Radio Waves," *Nature*, **131:** 840, 1933.

Thomas, J.O. and M.K. Andrews, "Transpolar Exospheric Plasma," *J. Geophys. Res.*, **73:** 7407–7417, 1968.

Thomas, J.O. and M.H. Rycroft, "The Exospheric Plasma During the International Years of the Quiet Sun," *Planet. Space Sci.*, **18:** 41–63, 1970.

Thome, G.D. and F.W. Perkins, "Production of Ionospheric Striations by Self-Focusing of Intense Radio Waves," *Phys. Rev. Lett.*, 32: 1238, 1974.

Thompson, R.L. and J.A. Secan, "Geophysical Forecasting

at AFGWC," *STPP*, **1**: 355 (see R.F. Donnelly, ed.), 1979.

Thomson, J.J., *Conduction of Electricity Through Gases*, p. 321, Cambridge University Press, London, 1906.

Thrane, E.V. (Chairman), "D-Region Predictions," *STPP*, **2:** 573 (see R.F. Donnelly, ed.), 1979.

Tinsley, B.A., A.B. Christensen, J.A. Bittencourt, P.D. Angreji, and H. Takahashi, "Excitation of Oxygen Permitted Emissions in the Tropical Nightglow," *J. Geophys. Res.*, **78:** 1174, 1973.

Titheridge, J.E., "Periodic Disturbances in the Ionosphere," *J. Geophys. Res.*, **73:** 243–252, 1968.

Trizna, D.B. and J.M. Headrick, "Ionospheric Effects on HF Over-the Horizon Radar," in *Effects of the Ionosphere on Radiowave Systems*, Proc. Ionospheric Effects Symposium 1981, Alexandria, Va. edited by J.M. Goodman, Naval Research Lab., Washington, D.C., 1982.

Tsunoda, R.T., "On the Spatial Relationship of 1-m Equatorial Spread F Irregularities and Plasma Bubbles," *J. Geophys. Res.*, **85:** 185, 1980.

Tsunoda, R.T., R.C. Livingston, J.P. McClure, and W.B. Hanson, "Equatorial Plasma Bubbles' Vertically Elongated Wedges from the Bottomside F-Layer," *J. Geophys. Res.*, **387:** 9171, 1982.

Turner, J.F., "The Development of the Ionospheric Index T," Ionospheric Prediction Science Series R Reports, IPS-R12, Sydney, Australia, 1968.

Turner, J.F. and P.J. Wilkinson, "A Weekly Ionospheric Index," *STPP*, **1:** 279 (edited by R.F. Donnelly), 1979.

UAG Report 10, *Atlas of Ionograms*, (avail. World Data Center A, U.S. Dept. of Commerce, NOAA/EDIS, Boulder, Colo.), 1970.

UAG Report 23, *URSI Handbook of Ionogram Interpretation and Reduction*, (avail. World Data Center A, U.S. Dept. of Commerce, NOAA/EDIS, Boulder, Colo.), 1972.

UAG Report 23a, *URSI Handook of Ionogram Interpretation and Reduction, Revision of Chapters 1–4*, (avail. World Data Center A, U.S. Dept. of Commerce, NOAA/EDIS, Boulder, Colo.), 1978.

UAG Report 50, *High Latitude Supplement to the URSI Handbook of Ionogram Interpretation and Reduction*, (avail. World Data Center A, U.S. Dept. of Commerce, NOAA/EDIS, Boulder, Colo.), 1975.

UAG-85 *Catalogue of Vertical Ionosphere Soundings Data*, (avail. World Data Center A, U.S. Dept. of Commerce, NOAA EDIS, Boulder, Colo.), 1982.

Uffelman, D.R. and L.O. Harnish, "Initial Results from HF Propagation Studies During Solid Shield," Naval Research Lab. Memorandum Report 4849, 1982

Utlaut, W.F., "An Ionospheric Modification Experiment using Very High Power, High-Frequency Transmission," *J. Geophys. Res.*, **75:** 6402, 1970.

Utlaut, W.F. and E.J. Violette, "Further Observations of Ionospheric Modification by a High-Powered HF Transmitter," *J. Geophys. Res.*, **77:** 6804, 1972.

Utlaut, W.F., E.J. Violette, and A.K. Paul, "Some Ionosonde Observations of Ionospheric Modification by Very High Power High-Frequency Ground-Based Transmission," *J. Geophys. Res.*, **25:** 6429, 1970.

Van der Pol, B. and H. Bremmer, "The Diffraction of Electromagnetic Waves from an Electrical Point Source Round a Finitely Conducting Sphere," *Phil. Mag. Series 7*, **24:** 141–176, 1937; **24:** 825–864, 1937; **25:** 817–834, 1938; and **26:** 262–275, 1938.

Vickery, J.F., C.L. Rino, and T.A. Potemra, "Chatanika Triad Observations of Unstable Ionization Enhancements in the Auroral F-Region," *Geophys. Res. Lett.*, **7:** 789–792, 1980.

Videberg, J.I. and G.S. Sales, "Long Range Survivable MF Radio Communication Study Using High Altitude Whispering Gallery Modes," AFCRL-TR-73-0552, AD774989, 1973.

Villard, O.G. Jr., "The Ionospheric Sounder and Its Place in the History of Radio Science," *Radio Sci.*, **11:** 847–860, 1976.

von Flotow, C.S., "Ionospheric Forecasting at Air Force Global Weather Central," in Effects of the Ionosphere on Space and Terrestrial Systems, edited by J.M. Goodman, U.S. Government Printing Office, Stock No. 008-051-00069-1, Jan 1978.

Wait, J.R., "Radiation from a Ground Antenna," *Can. J. Tech.*, **32:** 1–9, 1954.

Wait, J.R., "Mixed-Path Ground Wave Propagation, I. Short Distances," *J. Res. NBS*, **57:** 1–15, 1956.

Wait, J.R., "Excitation of Surface Waves on Conducting, Stratified, Dielectric-Clad, and Corrugated Surfaces," *J. Res. NBS*, **59:** 365–377, 1957.

Wait, J.R., "The Electromagnetic Fields of a Horizontal Dipole in the Presence of a Conducting Half-Space," *Can. J. Phys.*, **39:** 1017–1028, 1961.

Wait, J.R., *Electromagnetic Waves in Stratified Media*, Pergamon Press, New York, 1970.

Wait, J.R., (ed.), *Electromagnetic Probing in Geophysics*, The Golem Press, Boulder, Colo., 163–207, 1971.

Wait, J.R., (ed.) "Special Issue on Extremely Low Frequency (ELF) Communications," *IEEE Trans. Commun.*, COM-22, 1974.

Wait, J.R., "Propagation of ELF Electromagnetic Waves and Project Sanguine/Seafarer," *IEEE J. Ocean. Eng.*, **OE-2:** 161–172, 1977.

Wait, J.R. and H.H. Howe, "Amplitude and Phase Curves for Ground Wave Propagation in the Band 200 Cycles per Second to 500 Kilocycles," NBS Circular 574, National Bureau of Standards, Washington, D.C., U.S. Government Printing Office, 1956.

Wait, J.R. and A. Murphy, "Influence of a Ridge on the Low Frequency Ground Wave," *J. Res. NBS*, **58:** 1–5, 1957.

Wait, J.R. and A. Murphy, "Further Studies of the Influence of a Ridge on the Low Frequency Ground Wave," *J. Res. NBS*, **61:** 57–60, 1958.

Wait, J.R. and K.P. Spies, "Characteristics of the Earth-Ionosphere Waveguide for VLF Radio Waves," NBS Tech. Note 300, National Bureau of Standards, Washington, D.C., U.S. Government Printing Office, 1964.

Wait, J.R. and L.C. Walters, "Curves for Ground Wave Propagation Over Mixed Land and Sea Paths," *IEEE Trans. Antennas Propag.*, **11:** 38–45, 1963.

Wand, R.H. and J.V. Evans, "Morphology of Ionospheric Scintillation in the Auroral Zone," in *Effect of the Ionosphere on Space systems and Communications*, edited

by J. Goodman, NTIS CSCL 04/1 N75-30714, Naval Research Laboratory, Washington, D.C., 1975.

Wand, R.H. and M. Mendillo, "Incoherent Scatter Observations of an Artificially Modified Ionosphere," *J. Geophys. Res.*, **89:** 203, 1984.

Watson, G.N., "The Transmission of Electric Waves Round the Earth," *Proc. Roy. Soc.*, **95:** 546–563, 1919.

Watt, A.D., *VLF Radio Engineering*, Pergamon Press, New York, 1967.

Weaver, P.F., "Backscatter Echoes from Field-Aligned Irregularities in the F-Region," *J. Geophys. Res.*, **70:** 5425–5432, 1965.

Weber, E.J. and J. Buchau, "Polar Cap F-Layer Auroras," *Geophys. Res. Lett.* **8:**125, 1981.

Weber, E.J., J. Buchau, R.H. Eather, and S.B. Mende, "North-South Aligned Equatorial Airglow Depletions," *J. Geophys. Res.*, **83:** 712, 1978.

Weber, E.J., J. Buchau, and J.G. Moore, "Airborne Studies of Equatorial F-Layer Ionospheric Irregularities," *J. Geophys. Res.*, **85:** 4631–4641, 1980.

Weber, E.J., H.C. Brinton, J. Buchau, and J.G. Moore, "Coordinated Airborne and Satellite Measurements of Equatorial Plasma Depletions," *J. Geophys. Res.*, **87:** 503–510, 513, 1982.

Wernik, A.W., C.H. Liu, and K.C. Yeh, "Model Computations of Radio Wave Scintillation caused by Equatorial Ionospheric Bubbles," *Radio Sci.*, **15:** 559, 1980.

Weyl, H., "Propagation of Electromagnetic Waves Above a Plane Conductor," *Ann. Physik*, **60:** 481–500, 1919.

Whalen, B.A., et al., "Preliminary Results from Project Waterhole—An Auroral Modification Experiment," *Can. J. Phys.*, **59:** 1175,1981.

Whitney, H.E., C. Malik, and J. Aarons, "A Proposed Index for Measuring Ionospheric Scintillation," *Planet. Space Sci.*, **17:** 1069–1073, 1969.

Whitney, H.E., J. Aarons, R.S. Allen, and D.R. Seemann, "Estimation of the Cumulative Amplitude Probability Distribution Function of Ionospheric Scintillations," *Radio Sci.*, **7:** 1095–1104, 1972.

Wilkinson, P.J., "Prediction Limits for foF2," *STPP*, **1:** 259 (see R.F. Donnelly ed.), 1979.

Wilkinson, P.J., "A Comparison of Monthly Indices of the Ionospheric F-Region," Ionospheric Prediction Service Series R Reports, IPS-R41, Sydney, Australia, 1982.

Winkler, C., "Radio Wave Guidance at VHF Through Equatorial Plasma Bubbles," *J. Atmos. Terr. Phys.*, **43:** 307, 1981.

Wright, J.W., "Kinesonde Observations of Ionosphere Modification by Intense Electromagnetic Fields from Platteville, Colorado," *J. Geophys. Res.*, **78:** 5622, 1973.

Wright, J.W. and A.K. Paul, "Toward Global Monitoring of the Ionosphere in Real Time by a Modern Ionosonde Network: The Geophysical Requirements and Technological Opportunity," NOAA Special Report, Space Environment Lab., July 1981.

Wright, J.W. and M.L.V. Pitteway, "Real-Time Data Acquisition and Interpretation Capabilities of the Dynasonde, 1. Data Acquisition and Real Time Display, 2. Determination of Magnetosonic Mode and Echolocation using a Small Spaced Receiving Array," *Radio Sci.*, **14,** 815–835, 1979.

Wright, J.W. and M.L.V. Pitteway, "Data Processing for the Dynasonde: The Dopplionogram," *J. Geophys. Res.*, **87:** 1589, 1982.

Yeh, K.C., "Travelling Ionospheric Disturbances as a Diagnostic Tool for Thermospheric Dynamics," *J. Geophys. Res.*, **77:** 709–719, 1972.

Yeh, K.C. and C.H. Liu, "Ionospheric Effects on Radio Communication and Ranging Pulses," *IEEE Trans. Antennas Propag.*, **AP-27**(6), Nov 1979.

Yeh, K.C. and C.H. Liu, "Radio Wave Scintillations in the Ionosphere," *Proc. IEEE*, **70,** 1982.

Yeh, K.C., H. Soicher, and C.H. Liu, "Observations of Equatorial Ionospheric Bubbles by the Radio Propagation Method," *J. Geophys. Res.*, **84:** 6589, 1979.

Zacharisen, D.M., "Space-Time Correlation Coefficients for use in Short-Term Ionospheric Predictions," NBS Report 8811, 1965.

Zevakina, R.A. E.V., Lavrova, and L.N. Lyakhova, *Manual on Short-Term Predictions of Ionospheric Geomagnetic Storms and Radio-Propagation Forecasting*, (text in Russian) Nauka, Moscow, 1967.

Zinn, J. and C.D. Sutherland, "Effects of Rocket Exhaust Products in the Thermosphere and Ionosphere," *Space Solar Power Rev.*, **1:** 109, 1980.

Zinn, J., C.D. Sutherland, S. Stove, and L. Duncan, "Ionospheric Effect of Rocket Exhaust Products: HEAO-C and Skylab," *J. Atmos, Terr. Phys.*, **44:** 1143, 1982.

Chapter 11

SOLAR RADIO EMISSION

W. R. Barron
E. W. Cliver
J. P. Cronin
D. A. Guidice

Since the first detection of solar radio noise in 1942, radio observations of the sun have contributed significantly to our evolving understanding of solar structure and processes. The now classic texts of Zheleznyakov [1964] and Kundu [1965] summarized the first two decades of solar radio observations. Recent monographs have been presented by Kruger [1979] and Kundu and Gergely [1980].

In Chapter 1 the basic phenomenological aspects of the sun, its active regions, and solar flares are presented. This chapter will focus on the three components of solar radio emission: the basic (or minimum) component, the slowly varying component from active regions, and the transient component from flare bursts.

Different regions of the sun are observed at different wavelengths. At millimeter wavelengths, the radiation is from the photosphere. Centimeter wavelength radiation originates in the chromosphere and low corona. Decimeter and meter wavelengths have their origin at increasing heights in the corona: at meter wavelengths the observed radiation comes from heights ranging from 100 000 to 700 000 km above the photosphere. For receiving equipment on the earth, the low-frequency limit for observation is the frequency at which radio waves are reflected by the ionosphere (for practical purposes, around 20 MHz). The high-frequency limit is set by absorption of radiation by atmospheric oxygen and water vapor. Recently, radio experiments on satellites have observed hectometric wavelength ($\lesssim$2 MHz) emission that originate at heights $\gtrsim$10 solar radii ($R_\odot$). This leaves only the frequency range form 2 to 20 MHz, corresponding to emission heights of 2–10 $R_\odot$, unexplored by radio methods.

11.1 BASIC DEFINITIONS

If the sun radiated only as a thermal source, the emitted energy density would vary with frequency and temperature according to Planck's radiation law. In the radio region, the Rayleigh-Jeans approximation for blackbody radiation is valid; the brightness, radiance per unit bandwidth, is

$$B_f = 2kTf^2c^{-2} = 2kT\lambda^{-2}. \qquad (11.1)$$

If the frequency f is in cycles per second, the wavelength λ in meters, the temperature T in degrees Kelvin, the velocity of light c in meters per second, and Boltzmann's constant k in joules per degree Kelvin, then B_f is in W $m^{-2}Hz^{-1}sr^{-1}$. Values of temperatures T_b calculated from Equation (11.1) are referred to as *equivalent blackbody temperature* or as *brightness temperature* defined as the temperature of a blackbody that would produce the observed radiance at the specified frequency.

The radiant power received per unit area in a given frequency band is called the power flux density (irradiance per bandwidth) and is strictly defined as the integral of $B_f d\Omega_s$ between the limits f and $f + \Delta f$, where Ω_s is the solid angle subtended by the source. In solar radio astronomy the relationship used is

$$F_s = B_f\Omega_s = 2kT_d\Omega_s\lambda^{-2}, \qquad (11.2)$$

where the apparent or disk temperature T_d is that temperature which a uniform source of the same angular size as the solar optical disk must have in order to produce the power flux density F_s received from the sun. Values of power flux density are usually given in solar flux units (1 solar flux unit, sfu = 10^{-22} W m^{-2} Hz^{-1}).

The power P_s received at the antenna due to solar radiation is given by:

$$P_s = F_sA_e, \qquad (11.3)$$

where A_e is the effective area of the antenna. P_s is also conveniently expressed in terms of the effective antenna temperature T_A corrected for any RF losses. T_A is defined by

$$P_s = kT_A\Delta f. \qquad (11.4)$$

T_A is readily measured with suitably calibrated instruments. The equation for calculating the solar power flux density

from a given antenna temperature measured at a given installation is

$$F_s = 2kT_A \frac{L}{A_e}, \qquad (11.5)$$

where A_e is the effective area of the antenna in square meters, and L is a dimensionless correction factor related to the antenna response shape and to the diameter and temperature distribution across the source. The value of L is unity only if the antenna half-power beamwidth is large compared to the source. L exceeds unity when the ratio of the antenna half-power beamwidth to the solar angular diameter drops below about five; thus, it is desirable to use a parabola small enough so the half-power beamwidth is more than five times the solar angular diameter. Once the solar flux density is known, the apparent temperature and the brightness of the solar disk may be calculated from Equation (11.2).

11.2 THE MINIMUM (ZERO-SUNSPOT) COMPONENT

The standard method for obtaining the basic radio flux density of the unspotted sun is to make a scatter plot of solar temperature at a given frequency against the projected sunspot area; the extrapolation of the curve that best fits these data to zero sunspot area determines the minimum or basic flux density at that frequency. By doing this for all frequencies, one determines the spectrum of the basic component of solar radiation. Figure 11-1 shows the distributions of power flux density for the sun and for black bodies at various temperatures. It is only at millimeter and shorter wavelengths that solar radio emission approximates a 6000 K blackbody. At wavelengths longer than 1 cm, the equivalent blackbody temperature ranges between 10^4 to 10^6 K for the spotless sun and from 10^4 to 10^{10} K for the disturbed sun depending on the condition of the sun and the time in relation to the 11-year sunspot cycle. Meter wavelengths are characterized by much burst activity; thus, the basic (or minimum) sun temperatures at these wavelengths are determined by making observations over a period of weeks or months. The low temperature to which the sun periodically returns, but never goes below, during this period is taken to be the zero-sunspot value; it is of the order of 10^6 K

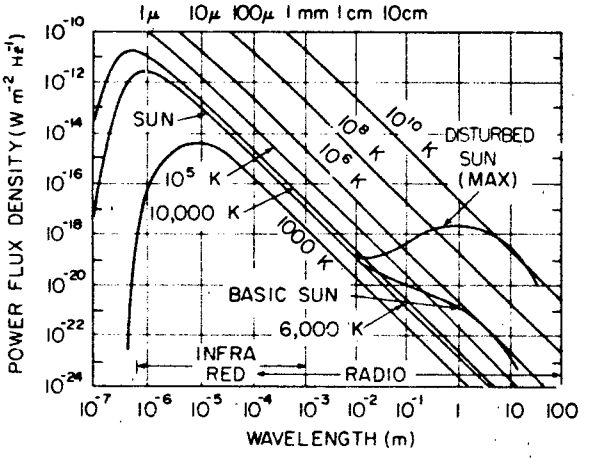


Figure 11-1 Solar spectrum and spectra of blackbody radiation at various temperatures. The solar power-flux density (power per unit area per unit bandwidth) is plotted against wavelength.

11.3 THE SLOWLY VARYING COMPONENT

The slowly varying component (SVC) exhibits a well defined 11-year cycle variation and a 27-day solar-rotation variation, since this emission originates principally in coronal condensations overlying active regions, and is well-correlated with sunspot number. The routine daily measurement of the combination of the basic (or minimum) component and the SVC of solar radio emission is referred to as the *quiet sun flux density*. The SVC of the sun as a whole is obtained by subtracting the basic component from the quiet-sun flux density. The SVC of individual active regions can be obtained by either eclipse observations or interferometric measurements.

11.4 THE BURST COMPONENT

During solar flares (Chapter 1) there may be large increases (bursts) in radio emission lasting anywhere from a few seconds to several hours. These bursts originate by bremsstrahlung, gyrosynchrotron, and plasma radiations. Characteristics of the bursts vary with wavelength. Bursts in the meter-wave range (12 m to about 50 cm) are classified by spectral type. No spectral classification exists for the decimeter or centimeter wave regions.

11.4.1 Meter-Wave Range (25 – 500 MHz).

Most information on solar bursts in the range from 12 m to about 50 cm is obtained from swept-frequency observations. Dynamic spectra are displayed on a cathode-ray tube and recorded photographically as a series of intensity-modulated traces that give intensity as a brightening in the frequency-time plane. Figure 11-2 is an illustration of idealized dynamic spectra of various types of bursts. These spectral types are discussed below in the order of their occurrence in the flare event.

Type III bursts are the most common type of solar radio

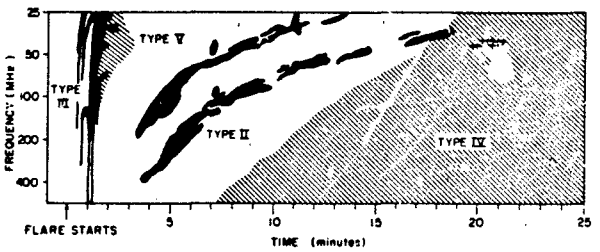


Figure 11-2. Idealized illustration of the record of a complete solar radio outburst as recorded by a dynamic spectrograph at meter wavelengths.

activity. They can occur either singly (duration ~5 s) or in groups. Only about one-third of Type III bursts are associated with flares. When associated with flares, however, the timing agreement with flare impulsive hard x-ray and microwave emission is often quite good (within seconds). Type III bursts are caused by streams of ~100 keV electrons propagating outward through the solar atmosphere and exciting plasma waves. Because of their relatively high drift rates (20 MHz/s) to lower frequencies, they are referred to as fast drift bursts.

Type V bursts consist of a broad band continuum of short duration (~1 min) that is preceded by a Type III burst and accompanied by centimeter-wave and hard x-ray bursts. The Type V burst may indicate the presence of a particularly rich stream of electrons, part of which is trapped in the corona and becomes visible either through gyrosynchrotron radiation or plasma waves.

The Type II burst, or slow-drift (~0.2 MHz/s) burst is presumed due to the presence of a shock wave propagating outward through the solar corona with a characteristic velocity of 1000 km/s. The disturbance excites plasma waves at the local plasma frequency. In the spectrograph record, these bursts appear as a narrow band of intense radiation that drifts gradually and often irregularly from high to low frequency. About 60% of all Type II bursts show emission at the second harmonic. Type II bursts are closely associated with solar flares.

Type IV emission has at least three distinct components; these components can not be separated on the spectrograph record and can only be distinguished by interferometers. *Flare continuum* is the broad band emission occurring at meter and decameter wavelengths during the flare impulsive phase. *Moving Type IV* bursts involve a variety of forms, although three basic types have been recognized. These are the magnetic arch, the advancing front, and the isolated source (ejected plasma "blob"). Both the flare continuum and moving Type IV emissions have durations ≤10–60 min. While the moving Type IV burst travels outward through the corona and can reach heights of $\sim 10^6$ km above the photosphere, the final stage of Type IV emission, the *storm continuum*, originates low in the corona near the corresponding plasma level and directly above the site of the optical flare. The storm continuum can last for several hours and often degenerates into the Type I noise storms whose durations range from hours to days. In contrast to the flare continuum and moving Type IV emissions that are only weakly polarized, storm continuum is strongly polarized in the ordinary mode. This suggests plasma radiation as the source of the storm continuum, while gyrosynchrotron radiation from energetic electrons spiraling in weak coronal magnetic fields are generally cited as the source of the flare continuum and moving Type IV emissions.

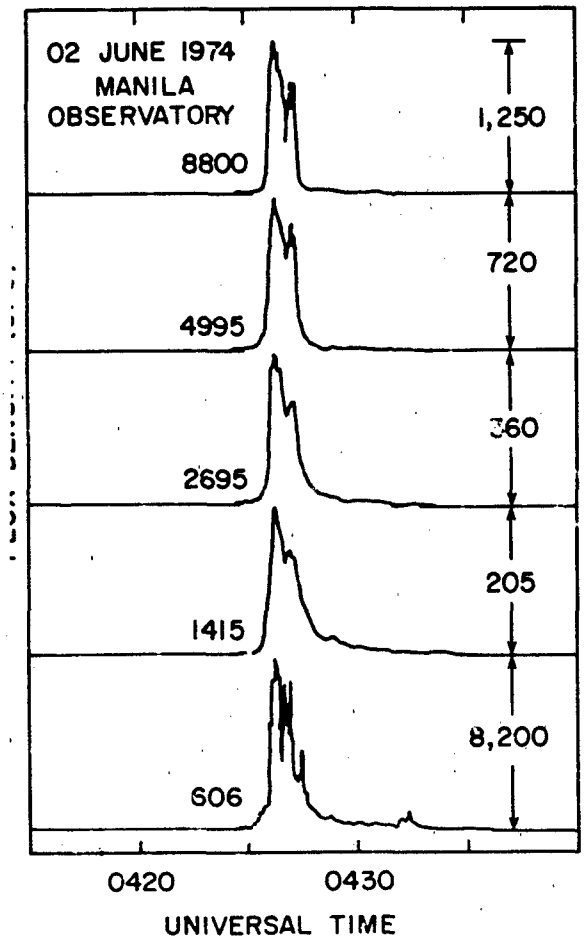
Type I events are distinguished from the relatively smooth broad-band Type IV emission by the presence of a great number of short (~1 s) intense bursts superimposed on the background continuum. These short intense bursts are distributed more or less randomly over the frequency range of the underlying continuum. Both the background continuum and the superimposed bursts are strongly circularly polarized, usually in the ordinary mode. Type I radiation appears to be more closely associated with certain active regions than with flares, although they can be flare-initiated. At present, the mechanism and origin of Type I emission are not well understood.

11.4.2 Decimeter-Wave Range (500 – 2000 MHz).

At decimeter wavelengths the emission is highly variable and complex. Rapid time structures (several peaks per minute) are often observed in the time profiles of decimetric radio bursts observed at discrete frequencies, and the relationship of these fast structures to the source of the smoother emission observed at centimeter wavelengths is not clear. Individual peaks in complex events are often strongly circularly polarized in the ordinary mode.

11.4.3 Centimeter-Wave Range (2000 – 35000 MHz).

Solar emission in the 15–1 cm range does not show as rapid fluctuations as emission in the meter and decimeter ranges. There appear to be at least two basic morphological types of centimeter-wave bursts. The first of these is the *simple impulsive event* that reaches a maximum peak-flux density ranging from 10^1 to $>10^4$ sfu in a few minutes. Impulsive bursts are interpreted in terms of gyrosynchrotron emission. Complex bursts may consist of several impulsive events in sequence, although the appearance of a relatively smooth broad band microwave Type IV component in the later stages of many complex events may indicate an additional or prolonged acceleration process. The second of these, the *gradual rise and fall* microwave burst, may occur by itself or may follow an impulsive event (monotonic decay only), in which case it is referred to as a *post-burst* event.



igure 11-3. The time-intensity profile of an impulsive microwave burst.

Peak flux densities of these gradual events seldom exceed 10 sfu; these events are generally explained in terms of bremsstrahlung from a Maxwellian distribution of electrons. Examples of the time profiles of microwave bursts observed at Sagamore Hill Solar Radio Observatory are contained in the burst atlas compiled by Barron et al. [1980]. A typical impulsive event is presented in Figure 11-3.

Instantaneous spectra of events in the centimeter frequency range are relatively smooth and tend not to have narrow-band emission features. For moderate sized events ($\geq$100 sfu), the peak-flux-density spectral maximum of the emission generally occurs at frequencies $\geq$9 GHz. For the largest events with centimeter wave peak flux densities $\geq$1000 sfu, the emission usually extends to the meter-wave range where it often exhibits a second spectral maximum with a minimum occurring at the intermediate wavelengths. This U-shaped spectral signature (Figure 11-4) is thought to reflect the fact that there are two different sources of burst radiation (one at centimeter waves and one at lower frequencies) with different electron energy distributions and different magnetic fields [Kundu and Vlahos, 1982].

11.5 CORRECTIONS TO QUIET-SUN AND BURST-FLUX DENSITIES

In the final two sections of this chapter, charts and tables of burst and quiet-sun flux densities from Sagamore Hill Solar Radio Observatory are presented. It is appropriate to discuss errors of absolute calibrations in these measurements. In 1973, a report [Tanaka et al., 1973] by an absolute calibration working group formed by Commission V of URSI was published. It contained corrections for Sagamore Hill for the years 1966–1971. For the years 1972–1976, correction factors were taken from the IAU's *Quarterly Bulletin of Solar Activity*. For 1977–1979, correction factors were derived by extrapolation from previous years and comparison of the Sagamore Hill quiet-sun flux densities with those of other solar-control observatories. For 245 MHz, a substantial correction factor (1.55) was applied to all the data as a result of an absolute calibrations measurement (using the radio source Cassiopeia A) carried out at Sagamore Hill in 1980. The multiplicative flux-density correction factors for the five frequencies for which data are presented are listed in Table 11-1. Before processing the data presented below, these correction factors were applied to the Sagamore Hill burst and quiet-sun data.

11.6 QUIET-SUN FLUX-DENSITY MEASUREMENTS

Solar microwave emission correlates well with solar EUV flux [Forbes and Straka, 1973], a fact which makes it useful as an input to ionospheric models in lieu of the more difficult to obtain sunspot number and EUV flux. For this reason, the 1965 edition of the *Handbook of Geophysics* included a table of the daily quiet-sun (non-bursting) flux densities observed at 2800 MHz (10.7 cm wavelength) by the National Research Council in Ottawa, Canada from 1947 to mid-1963. We continue the table in this handbook using Ottawa data through 1965 and Sagamore Hill data thereafter. In addition, the 8800 MHz quiet-sun flux-density values from Sagamore Hill are included.

In Table 11-2 we present the observed daily solar flux density value measured at 2800 MHz at Ottawa for the years 1963–1965 and the 2695 and 8800 MHz flux density values measured at the Sagamore Hill Radio Observatory from January 1966 through December 1979. It should be noted that these are observed values and not values adjusted to 1 AU. The values are taken at local noon at the Sagamore Hill meridian. This is done so that the radiation will have passed through the shortest path in the earth's atmosphere

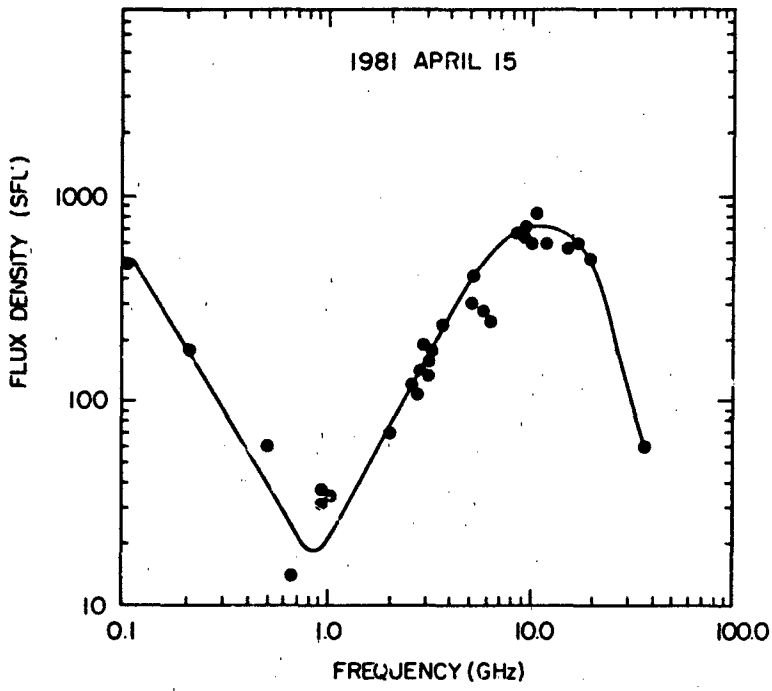


Figure 11-4. The peak flux-density spectrum of a large microwave burst. The U-shape is characteristic for large microwave bursts with peak flux density > 1000 sfu.

and be subject to the minimum atmospheric attenuation. In some instances this meridian-transit observation is not possible because a solar radio burst is in progress at that time. When this happens, the observation is taken as soon as it has been determined that the radio burst has ended. In virtually all cases this was within an hour of the time of meridian transit so that the added attenuation due to the increased atmospheric path length was negligible.

Note the asterisk (*) between the 2695 MHz flux-density value and the 8800 MHz flux-density value on some days. This indicates that there were adverse weather conditions, usually rain or snow, present when the observations were being made. These conditions will cause the signals received to be further attenuated. The asterisk indicates that some adjustment has been made to the observed value to compensate for this problem.

Table 11-1. Calibration correction factors, Sagamore Hill, 1966–1979.

Year	245 MHz	610 MHz	1415 MHz	2695 MHz	8800 MHz
1966	—	0.91	1.16	0.90	0.91
1967	—	0.91	1.16	0.90	0.91
1968	—	0.91	1.16	0.90	0.95
1969	1.55	0.91	1.16	0.94	1.00
1970	1.55	0.91	1.16	0.94	1.00(2)
1971	1.55	0.91	1.16	0.94	1.00(2)
1972	1.55	0.91	1.08	0.93	0.95
1973	1.55	0.91	1.03	0.91	0.90
1974	1.55	0.91	1.05	0.90	0.90
1975	1.55	0.91	1.05	0.90	0.87
1976	1.55	0.92	1.08	0.90	0.85
1977	1.55	1.00	1.06	0.88	0.88
1978	1.55	1.00	1.06	0.85	0.87
1979	1.55	1.00	1.06	0.85	0.84

Notes: (1) For the period 9 June to 31 August 1967, all 2695 MHz burst and daily flux-density values should be multiplied by 1.13.
(2) For the period 1 November 1970 to 31 August 1971, all 8800 MHz burst and daily flux-density values should be multiplied by 1.14.

CHAPTER 11

11.7 SOLAR RADIO-BURST CLIMATOLOGY BASED ON 1966-1978 SAGAMORE HILL OBSERVATIONS

In addition to its role in maintaining the ionosphere, the sun is the most powerful natural source of radio interference to ground-based or near-earth (aircraft or satellite) radio receivers. Therefore, it is of interest to designers and operators of radio receiving equipment to have information on what sort of interference one might expect from the sun at different frequency ranges in the radio spectrum and over what percentage of time the interference might occur (above a certain intensity level). Since radio burst emission from the sun varies unpredictably over a wide range of frequencies, one can only give the statistical likelihood of certain levels being exceeded at various frequencies (or frequency ranges) of interest. We call this body of statistics on solar radio-burst occurrence "climatology." Burst climatology is provided by Sagamore Hill observations at 245, 610, 1415, 2695, and 8800 MHz [Castelli et al., 1973]. Observations at 245 MHz began in March 1969; observations for the other listed frequencies began in January 1966.

To produce a useful solar-burst "climatology", one must first determine what intensity (flux-density) level a burst must reach to interfere with operational systems. Since radio-burst emission occurs over only a very small fraction of the time that any system is exposed to the sun's steady-state (quiet-sun) radio emission, the "disruptive" level of burst emission must be significantly greater than the quiet-sun flux-density levels which vary over the 11-year solar-activity cycle (Figure 11-5).

Since the quiet-sun flux densities are different at different frequencies, one cou'd choose a burst level greatly exceeding the quiet-sun level at any frequency covered in the climatology study (1000 sfu, for example). One could also choose a burst level at each frequency that is a fixed multiple of the mean monthly quiet-sun emission (for ex-

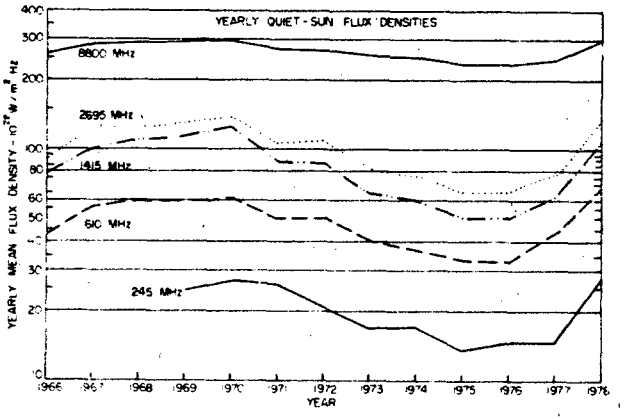


Figure 11-5. Yearly quiet-sun flux densities for several frequencies measured at Sagamore Hill.

Table 11-2a. Daily values of observed solar flux density at 2800 MHz recorded at the Algonquin Radio Observatory (ARO) of the National Research Council, Ottawa, Canada.

DAY	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
						1963						
01		87	74	73	82	84	76	87	74	68	87	78
02		86	75	74	82	81	77	87	73	69	85	78
03	77	85	78	74	81	81	78	87	74	70	83	77
04	79	84	80	70	82	79	78	86	75	71	83	76
05	77	83	82	72	84	78	78	86	74	73	80	76
06	77	85	85	78	87	77	77	85	74	73	78	76
07	77	83	84	80	86	80	77	85	78	75	76	77
08	76	82	83	81	86	80	77	81	77	75	75	78
09	78	79	82	82	88	93	77	80	77	86	76	78
10	80	79	80	82	87	95	86	77	76	87	75	78
11	81	76	78	82	84	103	75	72	72	87	76	80
12	78	74	77	83	87	109	74	73	77	84	77	82
13	79	74	74	83	89	107	76	74	80	84	77	81
14	86	75	80	87	85	100	77	71	80	84	78	81
15	85	76	80	88	80	96	78	72	80	85	81	81
16	82	77	79	87	100	89	78	74	105	87	81	82
17	82	79	80	87	100	86	74	80	99	84	80	82
18	80	81	80	88	98	82	74	80	97	83	82	82
19	78	79	77	84	95	78	74	79	103	85	84	82
20	78	77	77	78	91	75	77	81	105	86	84	82
21	76	74	76	74	88	73	75	84	90	84	86	82
22	75	76	76	72	89	72	73	86	105	84	86	82
23	74	75	75	71	93	72	72	90	99	84	84	82
24	73	76	75	73	88	72	72	87	95	84	83	
25	74	78	75	72	83	74	74	85	94	86	82	74
26	73	77	73	72	76	74	73	82	94	86	82	74
27	81	75	74	75	80	72	74	80	78	88	81	73
28	80	74	73	78	79	74	73	77	74	84	79	73
29	79		75	78	80	77	77	77	71	85	79	72
30	78		74	80	83	76	84	77	69	85	79	71
31	82		71		85		85	77		82		71
MEAN	78	79	78	79	86	83	76	81	85	84	81	77
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
						1964						
01		73	78	77	69	68	67	67	70	72	75	76
02	71	72	75	75	68	68	67	68	69	72	74	76
03	73	71	74	77	70	68	67	68	70	72	74	77
04	75	71	75	77	70	68	68	68	70	71	73	78
05	74	72	72	76	72	68	68	69	70	72	73	78
06	75	72	74	76	71	68	68	68	70	73	74	77
07	73	73	73	76	71	70	67	68	71	74	73	77
08	73	73	74	74	72	70	68	68	71	77	72	77
09	73	72	72	75	71	69	67	68	71	73	72	77
10	73	72	75	73	70	70	67	69	72	72	72	77
11	75	72	75	74	70	70	68	68	72	72	72	80
12	76	73	77	72	69	69	67	70	72	72	72	77
13	76	72	78	72	68	68	66	74	72	72	72	78
14	76	73	79	72	68	70	69	76	72	71	72	78
15	75	73	79	71	68	72	70	75	71	71	72	78
16	74	73	77	71	70	71	69	73	70	71	72	78
17	72	74	78	72	70	71	69	72	68	71	76	80
18	74	76	75	72	70	72	68	70	68	72	75	80
19	75	76	74	71	69	70	67	71	69	73	75	81
20	76	76	74	71	68	70	67	70	69	73	76	80
21	75	78	74	72	68	70	66	69	69	71	74	78
22	75	80	78	71	67	70	66	69	69	72	73	78
23	75	84	77	72	67	67	66	69	68	73	72	78
24	74	85	77	72	68	68	66	68	68	74	71	75
25	73	84	74	71	68	68	66	68	68	76	71	74
26	74	86	74	70	68	68	65	68	69	76	70	74
27	73	85	75	70	68	67	65	68	69	76	72	76
28	77	84	76	70	70	67	65	67	70	74	71	77
29	78	81	75	69	69	67	66	67	71	74	73	77
30	75		78	69	68	67	66	69	71	74	74	77
31	74		77		66		66	69		75		78
MEAN	74	76	76	72	69	69	67	69	70	73	73	78
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
						1965						
01	86	78	76	71	71	71	76	72	75	92	79	75
02	84	79	75	72	71	74	76	73	76	93	80	75
03	82	79	75	71	71	75	76	77	77	96	81	75
04	81	77	75	71	70	78	76	78	77	98	80	74
05	80	76	76	71	69	78	75	76	79	92	78	75
06	80	76	77	71	70	78	78	79	77	85	81	76
07	79	77	77	71	71	77	82	80	78	84	85	75
08	78	75	74	70	72	78	81	77	79	83	80	77
09	77	75	73	72	72	79	81	78	76	83	84	75
10	76	76	73	73	72	78	80	76	76	80	84	75
11	75	74	72	74	71	76	80	77	76	76	81	76
12	75	73	74	73	72	76	78	76	75	75	81	76
13	75	72	76	74	74	77	76	75	75	76	77	75
14	74	72	75	75	76	75	74	75	75	75	76	75
15	75	73	73	75	80	77	74	72	75	74	76	77
16	74	73	71	75	86	76	72	74	74	72	74	78
17	74	73	74	75	91	76	72	72	74	72	74	78
18	74	72	75	73	90	78	72	73	75	72	75	77
19	74	72	77	74	94	75	72	74	73	72	73	77
20	76	71	74	72	95	76	73	74	73	73	73	74
21	77	72	74	72	93	76	73	74	72	73	72	74
22	77	72	73	73	92	78	72	73	71	76	72	73
23	76	73	73	73	86	80	71	73	72	79	71	73
24	76	74	72	71	85	78	70	73	76	76	71	71
25	76	74	74	70	81	79	70	73	76	78	71	72
26	76	74	73	69	78	79	69	72	77	78	72	77
27	79	76	72	69	76	78	70	74	78	78	74	84
28	77	76	72	70	74	77	71	73	80	77	77	84
29	79		72	70	74	76	71	74	87	77	74	85
30	81		71	70	74	77	71	75	89	76	75	82
31	80		74		72		71	75		78		81
MEAN	78	75	74	72	78	77	74	75	76	80	77	76

Table 11-2b. Daily values of observed solar power flux densities at 2695 and 8800 MHz. Recorded at the AFCRL/AFGL Sagamore Hill Radio Observatory (SGMR): Hamilton, Massachusetts.

DAY	JANUARY 2695 MHZ	JANUARY 8800 MHZ	FEBRUARY 2695 MHZ	FEBRUARY 8800 MHZ	MARCH 2695 MHZ	MARCH 8800 MHZ	APRIL 2695 MHZ	APRIL 8800 MHZ	MAY 2695 MHZ	MAY 8800 MHZ	JUNE 2695 MHZ	JUNE 8800 MHZ	JULY 2695 MHZ	JULY 8800 MHZ	AUGUST 2695 MHZ	AUGUST 8800 MHZ	SEPTEMBER 2695 MHZ	SEPTEMBER 8800 MHZ	OCTOBER 2695 MHZ	OCTOBER 8800 MHZ	NOVEMBER 2695 MHZ	NOVEMBER 8800 MHZ	DECEMBER 2695 MHZ	DECEMBER 8800 MHZ
1966																								
1	74	255			72	262	97	278	79	250	93	257	86	247	103	261	96	270			86	259	87	247
2	77	253			69	252	92	273	81	253	90	249	84	247	103	260	90	267	92		88•	259	90	253
3	57				68	252	86	269	82	253	87	250	85	245	105	261	90	269	91		86•	258	92	256
4	72	258			69	251	85	265	82	256	86	250	87	255	101	258			93		86	260	102	259
5	77	256					85	263	80	250	85	250	86	266	93	256		263	92		87	264	115	274
6	75	240			66	252	94	263	77	250	95	253	98	288	89	251	82	256	96		93•		116	273
7	74	255			67	249	93	268	80	252	83	251	94	281	85	243	84	255	94		100	288	117	273
8					67	250	92	270	76•	236	86	251	94	272		252	84	256	89		107	285	119	271
9	74				68	252	88	266	74•	228	85	250	88	257	81	248	83	253	93		109	279	128	282
10	73	249			68	252	84	262	76	250	84•		91	255	79	248	84	253	95		112	278	146•	277
11	75	255			69	253	82	265	77	254	82	249	89	256	78	248	87	255			116	278	151	290
12	78	256			71	254	80	262	77	252	80	249	84	247	79	251	91	263			122•	277	148	275
13	81	261			73•	244	77	251	79•	240	84•	236	82	246	82	252					117	280	155	276
14	84	267			71	254	79	254	79	254	81	248	83	248	79	248			103		119	285	149	277
15	93	265	76	253	78	258	82	253	81	260	81	244	84	249	85	248			107		117	283	142	271
16	101	265	75	257	80	259	81	251	83	260	81	250	86	248	85	253	109		107•		114	280	136	267
17	96	267	73	262	82	271	83	252	82	254	86	248	84	248	83	250	114		109		109	278	119	263
18	101	268	75	260	93	271	79	251	81	257	80	249	83	249	83	256	120		108		106	273	105	261
19	101	268	76	262	90	270	75	248	89•	250	79	248	85	249	84	256	128		107•		106	269	110	266
20	94	265	75	264	94	270	81	253	93	264	80	249	89	253	87	257	125		108		104	257	106	259
21	88	264	75	268	99	278	79	250	94	271	81	244	86	252	88	258	120		109		104	257	101	261
22		264	77	268	89	262	82	251	96	276	82	248	91	254	94•	258	114		108		107	259	101	262
23	84	249	72	258	84	260	88	259	91	264	84	248	96	261	102•	262	115		95		107	258	102	265
24			76	258		260	90•	259	90	262	84	249	100	270	109	271	111		92		103	255	104	259
25	87	249			75	251	90	259	89	258	96	258	106	268	110	277	104		90		103	252		
26	82	255	74	268	74	255	91	259	92	263	95	256	110	271	112	278	100		87		104	254	108	257
27	78	252	73	260	74	252	86	254	89	257	82	248	107	268	115	287	103		82		108•	264	105	254
28	79	229	74	263	78	262	85•	254	87•		82	249			119	286	96		84	261	100	254	108	256
29	72	241			88	272	84	256	88•	249	84	248	112	274	108	280	88		89	267	92	258	103	256
30					88	277	79	250	85	252	85	249	109	268	107	280	84		92	267	92	254	112	263
31					95	291			92	255			116	268	102	278			90	258			116	262
MN	82	254	75	262	78	260	85	259	84	255	85	249	92	259	94	261	100	280	96	263	104	267	116	266
1967																								
1	120•	272	148	295	184	344	144	288	113	269	143	280	104	255	147	288	133	294	114	285	118	285	128	296
2	129	288	140•	292	188	347	117	280	114	269	128	282	103	255	135	281	130	291	108	273	120•	287	120	286
3	154	297	132	290	183	335	112	273	109	260	119	278	109	257	132	277	122	281	107	274	118	285	106•	279
4	197•	299	134	287	189	342	111	268	106	262	124	275	107	251	126	284	118	278	104	270	113	285	109	283
5	164	301	136	296	174	324	105	267	108	260	111	265	104	248	127	285	108	278	103	273	107	279	106	277
6	157	291	150	298	165•	313	108•	265	107	260	101	258	100	248	127•	280	107	268	107•	273	104	275	111	280
7	152	292	163•	300	150•	297	115•	268	105	260	93	251	97	246	115	274	105	266	104	271	99	274	115	281
8	134•	287	146	283	141	288	117	274	100•	248	87•	248	89	240	119	270	107	275	101	270	98	271	116•	278
9	126	289	140	288	140	282	116	264	96•	253	88	242	93	244	112	268	106	275	99	273	95•	275	116	283
10	142	283	137•	283	128	277	109	265	91•	244	79•	237	92	243	112	262	104•	277	104•	280	101	280	115	280
11	132	280	129•	283	120	270	115	269	88	249	79	240	92	240	112	266	105	278	116	280	110	281	125•	288
12	130	280	132	285	117	269	112	268	88	252	82	242	90	240	118	268	105	278	115	283	112	288	125•	290
13	120	274	127	280	114	275	107	266	85	249	85	249	93	248	118	265	103	276	109	275	117	289	128	294
14	128	272	120	282	111	271	106	275	89	255	88	248	104	259	106	263	101	275	107	278	121	288	136	303
15	116	264	112	279	119	278	110•	268	93•	251	91	251	109•	257	99	257	104	278	100	275	126•	289	147	315
16	116	262	106•	279	118	278	112•	269	94	255	92	253	106	259	103	261	97	269	98	274	134	299	173•	329
17	108	259	114	280	122	282	109	265	96	257	94	253	105	258	115	264	98	271	96	273	140	304	177	332
18	111	262	116	282	122	282	111•	262	103	260	100	253	111	262	123	279	102	263	96	272	140	307	198	355
19	128	266	111	281	121	280	113•	278	115	269	104	257	105	262	137	291	103	267	99	273	148	315	186	349
20	127	267	116•	283	123	279	107	268	120	275	104•	260	113	267	134	287	106	267	107	276	149	312	182	339
21	132	272	118•	285	129	288	111	270	142	297	108•		113	271	145	293	104	264	105•	271	142	306	173	337
22	135	276	132	299	137•	298	117	280	161	329	111	264	121	285	141	287	101	264	115	283	139•	305	159	322
23	139	281	136•	295	144•	296	119	285	171	348	119•	262	128	293	148	291	100	267	114	282	140•	298	144•	294
24	135	278	139	323	140	293	113•	284	177	337	107	244	141	307	146	295	108	274	113	279	134	298	129	290
25	138	282	152	329	142	287	113	270	184•	343	105	249	134	330	135	294	110	274	118	287	132•	285	145	299
26	139	289	158	337	152	294	106	264	201•	343	103	249	151	332	136	294	112	273	124	295	142	305	149	315
27	145	298	158	331	148	291	110	264	187	306	107	247	160	330	136	296	114	280	131	312	142	302	156	315
28	146	298	153	344	158	299	115	268	174	303	112	257	179	325			110	278	145	319	139	308	164•	326
29	151	302			162	316	115	269	161	294	112	257	179	308	141	303	107•	272	137	314	135	297	157•	313
30	155	298			155	308	112	268	151	289	114	255	162	301	130	295	109	272	136	306	128	298	144	302
31	147	294			142	298			153	290			149	291	132	295			129	298			142	309
MN	137	283	134	295	143	296	112	269	125	279	103	255	118	271	127	280	108	275	112	273	125	293	141	304
1968																								
1	152	337	205	466	137•	294	123	281	115	279	125	275	104	260	107	269	106	268	114	278	137	303	123	294
2	151	332	200•	391	138	295	120	280	122	285	121•	275	96	259	109	267	111	270	113	280	135	302	131	298
3	195•	337	194	370	132	292	114	279	128	282	121	275	95	257	113	280	117	276	120	286	126	292	127	295
4	164•	354	179	348	120	284	110•	275	129	286	121•	269	94•	256	110	278	119	278	125	283	121	280	126•	276
5	174	357	161	333	119	286	108•	275	127	281	115	269	96	265	109	280	116	277	125	281	115	276	135	293
6	181	362	144	307	119	286	104	264	122	281	121	279	98	269	117	286	113	279	125	285	109	275	126	275
7	192	370	143	302	111	275	104	269	117	280	129	280	106	276	112	288	115	278	123•	268	114	280	123	276
8	212•	364	134	294	100	272	111	274	117	281	126	282	112	293	118	293	124	285	121	282	118	284	122	277
9	209	353	134	298	103	276	116•	284	115	282	122	281	120	286	115•	280	124	283	115	273	120	289	128	282
10	191	351	137•	305	102	274	116	298	113•	282	124•	279	127	285	118	281	128	287	114	276	116•	273	131	287
11	190	348	138	302	113	279	115	284	107	275	124•	283	132	282	120	292	127	275	113	279	117	285	127	283
12	186	342	137	297	111	284	117	287	105•	272	114•	275	135	290	130	295	130	281	117	282	116•	286	122	276
13	182	340	129	289	110•	282	119	297	104	268	118	271	125	284	140	303	126	280	110	275	115•	284	113	272
14	174	337	126	284	110•	281	115	288	111	281	116	275	125	286	152	319	123	277	110	269	111	275	116	275
15	161	332	124	287	105	276	116•	283	113	284	113	278	122	275	148	311	112	273	117	277	116	287	119	276
16	152	321	122	285	107	274	116	284	122	287	114	286	123	281	146	305	108	269	117	281	112	289	112	267
17	150	313	118•	280	114	277	109	281	137	295	115•	276	115	278	139	292	108	270	123	290	124	294	119	276
18	144	313	117	281	110		106	288	145•	299	120	280	107	269	135	288	109	270	121	290	127•	289	119	273
19	135	305	118	285	112	277	107	278	152	305	119	285	140	269	131	282	105	268	131•	294	122	279	126	278
20	127	297	119	293	110	271	101	268	157•	305	124	289	169	266	134	282	105	266	124•	299	121	281	131•	275
21	122	297	126	296	120	276	102•	270	156	306	128	289	167	265	132	286	111	274	139	297	118	278	132	279
22	119	300	129	294	122•	277	97	262	151	312	124	289	112	270	133	285	104	267	143	299	115	278	133	284
23	120	302	131	294	127•	282	95	260	148	305	132	302	117	265	112	267	103	267	144	295	115	279	131•	275
24	121	314	141	300	136	285	100•	261	149	305	129	292	123	273	106	267	117	275	150	295	117	281	139	280
25	134	323	154	309	142	291	100	265	143	302	128	292	129	279	102	266	135	294	151•	296	119	288	142	288
26	144	329	155	309	137	280	94	259	142•	298	120•	277	124	275	98	265	124	286	152	303	118	291	141	286
27	154	334	148	304	141	285	97	265	138	289	123•	281	118	271	95	257	130	279	156	310	121	285	149	294
28	164•	386	149	307	125	286	102	266	128	287	119•	272	118	275	98	256	132	286	151•	310	115	279	138•	294
29	187•	386	142	301	134	286	105	275	126•	279	114	268	121	287	100	266	135	292	166	341	115	278	127	288
30	202	415			136•	290	107	272	126	278	107	264	112	274	104	267	119	292	149	312	115	284	128	291
31	204	417			134	286			126	277			109	267	101	266			129	306			124	284
MN	165	342	145	309	121	282	108	275	129	286	121	275	114	274	119	281	118	277	130	296	118	284	128	282

Table 11-2b. (Continued)

DAY	JANUARY 2695 MHz	JANUARY 8800 MHz	FEBRUARY 2695 MHz	FEBRUARY 8800 MHz	MARCH 2695 MHz	MARCH 8800 MHz	APRIL 2695 MHz	APRIL 8800 MHz	MAY 2695 MHz	MAY 8800 MHz	JUNE 2695 MHz	JUNE 8800 MHz	JULY 2695 MHz	JULY 8800 MHz	AUGUST 2695 MHz	AUGUST 8800 MHz	SEPTEMBER 2695 MHz	SEPTEMBER 8800 MHz	OCTOBER 2695 MHz	OCTOBER 8800 MHz	NOVEMBER 2695 MHz	NOVEMBER 8800 MHz	DECEMBER 2695 MHz	DECEMBER 8800 MHz
1969																								
1	130•	294	120•	290	148	309	160	335	108	285	94	258	123•	277	149	293	136	273	114	270	126	289	118	29•
2	130	294	126	293	138	290	163•	332	108	278	100	262	134	287	160	302	130	273	121	276	121	285	122	276
3	135	301	128	287	132•	293	167	325	113	274	110	270	144	301	155	293	128	272	1.5•	273	113•	278	127	275
4	144	302	127	275	131	298	154	308	114	271	126	289	139	298	161•	299	127	283	111	267	113•	271	129	279
5	152	304	131	286	130	297	148	306	131	282	161	317	138	291	154	295	116	268	120	276	112•	267	128	293
6	164	316	129	289	127	292	141	297	116	272	165	331	137	290	144	283	114	271	123	278	116•	272	122	281
7	175	328	131	291	123	289	133	290	112	271	178	329	136	285	135	279	107	268	126	286	116•	264	111	271
8	169	323	124	285	126	290	128	283	110•	273	200	342	137	288	128	274	104	259	124•	278	115	270	106•	257
9	174	340	122•	287	127	291	125	287	115•	271	195	351	136	282	122	272	100•	244	120	280	113	265	104	261
10	160	326			123	285	126•	289	117	274	200	369	134	280	116	264	98	257	116	283	107	262	103	265
11	156	321	119	284	123	285	128	288	129	278	200	369	131	276	112	267	99	258	110	276	106	271	104•	269
12	151	311	112	282	125	286	136	289	127	288	188	353	122	272	107	266	101	261	111	276	105	267	113	279
13	148	309	115•	278	128	284	151	304	128	279	186	353	121•	267	104	255	112	269	107	272	111	265	119	285
14	145	306	114	279	124	289	153	302	1 •	282	182	337	111	258	97	257	109	265	105	267	121•	270	121	282
15	141	308	116	274	143	318	158	315	136	283	164	314	106	257	97	252	115	272	105	263	123	278	120	285
16	148	302	117	279	158	320	142	287	132	280	144	298	105	256	98	248	115	264	96	263	125	296	127	285
17	141	299	123	283	181	337	134	282	136	289	132	280	103	255	98	247	113	267	106	273	145	318	132	288
18	135	290	127	282	182	332	132	281	140	292	126	275	99	258	88	246	117	269	111	276	161	358	134	296
19	121	287	135•	292	186•	340	130•	279	134	296	119	273	98	253	85	244	112	265	125	293	165	360	135	297
20	119	280	146	296	194	356	133	286	136•	297	117	271	96	250	91	250	117	273	141	306	165	344	142	295
21	122	279	154	294	207•	356	133	289	145	298	115	267	97	253	96	253	119	278	151	321	176	361	150	297
22	121	280	168	302	197	356	130	290	154	293	111	266	94	244	101	260	118	273	160•	320	186	362	149•	288
23	119	281	187	322	186	341	121	279	149	287	109	262	96	252	110	269	125	277	172	329	186	336	147	37•
24	120•	284	183•	333	174	335	124•	281	150	285	107	264	95	254	121	282	134	282	178	326	176	339	145	296
25	124	289	181•	338	162•	342	122	280	146	293	99	295	98	254	131	281	145	292	164	328	176	348	143	296
26	135	297			174	350	121	277	140	291	97	259	99	257	131	287	141•	282	180	330	158	325	149	299
27	132	288	176	332	188	338	117	268	130	288	94	252	105	258	140	289	130	277	180	325	159	319	135•	285
28	122	286	161	324	191	341	113	267	121	276	98	257	116	267	147	299	122	274	166	320	144	308	138	287
29	122	286			159	351	108	261	102	264	102	264	116•	265	151	289	118	277	155	312	128	285	140	289
30	116•	278			164	339	109	267	97	262	115	271	120	271	144	280	120	277	144	305	124	277	147	290
31	115	287			162	336			93	260			139	285	139	278			136	293			144•	282
32	130	299	137	294	153	318	135	291	127	281	130	295	117	269	122	272	110	271	133	292	136	300	129	284
1970																								
1	141	298	122	295	161	320	145	302	135	291	125	274	149	301	114	264	130	289	110	265	146	334	140	279
2	135	292	124	292	159•	318	149•	298	141	297	121	274	158	307	109	265	132	288	105	262	143	329	144	284
3	127	278	113•	285	151	305	145	305	138•	287	114	269	166	315	110	260	137	290	104	267	140	313	141	284
4	126	279	111	278	153	309	149	308	139	298	111	267	165	313	107	266	144•	302	118	280	133	307	154•	291
5	117	271	115	281	155•	277	164	324	142	297	113	269	160	303	107	264	148	302	113	264	132•	299	155	300
6	110	272	122	289	156	304	174	336	139	298	104•	255	156	304	107	262	134	288	111	268	138	298	158	303
7	108	271	128	287	156	295	195	365	140	299	117	275	141	284	107	261	139	286	114	270	137	297	153	302
8	107	267	141	307	159	306	181	339	135	292	114	269	138	285	110	265	132	279	121	277	128	291	153	295
9	115	272	142	312	149	300	188	343	133	285	117	268	125	276	117	267	134	283	128	277	130	297	157	324
10	130	284	140	347	142	299	197	365	136	290	124	274	129	288	115	262	122	274	126•	277	136	305	148	297
11	150	308	174	350	144	296	177	332	148	303	144	297	119	268	105•	252	115	269	122	273	144	312	160	301
12	164	333	184	338	151	299	174	331	154	305	143	301	110	265	120	279	106	263	119	270	152	316	168	310
13	164	298	189	374	140•	302	164	320	165	317	169	330	109	264	130	297	103	265	113	265	169	333	148•	340
14	167	313	187	333	135	289	153	310	169	322	177	338	107	260	155	326	102•	233	114	270	177	372	141	301
15	178	321	185•	327	131	286	141	301	170	324	176	330	102	257	143	287	100•	239	120	265	179•	367	136	293
16	166	308	181	334	124	287	131	295	175	322	163	312	101	260	140	298	100	252	117•	269	189	426	132	296
17	171	308	188	344	120	283	128	292	165•	302	162	307	103	262	151	281	98	258	113	271				
18	175•	305	179	333	119	288	122	292	164•	286	151	297	109	269	124	282	102•	259	120	275			142	301
19	168	309	185	334	116	281	117	279	166	301	149	293	113	277	123	281	114	271	120	275	181	393	143	298
20	165	309	182	327	127	282	114•	274	156	284	137	287	116	287	123	276	123	274		275	163	324	143	301
21	172	320	192	336	135	284	111	270	148	290	134	286	136	298	119	270	127	272	120	279	168•	332	135	304
22	157	304	173	326	138	289	111	273	144	288	128	278	134	292	118	270	122	277	124	283	146	287	127	297
23	143	288	178•	331	140	285	112	274	139	287	122	270	142	292	124•	269	139	279	129•	285	134	291	121•	254
24	133	289	175	332	151	287	110•	274	139	292	121	279	139	284	117	270	135	284	140	299	126	282	112•	281
25	142	288	169	326	150	293	115	280	137	292	126	273	143	285	115	268	132	278	135	312	121	276	113	290
26	154	305	171	327	150	291	114	278	140	290	123	279	150	290	116	268	133	282	169	319	122	278	111	284
27	151	301	162	311	139	289	118	278	133	282	131•	286	162	293	115	274	127•	277	168	310	132	290	107	279
28	151	305	160•	312	137	282	120	283	132	282	133	294	153	293	119	278	126	283	166	303	140	287	107	276
29	152•	308			140•	291	133	291	134	286	145	291	134	279	124	282	118	272	167	301	143	291	113	279
30	151	304			133	286	130	286	139	290	146•	289	132	276	126	287	114	267	156	294	147	291	120	289
31	151	291			131•	277			140	287			121	269	131	291			150	287			119	287
32	146	295	160	319	142	293	143	302	146	296	135	286	139	283	120	275	123	274	128	280	146	313	137	292
1971																								
1	119•	286	153	304	117	285	83	264	84	243	84	261	124	271	94	254	79	247	100	264	99	273	110	277
2	121	288	149	315	107	279	95	269	87	254	84	251	118	273	93	254	79	248	101	264	99•	273	107	273
3	124	294	138	308	104	281	94•	261	93•	246	83•	235	110	260	91	254	80	246	107	269	102•	274	104	277
4	127•	302	139	301	106•	273	96	268	99	258	84	236	115	276	90	252	82	248	97	269	104•	272	99	280
5	132•	301	137•	301	96	278	97	269	104	271	84	236	114	278	91	255	81	246	96		106	273	107	279
6	134	305	132	309	97	276	97	269	111	280	82	238	109	270	91	252	84•	248	93	263	101	269	106	284
7	138	306	124	294	90•	272	101	268	114	281	83	238	102	276	89	255	82	250	91	259	98	264	109	286
8	136	305	114	285	89	274	98	268	115	275	83	252	99	269	97	257	83	253	89	260	98•	268	107	281
9	141	305	105	273	85	273	97	268	116•	274	82	252	92	258	96	254	91	254	91	259	98	264	109	286
10	139	314	102	273	90	274	98	267	122	271	84	253	92	247	93	254	85	250	91	265	93	264	112	293
11	137	306	104	273	92•	246	110	276	122	276	82	247	94	261	91	255	85•	246	86•	264	97	269	115	290
12	139	308	105	274	99	278	117	276	114	271	81	252	94	265	88	253	79	246	84	258	97	268	115	293
13	134	302	108	276	98	271	126	282	111•	265	80	250	98	262	88	254	79•	241	85	261	98	267	117	291
14	141•	312	108	275	96	272	127•	280	107	264	82	251	98	267	86	254	86	244	84	255	94	269	118	287
15	140	318	110	277	92•	265	127	283	104	261	79	249	106	276	85	253	86	252	91	254	93	271	113	287
16	148	317	121	286	96	272	126	281	102•	256	78	250	106	275	85	251	102	259	78	249	96	270	112•	288
17	141	319	125	293	96	277	127	280	102	257	80	250	103	268	86	253	101	257	84	251	98	272	117	289
18	144	319	120	297	95	273	127	280	98	255	82	247	104	271	97	254	104	255	89	251	98	276	123	296
19	150	325	132	295	91	271	123	280	98	257	82	252	111•	267	109	264	99	260	97	267	97	278	124•	293
20	159	314	124•	277	91	274	122	279	96	252	81	252	101	265	123	289	98	257	110	273	100	281	124	294
21	164	327	126•		92	270	114	269	90	242	82	247	95	268	129	293	99	252	108	285	103	279	126•	292
22	165	321	134	296	93	273	108	268	76•	243	84	247	99	270	130	277	97	250	110	282	106	284	129	29•
23	159•	321	132•	288	89	266	109	279	76	247	82	244	104	279	129	282	97	259	109	286	113	287	131	295
24	163	320	133	293	90	270	98	275	75	246	84	246	114	281	130	289	96	259	109	28•	113	286	129	29•
25	154	319	130	297	91	272	92	258	79	246	85	253	105	269	124	273	97	264	116	291	116	29•	122	29•
26	151•	310	121	287	90	264	92	260	84	249	100	255	104	271	118	271	97	268	114•	286	116•	289	122	264
27	152	314	119•	275	90	268	89	259	91	257	105	262	101•	267	119•	264	103	275	105	276	12	284	115•	281
28	152	314	117	295	87	265	87	251	94	256	113	273	98	262	109•	257	103	274	99	266	124	284	11	278
29	146	303			85	262	83•	248	94	259	122	274	97	260	97	259	101	270	96	265	114	276	11	279
30	149	312			81	262	83	243	94•	259	126	279	106•	269	89	254	101	267	98	269	114	275	11	282
31	150	310			85	269			95	259			96	256	83	249			98	272			11	286
32	144	310	124	290	93	271	105	269	98	259	91	254	103	268	101	262	92	255	96	267	114	276	114	286

Table 11-2b (Continued)

DAY	JANUARY 2695 MHz	JANUARY 8800 MHz	FEBRUARY 2695 MHz	FEBRUARY 8800 MHz	MARCH 2695 MHz	MARCH 8800 MHz	APRIL 2695 MHz	APRIL 8800 MHz	MAY 2695 MHz	MAY 8800 MHz	JUNE 2695 MHz	JUNE 8800 MHz	JULY 2695 MHz	JULY 8800 MHz	AUGUST 2695 MHz	AUGUST 8800 MHz	SEPTEMBER 2695 MHz	SEPTEMBER 8800 MHz	OCTOBER 2695 MHz	OCTOBER 8800 MHz	NOVEMBER 2695 MHz	NOVEMBER 8800 MHz	DECEMBER 2695 MHz	DECEMBER 8800 MHz
										1972														
1	96	263	101	257	116	276	84	246	83	244	107	259	115	243	136	310	126	272	161	267	117	30[illegible]	84	271
2	93e	264	101	260	118e	277	86	250	82	238	114	265	118	247	132	312	126	270	97	262	112e	287	81	269
3	84	252	96e	256	118e	276	87	255	84e	237	122	271	121	244	128	305	110e	269	93	261	102	274	77	264
4	92e	258	93e	257	123	280	88	257	87e	251	132	277	131	254	128	304	109	257	91	258	95	275	75	262
5	84e	256	84	255	127e	281	84	256	84	250	143	279	139e	264	128	307	108	266	88	255	90e	27[illegible]	71e	265
6	88	252	84	256	121	281	105	266	102	256	136	277	127	255	121	294	104	263	88	256	84	262	71e	262
7	88	259	85	261	120	281	107	267	103	252	131e	276	125	253	136e	302	96	257	83e	254	77	254	8[illegible]	268
8	94	254	100	263	126e	285	107	266	115e	258	129	286	117	246	108	274	101	264	86	257	75e	251	86	272
9	91	255	98	260	123	275	116	266	113e	259	130	287	111	245	106	278	100	265	89	264	75	253	88	274
10	85	251	105	262	123	275	115	264	121	262	120	278	105	238	102	271	93	263	84	260	75	253	90e	268
11	88e	253	112	264	122	274	116e	261	124	275	129	277	109	247	96	264	89	259	84	256	80	259	99	268
12	92	253	114	271	116	273	115	260	129	277	129	276	104	245	92	253	87	254	84	256	8[illegible]	260	102	273
13	85	254	118	273	114	270	113e	255	140	288	121	257	104	240	89	247	84	252	83	261	8[illegible]	261	105	282
14	104	258	126	281	117	273	111	256	146	297	119	252	104	244	86	240	86e	254	85	261	75e	259	109	282
15	111	264	134	286	122e	273	111	252	146	288	139	270	100	245	85	242	90	256	9[illegible]	27[illegible]	77	262	108	285
16	115	265	142	293	119	281	109	256	145	284	121	259	96	239	86	249	89	256	94	27[illegible]	81	265	103e	289
17	107	260	141	309	118e	281	107	261	147	284	126	257	99	237	93	245	88	253	101	276	82	26[illegible]	104e	285
18	104	257	181	324	118	273	106	255	142	280	121	248	95	237	94	255	89	254	99	277	86	261	103	287
19	111	265	190e	329	116	273	105	256	134e	274	126	255	95	234	100	259	92e	262	106	287	92	262	106	286
20	114	263	189e	319	122	275	96e	256	129e	270	120	251	96	234	101	258	98	266	108	29[illegible]	95	269	102e	279
21	121	266	184	306	121	276	93	244	127	268	123	256	95	234	102	262	104	272	118	303	100	275	107e	277
22	118	270	167	307	120e	276	97	248	125	263	118e	256	95	234	112	264	104	272	125e	294	103e	282	99e	271
23	127	276	164	298	127	278	97	247	118	269	116	249	96	243	111	265	110	273	133	325	107	286	98	264
24	128	277	153	293	111	267	98	251	117	275	110e	249	98	246	108	265	112	275	137	337	110	288	93	263
25	122e	272	141	287	103	261	98	255	109	277	107	248	97	248	112	263	112	274	135	344	102	285	97	269
26	119	266	141e	287	99	260	98	255	100	268	103	244	106	255	124	273	113	272	153	357	99e	28[illegible]	91	264
27	114	265	126	276	92	259	96	262	98	256	107	244	109	262	126e	269	107	270	15[illegible]	36[illegible]	94	272	9[illegible]	265
28	113e	265	118	273	89	259	90	25[illegible]	100	269	112	247	117	275	133	277	104	269	15[illegible]	359	89	271	86	266
29	109	264	118	277	85	249	86	252	100	262	114	247	125	293	146	287	104	264	133	325	88	271	86e	270
30	110	270			86e	252	83	250	103	256	114e	247	128	304	133	276			128	302			88e	271
31	107	259			85	249			105	260														
32	105	262	130	281	114	272	100	256	115	266	122	262	109	290	113	273	101	264	108	29[illegible]	96	269	92	273
										1973														
1	86	245	79	263	88	264	97	266	98	274	72	259	74	242	73e	238	98	265	97	265	79	256	81	261
2	89	252	77e	261	89	263	105e	269	101	275	71	247	77	247	72e	235	111	275	94	265	77	257	81	266
3	87	260	84	262	87	265	113	279	101	270	69	243	77	247	72	231	116	271	92e	259	74	252	76	257
4	100e	260	84	262	85e	265	106	271	104	274	69	243	78	243	73	236	119	277	85	252	71	251	74	252
5	103	259	85	263	84	263	102	269	99	265	71	248	86e	244	72	240	120	269	80	248	67	252	70e	248
6	106	270	88	269	87	267	99	263	94	254	71	245	83	247	74	239	114e	263	78	252	67	250	66	248
7	102	265	89	272	91e	274	95	265	89	257	72	242	84	249	76	246	109	256	75	250	66	251	66	249
8	94	261	91e	270	87e	255	106	280	83	251	73	247	84	250	74	242	101	252	71	251	65	251	67	250
9	93	264	92	273	92	264	109	277	91e	247	74	247	86	257	73	239	100	248	70	251	64	252	62e	247
10	94	259	95e	274	97	270	102e	266	79	247	76	249	84	249	70	243	96	249	70	247	64	252	68	245
11	93	263	89e	274	102	271	99	269	73e	243	78	247	78	238	67	238	88	239	67	246	65	252	67e	25[illegible]
12	93	268	94	275	101	261	90	261	73	241	78	250	71	241	65	233	79	240	64	242	66	249	69	249
13	93	269	92	273	105	268	87	265	71	243	81e	249	67	235	63	233	74	244	63	242	67	248	69	249
14	97	258	86	275	106	274	81	260	76	245	82	249	66	236	64	233	74	244	65	245	68	245	69e	251
15	86	258	86e	273	96	268	76	259	73e	244	83	250	68e	236	60e	236	73	242	67	247	70	247	72	251
16	83	253	89	268	94	265	73	254	77	243	80	248	68	240	62	239	73	241	67	248	70e	249	75	260
17	85	255	87	264	88e	260	71	249	78	242	76	246	68	240	63	238	72	247	72	251	69	253	77e	264
18	88	255	81	265	89	263	69	249	81e	247	77	244	70	241	64	238	72e	247	70	25[illegible]	70	254	81	265
19	85	256	77	265	88	271	72	246	85	256	83	253	68	242	64	239	76	247	68	247	70	253	8[illegible]	261
20	87	254	80	260	80	262	77	250	88	251	89	251	66	238	65	237	79	247	69	252	72	256	83	265
21	88	258	78	257	77	259	83	254	87	249	92	247	66e	238	70	243	84	259	72	252	76	256	89e	272
22	88	261	79	262	80e	259	90	256	88	250	91	253	65	238	70	243	87	256	76	258	78e	251	92	269
23	88	257	81e	262	77	258	94	262	81	246	90	250	65	237	71	243	92	254	79	257	81	256	92	274
24	87	259	88	264	78	257	96	264	87	252	88	248	68	242	73	247	98e	257	90	264	8[illegible]	255	93	273
25	87	262	82	262	78	255	95	264	86	255	84	247	69	243	77	248	100	262	93	268			91	271
26	85	257	86	265	74e	256	94	265	84	254	82e	245	71	243	78	250	102	263	93e	268	85	258	84e	263
27	83	260	87	270	88	263	98e	274	84	255	78e	244	72e	242	81	251	102	267	96	278	85	253	8[illegible]	264
28	82e	257	90	265	83	264	97e	279	80e	250	79	244	70	241	84e	253	104	267	92	270	84e	256	75	256
29	80e	261			86	264	96	277	78	240	78	244	69e	245	87	256	104	266	91	263	89	256	71	251
30	81	258			90	263	94	275	76	241	81	247	71	242	87	256	103	267	93	264	93	260	67	249
31	79	262			97	265			73	242			73	240	95	265			84	260			66e	248
32	90	259	86	267	88	263	92	269	84	252	79	247	73	242	72	242	94	256	79	259	73	253	76	257
										1974														
1	69e	249	65	249	69	243	84	244	84	252	80e	246	89	264	71	240	67		82	252	8[illegible]	255	73	259
2	67	251	69e	256	69	246	63e	232	100	264	84	245	95	275	71	241	68e		82	254	84	261	68e	256
3	68	252	68	252	71	248	62	234	100	261	82	242	99	274	70	239	65e		83	255	87	264	69	254
4	66	256	66	252	69	240	66	237	102	256	80	245			72	241	69e		89	263	88e	256	68	252
5	66	249	70	250	71	242	67	236	102	257	79	239	107	287	71	241	69	236	93	274	86e	256	67	252
6	69	253	71	251	69	243	70	244	106	261	78	245	102	269	74	243	70	232	99	262	85	257	67	249
7	69	255	72e	252	68	243	70	244	103	256	77	242	95	261	72	243	73	232	1[illegible]	262	78e	252	64	65
8	74	259	74	257	69	249	71	242	100	257	91	256	86	256	75	247			112	274	75e	248	64e	245
9	76e	265	75	256	72	251	78e	251	96	255	77	241	78	248	74	243	82	250	113	274	7[illegible]	251	62	243
10	81e	266	75	254	76e	251	82e	256	89e	253	78	247	73	242	75	245	86	255	116e	272	7[illegible]	256	65	254
11	84e	265	73	252	77	254	8[illegible]	256	90	251	80	244	70	242	79	251	9[illegible]	264	127	297	69	252	69	25[illegible]
12	86	265	70	247	78	255	93	262	83e	250	74	245	70	238	83	256	9[illegible]	267	121	291	72e	254	71	24[illegible]
13	87	256	70	247	75	254	91e	267	79	244	73	243	69	237	86	251	98	275	113	28[illegible]	68	251	74	256
14	88	256	70	250	75	250	99	277	73	245	72	241	69	243	77	247	9[illegible]	265	10[illegible]	268	71	252	c[illegible]	261
15	87	249	76	259	73	252	107	295	66	239	69	238	71	246	76	250	88	261	93	256	73	255	77	262
16	84e	253	77	256	75	254	99	284	67	237	71	239	77e	250	81	246	92	274	79e	254	72	261	82e	261
17	83	259	77	254	74	251	88	267	63	232	71	236	76	251	84	254	86	250	74	248	77	265	81	260
18	81	255	78	256	73	251	83	262	62	233	70	231	71	245	8[illegible]	249	9[illegible]	251	71	25[illegible]	8[illegible]	265	80	267
19	80	249	74	253	70	247	77	255	61	236	66	234	73e	245	74	25[illegible]	8[illegible]	245	67	249	87	27	84	271
20	78	254	73	245	70	247	71	247	59	236	66	232	73	243	70	246	74	242	67	247	87e	267	81	26[illegible]
21	74	247	76	252	67e	247	67	241	60	238	67	228	75	252	70	237	72e	233	66	248	89e	272	79	262
22	78	257	78	248	71	248	63	238	59	237	68	229	80	251	67	24[illegible]	69		65	245	94	28[illegible]	82e	267
23	76	244			71	248	62	235	60e	236	68	236	82	252	69	238	65		64	245	88	274	77	256
24	72	246	73	248	67e	246	62e	240	60	235	66	235	78	248	65	233	64	234	66	248	88	266	73e	256
25	67	241	73	249	66	245	65	242	62e	232	67	236	75	243	64	239	64	235	69e	245	85	265	71e	255
26	64	238	68	247	63	240	64	239	68e	229	67	237	78	245	62	235	65	237	72	252	83	272	69	255
27	65	236	70	244	65	244	64	242	70	246	69	235	74	242	62	233	67	237	71	254	83	265	67	257
28	64	236	70	246	62	245	65	236	74	245	70	238	71	235	61e	23[illegible]	69	235	72	255	8[illegible]	261	68	252
29	65	242			62	245	68	241	74e	246	76e	241	70	25[illegible]	57e	229	77	245	72	252	76	259	69	251
30	66	245			61	239	76	250	75	246	79	244	71	238	61	226	79	249	73	248	76	261	68	25[illegible]
31	63	243			64	238			79	247			73	244	64	236			81	257			7[illegible]	252
32	74	251	72	251	70	247	75	250	78	246	74	240	79	250	71	242	77	248	86	259	8[illegible]	261	72	256

Table 11-2b. (Continued)

DAY	JANUARY 2600 0800 MMI MMI	FEBRUARY 2600 0800 MMI MMI	MARCH 2600 0800 MMI MMI	APRIL 2600 0800 MMI MMI	MAY 2600 0800 MMI MMI	JUNE 2600 0800 MMI MMI	JULY 2600 0800 MMI MMI	AUGUST 2600 0800 MMI MMI	SEPTEMBER 2600 0800 MMI MMI	OCTOBER 2600 0800 MMI MMI	NOVEMBER 2600 0800 MMI MMI	DECEMBER 2600 0800 MMI MMI
1975												
1	72 246	65 253	64 232	64 229	66 230	65 214	65 231	75 240	70 239	66 235	63 243	66 247
2	70 248	67 251	65 226	63 227	64 243	59 218	63 228	77 242	74 240	67 238	64 238	65 245
3	71 251	68 250	67 226		68 235	61 222	63 229	82 250	75 241	67 240	64 240	67 244
4	68 244	70 248	68 234	65 226	69 232	61 224	63 231	88 250	78 241	67 238	67 242	68 247
5	69 248	72 244	68 234	65 228	68 235	60 224	62 231	89 251	78 240	66 239	70 249	70 246
6	67 243	72 242	68 233	66 224	65 [illegible]	64 [illegible]	64 231	102 252	79 242	64 237	72 250	68 242
7	70 243	74 243	68 231	66 224	63 231	57 220	61 227	98 255	77 244	64 240	72 247	68 244
8	71 243	74 244	68 231	65 228	61 228	58 224	61 226	96 249	78 243	64 239	71 242	68 248
9	72 244	72 247	67 232	65 230	60 227	58 223	60 224	96 242	71 241	66 244	68 238	68 246
10	72 247	71 248	67 231	64 230	60 227	58 218	60 224	96 244	69 238	64 245	72 238	67 245
11	72 244	70 248	68 229	63 226	58 224	57 220	62 227	90 228	69 235	65 244	71 246	67 244
12	73 243	69 241	68 231	61 224	58 224	56 218	60 228	88 240	68 235	67 241	72 244	67 247
13	73 242	68 238	68 229	62 224	57 224	57 216	72 231	81 229	68 237	68 244	73 248	66 242
14	73 240	66 248	68 231	61 222	58 223	58 221	73 226	75 238	69 235	69 246	74 251	65 241
15	71 238	68 244	68 237	61 224	58 222	57 221	71 232	71 238	69 235	71 242	76 248	65 240
16	70 237	67 241	67 232	61 225	58 221	58 221	69 230	69 237	69 231	69 241	80 238	65 244
17	69 240	69 239	67 230	61 221	58 220	58 221	68 230	69 235	69 235	69 244	80 235	65 245
18	69 241	68 241	67 228	59 221	58 223	57 219	68 228	69 236	69 235	69 242	80 236	65 246
19	68 239	69 237	68 228	59 220	58 223	58 222	67 227	69 235	69 235	69 241	79 236	65 244
20	68 238	68 242	68 226		58 223	58 224	67 231	69 234	69 234	69 239	80 230	65 251
21	68 238	69 244	68 230	59 221	58 218	58 222	68 235	67 235	69 234	67 237	78 230	
22	69 236	69 243	68 228	59 220	59 219	57 220	68 233	68 228	69 230	69 238	74 247	64 246
23	67 238	69 240	68 228	60 218	60 213	58 222	68 234	69 231	69 230	70 244	69 242	63 243
24	68 238	69 242	68 227	60 219	60 218	61 226	68 229	69 231	69 230	69 241	68 241	66 249
25	69 239	68 238	68 228	62 223	60 223	63 231	68 229	67 236	68 230	69 243	68 243	66 246
		68 238	68 230	63 221	60 219	60 235	69 228	69 233	68 233	69 231	68 244	68 247
27	68 243	69 236	68 231	64 221	59 220	67 233	68 232	69 234	68 232	68 237	68 243	66 247
28	68 241	69 242	68 231	65 219	60 220	66 234	68 234	69 239	68 232	68 238	68 241	66 245
29	68 239		68 230	64 224	59 217	63 233	68 234	71 243	69 234	67 237	68 243	67 244
30	68 238		68 229	63 222	60 219	66 238	68 234	72 241	68 235	69 238	69 249	67 243
31	68 238		68 228		62 211		68 235	73 239		68 242		66 244
28	69 241	67 243	68 230	62 224	61 224	60 224	68 230	77 240	68 236	66 240	71 246	66 245
1976												
1	66 239	63 237	66 238	71 237	67 232	58 224	66 216	66 236	65 [illegible]	66 224	66 237	63 237
2	66 241	65 237	66 238	68 238	65 223	58 224	68 217	66 227	64 230	65 226	63 237	65 241
3	67 240	63 236	66 236	67 246	64 230	57 222	68 219	70 232	64 229	65 230	63 237	66 241
4	67 240	63 234	66 235	67 243	64 233	57 223	66 222	70 232	65 227	64 229	62 235	66 246
5	67 243	62 235	68 232	68 241	62 228	58 222	68 218	71 230	64 225	64 228	63 231	65 237
6	66 239	65 238	68 231	63 235	61 224		68 218	71 232	63 224	66 225	66 232	65 241
7	66 237	67 241	66 231	63 230	65 230	58 218	68 218	69 232	62 228	65 227	62 232	64 236
8	65 239	62 241	66 232	66 231	65 232	59 221	66 219	70 231	63 230	65 227	62 226	65 237
9	68 240	62 241	68 231	67 232	61 230	60 225	68 220	69 234	63 227	66 228	65 227	66 236
10	68 241	62 240	68 233	67 237	61 230	58 221	68 224	70 232	64 226	66 228	65 230	70 242
11	66 242	62 238	68 231	70 238	62 230	59 221	68 222	68 230	64 228	66 233	65 233	
12	71 246	61 239	66 232	71 239	61 229	61 226	66 220	68 232	63 228	66 235	65 227	72 236
13	74 245	64 239	66 230	71 237	63 229	62 224	66 219	66 230	64 228	67 231	64 235	
14		62 238	66 230	71 238	63 224	62 224	66 218	61 216	65 230	67 231	64 235	73
15	74 247	63 238	66 236	70 236	65 227	62 225	66 218	61 216	64 226	67 233	66 235	74 242
16	72 247	63 238	68 230	67 229	64 227	64 227	66 218	62 231	66 230	69 233	66 238	72 236
17	71 245	64 237	70 231	68 229	64 225	64 225	66 218	66 232	67 227	70 237	65 240	71 241
18	71 247	64 235	70 234	68 234	66 228	66 219	66 220	66 232	64 226	70 241	65 241	67 243
19	69 238	64 235	72 244	68 232	65 229	65 223	66 219	66 227	63 226	69 239	65 244	67 243
20	69 246	64 238	73 243	68 231	65 231	65 221	66 218	63 226	63 225	68 236	65 242	66 238
21	68 241	64 238	73 241	68 235	64 226	65 222	66 219	60 220	62 226	67 235	66 242	66 236
22		63 238	73 243	68 236	65 224	65 224	66 220	61 222	61 227	68 237	66 241	67 240
23	68 243	63 240	73 245	68 227	64 224	64 224	66 217	60 222	63 226	73 239	66 241	68 237
24	68 241	62 235	73 242	67 234	68 227	63 225	67 218	59 225	63 225	73 241	66 239	68 235
25	68 241	61 235	73 245	68 236	68 226	62 220	66 223	60 227	64 226	73 239	66 238	68 237
26	68 240	62 233	73 243	67 230	68 224	62 224	66 220	59 222	64 225	74 237	66 240	70 238
27	68 238	61 230	74 242	67 230	68 223	62 222	67 218	60 223	63 223	71 236	66 238	70 240
28	68 235	60 228	73 236	66 232	68 218	60 219	66 223	61 225	63 225	70 239	66 238	69 238
29	68 234	61 234	73 237	68 233	67 223	60 219	66 225	62 228	63 225	68 235	65 240	69 238
30	68 239		74 264	71 235	68 223	58 218	66 226	62 230	63 224	66 235	66 241	70 238
31	68 239		74 264		68 223		66 226	64 230		66 235		70 235
29	67 242	62 237	69 241	68 236	63 226	61 223	68 220	65 227	64 227	68 233	65 238	68 239
1977												
1	68 240	73 243	67 242	66 241	65 238	77 238	66 241	69 233	72 237	61 231	67 250	66 [illegible]
2	71 249	77 [illegible]	69 245	67 238	66 [illegible]	77 242	65 243	69 237	74 238	62 244	66 248	67 252
3	73 246	76 243	70 240	66 237	72 242	80 244	67 240	68 230	70 241	66 245	67 244	68 252
4	71 245	72 252	70 241	67 246	72 248	80 234	69 237	71 [illegible]	72 239	68 254	67 [illegible]	68 262
5	69 240	73 248	71 237	68 243	68 235	77 231	72 238	73 237	73 237	69 262	69 [illegible]	69 [illegible]
6	68 244	72 [illegible]	70 [illegible]	68 242	66 231	78 241	72 243	74 222	73 237	66 [illegible]	68 243	68 [illegible]
7	66 [illegible]	69 [illegible]	70 [illegible]	68 243	67 238	81 237	73 235	73 231	77 245	67 [illegible]	68 260	103 276
8	67 [illegible]	69 [illegible]	70 [illegible]	68 237	67 240	82 239	69 232	71 230	69 246	70 [illegible]	68 242	101 276
9	68 [illegible]	70 [illegible]	72 [illegible]	68 243		78 237	68 230	73 232		70 [illegible]	68 242	109 282
10	68 [illegible]	70 [illegible]	72 [illegible]	68 244		78 234	68 228	74 243	68 238	69 [illegible]	68 244	112 290
11	68 244	70 [illegible]	72 [illegible]	70 239	70 233	73 231	69 230	74 238	68 240	68 249	68 252	107 282
12	70 246	70 [illegible]	69 [illegible]	69 241	69 234	71 230	68 230	74 236	68 243	69 251	69 [illegible]	109 284
13	71 244	72 [illegible]	68 [illegible]	71 241	69 235	71 231	68 231	77 238	109 263	69 [illegible]	69 [illegible]	108 [illegible]
14	68 [illegible]	74 [illegible]	69 [illegible]	73 244	69 242	69 236	69 230	73 233	104 262	68 [illegible]	69 [illegible]	69 [illegible]
15	69 [illegible]	73 [illegible]	69 [illegible]	74 243	71 246	68 229	69 231	73 238	108 263	68 [illegible]	70 [illegible]	68 [illegible]
16	70 [illegible]	72 [illegible]	69 [illegible]	74 241	73 241	68 234	69 230	73 240	110 263	68 [illegible]	70 [illegible]	68 [illegible]
17	72 [illegible]	71 [illegible]	69 [illegible]	72 245	70 237	69 234	69 228	73 236	109 257	68 [illegible]	69 [illegible]	68 [illegible]
18	69 [illegible]	71 [illegible]	69 [illegible]	69 236	70 244	71 [illegible]	68 230	74 236	105 254	68 [illegible]	69 [illegible]	69 [illegible]
19	72 [illegible]	72 [illegible]	69 [illegible]	71 240	70 239	70 233	68 233	74 243	91 256	68 [illegible]	69 [illegible]	69 [illegible]
20	70 [illegible]	74 [illegible]	68 [illegible]	71 240	68 241	69 232	67 231	76 244	90 [illegible]	68 [illegible]	69 [illegible]	69 [illegible]
21	68 [illegible]	73 [illegible]	68 [illegible]	69 238	68 [illegible]	68 234	69 229	76 244	90 [illegible]	67 [illegible]	69 [illegible]	69 [illegible]
22	70 [illegible]	71 [illegible]	68 [illegible]	68 234	68 [illegible]	69 242	70 235	77 240	94 253	68 [illegible]	69 [illegible]	69 [illegible]
23	72 [illegible]	70 [illegible]	68 [illegible]	68 [illegible]	68 [illegible]	68 [illegible]	72 243	73 234	91 244	70 [illegible]	69 [illegible]	69 [illegible]
24	72 [illegible]	68 [illegible]	68 [illegible]	68 [illegible]	68 [illegible]	68 [illegible]	72 240	73 235	92 [illegible]	68 [illegible]	69 [illegible]	68 [illegible]
25	69 [illegible]	68 [illegible]	68 [illegible]	68 [illegible]	68 [illegible]	68 [illegible]	73 [illegible]	73 236	92 [illegible]	69 [illegible]	69 [illegible]	69 [illegible]
26	71 [illegible]	67 [illegible]	68 [illegible]	68 [illegible]	68 [illegible]	101 [illegible]	73 [illegible]	74 240	92 [illegible]	69 [illegible]	69 [illegible]	69 [illegible]
27	70 [illegible]	67 [illegible]	68 [illegible]	68 [illegible]	67 [illegible]	99 [illegible]	73 [illegible]	73 240	87 [illegible]	69 [illegible]	69 [illegible]	69 [illegible]
28	71 [illegible]	67 [illegible]	68 [illegible]	68 [illegible]	67 [illegible]	95 [illegible]	74 [illegible]	73 [illegible]	88 [illegible]	69 [illegible]	69 [illegible]	69 [illegible]
29	71 [illegible]		68 [illegible]	68 [illegible]	67 [illegible]	93 [illegible]	74 [illegible]	74 [illegible]	88 [illegible]	69 [illegible]	69 [illegible]	69 [illegible]
30	71 [illegible]		68 [illegible]	68 [illegible]	67 [illegible]	91 [illegible]	74 [illegible]	74 [illegible]	88 [illegible]	69 [illegible]	69 [illegible]	69 [illegible]
31	73 [illegible]		68 [illegible]		67 [illegible]		69 [illegible]	74 [illegible]		69 [illegible]		69 [illegible]
28	69 [illegible]	74 [illegible]	68 [illegible]	68 [illegible]	69 [illegible]	80 [illegible]	71 [illegible]	73 [illegible]	89 [illegible]	67 [illegible]	68 [illegible]	68 [illegible]

Table 11-2b (Continued)

DAY	JANUARY 2695 MHZ	JANUARY 8800 MHZ	FEBRUARY 2695 MHZ	FEBRUARY 8800 MHZ	MARCH 2695 MHZ	MARCH 8800 MHZ	APRIL 2695 MHZ	APRIL 8800 MHZ	MAY 2695 MHZ	MAY 8800 MHZ	JUNE 2695 MHZ	JUNE 8800 MHZ	JULY 2695 MHZ	JULY 8800 MHZ	AUGUST 2695 MHZ	AUGUST 8800 MHZ	SEPTEMBER 2695 MHZ	SEPTEMBER 8800 MHZ	OCTOBER 2695 MHZ	OCTOBER 8800 MHZ	NOVEMBER 2695 MHZ	NOVEMBER 8800 MHZ	DECEMBER 2695 MHZ	DECEMBER 8800 MHZ
1978																								
1	122•	277	133	293	133	300	118•	282	163•	318	122	278	119	283	84•	250	143	294	121	280	139	303	153	316
2	127•	276	148	305	135	312	117•	287	166	324	122	276	104	267	84•	247	144	318	119	276	145	302	156	314
3	122	286	159	321	138	309	122	290	156	321	106	264	96	259	84•	250	149	319	119	283	152	316	154	317
4	111	287	148	305	138	318	133	297	149	322	99	266	94•	261	92•	260	158	328	117•	279	164•	318	152•	314
5	113	278	153	311	144	317	134	292	159	316	96•	258	99	263	97	257	152	328	122•	276	168•	311	153	311
6	112	280	151•	315	154	317	137	288	154	311	89	251	99	265	102•	255	159	303	135•	281	157	306	159	316
7	106•	272			157	319	134	283	146	322			103	264	108	259	155	313	127	290	156•	304	165	325
8	96•	265			161	326	136	284	125	286	88•	253	110	279	114	255	149	296	133	299	156•	304	170•	337
9			148	315	170	331	142	298	121•	279	89•	256	121	299	116	264	140	293	142	305	155	306	176•	347
0			141	306	171	322	140	296	117	272	89	258	129	319	107	257	130	289	142	304	153	316	183	358
1	87	262	146	309	165	322	156	307	118	277	92	257	136	321	101	258	123•	284	157	312	148	313	195	383
2	85	261	153	324	156	310	138	286	122	275	95	256	144	336	101•	260	118•	285	158	307	141	310	200	388
3	83•	265	137	317	147	303	126	278	122	274	101	261	133	317	111	263	121	284	156	302	135	303	196	377
4			139	323	145•	298	125	287	129	278	114	270	134	311	107	266	127	286	164•	308	122	290	185	376
5	81	253	126	310	144	291	127	290	127•	276	110	278	137•	307	109	271	132•	298	172	314	123	299	180	357
6	79	255	122	300	134	285	123	284	126•	273	117	278	136•	302	100	261	139	296	158	309	118	289	166	347
7	78•	253	118	285	124	288	117	277	126	271	122	274	127•	298	98	252	146	314	159	310	117•	288	161	339
8	75•	251	117	281	121	286	116	281	120	264	131	282	128	299	9-		151	313	153	307	113	285		
9	78	254	112	279	115	278	133	290	113	253	137	298	120	290	91		146	310	155	307	119	292	144	321
0			115	284	119	280	130•	281	114	269	149	295	116	284	87		148	298	153	306	121	303	123	306
1	84	258	114	285	107	268	125	277	114	267	164	298	110	281	87		146	293	158	313	116•	291	127	305
2	88	256	119	286	106	271	130	285	119	283	165•	302	102	264	85		155•	302	141•	311	116•	290	121	309
3	91	264	125	288	105	268	147	300	126	284	167	305	98	262	85		149	305	142•	307	113•	288	120•	299
4	93	269	126	289	108	272	139	307			170	304	95	257	89•	240	136	293	144	305	113•	291	126•	304
5	92•	271	133	290	107	273	149	305	124•	284	150	292	92	256	86	242	139	291	143	300	115	291	126•	301
6			130	288	101•	269	146	308	126	286	158	291	88	261	84	253	133	289	140•	294	124	301	134	298
7	110	278	130	290	100•	266	156	319	127	278	151	278	88	253	91	259	137	298	138	291	133•	309	143	318
8	121	280	128	298	100	264	181	364	129	294	146	282	87	260	92•	258	131•	295	135	295	143•	316	148	319
9	122	278			98	264	164	331	122	274	141	282	87		94	258	129	289	137	300	146	313	150	321
0	126	285			107	272	184	395	125	278	133	284	89	252	106	268	124	294	133	302	154	318	164	336
1	128	291			109				122	274			88	257	117•	280			131	295			178•	341
M	109	270	134	300	130	293	137	297	130	286	125	277	110	281	97	258	140	300	142	299	136	302	157	330
1979																								
1	187•	331	173	324	146	305	175•	314	160	297	160	343	136	324	118	282	144	321	174	334	186	366	149	328
2	185•	333	169	323	148	296	178•	322	157	299	171	354	181	335	115	278	148	307	178	339	186	336	159	339
3	198	336	172	318	147	303	171	308	143	295	180	344	179	328	116	291	153•	311	177•	345	186•	344	190	365
4	178	333	170	318	155	307	163	314	152	297	192	333	184	339	112	279	149	324	176	357	186	354	201	369
5	186	342	185	316	173	317	158	307	144	287	194	344	166	333	113	276	149	303	163	334	220	391	208	371
6	182•	335	193	325	161	304	159	306	149	303	199	339	174	342	117	279	150•	310	163	330	239	438	196	352
7	178•	328	193	321	155	303	144	308	157	297	197	338	169	327	119	297	151	316	173	329	255	435	196	360
8	176•	328	199	331	150	302	151	302	154	303	202	339	169	329	126	313	152	302	179•	341	273	462	219	372
9	175	327	178	336	150	307	151•	306	151	307	199	344	162	329	129	307	158	314	168	339	285	445	210	370
0	173	323	181	361	162	308	152	308	145	302	197	349	151	319	125•	301	152	315	169•	337	319	479	206	370
1	167	324	189	370	156•	317	151	304	148	311	188	338	143	307	125•	299	154	332	186	335	295	459	208	364
2	159	332	180	350	174	326	157	306	150	310	181	336	134	297	117•	301	153	312	193•	347	252•	423	201	379
3	177•	330	186	345	158	323	154	310	156•	307	157		130	299	127	303	155	322	205	345	221•	396	217	365
4	178•	330	202	373	164	318	148•	303	153•	304	157	329	126	291	141	311	160	318	207	372	217	370	218	378
5	176	330	199	348	154•	314	151•	297	152•	303	144	306	119	284	136	307	166	340	208	376	225	380	218	373
6	182	324	205	365	162	325	155•	293	153•	303	137	314	113	281	134	306	166	351	200	385	209	371	201	371
7	154•	318	200	364	163	318	144	304	164	318	131	302	112	282	131	312	170	351	203	386	209	371		
8	168	339	223	386	175	325	144	296	144•	304	129	296	112	278	143	327	181	353	203	373	211	380	173	346
9	180	339	217	382	164	321	142	285	132•	291	119	298	115	289	151•	344	193	374	212	386	196	364	166	359
0	179	344	202	363	158	318	135	291	131	289	129	300	118	286	166	357	191	360	207	369	183•	358	173	369
1	188•	339	203	355	157	312	127	292	130	285	121	299	120	282	170	344	186	340	191	349	162	349	168	349
2	199	333	185	361	152	318	143	294	133	294	118	286	122	297	179	360	185•	344	188	349	166	340	152	327
3	207	349	169•	331	160	318	137	302	134•	282	115	286	132	291	180	352	190	347	185	345	160	335	148	315
4	190	325	155•	315	167	333	134	300	130•	277	120	291	135	308	179	340	199	343	200	354	155	321	144	307
5	182•	323	144	297	163•	328	140	286	127•	276	124	299	131	301	190	360	195	338	188	357	146	307	142•	307
6	175•	321	134•	281	179	335	156	295	133•	276	123	307	133	312	183	351	193	329	179	353	143•	318	140	307
7	178•	324	135	279	165	333			127	277	127	307	129	289	176	340	197	346	186	344	138	313	140	318
8	181	328	147	301	163	329	161	312	123•	274	135	302	129	288	172	334	189	355	169•	345	134	311	142	323
9	187	339			173	318	158	299	127	275	137	312	126	292	159•	316	180	341	188	350	138	297	158	342
0	186•	336			160•	323	162	308	135•	283	141	315	129	294	154	302	177•	318	193•	353	139	320	166	324
1	188•	336			179•	329			143	291			121	291	147	312			192	352			173	353
4	181	332	182	337	161	317	152	303	143	294	154	319	139	305	149	315	167	331	188	352	201	371	179	349

ample, five times the quiet-sun level). This climatology presents results using both approaches.

With regard to the statistics of burst levels over a certain threshold (flux-density level), one might want to know how many level-exceeding bursts occur in a given period (in a 24-hr day, for example). One might also like to know for what fraction of time (min per 24-hr day, for example) do bursts remain above a certain threshold. Again, both kinds of statistics are presented in this climatology.

The radio-burst climatology presented here is based on routine, daily, whole-sun patrol observations made at the Sagamore Hill Radio Observatory over the years 1966 to 1978. This period covers the decline from the maximum of the 20th solar cycle (a relatively low-amplitude cycle) through the minimum (around 1976) to the rising portion of the new (and higher-amplitude) 21st solar cycle. Data from the 1966–1978 period should provide a reasonable measure of the intensity/duration parameters of solar-burst radio emission during the least and most disturbed portions of the solar cycle.

The Sagamore Hill data on radio bursts, published routinely in *Solar Geophysical Data* (SGD) are given in terms of the following parameters: the peak flux density S_p (sfu), the mean flux density S_{mc} (sfu) and the duration (min) T_d from start to end. These parameters are illustrated for the idealized burst in Figure 11-6. The area under the curve is

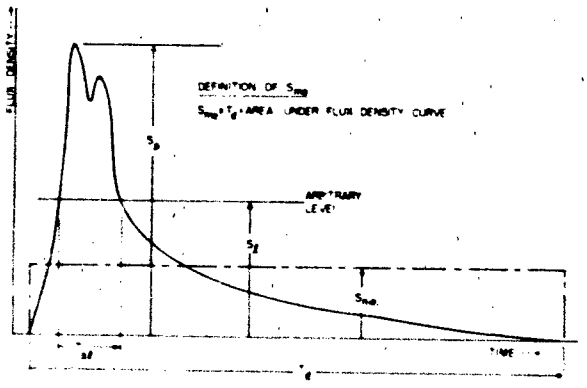


Figure 11-6 Burst parameters depicted for a typical microwave event. The symbols are explained in the text.

the time-integrated flux density. The mean flux density S_{mc} is defined as the amplitude (flux density) of a rectangle with area equal to the time-integrated flux density and duration equal to the burst duration T_d.

In Figure 11-6, we have also shown an arbitrary flux density level, S_t. In the climatology study, we eventually want to derive (from the SGD-listed data alone) the time that bursts exceed a certain level. To illustrate the discussion that follows, we have shown the level S_t between S_p and S_{mc} (if $S_t > S_p$, the burst would never exceed the arbitrary level). We have also shown the "above-the-level" time, T_{at}.

For $S_{mc} < S_t$, the longest possible length of time that a burst with mean flux density S_{mc} could exceed the S_t threshold is $S_{mc}/S_t \times T_d$. This is the worst case; it comes nowhere near representing the results for actual bursts.

In similar fashion if $S_{mc} > S_t$, the longest possible length of level-exceeding time would be T_d. Obviously, this maximum result is not representative of T_{at}. A burst with an extremely large very-steep peak and long low-level tail could have its mean S_{mc} greater than S_t, yet its flux density could lie below S_t for the overwhelming majority of the burst duration time T_d.

For each burst there is a factor by which $(S_{mc}/S_t)T_d$ [for $S_{mc} < S_t$] or T_d [for $S_{mc} > S_t$] exceeds the actual "above-the-level" time T_{at}. We do not have the resources to look over the time profiles of every burst at each frequency. Therefore we had to look at a representative sample of bursts and derive an overall factor to multiply the $(S_{mc}/S_t)T_d$ or T_d obtained from the SGD data.

Analysis of a limited number of relatively large events from a Burst Atlas compiled by Barron et al. [1980] shows that the factor by which $(S_{mc}/S_t)T_d$ or T_d overstates T_{at} varies between 1.5 and 10, independent of frequency, with a mean of about 3. Therefore, we have used the following relationships to derive the "above the level" times of bursts from the flux-density and whole-burst duration data listed in *Solar Geophysical Data*:

$$T_{at} = \frac{1}{3}(S_{mc}/S_t)\,T_d \text{ for } S_{mc} < S_t$$
$$T_{at} = \frac{1}{3}\,T_d \qquad \text{for } S_{mc} > S_t. \tag{11.6}$$

Having discussed the data limitations and the approximating adjustments, we now present the results of the burst climatology study. In Figure 11-7, we show the number of bursts (statistically) per 24-hr day exceeding 500% of the quiet-sun (flux-density) level. The numbers are given for five frequencies (245 MHz, 610 MHz, 1415 MHz, 2695 MHz, and 8800 MHz) over the years 1966 to 1978. In Figure 11-8 we show the number of bursts (statistically) per 24-hr day exceeding a flux-density level of 10^{-19} W/m^{2}Hz (1000 sfu).

In Figure 11-9, we show the time duration (min per 24-hr day) that bursts exceed (statistically) 500% of the quiet-sun flux density level. These durations are given, as in the previous graphs, for the five patrol frequencies over the years 1966 to 1978. In Figure 11-10 we show the duration (min per 24-hr day) that bursts exceed (statistically) the 10^{-19} W m^{-2} Hz^{-1} flux-density level.

The burst number and duration statistics (in Figures 11-7 through 11-10) are given on a "per 24 hr day" basis. Obviously, the sun is not in view (available to provide

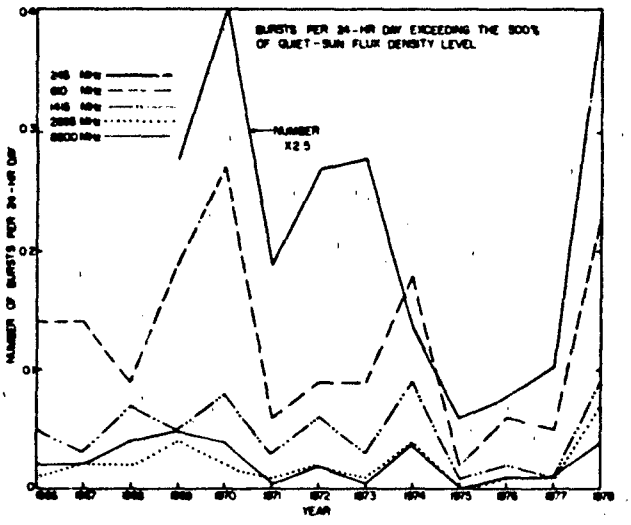


Figure 11-7. The number of bursts per day that exceeded 500% of the quiet-sun flux-density level for the years 1966–1978.

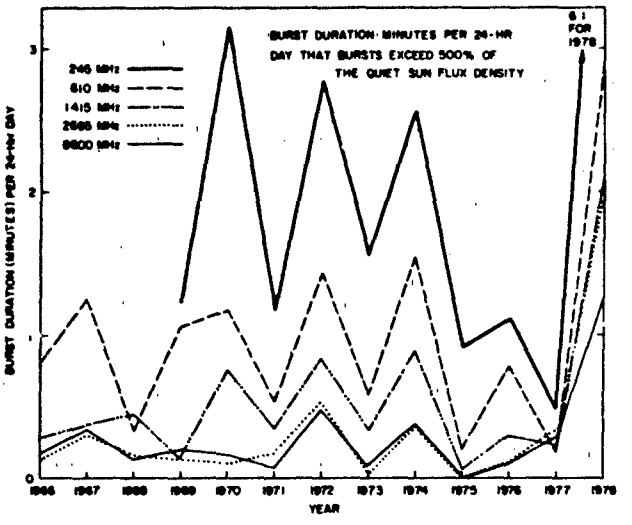


Figure 11-9. The number of minutes per day that solar burst radiation, at various frequencies, exceeded 500% of the quiet-sun flux density.

interference) 24 hours a day from any ground-based operating location. On sites far from the equator (high latitudes), sun-in-view hours per day vary widely with season. Also, the system susceptible to solar radio interference may be in operation only during certain hours of the day (or night). Hence, the best method of presentation is to give the statistics on a "per 24-hr day" basis and let the user determine what fraction of his operating time the sun is actually above the horizon to provide interference to his system. Finally, it is important to point out that the relatively low average "outage" times obtained (~ 1 min/day) result from sporadic events during which the burst radiation from the sun may exceed the indicated threshholds for durations ranging from ~10 min to 2 hours.

At the higher radio frequencies (above 500 MHz), strong solar radio interference (above the quiet-sun level) occurs as distinct, individually-identifiable radio bursts. The bursts have different spectral (frequency) characteristics, but at a given frequency the burst has a start time, a peak (or several peaks), and a time of decay. At lower frequencies (245

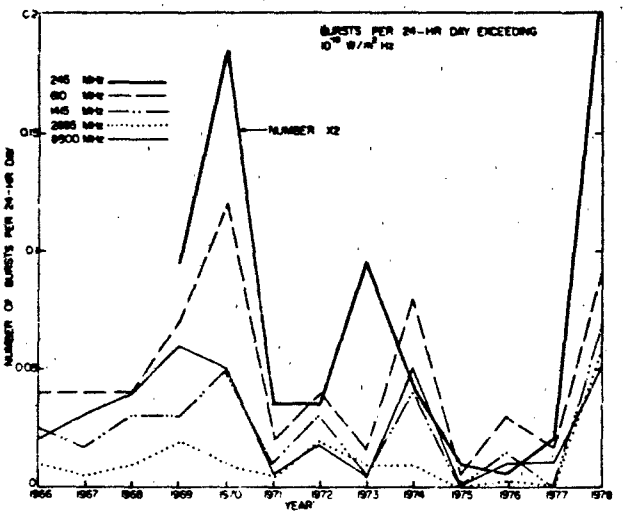


Figure 11-8. The number of bursts during a 24-hr day that exceeded, at various frequencies, 1000 sfu.

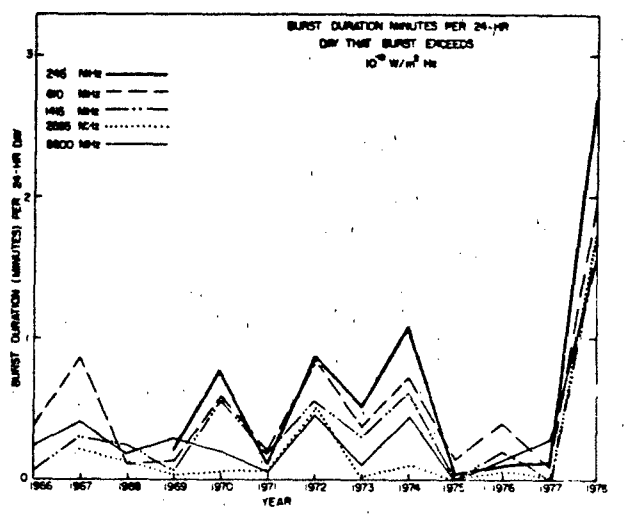


Figure 11-10. The number of minutes per day that solar-burst radiation, at various frequencies, exceeded 1000 sfu.

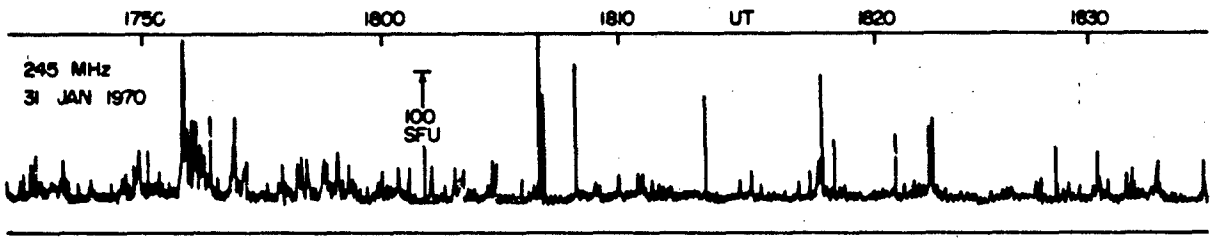


Figure 11-11. The time-intensity profile of a typical 245 MHz noise storm observed at Sagamore Hill.

MHz, for example) there is, in addition to distinct radio bursts, another phenomenon known as solar noise storms. These storms, more prevalent during the sunspot-maximum portion of the 11-year cycle, often last for many hours and sometimes last for a period of days. The noise storms consist of very irregular low-to-moderate increases in solar flux density above the quiet-sun level. For only a very small portion of the noise storm's duration does the added flux density exceed the moderate level (for example, 100 sfu). Figure 11-11 shows a typical example of such a noise storm observed at 245 MHz.

Since these noise storms constitute another significant type of solar radio interference for systems with operating frequencies below 500 MHz, we have also included a statistical graph on noise storm duration, based on our 245 MHz data. See Figure 11-12. The statistical duration is given in terms of minutes of noise storm occurrence per 24-hr day. The shortened period plotted (only 1970 to 1978) gives some indication of the sunspot-cycle variation.

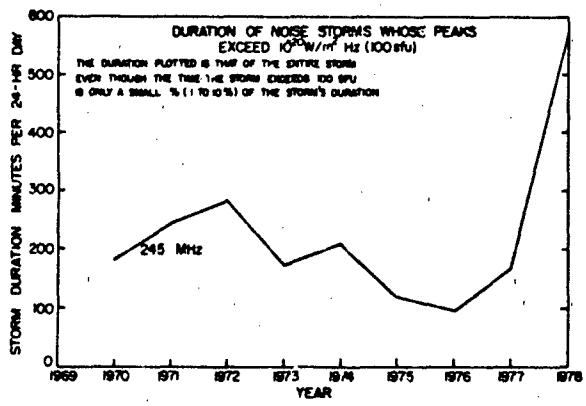


Figure 11-12. The average number of minutes per 24-hr day that a noise storm was reported by Sagamore Hill observatory for the years 1970–1978. The large peak in 1978 may be due (at least partially) to a change in reporting procedures.

REFERENCES

Barron, W.R., E.W. Cliver, D.A. Guidice, and V.L. Badillo, "An Atlas of Selected Multi-Frequency Radio Bursts from the Twentieth Solar Cycle," AFGL-TR-80-0098 ADA088220 1980.

Castelli, J.P., J. Aarons, D.A. Guidice, and R.M. Straka, "The Solar Radio Patrol Network of the USAF and its Application," *Proc. IEEE*, **61:** 1307, 1973.

Forbes, J.M. and R.M. Straka, "Correlation Between Exospheric Temperature and Various Indicators of Solar Activity," AFCRL-TR-73-0378, AD766421 1973.

Kruger, A., *Introduction to Solar Radio Astronomy and Radio Phys.*, D. Reidel, Dordrecht, Holland, 1979

Kundu, M.R., *Solar Radio Astronomy*, Interscience, New York, 1965.

Kundu, M.R. and T.E. Gergely (eds.), *Radio Physics of the Sun*, IAU Symp. No. 86, D. Reidel, Dordrecht, Holland, 1980.

Kundu, M.R. and L. Vlahos, "Solar Microwave Bursts—A Review," *Space Sci. Rev.*, **32:** 405, 1982.

Tanaka, H., J.P. Castelli, A.E. Covington, A. Kruger, R.L. Landecker, and A. Tlamicha, "Absolute Calibration of Solar Radio Flux Density in the Microwave Region," *Solar Phys.*, **29:** 243, 1973.

Zheleznyakov, V.V., *Radio Emission of the Sun and Planets* (in Russ.) Izdat. 'Nauka', Moscow 1964 Engl. Transl. Pergamon Press, 1970.

Chapter 12

THE AURORA

Section 12.1 J. A. Whalen
Section 12.2 R. R. O'Neil and R. H. Picard

12.1 PHENOMENOLOGY, MORPHOLOGY, AND OCCURRENCE

12.1.1 General Characteristics of the Aurora

The aurora is the name given to the light resulting from the precipitation of electrons and protons from the magnetosphere into the earth's atmosphere. This light consists of atomic line spectra and molecular band spectra characteristic principally of oxygen and nitrogen, the chief constituents of the upper atmosphere ionized or excited by collisions with these precipitating particles.

Associated with auroral precipitation and ionization, currents called auroral electrojet currents can flow in the atmosphere. These currents produce magnetic fields detectable at ground level by magnetometers and form the basis for magnetic indicators of auroral activity (see Chapter 4).

Subsection 12.1.1 is a qualitative introductory discussion of the general classes of aurora in terms of the precipitating particles, the source of each type of particle in the earth's magnetosphere, the energy spectrum of each, the height at which each produces maximum ionization and excitation in the earth's atmosphere, the resulting electron density for each, the spatial structure of each, and a map of the general global morphology of each. The succeeding subsections provide tables, graphs, and definitions that specify and characterize auroral properties in as quantitative terms as are currently available.

12.1.1.1 Particle Energy—Ionization and Excitation Height Relation. Particles penetrate into the atmosphere to a height dependent on their initial kinetic energy, the higher the energy the deeper the penetration and therefore the lower the altitude of the resulting ionization and excitation.

The condition for electrons (monoenergetic and isotropic) is shown in a graph of height of maximum production of ionization versus electron energy in Figure 12-1. Note that the height scale is divided into three sections which are the three fundamental altitude regimes: the D region (50–90 km), E region (90–160 km), and F region (160–500 + km). The particle energies and morphologies of each are dependent on the source regions in the magnetosphere.

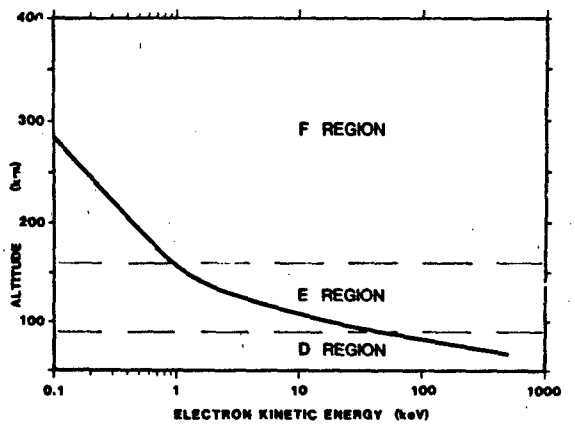


Figure 12-1. Altitude of maximum production of ionization in the upper atmosphere by auroral electrons as a function of incident electron kinetic energy. Calculation by Rees [1964] assumes mono-energetic electrons of isotropic incidence. Dashed lines are the approximate boundaries of the D, E, and F regions which are the principal ionospheric altitude regions.

12.1.1.2 Particle Energy—Magnetospheric Regime Relation. These general source regions are the magnetosheath, the plasmasheet, and the Van Allen or so-called trapped radiation belts. These regions are occupied by electrons of kinetic energies of ~0.1 keV, 1–10 keV, and >40 keV, respectively. (Protons also exist in these regions, the energies of which are greater than electron energies by a factor of ~5).

Quite fortuitously, electrons that precipitate into the atmosphere with energies characteristic of these three magnetospheric regimes produce maximum ionization and excitation in the three different altitude regimes; that is, ~0.1 keV electrons ionize principally at heights in the F region, 1–10 keV in the E region, and >40 keV in the D region.

A cross section of the magnetosphere through the noon-midnight meridian schematically defines these regions to-

gether with their particle populations and associated ionospheric height regimes (Figure 12-2).

The origin of the magnetosheath plasma is the solar wind proton-electron plasma, which is approximately thermalized in passing through the bow shock on the solar side of the

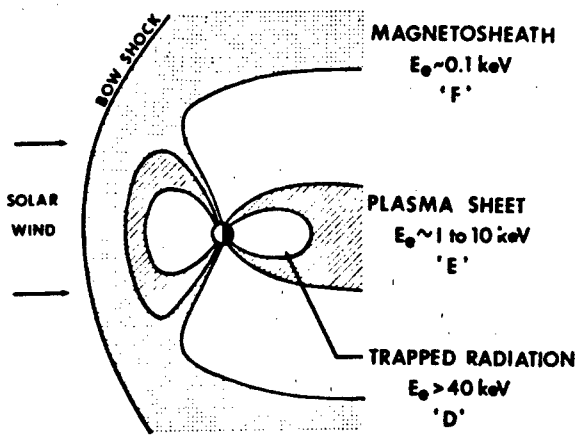


Figure 12-2. Schematic of the magnetosphere of the earth in noon-midnight cross section showing principal magnetospheric domains, the associated particle populations via representative electron kinetic energy, Ee, and the ionospheric region (F, E or D) in which these particles produce maximum ionization.

magnetosphere. These magnetosheath particles are thought to be the source of the plasma sheet particles and they, in turn, to be the source of the Van Allen radiation belt particles. However, the route of entry of the particles into these two regimes and the processes by which they are accelerated are not understood. As such, they remain the fundamental unresolved questions of magnetospheric physics and the subject of much active research (see Chapter 8). It is not the purpose here to deal with these questions, but it is well to note that the fundamental links between the magnetosphere and the ionosphere make the ionospheric measurements important ways of probing the magnetosphere in pursuit of these answers.

12.1.1.3 Spatial Structure. Another key distinction that one must make between auroras is their spatial structure, that is, whether they are relatively structured or unstructured. This distinction is most evident in the 1 to 10 keV plasmasheet precipitation, which represents the bulk of the energy of auroral precipitation. Here the structured component is known as the discrete aurora. Since it is the most visible, it is popularly known as "the aurora". The unstructured component is known as the continuous aurora (also called mantle and diffuse aurora). Although often neglected because its faintness and absence of contrast make it difficult to observe, this unstructured aurora comprised 80% of the total energy input into the polar region at solar maximum (IGY) and 50% in solar minimum (IQSY) [Sandford, 1968].

These two structural types can be distinguished in the large scale photographs of the aurora by the Defense Meteorological Satellite System (DMSP) (Figure 12-3). The bright structured forms at high latitude are the discrete au-

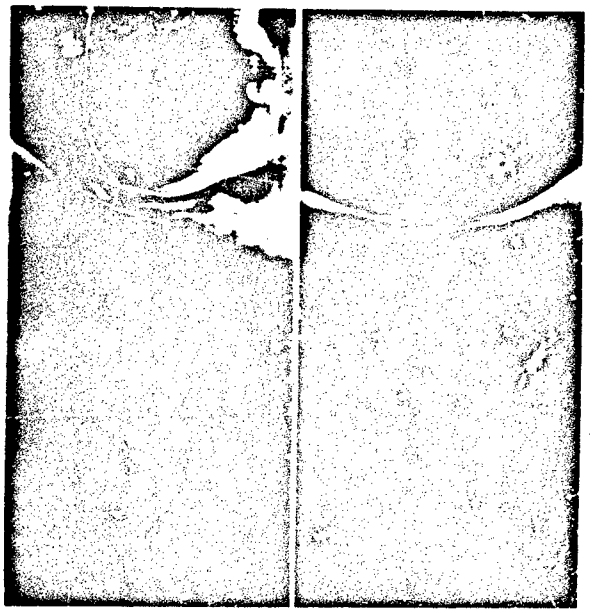


Figure 12-3. Two DMSP photographs showing the midnight aurora across the North American continent defined by city lights.

roras; the relatively faint glow at lower latitudes is the continuous aurora. City lights show the location and scale of the photographs.

12.1.1.4 Morphology of the E Region. The most familiar model of the aurora is the Feldstein and Starkov [1967] auroral oval, a statistical compilation as determined from auroral all-sky camera photographs taken during the IGY. The oval is a band encircling the magnetic pole but displaced eccentrically to higher latitude in the day sector as a consequence of the distortion of the earth's magnetic field by the solar wind. Figure 12-4 shows the oval projected on a map of the northern hemisphere in magnetic coordinates [Whalen, 1970]. The pattern, fixed with respect to the sun, is shown in four different orientations with respect to the earth at the indicated values of UT. With increasing magnetic activity, the oval expands and moves equatorward.

The oval typical of moderate conditions (that designated by the magnetic index Q = 3) will be used to form a reference system for other types of auroras. The coordinate system of the oval, Corrected Geomagnetic Latitude and Corrected Geomagnetic Local Time, is shown in six different plots in Figure 12-5. The Q = 3 oval is drawn in

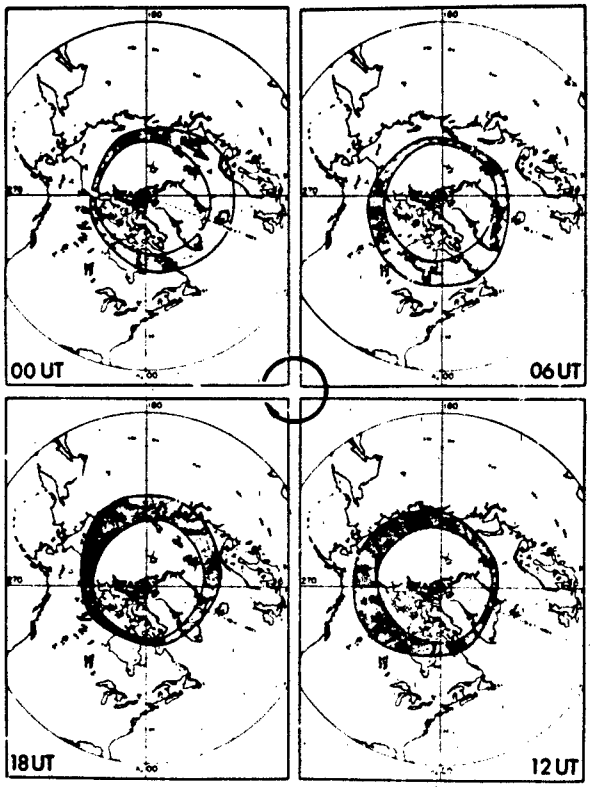


Figure 12-4. The auroral oval, the locus of maximum discrete (E) auroral activity, projected on the north polar region as mapped in Corrected Geomagnetic latitude and longitude. The orientations of the oval at four different values of UT are shown.

each plot bounded by the heavy lines. The shaded areas schematically map the types of auroras differentiated as to structured (top row) and unstructured (bottom row), and to particle energy/ionospheric regime by columns.

The E type auroras appear in the center column of Figure 12-5, the discrete auroral oval itself (above) and the continuous aurora (below). The continuous aurora tends to overlap the oval in the night sector [Whalen et al., 1977], but to locate equatorward of the oval in the day sector [Whalen et al., 1971]. In quiet times the discrete aurora disappears entirely and the continuous aurora remains [Wagner et al., 1973]. Although the continuous aurora forms a band that encircles the pole for all levels of activity, its location can change and the total energy fluxes can vary by several orders of magnitude. In terms of the magnetosphere, the continuous aurora appears to be precipitation from the plasma sheet, both having energy spectra which can be described as Maxwellian [Rearwin and Hones, 1974].

The discrete auroral particles also apparently arise from the plasma sheet but undergo additional accelerations. A further distinction is that the continous aurora can result from both electron and proton precipitation. The discrete aurora results from electrons alone since protons effectively diffuse throughout distances of a few hundred kilometers in the atmosphere. The continuous aurora will be examined in detail in a subsequent section.

12.1.1.5 Morphology of the F Region. The F precipitation appears to have structured and unstructured components analogous to the E region (for example, all-sky photometric observations of Mende and Eather [1976]). The location of each appears to be statistically well represented by the oval—whether there are differences in the morphology of the two is not known. The ionosphere at F layer heights exhibits irregular and spread conditions throughout these regions of soft precipitation. In addition, the polar cap (the circular area poleward of the oval) is also a region of spread F.

An additional ionospheric region of importance is the F layer trough, a region of low ionization 5°–10° wide located equatorward of the oval in the night and evening sectors [Muldrew, 1965]. The poleward edge or wall of the F layer trough is a region of large electron density gradients that can cause large refractions and thus large departures from great circle propagation paths for HF. Buchau et al. [1978] treat a specific case. Pike [1976] has developed an analytical model of the F layer trough wall.

These F region precipitation patterns are shown in the left column of Figure 12-5. Thermalized solar wind from the magnetosheath impinges directly on the atmosphere to produce the dayside region but not the night sector regions, the field lines for which map down from the plasma sheet. In any case the regions are continuous in local time as shown.

One of the complications in the F layer is that the regions of production of ionization are not necessarily the bounds of the location of the ionization. This is because ions in the F-layer have relatively long lifetimes during which they can be transported considerable distances from their origin.

12.1.1.6 Morphology of the D Region. The unstructured component of the D region precipitation exists as a band generally a few degrees wide that extends from midnight to morning and sometimes is entirely circumpolar [for example, Hook, 1968]. In quiet times its locus near midnight is the equatorward edge of the continuous aurora; in active times the distribution can be quite wide. The temporal characteristics have been the subject of many studies [for example Hartz and Brice, 1967]. The occurrence of this precipitation causes D region HF absorption of relatively short duration (15 minutes) near midnight, longer (1–2 hours) in the morning sector. Structured D region auroras seem to be associated with activity in the night sector as, for example, with poleward expanding arcs during substorms. Thus they are transient in nature and occupy the midnight sector of the auroral oval. The patterns for the D region are shown schematically in the righthand column of Figure 12-5.

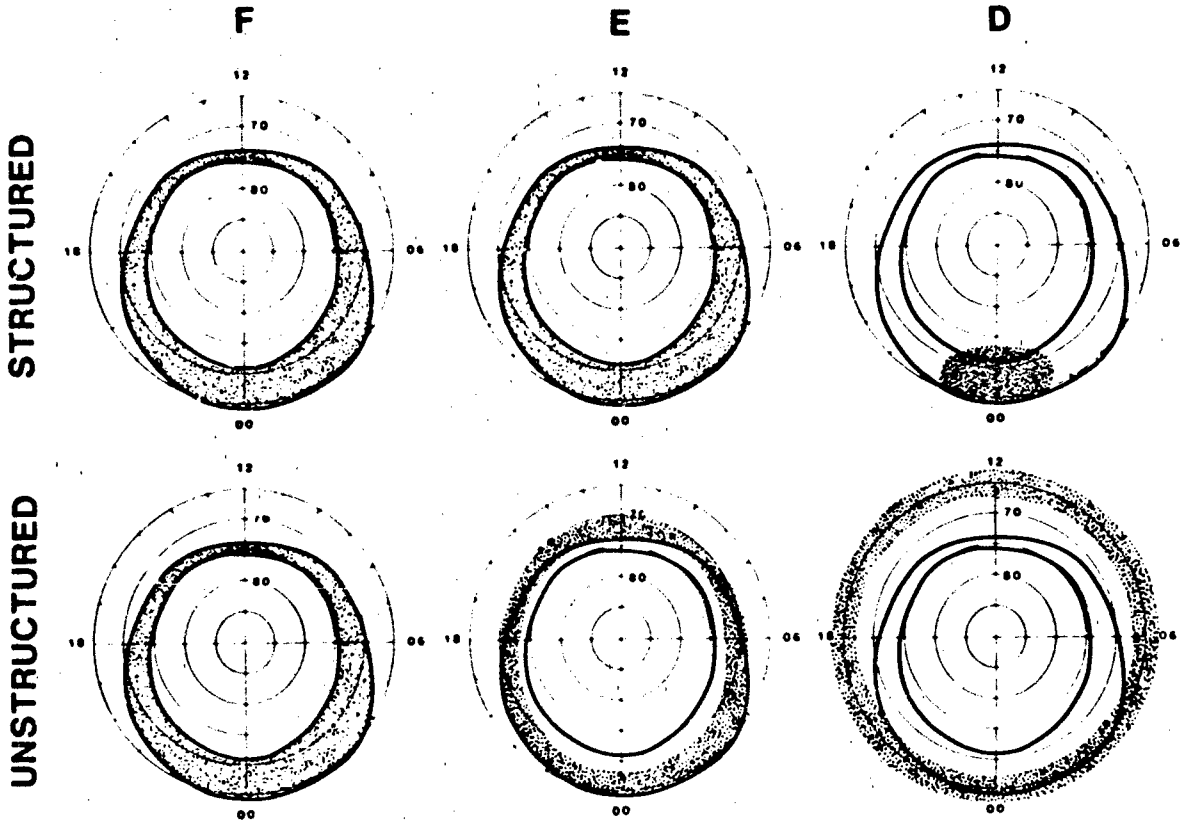


Figure 12-5. A schematic representation of the regions of structured (top row) and unstructured (bottom row) auroras differentiated as to ionospheric regime (particle energy) from left to right, F(~0.1 keV), E(1-10 keV) and D(>40 keV). Coordinate systems are CG latitude and CG local time. The Q = 3 auroral oval is defined by the heavy solid lines; the location of the specific auroras, by the shaded areas.

The unstructured pattern marks the region in which radiation belt particles "un-trap" so as to precipitate into the atmosphere. Berkey et al. [1974] have described a number of events in which the dynamics of this precipitation are detailed.

12.1.1.7 Composite. In Figure 12-5 two generalizations can be made regarding the unstructured auroras. First, in the day sector the D, E, and F precipitation and ionization production are at different latitudes (D lowest, E intermediate, and F highest) so that there should be little question as to their separate identities. Second, in the night sector the three overlap, a condition which had led to the confusion that they are a single phenomenon since they can be in the same place.

For structured auroras, E and F categories are apparently collocated at all local times and with D near midnight. Two features are of note regarding the relation between structured and unstructured auroras. All six categories overlap in the night sector. In the day sector, unstructured F auroras due to particles of magnetosheath origin are collocated with structured E and discrete aurora, the energies of which correspond to those of the plasma sheet. This has lead to the paradoxical situation that the F region has been correctly defined as being the Feldstein oval in the dayside because the photographs have detected structured E auroras (unaccountably) in the same location as the unstructured F auroras.

The foregoing is not intended to be a complete description of the auroral ionosphere but is intended to identify the *minimum* distinctions necessary for such a description. Figure 12-5 presents, schematically, all six of these phenomena together in hopes of clarifying their existence and interrelation about which a great deal of confusion exists.

12.1.2 Distribution of Auroras in Magnetic Latitude and Local Time

The global distribution of auroras has proved to be a difficult subject to quantify. A principal reason is that most studies have failed to distinguish among the large number of auroras, each of which can exhibit quite independent brightness and location on a small and large scale, short and long term.

In addition, studies have been necessarily piecemeal,

ground based studies being too small in scale and satellite studies being too short in time. The Dynamics Explorer (DE) imager addresses these contradictory elements, and if it can distinguish between the classes of auroras, promises to provide valuable information in this area.

At the present time, the available distributions are generally statistical in nature and fall into two categories: morphology of the aurora on the one hand and average values of parameters on the other. However, the former is accomplished without regard to intensity and the latter confuses intensity with location or frequency of occurrence so that the numerical values do not represent actual values. In addition these statistical measurements are ordered through a magnetic index (usually Kp) that encompasses a large variety of auroral conditions.

A second approach has shown progress in defining quantitatively the instantaneous distribution of one of these classes of aurora—the continuous aurora—throughout nearly its entire latitude-local time extent. It finds that no single parameter, magnetic or otherwise, can quantitatively describe the intensity and location of this aurora since both can vary quite independently. Rather this approach measures both intensity and location and takes these parameters to be fundamental auroral indices.

Unfortunately neither of these approaches can be considered to give adequate quantitative descriptions of the aurora in this scale at the present time: the statistical approach because of inherent limitations and the instantaneous approach because it has been demonstrated to date only on a limited basis. Nevertheless, this section will endeavor to present in each of these categories the best results which are available at the present time.

12.1.2.1 Statistical Distributions.

Frequency of Occurrence—The frequency of occurrence of aurora in the zenith as determined by all-sky camera (ASCA) in one-half hour intervals during the IGY is shown in Figure 12-6. The 70% contours define a band which was originally named the auroral oval [Feldstein, 1966] and has since been called the auroral oval zone or the statistical auroral oval. This is apparently the discrete aurora since the continuous aurora has generally too low an intensity and presents too little contrast to be seen in the zenith.

The Auroral Oval—The auroral oval (originally called the auroral oval belt) is a band encircling the pole, the high and low latitude borders of which are the average locations of the edges of auroral luminosity seen in ASCA photographs during the IGY. Such a band has been determined for each value of magnetic index Q, where Q was determined within the same 15 minute intervals of UT in which the ASCA photographs were taken but from magnetometer measurements in the midnight sector near 65°CG latitude [Feldstein and Starkov, 1967]. Figure 12-7 shows examples of ovals for three values of Q.

The ovals are thinner and displaced poleward in the day sector. In addition, the day and night sectors differ in their

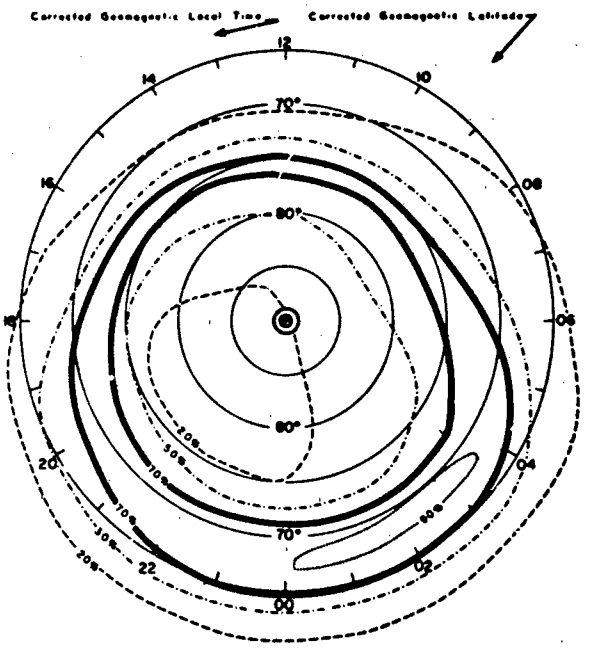


Figure 12-6. Percentage of occurrence of discrete aurora in the zenith as recorded by ASCA in the Northern Hemisphere during the period 1957–1959.

response to increasing activity, the night sector expanding in width and day sector moving equatorward. The standard deviation of these boundaries is about 1°. No clear distinction is made between discrete and continuous auroras. However, they appear not to be exclusively discrete auroras. Since the measurements were generally not in the zenith, the continuous aurora would tend to be visible particularly at large zenith angles due to VanRhijn enhancement. Thus the oval borders could be the continuous aurora and probably are in the equatorward region since the discrete aurora tends to be absent there [Lui et al., 1977; Whalen, 1983].

Although no intensity information was recorded, the oval pattern has been found in a number of cases to be quite representative of the instantaneous pattern of the aurora (for example, the DE image—Figure 12-25 in Section 12.2).

The solar cycle dependence of the oval is such that at solar minimum the midnight sector (the only sector reported) was displaced 1° to 2° poleward of its location at solar maximum [Starkov and Feldstein, 1970].

The equatorward boundary Λ_{eq}, of the auroral oval can be represented by the expression

$$\Lambda_{eq} = 72° - 0.9°Q - 5.1°\cos\left(\frac{360°}{24}\, t_{CG} - 12°\right) \qquad (12.1)$$

for $Q > 1$ where Q is the magnetic index, t_{CG} is the CG local time [Starkov, 1969]. This border can be closely ap-

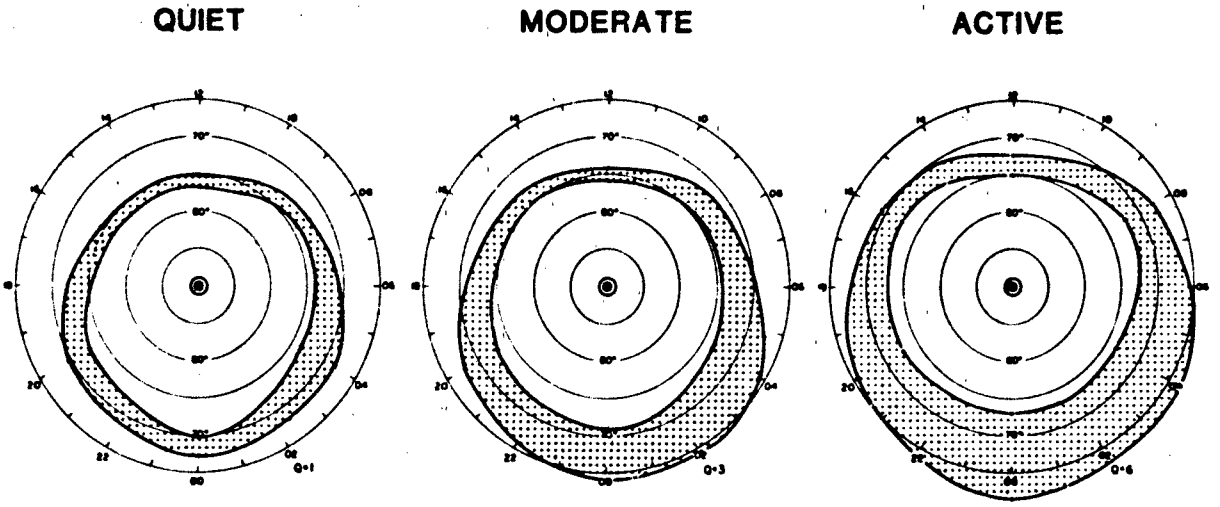


Figure 12-7. Variation in the size of the auroral oval with activity as denoted by the magnetic index Q. Coordinate system is CG latitude and CG local time.

proximate by a circle centered at 84.9°CG latitude 00:48 CG local time having a radius R = 18° + 0.9°Q [Gassman, 1973]. Holzworth and Meng [1975] have derived similar expressions for the ovals.

An example of the nature of the fit of such an offset circle will be shown in Figure 12-11 for the Q = 2 oval. Also shown in Figure 12-11 are other examples of other auroral parameters that can be ordered in such an offset pole reference frame and which will be discussed below.

The magnetic Q index is not generally available, an unfortunate circumstance because it has been the most successful single parameter in ordering auroral morphology. In the absence of the magnetically measured Q, it is possible to infer Q by taking a measurement of Λ_{eq} at a single t_{CG}, substituting these values into Equation (12.1) and solving for Q. By substituting this effective Q into Equation (12.1) one can calculate Λ_{eq} at other local times.

Average Values—A number of studies have provided distributions in magnetic latitude-local time of the average values of auroral parameters. One of the difficulties in this approach is that the measurements are ordered in terms of magnetic indices that typically span three hours. During such an interval auroral intensity and location are never fixed so that a given Kp represents a large variety of auroral conditions. In particular the aurora can move throughout a range that is generally much larger than its instantaneous extent. As a result instances of non-occurrence of aurora within a spatial interval are considered to be zero intensity and these are averaged together with cases of finite intensity within that interval. Accordingly the resulting average can produce values of intensity that never occur in nature; the average spatial distribution is one than can never occur instantaneously.

Nevertheless these average distributions have been useful in revealing some overall trends related to activity and in producing large scale spatial integrals. However, it is important to keep in mind Sandford's [1968] caveat that such distributions "give an average gross representation of parameters that show rapid and large variations in space and time."

The most extensive set of distributions of important optical auroral parameters are due to Sandford [1968]. Included in Figures 12-8 and 12-9 are contours in CG latitude—CG local time of intensity of emissions at 6300 Å, 5577 Å, and 3914 Å, and frequency of occurrence of discrete aurora for three levels of activity (quiet, moderate and disturbed based on magnetic local K indices of 0–2, 3–4, and >4 respectively) for solar cycle maximum (1958–1959) and minimum (1963).

The latitudinal distributions of these same parameters are shown at midnight and at noon for values of K from 0 to 8 for solar maximum (Figure 12-10, left) and solar minimum (Figure 12-10, right).

In addition, Sandford [1968] has shown conclusively the existence of a non-discrete form of the aurora which he termed mantle aurora. The term continuous aurora [Whalen et al., 1971] was coined to distinguish the E-region component of this aurora from the possible D-region component which Hartz and Brice [1967] had named diffuse aurora.

Sandford found that at solar maximum the mantle aurora contributed 80% of the total overall auroral energy flux and the discrete aurora accounted for only 20%. At solar minimum, on the other hand, the two energy fluxes were equal. The principal solar cycle difference in auroral energy fluxes was that the mantle auroral energy flux was lower by a factor of four at solar minimum than it was at solar maximum. Energy flux in the discrete aurora on the other hand was the same at solar minimum as it had been at solar maximum although the spatial distributions at solar minimum were narrower and more contracted toward the pole.

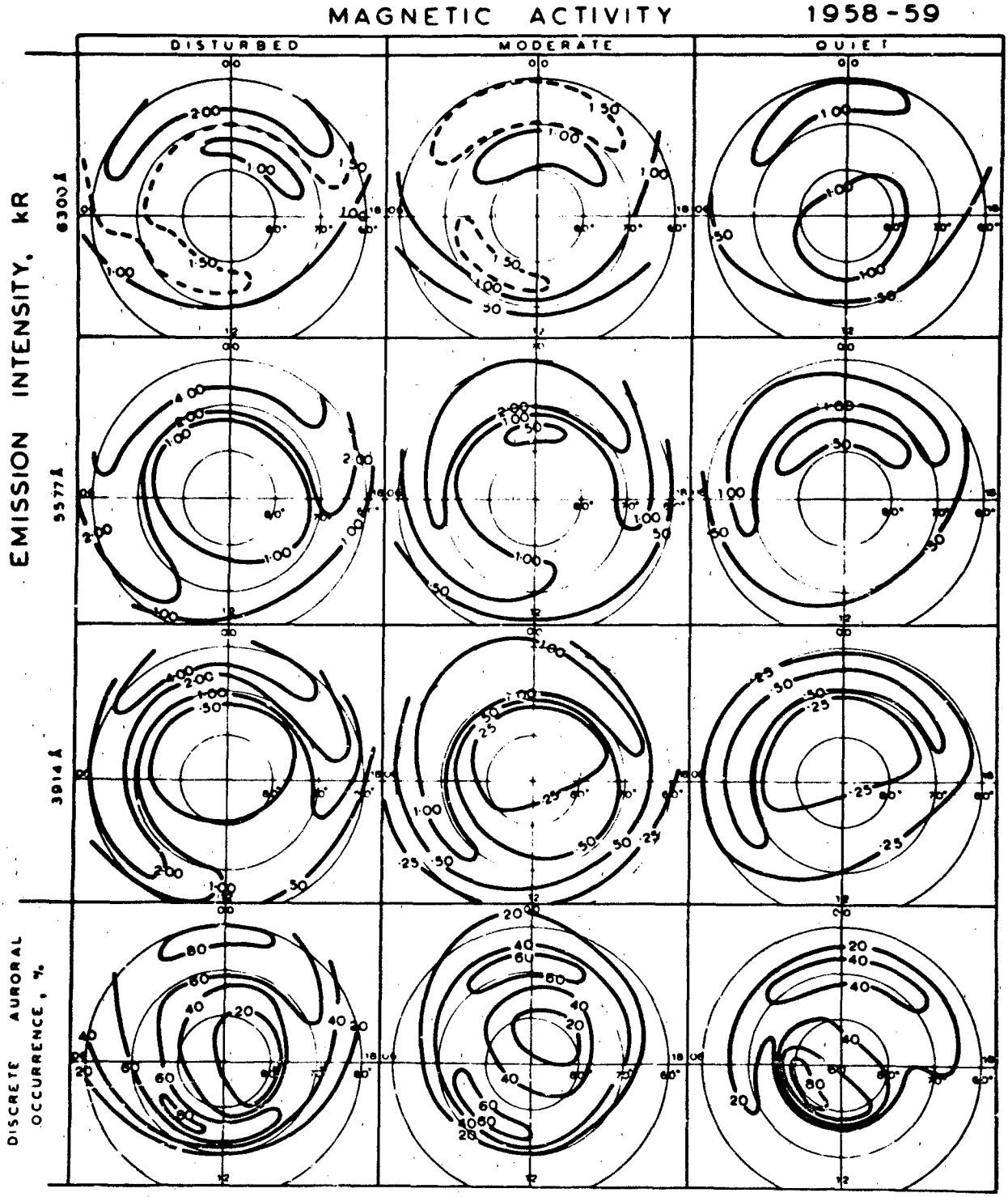


Figure 12-8. Contours of occurrence of discrete aurora (in percent) and average intensity of auroral emissions (in kilo Rayleighs) for 3 levels of activity from Southern Hemisphere measurements in years 1958-1959. Coordinates are CG latitude and CG local time. [Sanford, 1968] (Reprinted with permission from Pergamon Press, Ltd. © 1968)

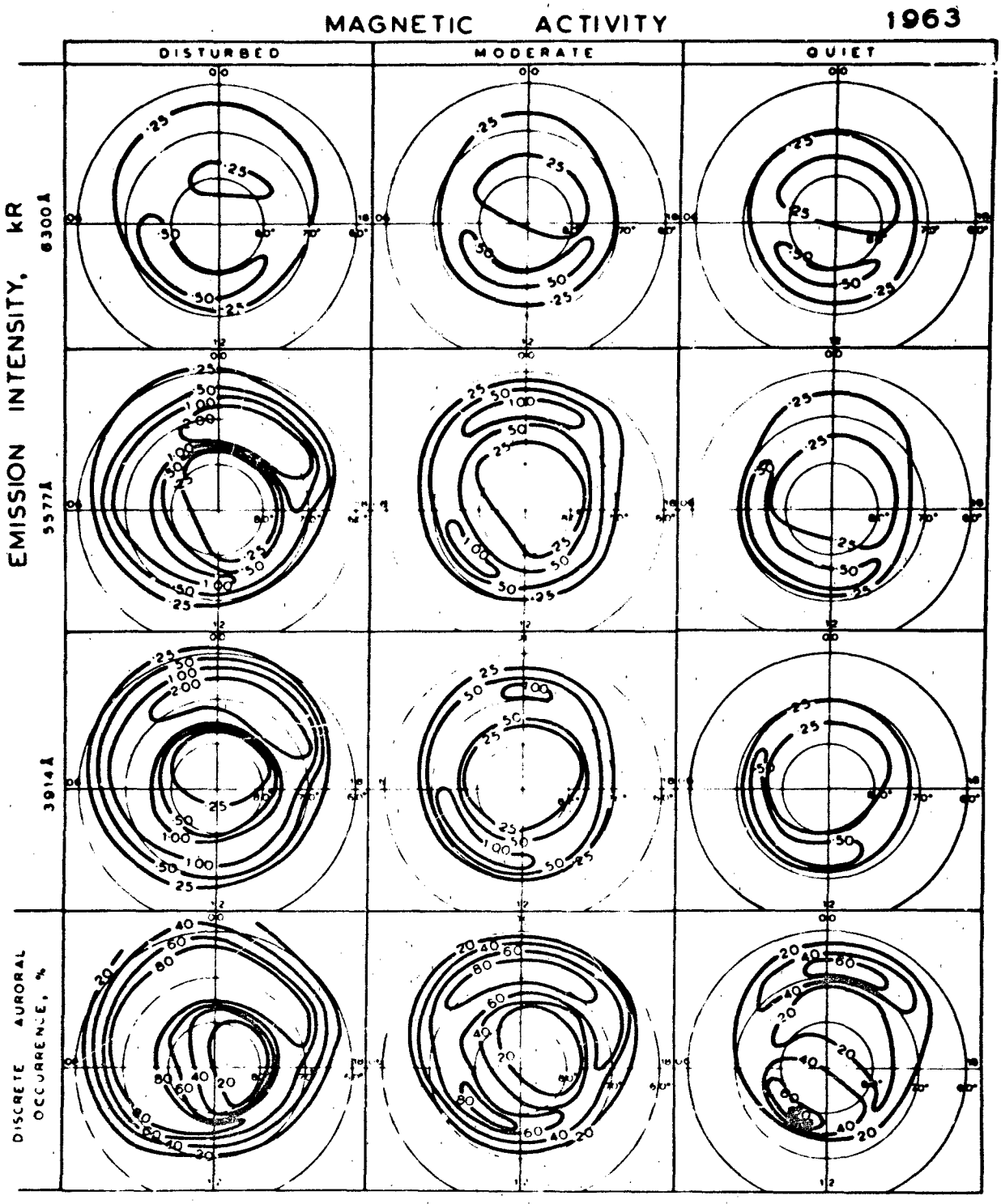


Figure 12-9. Contours of occurrence of discrete aurora (in percent) and average intensity of auroral emissions (in kilo-Rayleighs) for 3 levels of activity from Southern Hemisphere measurements in 1963. Coordinates are CG latitude and CG local time. [Sanford, 1968] (Reprinted with permission from Pergamon Press Ltd. © 1968.)

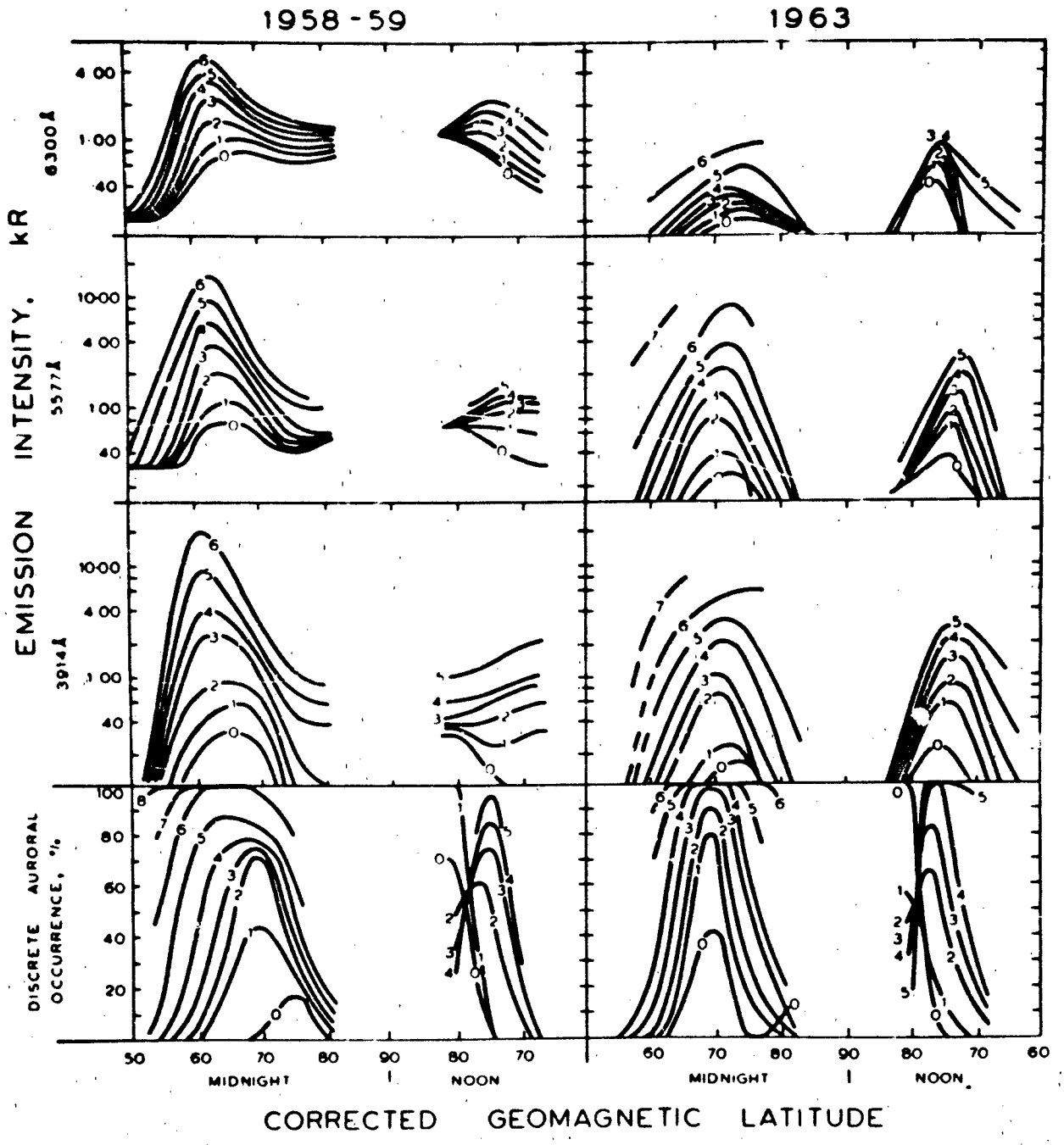


Figure 12-10. The percentage occurrence of discrete aurora and the median intensities of 3914, 5577 and 6300 Å emissions as a function of local K-index and CG latitude along the noon-midnight CG time meridian for the years 1958–1959 and 1963. [Sanford, 1968] (Reprinted with permission from Pergamon Press Ltd. © 1968.)

Other examples of distributions of average parameters will be shown in Figure 12-11.

Equatorward Boundaries—The equatorward boundary of precipitating electrons measured by DMSP satellites has been related to magnetic index Kp via

$$\Lambda_{eq} = \Lambda_0 + \alpha Kp \qquad (12.2)$$

where Λ_{eq} is the CG latitude of this equatorward boundary, Λ_0 is the latitude at Kp = 0 and α is the straight line slope [Gussenhoven et al., 1983]. Values of these parameters are

Table 12-1. Parameter values measured by DMSP satellites

	$\Lambda = \Lambda_0 + \alpha Kp$			
MLT	Number*	Λ_0	α	Cc†
0000–0100	312	66.1	−1.99	−0.80
0100–0200	220	65.1	−1.55	−0.68
0400–0500	267	67.7	−1.48	−0.57
0500–0600	1123	67.8	−1.87	−0.71
0600–0700	2462	68.2	−1.90	−0.74
0700–0800	3159	68.9	−1.91	−0.76
0800–0900	2159	69.3	−1.87	−0.73
0900–1000	1178	69.5	−1.69	−0.66
1000–1100	864	69.5	−1.41	−0.57
1100–1200	513	70.1	−1.25	−0.52
1200–1300	353	69.4	−0.84	−0.35
1500–1600	63	70.9	−0.81	−0.34
1600–1700	204	71.6	−1.28	−0.66
1700–1800	526	71.1	−1.31	−0.69
1800–1900	997	71.2	−1.74	−0.82
1900–2000	2469	70.4	−1.83	−0.82
2000–2100	3309	69.4	−1.89	−0.82
2100–2200	3092	68.6	−1.86	−0.79
2200–2300	1482	67.9	−1.78	−0.77
2300–2400	461	67.8	−2.07	−0.81

*Number of boundaries
†Correlation coefficient

given in Table 12-1. This relation is plotted for the 2300–2400 MLT sector in Figure 12-12 where the points are mean values and the error bars are standard deviations. Also shown in this figure are the midnight sector boundaries determined by the following means: (1) ASCA [Feldstein and Starkov, 1967], (2) DMSP Images [Sheehan and Carovillano, 1978], and (3) 6300 Å at 100 R [Slater et al., 1980]. The trend in all four cases is that latitude decreases with increasing Kp. However, there is considerable scatter which is due primarily to the fact that Kp encompasses a large range of auroral conditions [Feldstein and Starkov, 1967]. An additional uncertainty can arise in particle data since the electrons are measured at a single pitch angle and pitch angle distributions can depart appreciably from isotropy [for example, Sharber, 1981].

This boundary has been found to be solar cycle dependent as noted earlier in relation to the auroral oval. In addition it has been reported to vary by several degrees as a function of UT, being at highest latitude near 0600 UT and lowest latitude near 1800 UT [Sheehan and Carovillano, 1978; Meng, 1979; Whalen, 1983].

Finally the instantaneous distribution of energy flux typically has a gradient with latitude that makes the location of an equatorward boundary a function of detector threshold. Strickland et al. [1983] show a case where a sensitive detector would place the apparent edge as much as 4° equatorward of that determined by a less sensitive detector. For statistical distributions the situation is worse judging by the 3914Å distributions of Sandford [1968] shown in Figure 12-10 (right).

Offset Pole Coordinates—A number of auroral phenomena have been found to be ordered by a coordinate system that lies within the frame of reference of magnetic latitude—magnetic local time but has an effective pole offset from the magnetic pole. This offset pole lies near the midnight meridian and has a colatitude near 4°.

The auroral oval was noted earlier as having an equatorward boundary which is such an offset circle. The Q = 2 oval is shown in Figure 12-11a together with a circle whose circumference is defined by the small circular points and whose center is the circled cross. As noted earlier, this is better approximated by an analytical expression then by a circle.

Three other examples are shown in Figure 12-11 (b) average intensity of auroral emissions at 3914 Å determined by Sandford [1968] for IGY, (c) blanketing frequencies, fbEs, of the ionized auroral E layer determined by Besprozvannaya and Shchuka [1976] for Kp 0 and 1 in IGY, and (d) average precipitating electron energy flux measured by satellites AE-C and AE-D in years 1974–1976 for AE ≤ 100 γ [Spiro et al., 1982].

That the offset pole reference frame is not an artifact of the statistical averaging process is indicated by the existence of auroral arcs seen nearly instantaneously (within approximately 10 min). Meng et al. [1977] have found a group of about 50 quiet auroral arcs imaged by DMSP satellites which are offset circles. These arcs extend continuously over 12 hours of magnetic local time and occur at all local times. Average radius of the arcs was 19° ± 5°. Average offset pole location was 85.8° CG latitude, 00:15 h CG local time.

12.1.2.2 Instantaneous Distributions. Recently progress has been made in quantitatively formulating the instantaneous distribution of the energy flux in the continuous aurora throughout much of its entire latitude-local time extent. This occurred in an 8-hour case study in which this energy flux was found to be constant instantaneously along contours that were defined analytically and approximated by offset circles. Changes that took place did so simultaneously (within 15 minutes time resolution) throughout local time, the extent of which was from evening through midnight to noon [Whalen, 1983].

The latitudinal distribution was therefore the same instantaneously at all local times within the observed extent. Two distributions were defined, the principal one being Gaussian. Furthermore, the Gaussian scale factor was constant in all 32 consecutive determinations. Therefore, the dynamics of the entire nearly global distribution could be described by the variation of only two parameters, the Gaussian maximum energy flux and its offset pole latitude. These

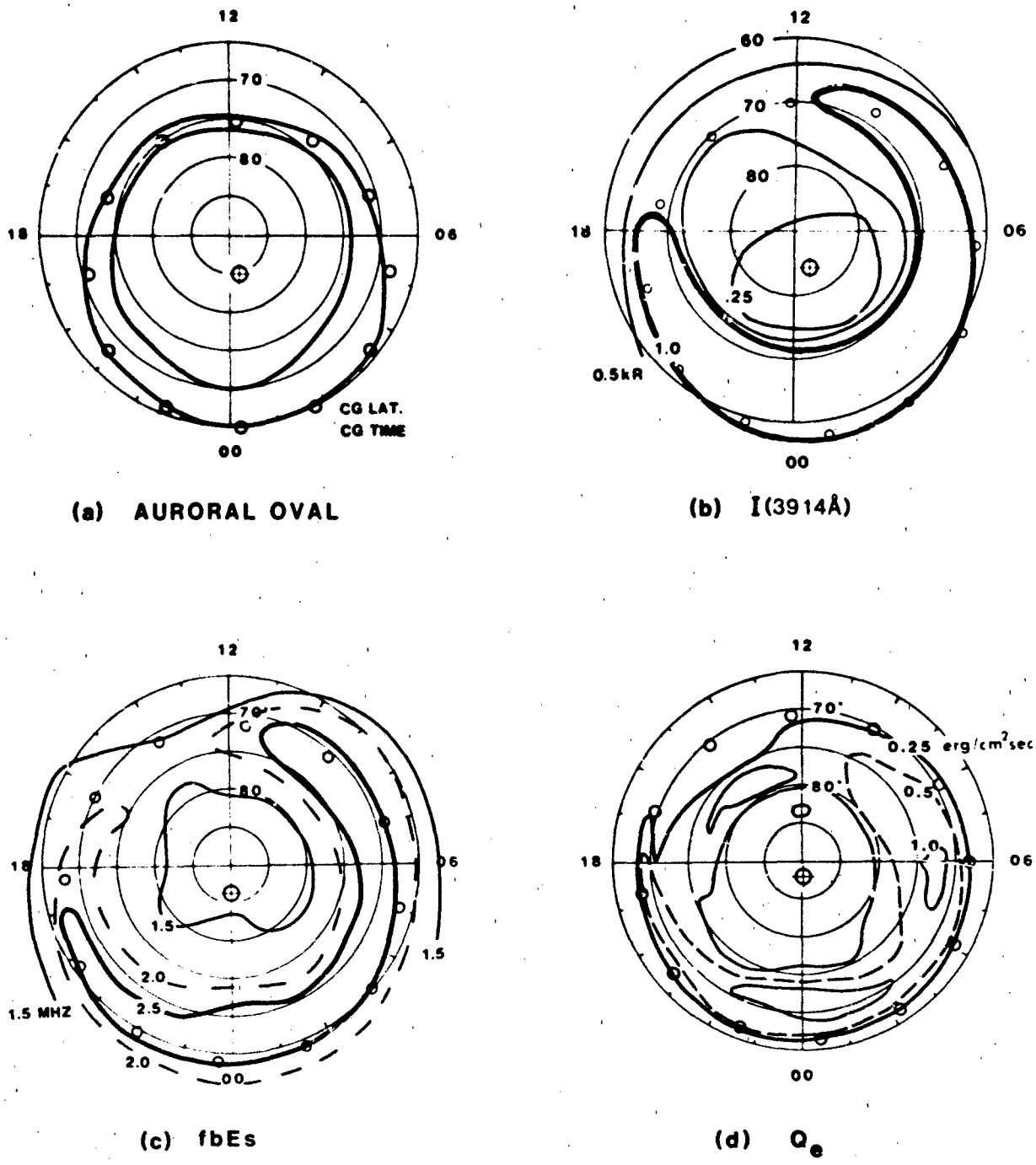


Figure 12-11. Examples of iso-energy flux contours of the continuou. aurora which are approximately circles about a pole offset from the CG pole. Circumference of each offset circle is defined by the 12 small circular points. and the offset pole, by the circled cross. (a) The auroral oval for magnetic index Q = 2 [Feldstein and Starkov, 1967] (b) Median intensity of 3914 Å auroral emissions for IGY [Sandford, 1967] (c) Auroral E ionization measured by median values of fbEs for Kp of 0 and 1 for 1958 [Besprozvonnaya and Shchuka, 1976]. (d) Average precipitating electron energy flux for AE ≤ 100 for 1974–1976 [Spiro et al., 1982].

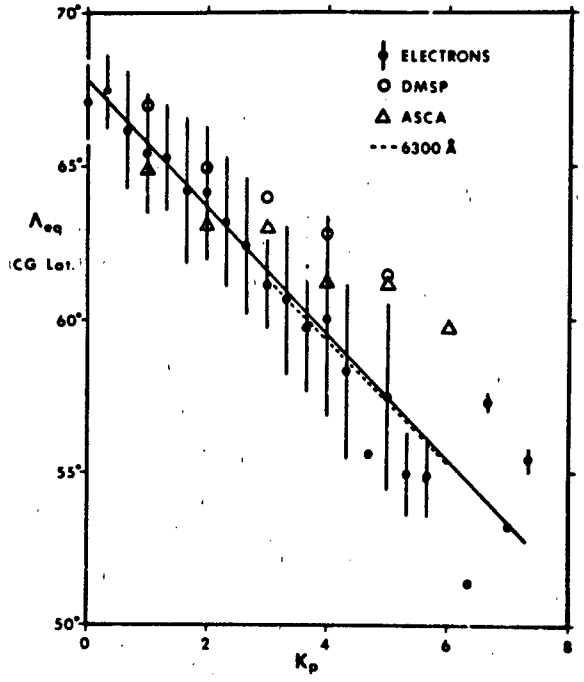


Figure 12-12 Examples of auroral equatorward boundaries near midnight. Electrons: Precipitating electrons [Gussenhoven et al., 1983] ASCA: All-sky camera photographs [Feldstein and Starkov, 1967] DMSP: DMSP images [Sheehan and Carovillano, 1978] 6300 Å: Photometer [Slater et al., 1980]. Error bars indicate standard deviations for the precipitating electrons but are also comparable for the other measurements.

parameters are therefore global indices of this aurora expressing independently the location and intensity of this aurora.

A second latitudinal distribution was found at higher latitude. Gaussian and high latitude distributions corresponded to the regions of particle precipitation named Central Plasma Sheet and Boundary Plasma Sheet by Winningham [1975]. The distributions in magnetic latitude and local time are shown schematically in Figure 12-13.

These distributions have additional significance in that the latitude of the Gaussian maximum, Γmax, is at or near the interface between positive and negative field aligned currents [Robinson et al., 1982]. In addition, Γmax is the equatorward limit of discrete auroras for a large range of conditions [Lui et al., 1977; Whalen, 1983]. Thus the Meng et al. [1977] auroral circles apparently also locate Γmax. Finally, Heelis et al. [1980] find the convection reversal region near the CPS-BPS boundary.

The latitudinal coordinates used in this case study are those approximated by the offset pole at 85.8° CG latitude and 00:00 hr CG time. This pole is determined from the ionospheric data of Besprozvannaya and Shchuka [1976] (Figure 12-11c), and is also very close to the average value which Meng et al. [1977] determined from the DMSP images (85.8°CG latitude, 00:15 h CG time). Although named the offset pole coordinate system, the formulation is an analytical expression similar to that found by Starkov [1969] [Equation (12.1)] for the low latitude boundary of the auroral oval to which the offset circles themselves are approximations. In this formulation the latitude Γ of a point on the earth is given by

$$\Gamma = \Lambda_{CG} + \Theta p \cos \frac{2\pi}{24} (UT - UT_{MN}) \quad (12.3)$$

where Λ_{CG} is the average latitude and has the value of the CG latitude of the point, Θp is the amplitude of the latitudinal variation and has the value of the CG colatitude of the offset pole (4.2° here), UT is universal time and UT_{MN} is the universal time corresponding to CG local midnight for the point considered. During this case study, contours of equal latitude Γ calculated by this expression were measured to be contours of equal energy flux Q, which were accurate to within 0.4° over the CG local time extent of 10 hours.

The function that describes the distribution of energy flux, Q, in offset pole latitudes Γ is

$$Q(\Gamma) = Qmax \exp\left[-1/2\left(\frac{\Gamma - \Gamma max}{\sigma}\right)^2\right] \quad (12.4)$$

where Qmax and Γ max define the peak value as shown in Figure 12-13.

The Gaussian scale parameter σ has been found to have a typical value of 1.4° (3.2°FWHM) and extremes of 1.2° and 2.0° [Whalen, 1981, 1983]. The maximum energy flux Qmax has a typical value of 1 erg cm $^{-2}$s $^{-1}$ and extremes of 0.25 to 12 erg cm $^{-2}$s $^{-1}$ [Omholt, 1955; Hilliard and Shepherd, 1966; Whalen, 1981, 1983].

The latitude of the Gaussian maximum, Γmax, has a median vlaue of 71° and extremes of 64° to 74° [Whalen 1981, 1983]. The circular auroral arcs of Meng et al. [1977] had an average latitude of 71° ± 5°, which is consistent with their being interface arcs [Robinson et al., 1982] at or near Γmax.

The spectrum of precipitated (that is, loss cone integrated) electrons in the Gaussian/CPS region is Maxwellian, so that the differential number flux (in electrons cm $^{-2}$s $^{-1}$ keV $^{-1}$) is

$$\phi(E) = \frac{Q}{2\,E_M^3} E \exp\left(\frac{-E}{E_M}\right), \quad (12.5)$$

where E is electron kinetic energy and E_M the characteristic energy.

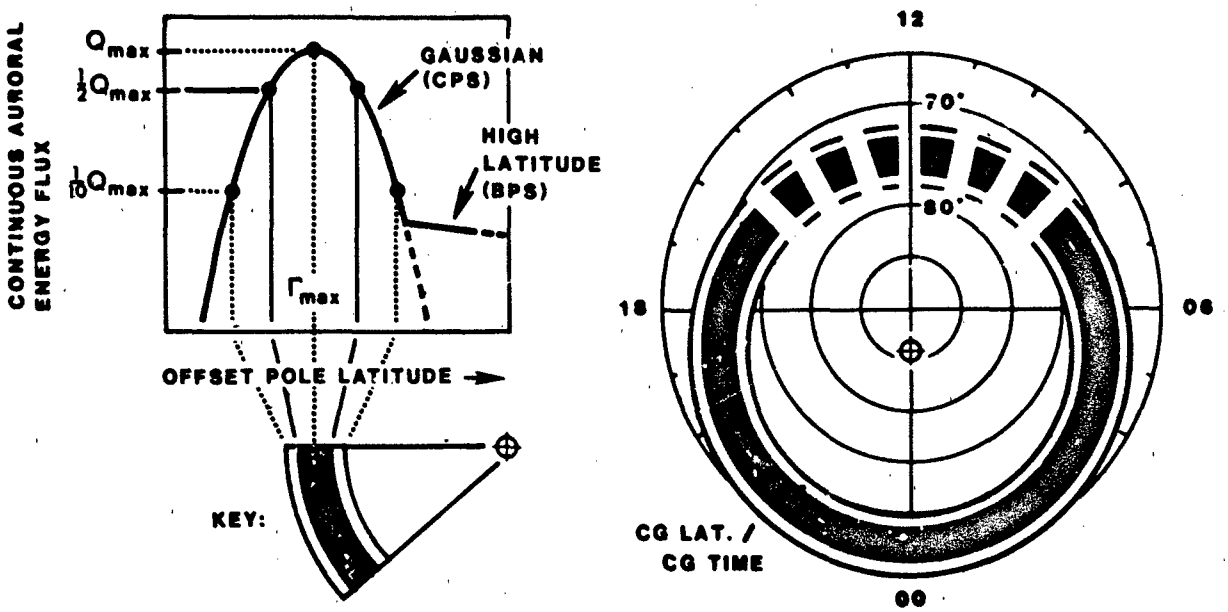


Figure 12-13. Schematic diagrams illustrating the instantaneous distribution of energy flux in the continuous aurora as described in the text. *Left*: Energy flux, Q, as a function of offset pole latitude, Γ, showing the Gaussian/CPS and high latitude/BPS distributions. *Right*: Contours of constant energy flux approximated as circles centered at the offset pole (the circled cross) in CG latitude and CG local time. The offset pole latitudinal distribution, which is therefore the same at all local times except for a region in the day sector, is the Gaussian/CPS distribution shown at the left. The contour of the Gaussian maximum energy flux, Q_{max}, as well as the contours of 1/2 Q_{max} and 1/10 Q_{max} are identified via the sector marked "Key" in the lower left. The dashed contours in the day sector indicate the approximate region in which statistical studies find energy fluxes to be lower than at other local times.

E_M has been found near midnight also to be Gaussian in latitude with a scale parameter close to that of Q.

$E_M$max has been measured at 1 keV [Sharber, 1981]. Inferring particle energies from the production heights measured by Hunten et al. [1963] for this aurora (which they termed diffuse surfaces), $E_M$max was generally between 0.4 and 1.5 keV with extremes of 0.3 and 9 keV. The dependence of $E_M$max on local time however has not been established.

Because of the equivalence of the particle, optical, and ionospheric parameters for the continuous aurora and auroral E-layer (as described in Section 12.1.3.2 and Figure 12-19) the following is true: contours of equal loss-cone integrated electron energy flux Q_c are also contours of equal column integrated N_2^+ emission at 3914 Å, I(3914 Å), as well as contours of equal $foEa^4$, where foEa is the auroral E layer critical frequency. Furthermore, because these three parameters are proportional to one another, for a Gaussian latitudinal distribution, the three have the same Gaussian scale parameter.

Contours of equal $foEa^4$ are also contours of equal maximum electron density, Ne_{max}, as well as contours of equal foEa. However, for a Gaussian latitudinal distribution, the scale parameters are related by the following multiple proportionality:

$$\sigma\,(foEa^4):\ \sigma\,(Ne_{max}):\ \sigma\,(foEa) = 1:\ \sqrt{2}:\ 2. \quad (12.6)$$

12.1.3 Empirical and Model Relationships

12.1.3.1 Auroral-Ionospheric Profiles. This section describes the ionospheric effects of auroral particle precipitation calculated by a number of workers. The effects of continuous and discrete electron auroras are presented as well as those of proton auroras. In general, isotropic pitch angle distribution of particles is assumed. Other distributions have been studied as well as energy spectra that are different from those assumed here [for example, Rees, 1963].

Continuous Aurora—The electron spectrum of the continuous aurora has been found to be Maxwellian. Sharber [1981] has found good agreement between precipitating electrons fluxes and their simultaneously measured ionospheric E-layer effects provided the particles are integrated over the actual loss cone and averaged over intervals of time and space appropriate to the response of the ionosphere itself and of the characteristics of the ionospheric measurements. The energy spectrum which results from this procedure is well represented by a Maxwellian function.

The Maxwellian differential electron number flux as a function of energy has the form

$$\phi(E) = \frac{Q_M}{2E_M^3} E \exp\left(\frac{-E}{E_M}\right) \text{ electrons cm}^{-2}\text{s}^{-1}\text{eV}^{-1}, \quad (12.7)$$

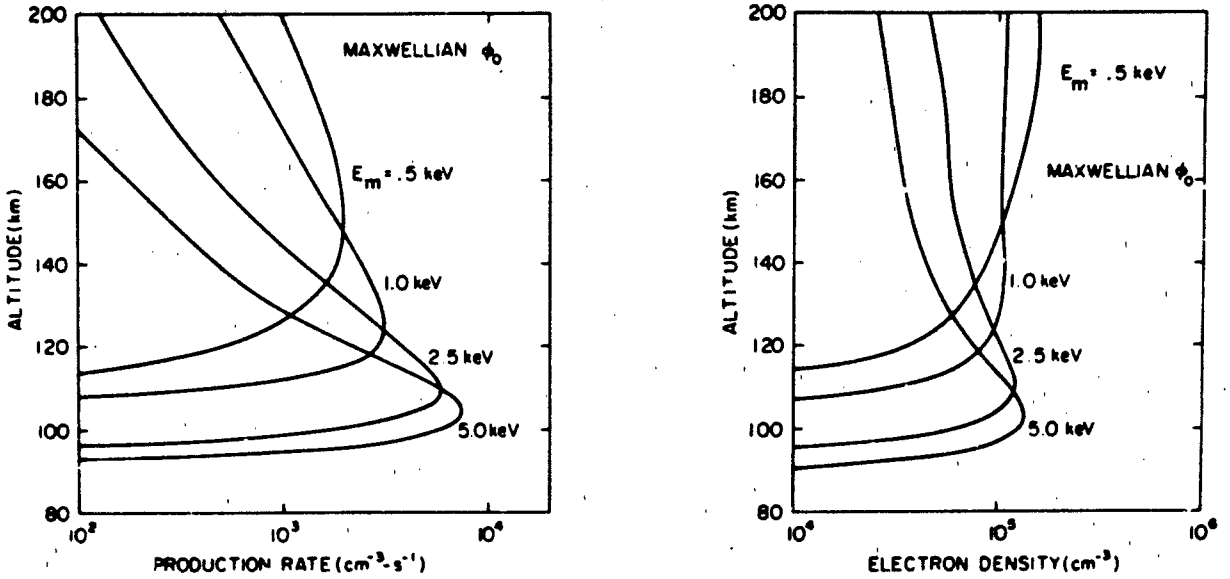


Figure 12-14. Altitude profiles of ion production rates (left) and electron densities (right) resulting from electron precipitation which produces the continuous aurora. Calculations [Strickland et al., 1983] assume Maxwellian spectra, isotropic incidence and energy flux of 1 erg $cm^{-2}s^{-1}$. Curves for 4 values of Maxwellian characteristic energy, E_M, are shown.

where energy E is in eV, energy flux Q_M is in eV $cm^{-2}s^{-1}$ and E_M is the characteristic energy (or temperature).

Since the average energy for the Maxwellian is $2E_M$, the total integral number flux, N_m (electrons $cm^{-2}s^{-1}$) is given by $N_M = Q_M/2E_M$.

Other useful relations can be derived from $N_M = eE_M\phi(E_M)$ where $\phi(E_M)$ is the maximum value of $\phi(E)$.

Altitude profiles of the rates of ion-electron pair production per unit volume resulting from this Maxwellian spectrum have been calculated by Strickland et al. [1983] for four values of E_M and are shown in Figure 12-14 (left). An energy flux of Q_M = 1 erg $cm^{-2}s^{-1}$ has been assumed; profiles for other values of Q_M can be determined from those shown since ionization rate is proportional to Q_M. In addition these profiles can be converted to energy deposition per cm^3 by multiplying the ion production rate by the factor 34 eV per ion pair.

The electron density profiles resulting from these production profiles have been calculated by Strickland et al. [1983] and appear in Figure 12-14 (right). In these calculations equilibrium conditions have been assumed since they are appropriate to the slowly varying nature of this aurora and the E-region.

A plot of the altitude of the peak ionization rate as a function of characteristic energy is shown in Figure 12-15 displaying the agreement in the results of several workers.

Discrete Aurora—The electron spectra that produce the discrete aurora have been modeled by either monoenergetic or Gaussian spectra. Although the actual spectra are more accurately modeled by a Maxwellian spectrum that has been accelerated through a fixed potential, the Gaussian is an adequate approximation for the calculation of the ionospheric response.

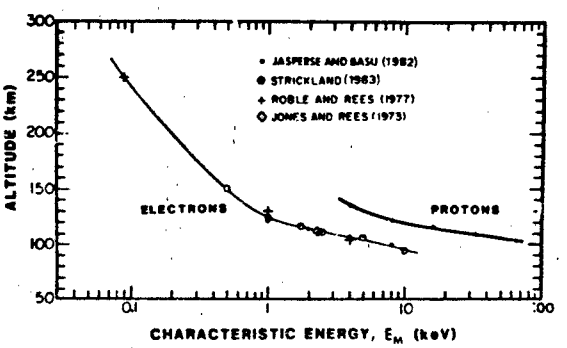
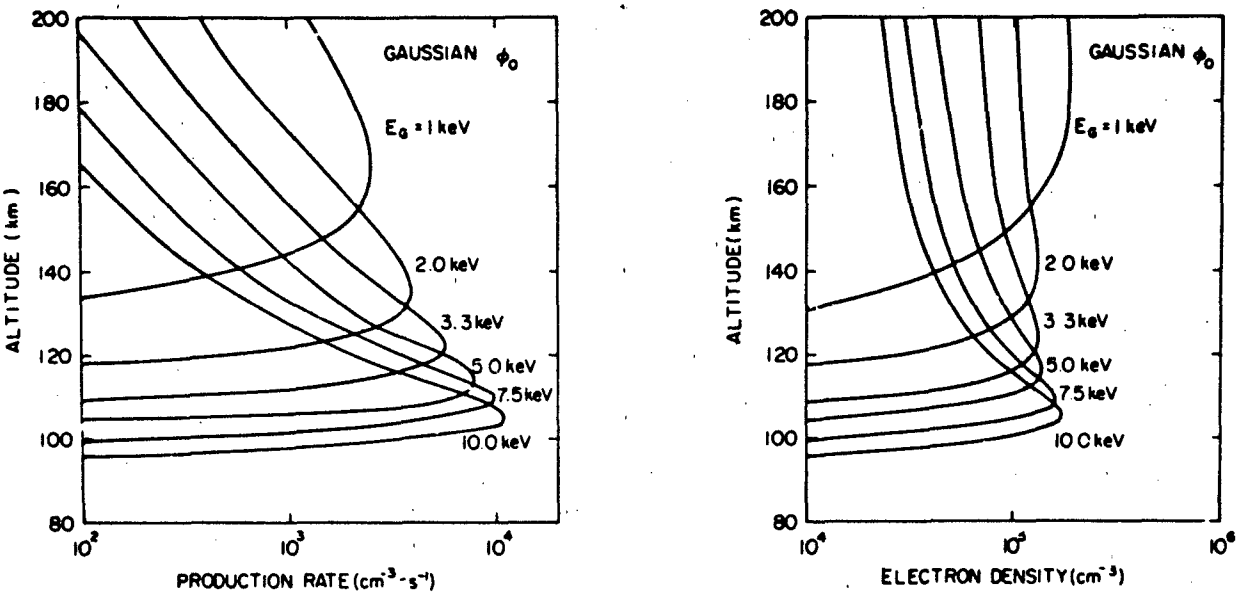


Figure 12-15. Altitude for the maximum ion production rate versus Maxwellian characteristic energy for ionization resulting from the precipitation of particles with Maxwellian spectra and at isotropic incidence.
Left: Electrons which produce the continuous aurora.
Right: Protons



ELECTRON PRECIPITATION
GAUSSIAN SPECTRA (DISCRETE AURORA)

Figure 12-16. Altitude profiles of ion production rates (left) and electron densities (right) resulting from electron precipitation which produces the discrete aurora. Calculations [Strickland et al., 1983] assume Gaussian spectra, isotropic incidence and energy flux of 1 erg $cm^{-2}s^{-1}$. Curves are shown for 6 values of Gaussian maximum energy, E_G.

The representation of the discrete differential number flux is

$$\phi(E) = \frac{Q_G}{\pi^{1/2}\sigma E_G} \exp\left[-\left(\frac{E - E_G}{\sigma}\right)^2\right] \text{electrons cm}^{-2}\text{s}^{-1}\text{eV}^{-1} \qquad (12.8)$$

where E_G is the maximum energy in eV. The Gaussian scale parameter σ has values near 0.2 E_G.

Altitude profiles of ionization rates for values of E_G and for Q_G = 1 erg $cm^{-2}s^{-1}$ appear in Figure 12-16 left and the resulting electron density profiles are shown in Figure 12-16 right [Strickland et al., 1983]. Again electron density has been calculated assuming equilibrium and so applies specifically to auroras that are constant in all respects for periods greater than about 1 min. For auroras that are more transitory, these electron density profiles can be considered to be upper limits.

The altitude of maximum ionization rate (points) and the altitudes at which the rate has fallen to 10% of maximum (bars) is shown in Figure 12-17 for the case of monoenergetic electrons of isotropic incidence [Rees, 1964].

Proton Aurora—Proton spectra have been found to be consistent with Maxwellian spectra [for example, Sharber, 1981]. Ionization rate profiles for precipitating protons with Maxwellian spectra have been calculated by Jasperse and Basu [1982] assuming isotropic incidence. Figure 12-18 shows a profile for each of five values of characteristic energy, E_M for energy flux of 1.0 erg $cm^{-2}s^{-1}$. A plot of the altitude of the peak ionization rate as a function of characteristic energy is shown in Figure 12-15.

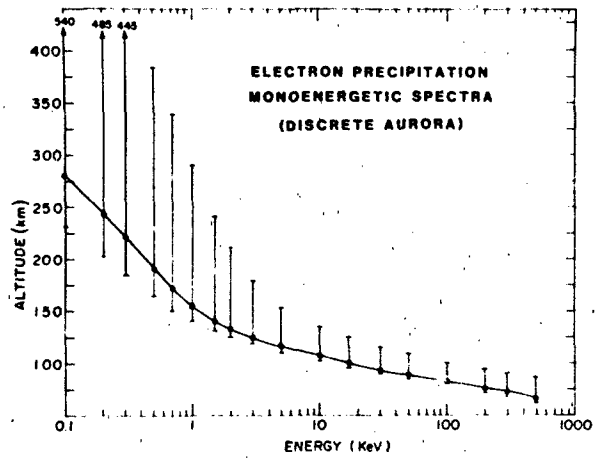


Figure 12-17. Altitude of the maximum ionization rate (points) and the altitudes at which this rate has fallen to 10% of the maximum (bars) versus energy, for ionization produced by electrons which form the discrete aurora as modeled by monoenergetic spectra of isotropic incidence [Rees 1964].

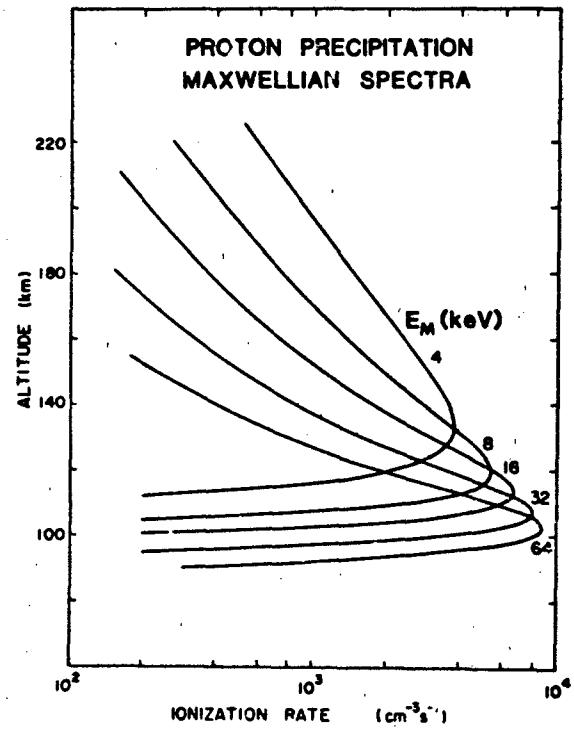


Figure 12-18. Altitude profiles of ion production rates resulting from precipitation of protons of isotropic incidence and Maxwellian spectra of characteristic energy E_M. Energy flux is 1.0 erg cm^{-2}s^{-1} [Jasperse and Basu, 1982].

12.1.3.2 Relations Between Auroral Electron, Optical, and Ionospheric Parameters. Auroral studies are conducted in three major disciplines: particle, optical and ionospheric. Under some conditions parameters in all three are closely enough related so that a measurement in one discipline permits inference of parameters in the other two. This section presents a graph interrelating these parameters in a convenient form as well as a description of the conditions under which these relations apply.

The parameters plotted in columns in Figure 12-19 are as follows:

IBC—The International Brightness Coefficient, a semi-quantitative index for visible auroras.

Q—Energy flux in erg cm^{-2}s^{-1} of precipitated electrons. These are loss cone integrated and subject to the space and time constraints described in Section 12.1.3.1.

I(3914Å)—Column integrated intensity of the N_2^+ emission band at 3914 Å produced by auroral particle precipitation.

The above relationships are well established [Dalgarno et al., 1965] and apply in the E-region to both discrete and continuous auroras. In addition, the relation between Q and I(3914 Å) is correct for the E-region although IBC is not defined there.

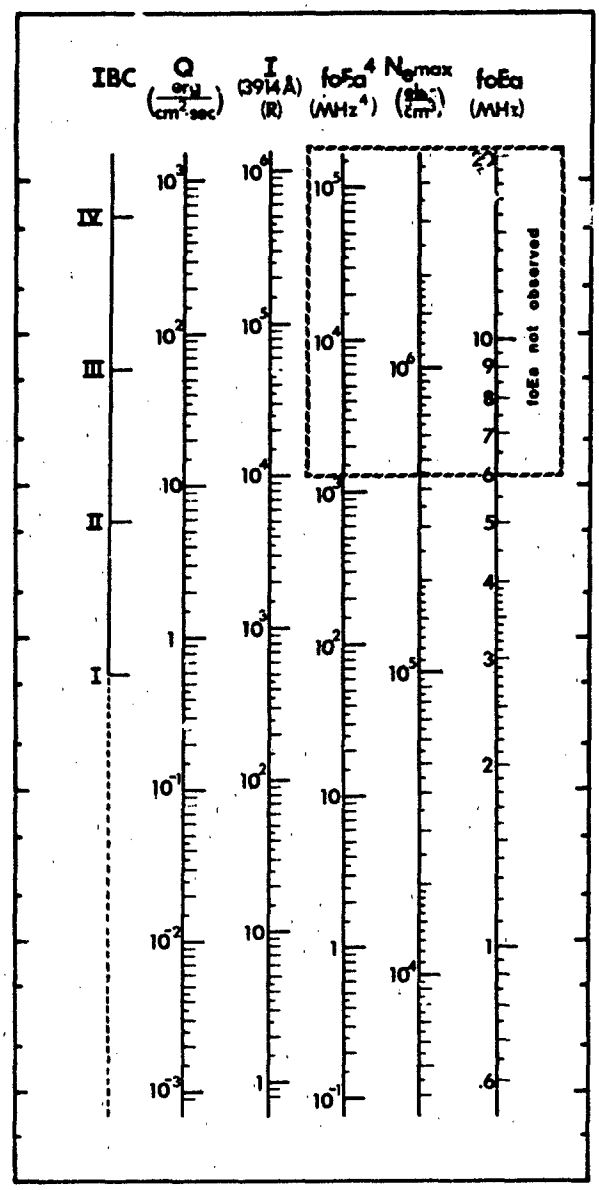


Figure 12-19. Inter-relations of particle, optical and ionospheric parameters for the continuous aurora and auroral E layer.
IBC: International Brightness Coefficient.
Q: Energy flux of precipitating auroral electrons.
I(3913 Å): Column integrated intensity of the N_2^+ emission band at 3914 Å.
foEa^4: Fourth power of auroral E layer critical frequency for the ordinary ray.
Ne_{max}: Electron density of the auroral E layer at its maximum in altitude.

The ionospheric parameters obey these quantitative relations only for the auroral E layer, which is produced by the continuous (or diffuse) auroral precipitation. This layer is extensive, highly uniform and is in equilibrium. foEa, the critical frequency of the ordinary ray measured by the

ionospheric sounder, is the maximum plasma frequency of the E-layer and has been related empirically to both Q and I(3914 Å) via its fourth power, $foEa^4$ [Omholt, 1955; Sharber, 1981] as shown in column 4.

Ne_{max} is the peak electron density in the auroral E-layer, being derived via $foEa = 8.98 \times 10^3 \sqrt{Ne_{max}}$ where foEa is in Hz and Ne_{max} in electrons per cm^3.

12.1.3.3 Auroral Activity. The discrete aurora on the small scale is frequently very unstable and chaotic. However, certain ordering of brightness and extent has been achieved on the large scale using DMSP auroral images the breadth of which is 2200 km, and airborne ASCA photographs, with a diameter of 900 km. The two data bases represent very different time scales, DMSP images being separated in time by 100 min and the ASCA by 1 min.

Auroral images by the DMSP satellite have been categorized as active, moderate, or quiet as described in Table 12-2 [Sheehan et al., 1982]. The distribution of frequency of occurrence appears in Figures 12-20a, b, c, and d showing the relation of occurrence frequency to magnetic indices Kp and AE and dependence on UT. In Figure 12-20e the columns are headed by activity observed on a given orbit; the rows are the occurrence frequencies observed on the following orbit.

Auroral images have been recorded by ASCA at 1-min intervals aboard an aircraft flying at local midnight for durations of 5–10 h [Krukonis and Whalen, 1980]. This data has permitted analysis of auroral activity over a considerable range of time durations as distinct from the DMSP images, the sampling rate of which is limited by the orbital period.

Twelve flights totaling 93 hours provided 5600 photographs in the midnight sector near 69° (±4°) CG latitude. When mounted in a montage format these photographs revealed individual events of auroral activation with measurable lifetimes. Categorizing the auroras in the above manner produced the histograms of the lifetimes or auroral activity shown in Figure 12-21. The combined A, M, and Q categories indicate a characteristic lifetime of auroral activity near 15 min (Figure 12-22).

Table 12-2. Description of the criteria used to rate auroral activity in DMSP images.

Category	Auroral Morphology	Latitudinal Extent	Longitudinal Extent	Comments
N	No aurora visible in a properly exposed image having adequate auroral zone coverage.			Seldom observed in a good quality image
Q	Quiet auroras consiting of a single, thin, structureless discrete arc or broader band of diffuse aurora [Lui et al., 1973] also called continuous aurora [Pike and Whalen, 1974].	≤ 1° CGL (discrete arc)	Up to the entire longitudinal extent of the image (≈ 3–12 hours of local time depending on coordinate orientation)	
M	Moderate state of auroral activity typified by patches, knots, or loops in discrete auroras, including 'westward surges' and bulges along otherwise thin arc structures: also, polar cap arcs extending from oval auroras. Gradations in the moderate range (M−, M, and M+) attempt further refinement, but distinctions are often difficult to make, particularly between M+ and A.	≤ 5° CGL	A few hours of magnetic local time, usually not the entire oval	Statistically most common situation
A	Very active auroral conditions with (apparent) large westward surges, poleward bulges, and a broad band of discrete aurora. With linear amplifiers employed in older satellites, bright auroras can saturate and wash out all detail. Newer satellites with logarithmic detectors can distinguish detail over a wide range of intensities.	≥ 5° CGL	The entire longitudinal extent of the oval visible in an image	M+ and A categories often difficult to distinguish

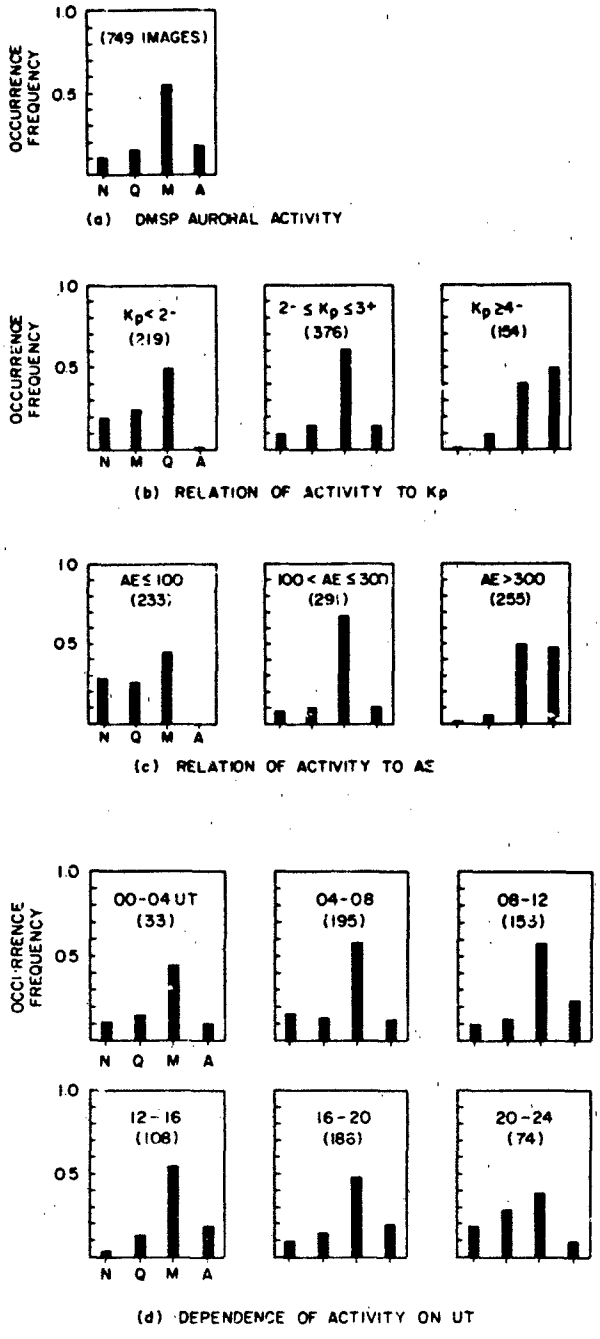


Figure 12-20. Frequency of occurrence of auroral activity observed in the DMSP images in categories of Active, Moderate, Quiet and No visible aurora as defined in Table 12-2 [Sheehan et al., 1982].
a. All cases.
b. Relation to Kp.
c. Relation to AE.
d. Dependence on UT
e. Recurrence of activity where columns are headed by activity observed on initial orbit; the 3 rows are the distribution of occurrence frequency on the following orbit.

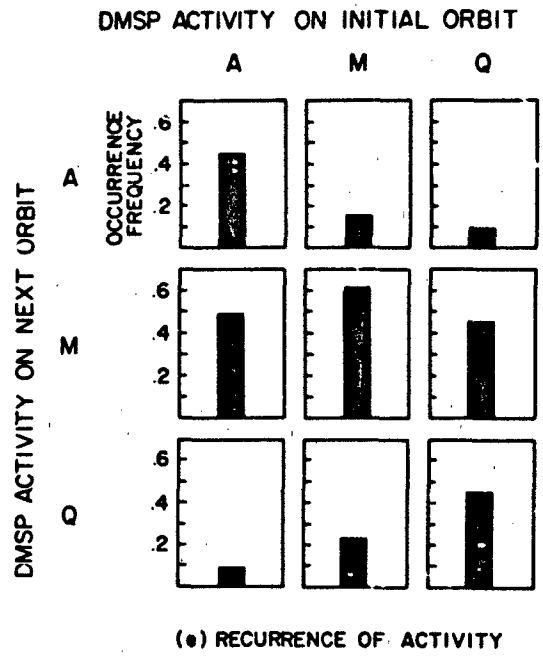


Figure 12-20. (*Continued*)

12.1.3.4 Heights of Discrete Auroras.

Dependence of Heights on Magnetic Latitude and Local Time—The altitude and latitude of peak emission at 5577 Å have been measured for homogeneous arcs by photometers using triangulation methods [Boyd et al., 1971]. The resulting heights versus magnetic latitude are plotted in Fig-

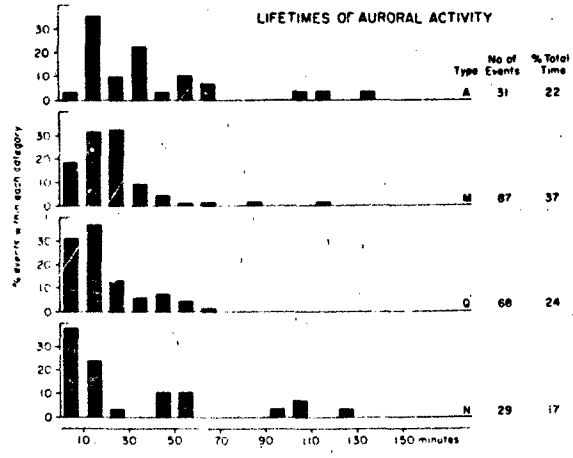


Figure 12-21. Distributions of lifetimes of auroral activity within categories described in Figure 12-20 as determined by airborne all-sky photographs near midnight. Number of events within each category and the percent of the total time represented by each category are listed at the right [Krukonis and Whalen, 1980].

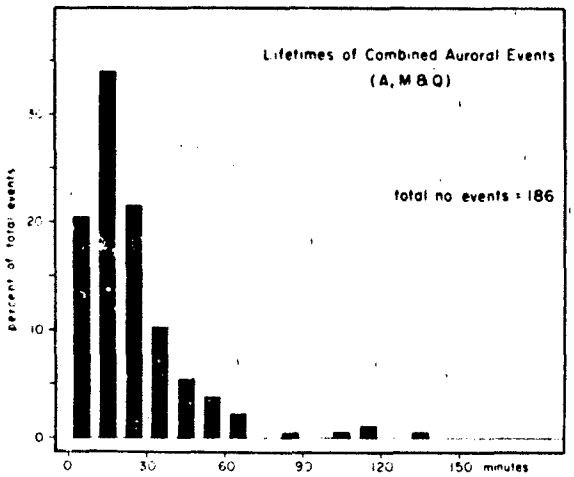


Figure 12-22. Percentage occurrence of lifetime of auroral activity for combined active, moderate and quiet categories (from Figure 12-21) showing Poissonian lifetime of ~15 minutes.

ure 12-23 in four 3-hour intervals of local time encompassing evening through midnight to morning. Heights have been related to the energy at which monoenergetic electrons produce maximum emission via the relations of Rees [1963] and these energies are plotted on the right hand scale.

Two relations are apparent: heights decrease with latitude within each local time interval and heights decrease as local time increases from evening to morning at each latitude where the aurora is observed.

The limitation of this study is that the triangulation technique requires relatively stable arcs, hence the results apply to relatively quiet conditions.

Height-Brightness Relations—A number of workers have reported that discrete auroras increase in brightness with decreasing heights as determined by optical triangulation [for example, Starkov, 1968 and Harang, 1951; Boyd et al., 1971]. Evidence for this effect, which is more quantitative in nature and not as dependent on auroral stability, has resulted from a study of discrete aurora by airborne ionospheric sounder and ASCA during 11 north-south crossings of the auroral oval near midnight [Wagner and Pike, 1972; Gassmann et al., 1972]. Auroral sporadic E echoes recorded by the ionospheric sounder were selected for times when the ASCA films showed aurora to be in the zenith. Virtual height for these echoes, h'Esa, is plotted versus the frequency fEsa in Figure 12-24, together with the qualitative designation of brightness from the ASCA photographs. The trend that brightness increases with decreasing height can be described in a more quantitative way by the approximate translation of frequency to intensity of 3914 Å emission via Figure 12-19. Height can be translated into approximate monoenergetic electron energy via Figure 12-17.

12.2 AURORAL OPTICAL AND INFRARED EMISSIONS

The optical emissions that characterize an aurora are produced by the deposition of energy in the atmosphere by energetic particles, primarily electrons and, in some instances, protons. The energetic particles are derived from

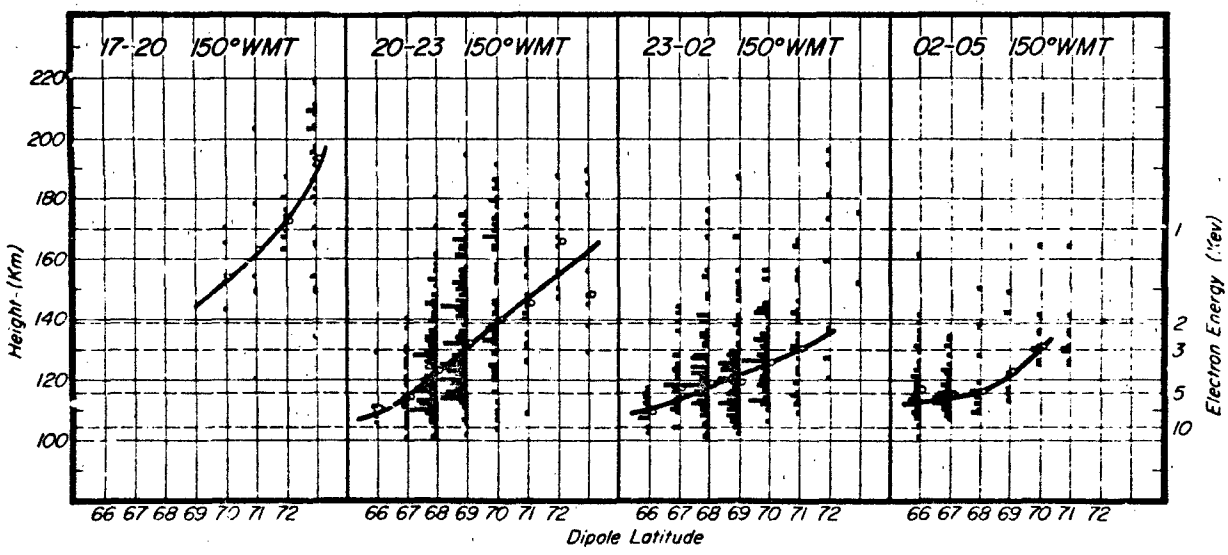


Figure 12-23. Distributions of the height of the peak emission in 5577 Å as a function of dipole latitude for four 3-hour time periods shown as 150 WMT. Local magnetic midnight is approximately 0115 150°WMT. The median point of the distribution in each latitude interval is shown by a circle.

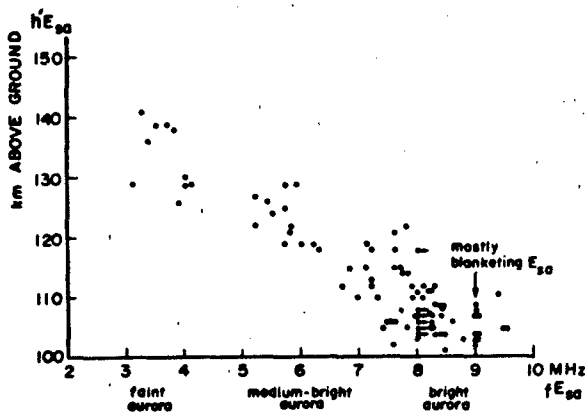


Figure 12-24. Relations between brightness (from airborne ASCA), auroral sporadic E frequency, fEsa, and virtual height of auroral sporadic E layer, h'Esa (from airborne ionospheric sounder) for discrete aurora in the zenith, from 11 latitudinal scans across the oval near midnight by aircraft [Wagner and Pike, 1972].

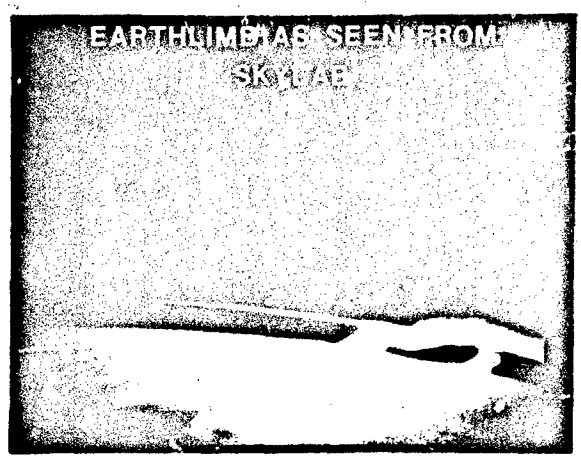


Figure 12-26. Photo of the earthlimb seen from Skylab 3 looking toward the southern auroral zone. The airglow layer runs in a thin band across the center of the photo, while a bright auroral arc crosses the earthlimb at lower right and ends in the foreground in the middle of the scene [Garriott, 1979]

the interaction of the solar wind, a gaseous plasma continuously streaming from the sun's surface, and the earth's magnetic field [Chapter 8]. The currents of charged particles generated by the energy coupling between the solar wind and the magnetosphere produce energetic electrons surging earthward along geomagnetic field lines in the auroral zone. Collisions of the precipitating energetic auroral particles with the increasingly denser atmospheric neutral atoms and molecules in the 200 to 100 km altitude range produce the luminosity associated with auroras illustrated in Figures 12-25, 12-26, and 12-27. As reviewed in Section 12.1, satellite images of auroral luminosity in the polar regions have contributed significantly to the understanding of the interaction of the solar wind and the magnetosphere, auroral substorms, and the morphology of the diurnal cycle of the auroral ovals surrounding the geomagnetic poles [Akasofu, 1981]. Auroral optical emissions have been examined by instrumentation located in ground-based observatories [Vallance Jones, 1974], research aircraft [E.R. Huppi et al., 1974], rocket probes [Stair et al., 1975; K.D. Baker et al., 1977; Stair et al., 1980; Feldman and Gentieu, 1982], and satellites [Hanson, 1973; Huffman et al., 1980; Frank et al., 1982]. These studies involve measurements of spectral distributions, photon emission rates, and altitude profiles. The data are then analyzed to identify the excited states of

Figure 12-25. View from space of intense auroral emission distributed around entire auroral oval. Image was taken at 0241 UT, 8 Nov 1981, with University of Iowa's ultraviolet auroral imaging instrumentation aboard Dynamic Explorer 1 while orbiting 20 000 km over North America. (Courtesy of L.A. Frank.)

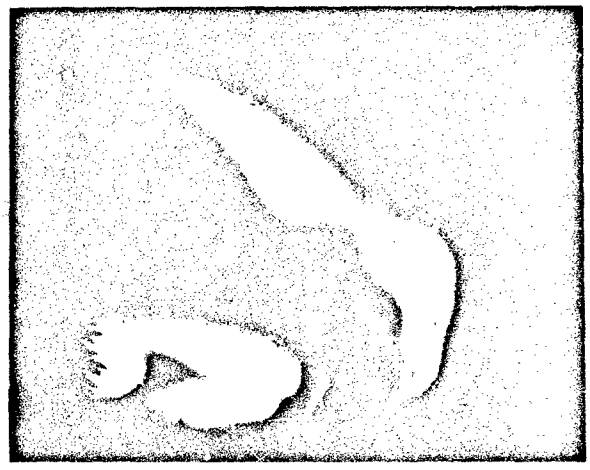


Figure 12-27. Bright auroral arcs recorded from the ground at Poker Flat, Alaska. (Courtesy of A.L. Snyder.)

the radiating atomic and molecular species and to determine their auroral production and loss processes.

Auroral phenomenology, morphology, and occurrence are the subject of Section 12.1. The current section will concentrate on the overall photon emission rates, spectral distributions, and altitude profiles of auroral optical emissions and the identification of the radiating species and their production and loss mechanisms. Particular emphasis will be placed on the infrared portion of the spectrum. Only emissions from aurorally disturbed atmospheres will be discussed; quiescent atmospheric emissions are treated in Chapter 13.

This section begins with a discussion of laboratory studies and their impact on auroral models and continues with a general discussion of emission rates in terms of production and loss mechanisms for auroral species. Processes giving rise to various auroral visible and ultraviolet emissions are discussed, and the section concludes with a discussion of the dominant aurorally enhanced infrared bands, especially the important CO_2 4.3 μm emission and the NO 2.7 μm and 5.3 μm emissions.

12.2.1 Laboratory Studies and Auroral Models

Auroral mechanisms are studied in laboratory experiments involving particle beams incident on atmospheric gases or gas mixtures. Such experiments measure specific parameters that quantify the incident electron energy loss, energy deposition profile, and excited-stated production [Cohn and Caledonia, 1970], as well as the spectral, temporal, and spatial radiative characteristics of the gases. Gas discharges have been used in flowing afterglow experiments [Piper et al., 1981a, b] and chemiluminescent reaction cells [Rawlins et al., 1981b] to determine rate coefficients and photon yields for specific auroral reactions. The laboratory experiments provide an opportunity to study a specific production or loss process for an excited species under controlled conditions of temperature and gas density.

Measured parameters, including cross sections for energy loss by incident electrons and for production of specific new species or excited states [Wadzinski and Jasperse, 1982], have contributed to the generation of a number of auroral models. Such models typically include descriptions of electron energy deposition and excited-state production, usually followed by chemistry codes to describe production of additional species and perhaps by an atmospheric motion code and a radiative transfer code. Descriptions of electron interaction with the ambient atmospheric gas have used a variety of techniques. Simple semi-empirical models of electron energy deposition have been constructed [Rees, 1963] using energy dissipation functions measured in the laboratory. The secondary electron production rate in such models is calculated from the observation that an electron-ion pair is produced for every 34 eV of energy deposited by the primary electron. Once the ionization rates are known, the production rate of secondaries as a function of energy and the secondary electron flux can be calculated [Rees et al., 1969], using analytic or experimental values for secondary electron production cross sections and electron-neutral inelastic cross sections. From the secondary electron flux, excited-state production rates may be deduced. Other descriptions of electron energy deposition and excited-state production have used range-theoretic methods [Stolarski and Green, 1967; Peterson et al., 1973], Monte-Carlo techniques [Berger et al., 1970], and Fokker-Planck [Banks et al., 1974] and Boltzmann [Strickland et al., 1976] transport equation methods. In addition to the above time-independent approaches, recent time-dependent models taking atmospheric motion into account have been formulated [Roble and Rees, 1977].

The models have been tested against the results of field experiments. Advances in instrumentation design and sensor technology have resulted in steady improvement of the precision, spectral resolution, and minimum detectable emission level in the field and laboratory experiments. The resulting improved and expanding data set has broadened the scope and increased the precision of our understanding of auroral processes.

12.2.2 Artificial Auroral Experiments

In recent years, rocketborne electron accelerators have been used in a variety of active experiments in the upper atmosphere [Winckler, 1982], including the investigation of wave-particle interactions, beam-plasma discharge, particle trapping in the geomagnetic field, generation of radio-frequency electromagnetic waves, and artificial auroral experiments (Figure 12-28). A series of artificial auroral experiments using rocketborne electron accelerators has been conducted at AFGL under the EXCEDE program [O'Neil et al., 1978a and b, 1979, 1982]. In artificial auroral ex-

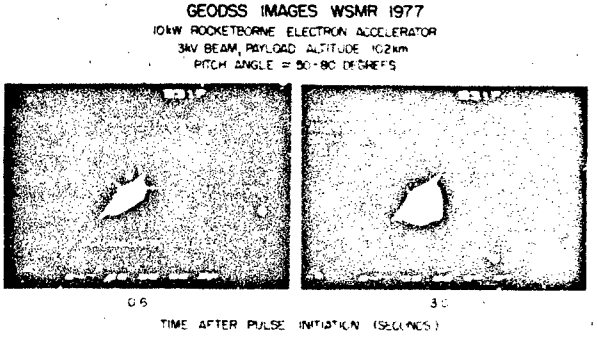


Figure 12-28. Ground-based image of emission from EXCEDE artificial auroral experiment taken through 31-inch GEODSS telescope at White Sands, New Mexico. The prompt emissions and afterglow are shown at two different times after pulse initiation.

periments such as EXCEDE, the dose conditions (primary electron energy, ionization rate, dose time, and dose altitude) are carefully controlled and defined. By contrast, in natural auroras the dose conditions must be inferred from the observations and may be less certain because the effects are integrated over a range of altitudes and dose levels. Consequently, artificial auroral experiments provide a useful adjunct to observations of naturally occurring auroras in the attempt to understand auroral phenomena.

12.2.3 Emission Rates

The kinetic energy of the auroral particles is converted to optical and infrared emissions of discrete wavelengths through processes initiated by inelastic collisions with ambient atmospheric species. These species may spontaneously relax to lower energy states directly while emitting radiation. Alternatively, they may undergo one or more chemical reactions, finally resulting in an excited product which also emits energy by spontaneous radiative decay. The specific auroral processes involved in the production of various excited species continue to be the focus of much aeronomic study. Excellent general reviews of auroral processes are given by Chamberlain [1961], Vallance Jones [1974], and Bates [1982]. The reader is also referred to the textbooks on atmospheric radiation by Goody [1964] and on infrared and optical spectroscopy by Herzberg [1950, 1945, 1967] and Huber and Herzberg [1979].

The intensity of atomic and molecular transitions from upper state u to lower state ℓ is given by the relation

$$I_{u\ell} = N_u A_{u\ell}. \qquad (12.9)$$

where $I_{u\ell}$ is the optical intensity in photons $cm^{-3} s^{-1}$, N_u is the population of the upper state in molecules cm^{-3}, and $A_{u\ell}$ is the Einstein A coefficient, or radiative transition probability, in s^{-1}. If auroral activity has continued long enough for the production and loss processes of a given state to come to equilibrium, the number of excited molecules per cm^3 at steady state, $N_{u_{ss}}$, is given by

$$N_{u_{ss}} = P_u/L_u. \qquad (12.10)$$

where P_u is the production rate per unit volume in $cm^{-3}s^{-1}$ and L_u the loss rate in s^{-1}. The loss rate of a given state may be described by the relation

$$L_u = \sum_{\ell} A_{u\ell} + K_{uO_2}[O_2] + K_{uN_2}[N_2] + K_{uO}[O] + \ldots, \qquad (12.11)$$

where K_{uM} is the rate coefficient in cm^3s^{-1} molecule^{-1} for collisional deactivation (or quenching) of the excited state by species M.

The production of a given excited state may be the result of electron impact in inelastic collisions of the form

$$e + N \rightarrow N^* + e, \qquad (12.12)$$

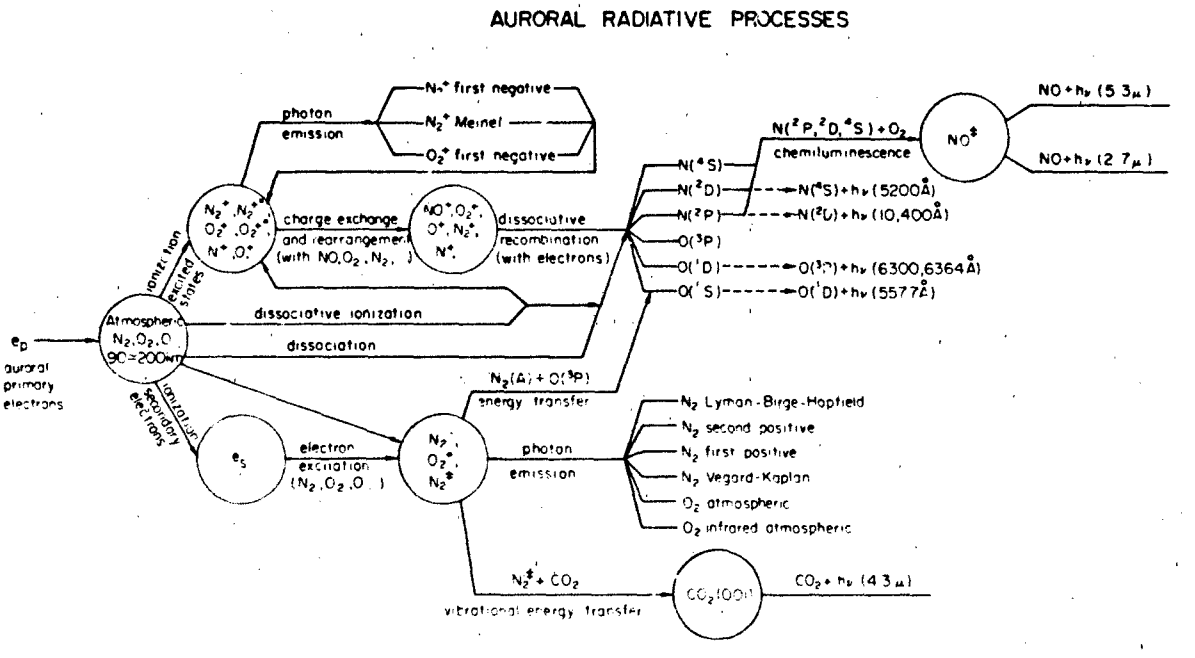


Figure 12-29. Schematic diagram of auroral radiative processes, showing sources of most important infrared emissions at 2.7, 4.3, and 5.3 μm.

the result of energy transfer

$$M^* + N \rightarrow M + N^*, \qquad (12.13)$$

or the result of a chemiluminescent reaction

$$A + BC \rightarrow AB^* + C. \qquad (12.14)$$

Here the asterisk indicates an excited state of the atom or molecule. In the case of energy transfer, production of the excited state N* has resulted from the collisional deactivation or quenching of the excited M* to some lower energy state. For the chemiluminescent reaction, some fraction of the exothermicity of the reaction has been converted to internal energy of the product molecule. A comprehensive review of auroral production and loss processes is given in the monograph by Vallance Jones [1974]. Production and loss processes for metastable states were recently reviewed by Torr and Torr [1982].

12.2.4 Auroral Processes

A flow diagram indicating some physical and chemical processes leading to selected ultraviolet, visible, and infrared auroral emissions is presented in Figure 12-29. This figure is schematic and indicates only examples of selected auroral production and loss mechanisms. All primary electrons with energies in excess of 100 eV partition their kinetic energy in the atmosphere in approximately the same branching ratios. That is, regardless of the primary electron energy, approximately the same fraction of the initial kinetic energy is spent in forming various ions (N_2^+, N_2^+*, N^+, N^+*, O_2^+, O_2^+*, O^+, . . .) or various neutral excited states (N_2^*, O_2^*, $N_2^\ddagger$, $CO_2^\ddagger$, $NO^\ddagger$, . . .). Here and in the remainder of this chapter, the asterisk is used to indicate electronically excited species, while the symbol ‡ indicates vibrational excitation.

Excited-state ions produced when the primaries ionize atmospheric neutrals may relax by spontaneous decay to a

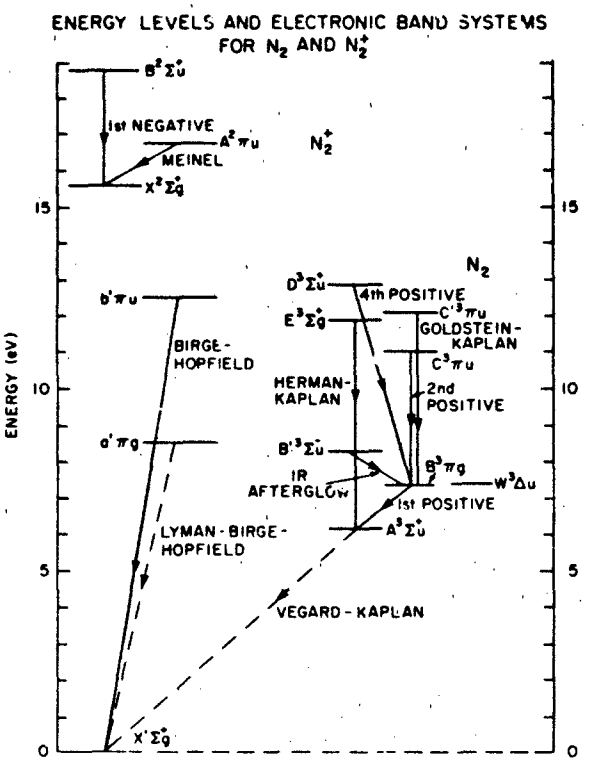


Figure 12-30. Electronic energy levels and principal electronic band system of N_2 and N_2^+ [adapted from Vallance Jones, 1974]. Forbidden transitions are indicated by dashed lines.

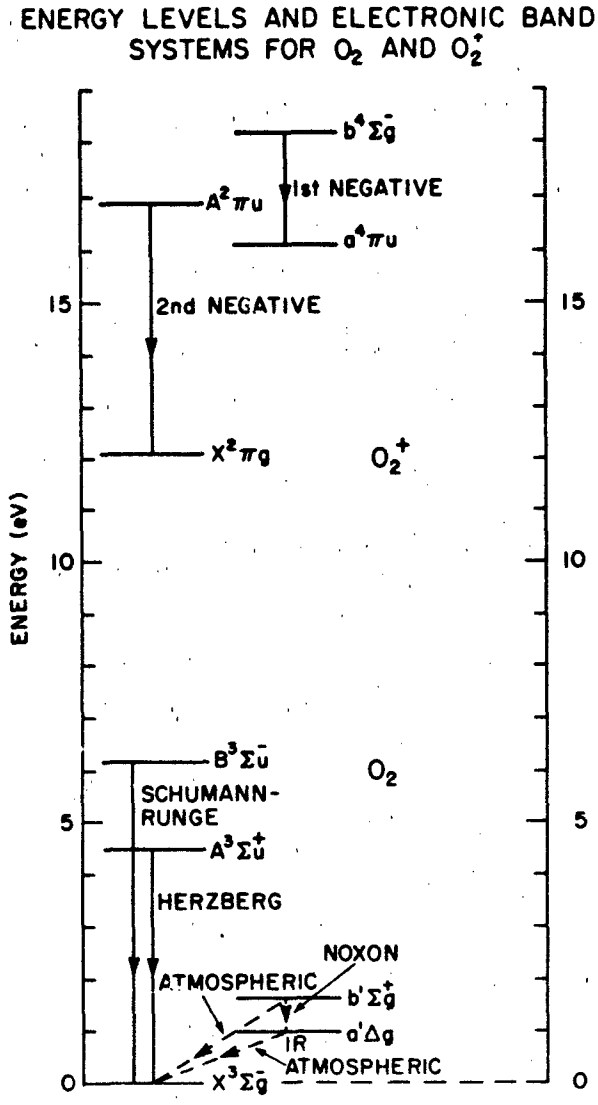


Figure 12-31. Electronic energy levels and principal electronic band systems of O_2 and O_2^+ [Vallance Jones, 1974]. Forbidden transitions are indicated by dashed lines.

lower state emitting a quantum of energy or photon in the process. The N_2^+ first negative ($B^2\Sigma_u^+ - X^2\Sigma_g^+$), N_2^+ Meinel ($A^2\Pi_u - X^2\Sigma_g^+$), and O_2^+ first negative ($b^4\Sigma_g^- - a^4\Pi_u$) systems (Figures 12-30 and 12-31) are examples of electronic transitions of molecular ions producing significant levels of auroral emissions. The secondary electrons resulting from ionization of the ambient atmosphere account for a large fraction of the primary electron's energy as it is slowed and stopped in the atmosphere by successive ionizing collisions. The secondary electrons lose most of their energy in inelastic collisions which result in excited states of neutral atmospheric species, N_2*, O_2*, and N_2‡. The N_2 ($a^1\Pi_g$) state, giving rise to the Lyman-Birge-Hopfield system, and vibrationally excited N_2‡ are representative of the excited-state neutral species produced by secondary electrons, as illustrated in Figures 12-29 and 12-30. Thus, the neutral excited species are produced by impact with neutrals of secondary electrons, which themselves result from ion-pair production by the more energetic primary auroral electrons. These excited neutral species, in turn, radiate, giving rise to aurorally enhanced emissions. Some of the affected band systems of N_2 and O_2 are shown in Figures 12-30 and 12-31.

The fraction of the primary electron energy radiated in a given optical transition as an energetic electron and its secondary electrons are stopped in the atmosphere is defined as the electron-induced radiant efficiency for that transition. The electron-induced radiant efficiency is independent of electron energy for energies greater than approximately 100 eV. Radiant efficiencies for selected auroral emissions are presented in Table 12-3, which indicates that 21 to 34% of the electron energy radiates in these auroral emissions, with the infrared emissions of CO_2 and NO accounting for large fractions of the incident electron energy. Table 12-3 is based

Table 12-3. Auroral IBC III intensities and electron-induced radiant efficiencies.

Emission	Zenith brightness[a](kR)	Electron-induced radiant efficiency (X 10^{-2})
N_2 first positive system[b]	900	1.6
N_2 second positive system[c]	90	0.5
N_2 Lyman-Birge-Hopfield system	380	4.4
N_2 Vegard-Kaplan system	55	0.3
N_2^+ first negative system	150	0.7
N_2^+ first negative (0-0) band	98	0.45[d]
N_2^+ Meinel system[e]	770	1.3
O_2 atmospheric system	1300	3.0
O_2 infrared atmospheric system	2500	3.5
O_2^+ first negative system	26	0.1
NO γ band	60	0.5
NI 8680 Å (^4P–$^4D^o$)	10.5	0.022
NII 2144 Å (^3P–$^5S^o$)	32[f]	0.3
OI 7774 Å ($^5S^o$–^5P)	9.6	0.022
OI 8446 Å ($^3S^o$–^3P)	11.5	0.024
OI 5577 Å (^1D–^1S)	100[d]	0.32
OI 6300, 6364 Å (^3P–^1D)[g]	2–100	0.0056–0.28
OI 1302–1306 Å (2^3P–3^3S^o)	30	0.4
NO first overtone 2.7 μm[h]	0.5–1.4 × 10^3	0.4–1.0
CO_2 (001–000) 4.3 μm[i,j]	3–30 × 10^3	1–12
NO fundamental 5.3 μm[k]	6–12 × 10^3	2–4
		Total 20.8–34.7

[a]Intensities are from Vallance Jones [1974], unless otherwise indicated.
[b]Vallance Jones and Gattinger [1976].
[c]Vallance Jones and Gattinger [1975].
[d]Reference value.
[e]Gattinger and Vallance Jones [1981].
[f]Sharp [1978].
[g]Intensity and efficiency are strongly altitude (or primary electron energy) dependent.
[h]R.J. Huppi and Stair [1979].
[i]Values are typical for rocketborne measurements of this strongly self-absorbed band. The primary production process, vibrational energy transfer from N_2‡, is a slow process, and steady-state radiance efficiencies are difficult to establish. See Stair et al. [1975].
[j]See also Gordiets et al. [1978].
[k]Inferred from Caledonia and Kennealy [1982].

principally on the auroral multiplet and band system intensities presented by Vallance Jones [1974] for an IBC III aurora and assumes an electron-induced radiant efficiency of 4.5×10^{-1} for the N_2^+ first negative (0-0) band at 3914 Å.

The band system intensities of Table 12-3 are given in Rayleighs (R). The unit, proposed by Hunten et al [1956], is an equivalent column radiance of 10^6 photons cm^{-2} s^{-1} and was named in honor of the fourth Lord Rayleigh, who made the initial measurement of the absolute intensity of the night airglow. Standards of auroral optical intensity were proposed by Seaton [1954] and Hunten [1955] based on the brightness of the most intense visible feature of the aurora, the $O(^1S - {}^1D)$ 5577 Å auroral green line (Figure 12-32).

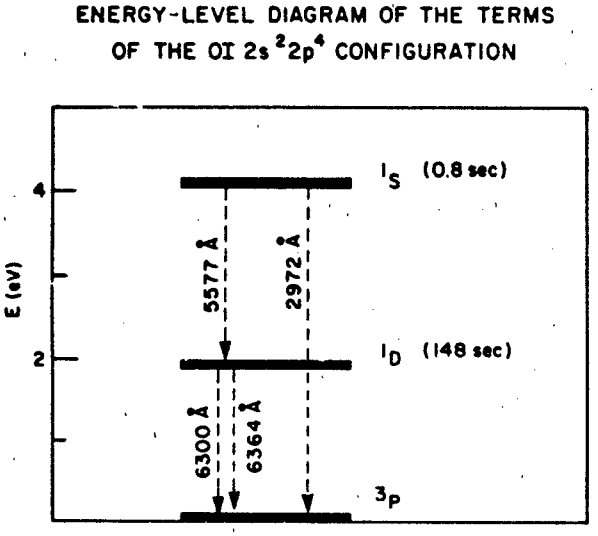


Figure 12-32. Lowest-lying energy levels and forbidden transitions in atomic oxygen. The radiative lifetimes of the ^1S and ^1D levels are indicated.

The result was the establishment of a logarithmic scale, the International Brightness Coefficient (IBC), defined in Table 12-4. In Figure 12-19 the IBC is related to other auroral parameters, such as the electron energy flux, peak E region electron density at steady-state, and optical radiance at 3914 Å.

The probability of radiative decay, or Einstein transition probability $A_{u\ell}$, is governed by quantum-mechanical selection rules [Herzberg, 1950], the most important of which are the electric-dipole selection rules. Transitions obeying electric dipole selection rules have high probability for decay (large $A_{u\ell}$ values) and are designated as allowed or permitted transitions. The N_2^+ first negative, N_2^+ Meinel (Figure 12-30), and O_2^+ first negative (Figure 12-31) transitions are examples of allowed electronic transitions. Transitions violating the selection rules have small transition probabilities and are denoted as forbidden transitions. Excited states from which only forbidden transitions originate are metastable. Forbidden transitions in atomic oxygen, atomic nitrogen, molecular oxygen, and molecular nitrogen are important auroral transitions (Figure 12-29), and their intensities and altitude profiles have been used as indices of both auroral brightness and auroral color type.

Table 12-4. International Brightness Coefficient (IBC).

IBC Class	Brightness (kR), $O(^1S–{}^1D)$ 5577 Å
I	1
II	10
III	100
IV	1000

The color of an aurora varies markedly and has been used to classify auroras into color classes according to the International Auroral Atlas (IAA), published by the International Union of Geodesy and Geophysics in 1963. Since energetic electrons penetrate to lower altitudes, the color variations are primarily due to differing mean energies and energy distributions for the primary electrons incident on the atmosphere. Metastable species with small Einstein transition probabilities are vulnerable to collisional deactivation at relatively low gas densities (high altitudes). The $O(^1D)$ state (Figure 12-32) with an A value of 6.8×10^{-3} s^{-1} is quenched in collisions with N_2 at altitudes less than 300 km. Thus, at altitudes less than 300 km, the 6300 Å auroral red line emissions from this state to $^3P_2$ and $^3P_1$ levels are diminished relative to other visible emissions. As a consequence, ground-based measurements of the 4278 Å N_2^+ first negative (0 – 1) band intensity I(4278) and of the atomic oxygen red line intensity I(6300) can be used to determine the total auroral electron flux and characteristic electron energy, or equivalently the altitude profile of energy deposition [Rees and Luckey, 1974; Shepherd and Eather, 1976]. While I(4278) is proportional to the total energy deposited, I(6300)/I(4278) is sensitive to the altitude profile of energy deposition and decreases in value for lower altitude auroras (more energetic primary electrons).

Similarly, the $O(^1S)$ state, which is the source of the green line (Figure 12-32), has an A coefficient of 1.25 s^{-1} and is quenched by O_2 at altitudes less than 100 km. The collisional destruction of this species contributes to the red lower border of the type-B red aurora. The $N_2(A^3\Sigma_u^+)$ state (Figure 12-30) is the major source of $O(^1S)$ in the aurora through the energy-transfer process (Figure 12-29)

$$N_2(A^3\Sigma_u^+) + O(^1P) \rightarrow O(^1S) + N_2, \quad (12.15)$$

as reported by O'Neil et al. [1979] and Piper [1982]. Einstein A coefficients for molecular transitions of auroral interest, in some cases summed over all possible lower states, are given in Table 12-5, while A coefficients for atomic oxygen transitions are the reciprocals of the lifetimes listed in Figure 12-32. Ultraviolet and visible auroral spectra re-

Table 12-5. Einstein coefficients for auroral molecular transitions.[a]

System	v′	$\sum_{v''} A_{v'v''}(s^{-1})$
N_2 first positive ($B^3\Pi_g$-$A^3\Sigma_u^+$)	0	1.12 (+5)[b]
(2.53–0.478 μm)	1	1.29 (+5)
	2	1.43 (+5)
	3	1.54 (+5)
	4	1.64 (+5)
	5	1.73 (+5)
	6	1.81 (+5)
	7	1.98 (+5)[c]
	8	2.12 (+5)[c]
	9	2.27 (+5)[c]
	10	2.31 (+5)[c]
	11	2.39 (+5)[c]
	12	2.43 (+5)[c]
N_2 second positive ($C^3\Pi_u$-$B^3\Pi_g$)	0	2.73 (+7)
(0.546–0.268 μm)	1	2.75 (+7)
	2	2.73 (+7)
	3	2.67 (+7)
	4	2.50 (+7)
N_2 Wu-Benesch($W^3\Delta_u$-$B^3\Pi_g$)	0	6.00 (−2)
(2.2–4.3 μm)	1	5.00 (+2)
	2	2.04 (+3)
	3	4.28 (+3)
	4	7.26 (+3)
	5	1.08 (+4)
	6	1.45 (+4)
	7	1.84 (+4)
N_2 Vegard-Kaplan ($A^3\Sigma_u^+$-$X^1\Sigma_g^+$)	0	5.30 (−1)
(0.532–1.25 μm)		
N_2 Lyman-Birge-Hopfield ($a^1\Pi_g$-$X^1\Sigma_g^+$)[d]		6.7 (+3)
(0.260–0.100 μm)		
N_2 Herman-Kaplan ($E^1\Sigma_g^+$-$A^3\Sigma_u^+$)[e]		3.0 (+3)
(0.274–0.213 μm)		
N_2^+ first negative ($B^2\Sigma_u^+$-$X^2\Sigma_g^+$)	0	1.41 (+7)
(0.587–0.286 μm)	1	1.38 (+7)
	2	1.4 (+7)
N_2^+ Meinel ($A^2\Pi_u$-$X^2\Sigma_g^+$	0	6.04 (+4)
(0.177–0.550 μm)	1	7.20 (+4)
	2	8.30 (+4)
	3	9.34 (+4)
	4	1.03 (+5)
	5	1.12 (+5)
	6	1.21 (+5)
	7	1.30 (+5)
	8	1.37 (+5)
	9	1.45 (+5)
	10	1.52 (+5)

System	v'–v"	$A_{v'v''}$ (s^{-1})	Band origin (μm)
O_2 atmospheric ($b^1\Sigma_g^+$-$X^3\Sigma_g^-$)[f]	0–0	8.5 (−2)	0.7619
	1–0	6.9 (−3)	0.6882
	2–0	(1.636 ± 0.040) (−4)	0.6287
	1–1	(7.04 ± 0.42) (−2)	0.7708
	0–1	4.16 (−3)[g]	0.8645
O_2 IR atmospheric ($a^1\Delta_g$-$X^3\Sigma_g^-$)[f]	0–0	2.58 (−4)[g]	1.269
	0–1	1.75 (−6)[g]	1.580
O_2 Noxon ($b^1\Sigma_g^+$-$a^1\Delta_g$)[f]	0–0	1.5 (−3)	1.908

[a]All data are from Loftus and Krupenie [1977] unless otherwise indicated.
[b]The notation "a(n)" means "ax10^n".
[c]Eyler and Pipkin [1983].
[d]Corrected for a → a' cascade [Freund, 1972].
[e]Using relative transition probabilities for E → A, E → C, and E → B from Freund [1969] and lifetime of E state from Borst and Zipf [1971].
[f]Krupenie [1972].
[g]Calculated from upper-state lifetime and Franck-Condon factor.

sulting from the EXCEDE artificial auroral experiment are presented in Figures 12-33a and b. Natural auroral spectra in the ultraviolet and visible are shown in Figures 12-34 through 12-37.

Caution is urged in applying artificial auroral spectra to situations involving natural aurora. The former are biased toward prompt emitters, and delayed emitters such as the CO_2 4.3 μm band may not be measured efficiently. This is true of ground-based spectra such as Figure 12-33b due to continual motion of the electron gun through the atmosphere into undosed regions. However, the remark applies especially to rocketborne spectra, where the spectrometer can only observe a dosed region for a short while.

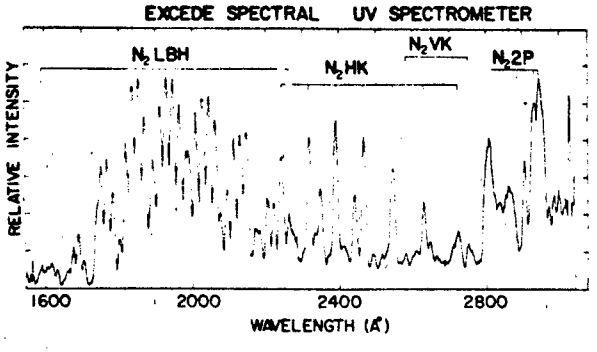


Figure 12-33a. Ultraviolet spectrum taken during EXCEDE SPECTRAL artificial auroral experiment at 125 km altitude. Principal band systems seen are the Lyman-Birge-Hopfield (LBH), Herman-Kaplan (HK), Vegard-Kaplan (VK), and second positive (2P) systems of N_2 [O'Neil et al., 1982].

12.2.5 Infrared Auroral Emissions

The development of infrared cryogenic instrumentation and sensors has facilitated the measurement of infrared auroral emissions in recent years [Stair et al., 1983]. Infrared artificial auroral spectra from EXCEDE are shown in Figures 12-38a and b, while natural airglow and auroral emissions in the 0.88 to 5.4 μm range are illustrated in Figures

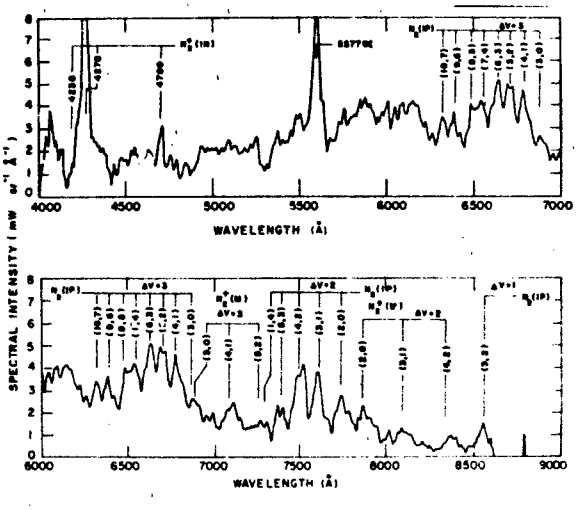


Figure 12-33b. Visible spectrum taken from the ground with an image-intensified spectrograph during the PRECEDE artificial auroral experiment with payload at 92 km. Principal band systems seen are N_2^+ first negative (1N) and Meinel (M) systems and the N_2 first positive (1P) system [O'Neil et al., 1978a].

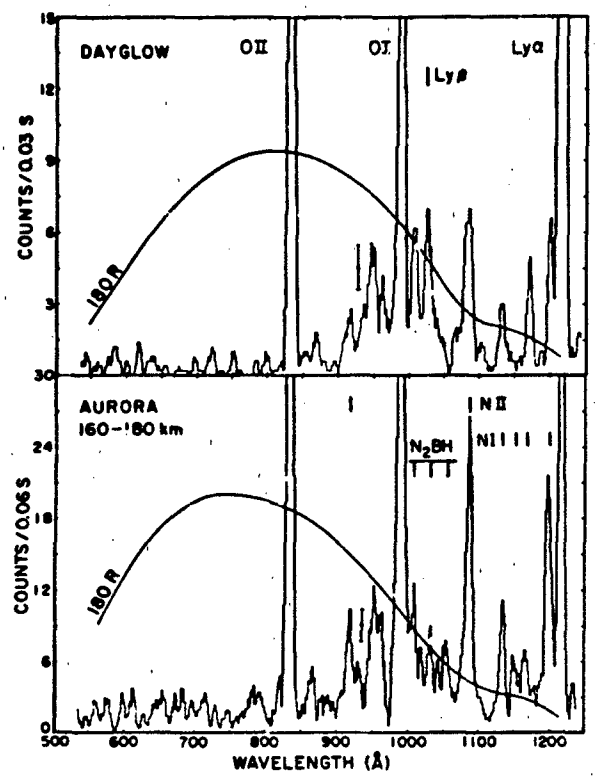


Figure 12-34a. Extreme ultraviolet side-looking spectrum (525–1250 Å) of a natural aurora taken at 160–180 km altitude (bottom), with a comparable measured dayglow spectrum (top). Principal band system is the N_2 Birge-Hopfield (BH) system. Prominent N I and N II lines and off-scale O I, O II, and Ly-α lines are also present [Feldman and Gentieu, 1982].

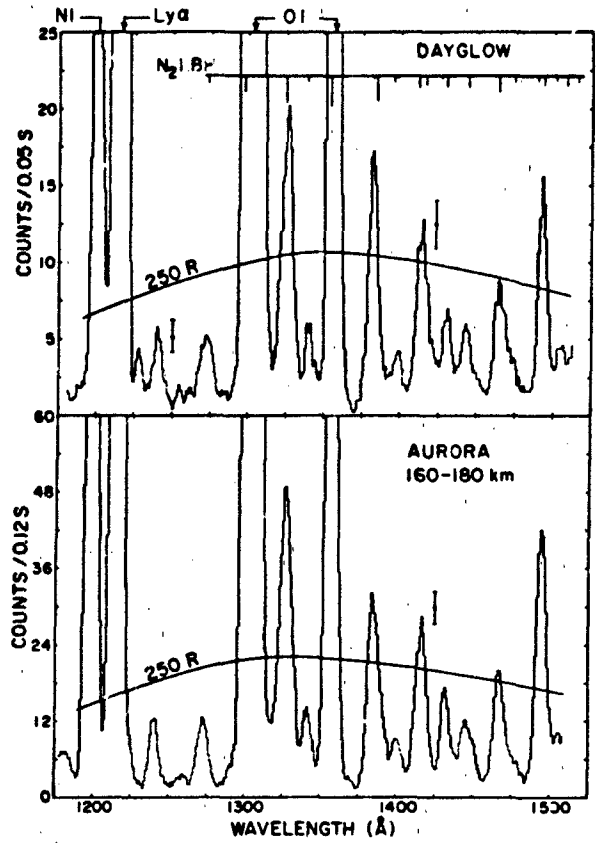


Figure 12-34b. Far ultraviolet auroral spectrum (1175–1520 Å) under the same conditions as Figure 12-34a. Dominant features are the N_2 Lyman-Birge-Hopfield (LBH) band system and atomic emissions from N I (1200 Å), Ly α (1216 Å), and O I (1304 and 1356 Å) [Feldman and Gentieu, 1982].

12- 39, 12-40, and 12-41. Spectra of the near-infrared portion of this range are also shown by D. Baker et al. [1977]. Emissions in this wavelength range correspond to energies varying from 1.4 eV at 0.88 μm to 0.23 eV at 5.3 μm. The energy levels involved in these less energetic transitions producing infrared emission are both electronic states of N_2 and O_2 (Figures 12-30 and 12-31) and vibrational transitions within the ground electronic state of minor atmospheric species. Pure vibrational transitions are forbidden in the homonuclear diatomic molecules N_2 and O_2 forming the major constituents of the atmosphere. Thus infrared-active minor atmospheric species, such as the diatomics OH, CO, NO, and NO^+ and the triatomic species CO_2, N_2O, NO_2, and O_3 become the principal sources of infrared airglow [Chapter 13] and auroral emissions at longer wavelengths. The potential energy curve and fundamental and first-overtone vibrational transitions of NO are shown in Figure 12-42, and the band origins and Einstein coefficients are given in Table 12-6. The vibrational energy levels of CO_2

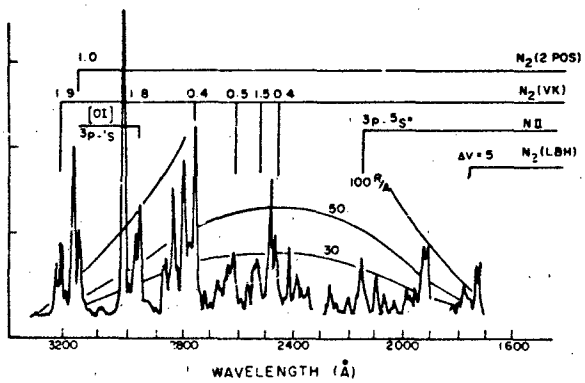


Figure 12-35. Far ultraviolet auroral spectrum from a nadir-looking satellite (1700–3200 Å). Principal band systems are the N_2 Lyman-Birge-Hopfield (LBH), Vegard-Kaplan (VK), and second positive (2POS) systems [Sharp and Rees, 1972; Sharp, 1978]. The dotted curve is a synthetic spectrum.

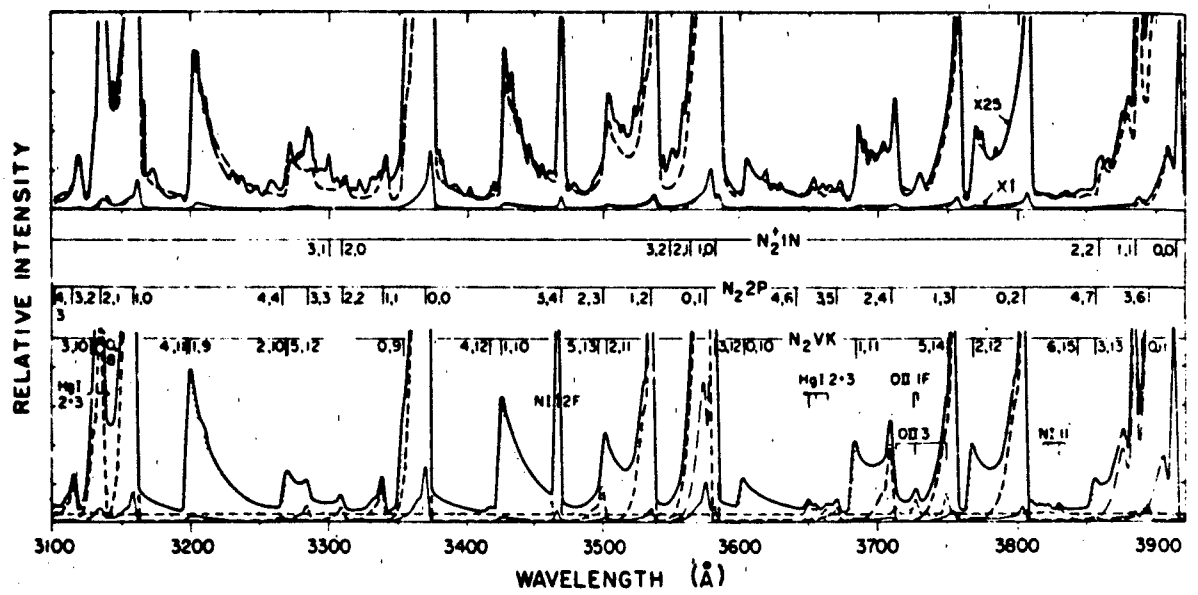


Figure 12-36a. Observed ultraviolet ground-based spectrum of an aurora from 3100–3920 Å (top) along with synthetic spectrum (bottom). The synthetic spectrum is also shown as a dotted line in the top panel, and various components of the synthetic spectrum are shown in the bottom panel. Several band systems are shown with the same notation as in Figures 12-33a,b [Vallance Jones and Gattinger, 1975].

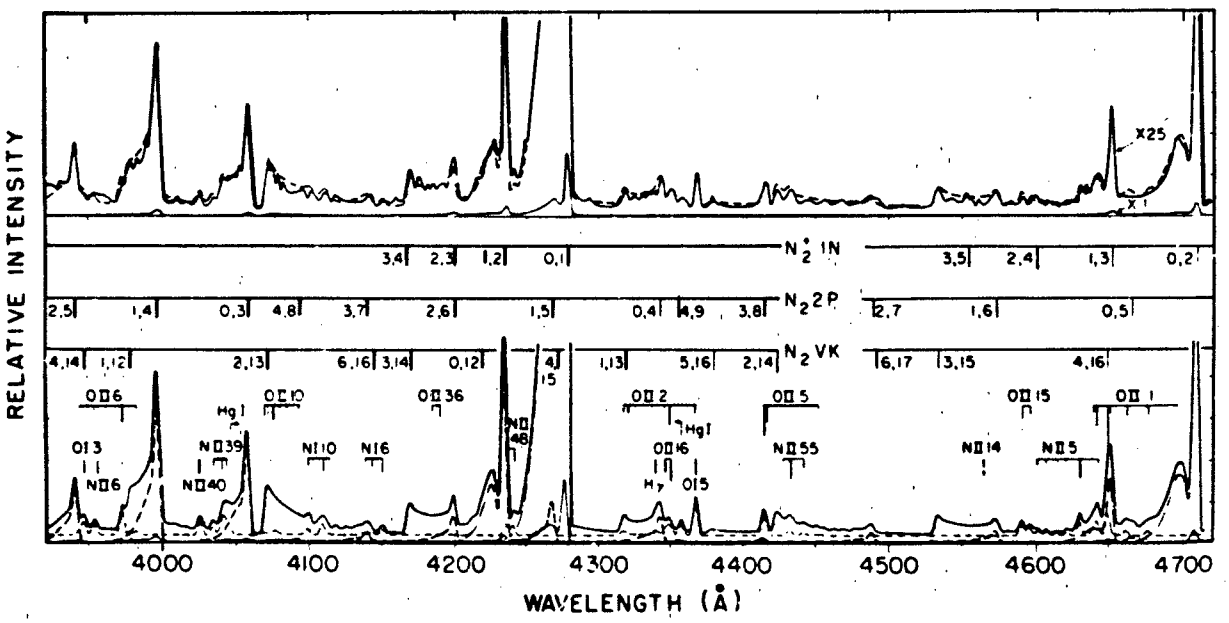


Figure 12-36b. Ground-based visible auroral spectrum from 3920– 4720 Å. Details are as in Figure 12-36a [Vallance Jones and Gattinger, 1975].

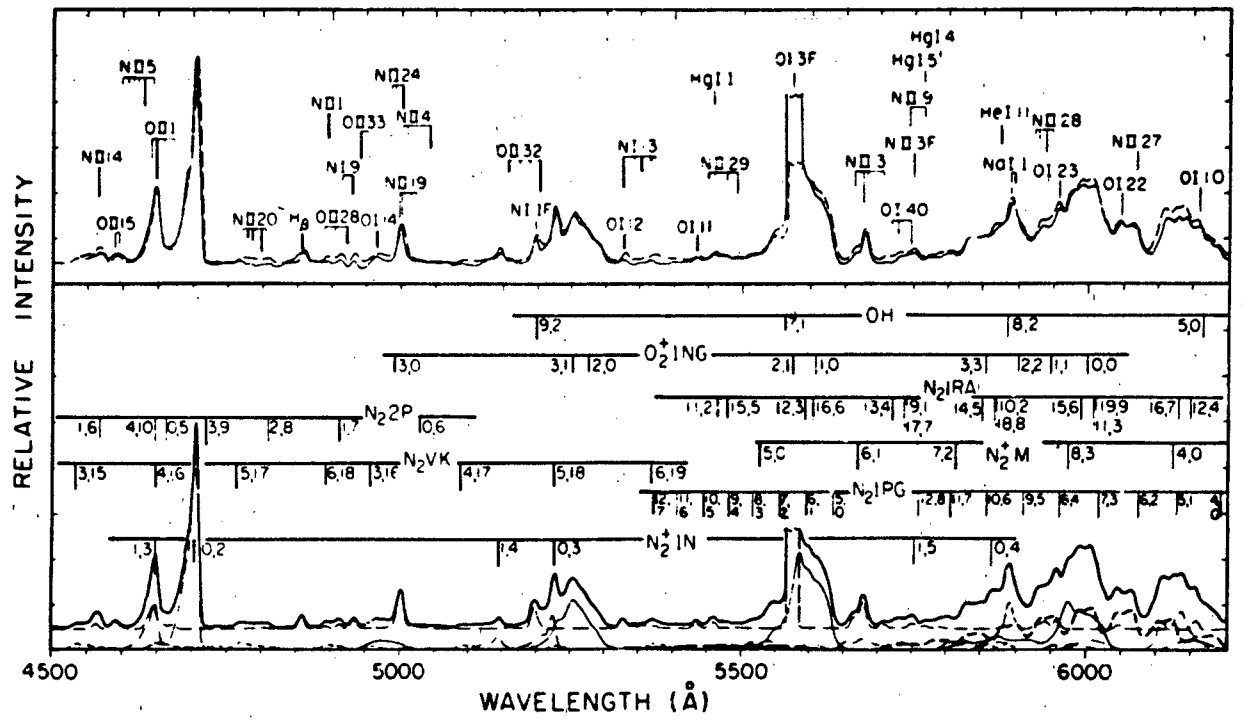


Figure 12-37a. Ground-based visible auroral spectrum from 4500– 6200 Å. Details are as in Figure 12-36a. Additional band systems shown are the O_2^+ first negative (1N) and the N_2 infrared afterglow (IRA) [Gattinger and Vallance Jones, 1974]

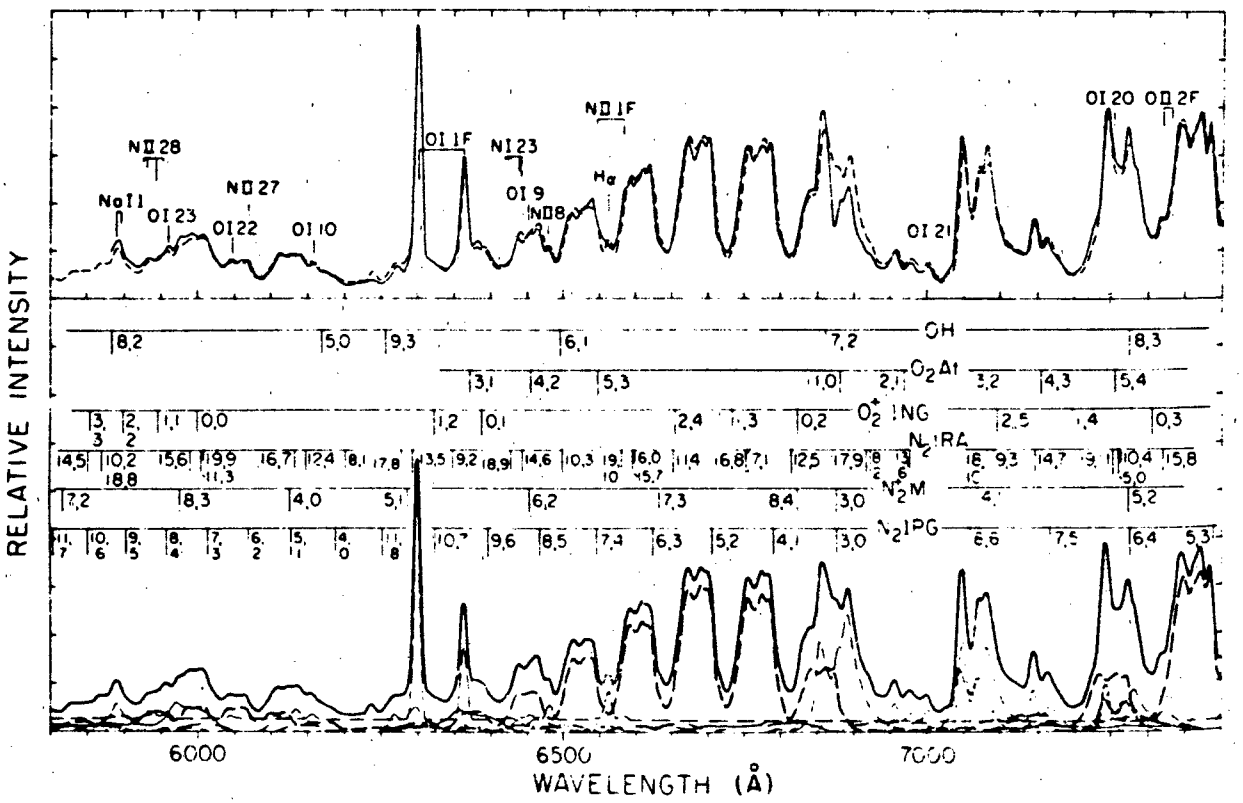


Figure 12-37b. Same as Figure 12-37a, but for spectral range 5800– 7400 Å. One additional band system, the O_2 atmospheric (Atm), is present.

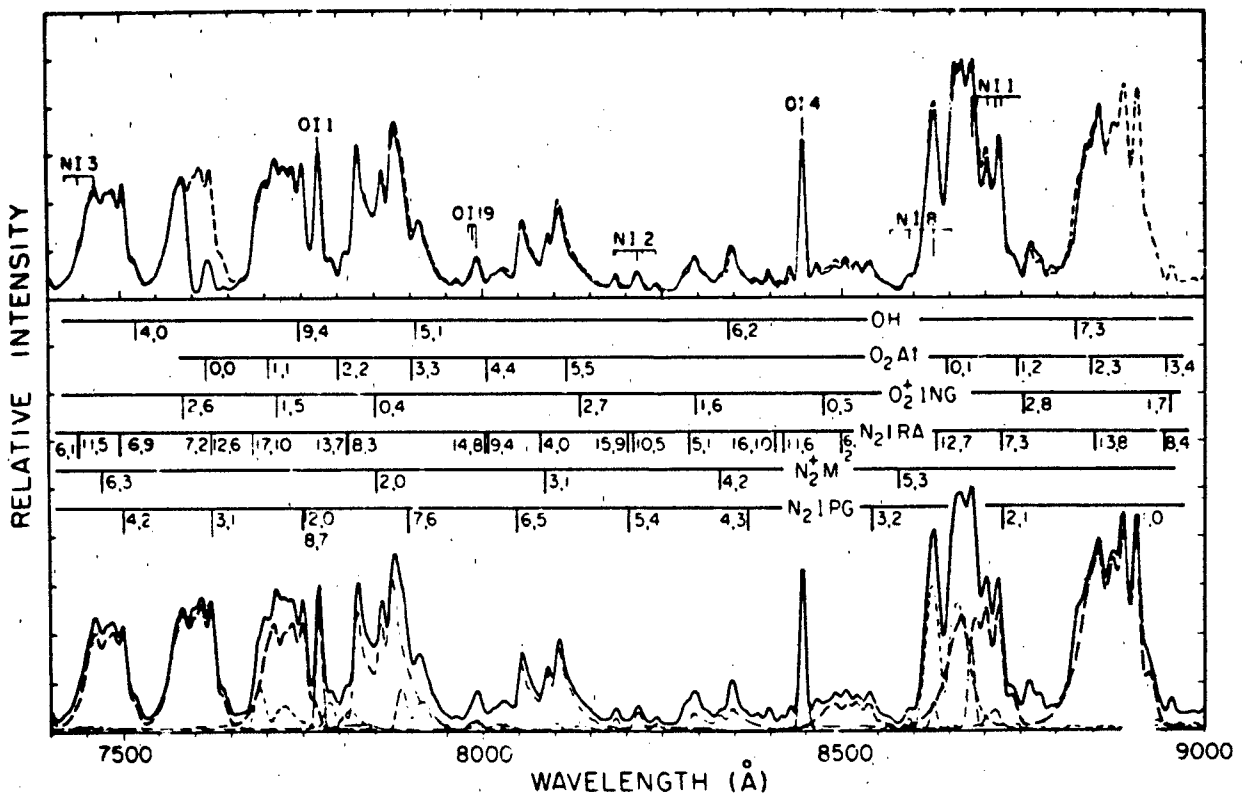


Figure 12-37c. Same as 12-37a, but for the spectral range 7400–9000 Å.

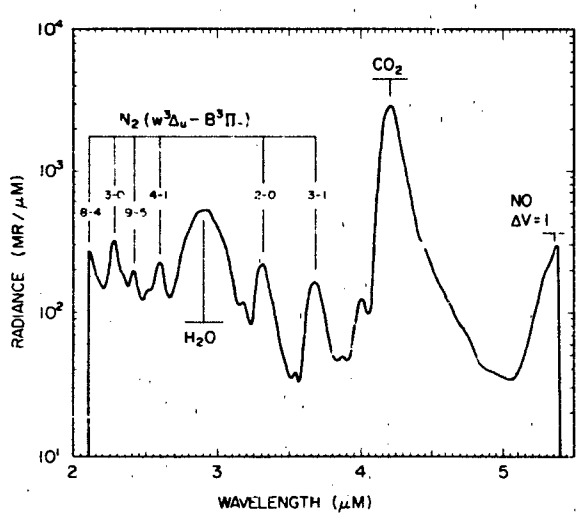


Figure 12-38a. Short-wave infrared spectrum taken of electron-induced emission with liquid-nitrogen-cooled spectrometer during EXCEDE artificial auroral experiment at altitude of 74 km. Dominant emissions are the N_2 Wu-Benesch band system, as well as NO and CO_2 vibrational bands [O'Neil et al., 1982].

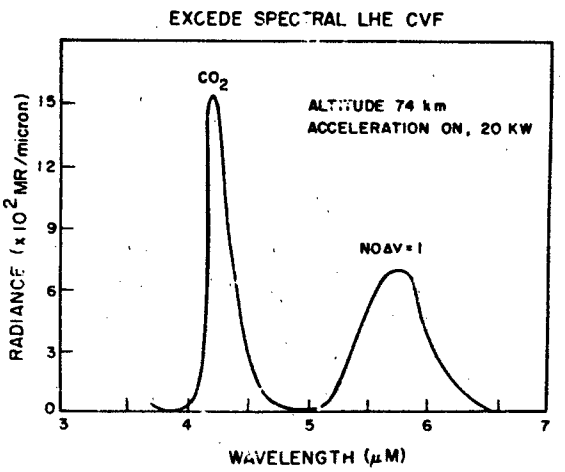


Figure 12-38b. Electron-induced infrared spectra taken with liquid-helium-cooled spectrometer at 74 km during EXCEDE artificial auroral experiment [O'Neil et al., 1982].

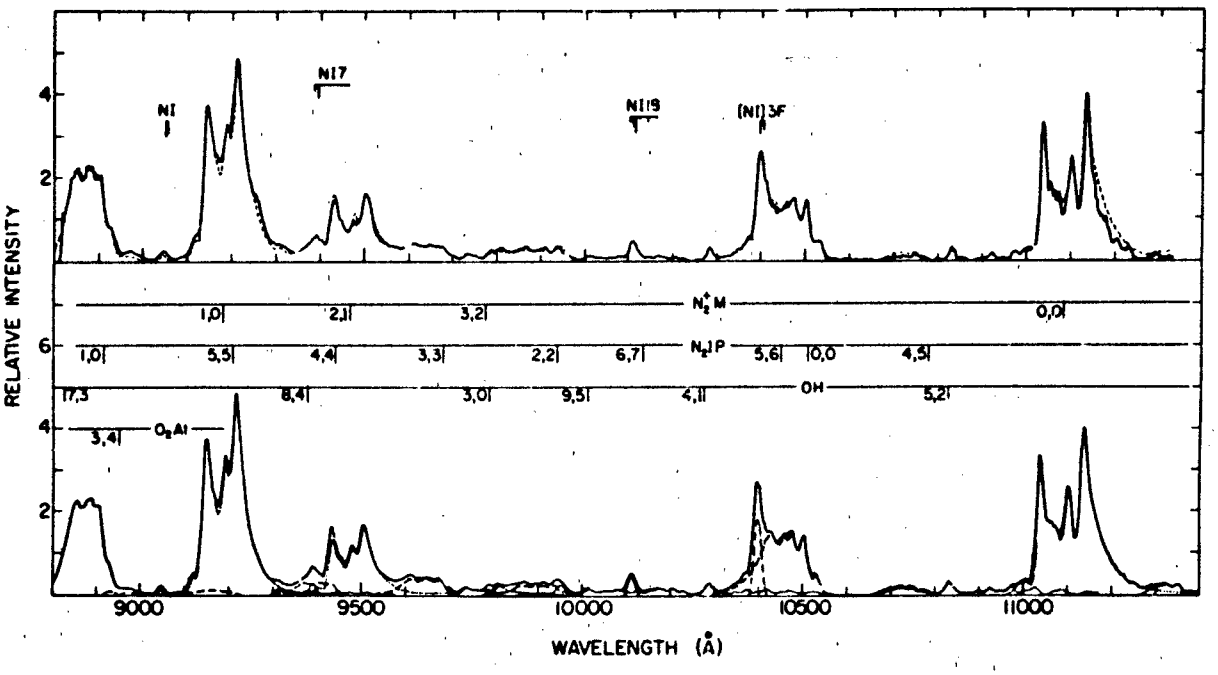


Figure 12-39. Observed ground-based auroral near-infrared spectrum from 8800–11400 Å (top), along with synthetic spectrum (bottom). Details are as shown in Figure 12-36a [Vallance Jones and Gattinger, 1976].

are schematically represented in Figure 12-43, and the band origins and transition probabilities for CO, CO_2, N_2O, NO_2, H_2O, O_3, and NO^+ are presented in Table 12-7.

Not all the above atmospheric infrared radiators give rise to emissions that are aurorally enhanced. Modest two-fold enhancements of the 1.27 μm emission from the O_2 (0—0) infrared atmospheric band ($a^1\Delta_g - X^3\Sigma_g^-$) (Figure 12-31) have been reported in a very strong IBC III$^+$ aurora, and the enhancement has been explained through direct electron excitation of O_2 [D.J. Baker et al., 1978]. Due to the long lifetime (3900 s) of the $O_2(a^1\Delta_g)$ state, horizontal winds can transport the excited species downwind by several hundred kilometers before it decays [Rees and Roble, 1980], serving to complicate the analysis. The Wu-Benesch bands of N_2 ($W^3\Delta_u - B^3\Pi_g$) (Figure 12-30) are strong features of the EXCEDE artificial auroral spectra (Figure 12-38a) in the 2–4 μm region, but have not been observed in natural auroras, probably due to lack of sensor sensitivity. Recent observations [Stair et al., 1983] during the HIRIS experiment indicate enhancements of the 15 μm CO_2 ν_2 band emission as well as the 9.6 μm O_3 ν_3 band emission in a very strong (IBC III$^+$) aurora. However, no mechanism for these enhancements is known at the present time. Consequently, these observations await elaboration through further experiments and data analysis. The remainder of this section deals with the most firmly established and best understood infrared auroral enhancements, namely the 2.7 μm and 5.3 μm emission from NO and the 4.3 μm emission from CO_2.

Nitric oxide is an important minor constituent in the thermosphere. It is infrared-active, radiating near 5.3 μm in the fundamental $\Delta v = 1$ sequence and near 2.7 μm in the first overtone ($\Delta v = 2$) (Figure 12-41). The 5.3 μm fundamental band emission is an important cooling mechanism for the thermosphere [Chapter 13; Kockarts, 1980]. The $v = 1$ state is predominantly produced, in the nighttime quiescent atmosphere, by collision with atomic oxygen and by absorption of earthshine [Chapter 13],

$$NO\ (v = 0) + O \rightleftarrows NO\ (v = 1) + O \qquad (12.16)$$

$$NO\ (v = 0) + h\nu\ (\text{earthshine}) \rightarrow NO\ (v = 1). \qquad (12.17)$$

These processes primarily produce the NO ($v = 1$) level with insignificant population of levels $v \geq 2$.

Vibrationally excited NO is also produced by chemical reactions initiated by precipitating auroral electrons. Laboratory studies have shown that the major chemiluminescent auroral process that produces NO (Figure 12-29),

$$N(^2D) + O_2 \rightarrow NO(v) + O, \qquad (12.18)$$

populates NO levels up to $v = 12$ [Kennealy et al., 1978]. The $N(^2D)$ is produced directly from N_2 by precipitating

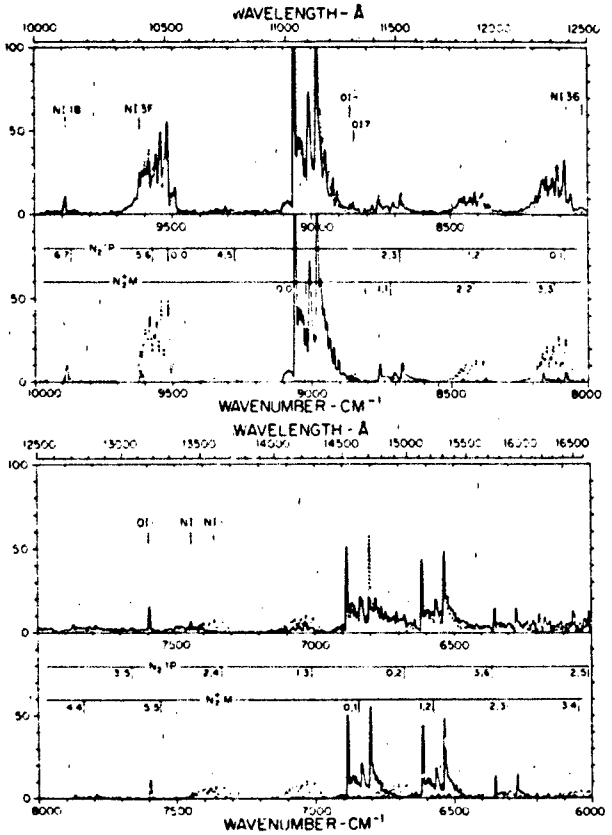


Figure 12-40. Infrared spectrum from 1.0 to 1.65 μm taken from the ground (top panels) along with corresponding synthetic spectra (bottom panels) [Gattinger and Vallance Jones, 1981].

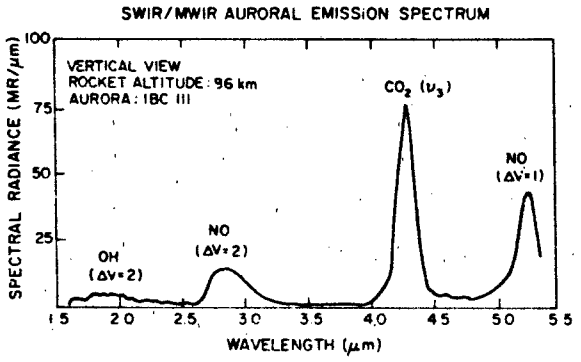


Figure 12-41. Zenith spectrum of natural aurora in the 1.5 to 5.4 μm [illegible] taken from 86 km showing principal emitters [illegible] 1975].

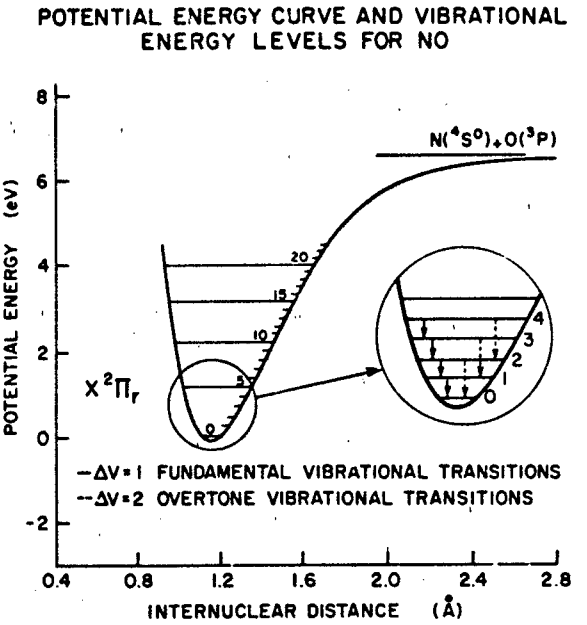


Figure 12-42. Potential energy curve for lowest electronic state of nitric oxide, showing vibrational energy levels as well as fundamental and first overtone vibrational transitions [adapted from Gilmore, 1965].

electrons and by dissociative recombination and charge-exchange reactions of aurorally produced ionic species,

$$
\begin{aligned}
N_2 + e &\rightarrow 2e + f_1 N(^2D) + (1 - f_1) N(^4S) + N^+ \\
N_2 + e &\rightarrow e + 2f_2 N(^2D) + 2(1 - f_2) N(^4S) \\
N_2^+ + e &\rightarrow 2f_3 N(^2D) + 2(1 - f_3) N(^4S) \\
NO^+ + e &\rightarrow 0.76\ N(^2D) + 0.24\ N(^4S) + O \\
N_2^+ + O &\rightarrow f_4 N(^2D) + (1 - f_4) N(^4S) + NO^+ \\
N^+ + O_2 &\rightarrow f_5 N(^2D) + (1 - f_5) N(^4S) + O_2^+ .
\end{aligned}
\qquad (12.19)
$$

The branching ratios f_i are of crucial importance, but they are uncertain for all reactions except NO^+ recombination (f_4 and f_5 are probably near unity). Nevertheless, the proportion of $N(^2D)$ that reacts with O_2 to produce NO(v) should vary with altitude as the reaction of Equation (12.18) competes with quenching of $N(^2D)$ by atomic oxygen. The variation of the $[O]/[O_2]$ ratio with altitude above 100 km then leads to an altitude-dependent efficiency for auroral NO(v) production. The altitude dependence of this production efficiency can then account for the variability of 5.3 μm auroral emission with auroral strength since auroras produced by more energetic electrons have lower peak ionization altitudes.

Table 12-6. Band origins and thermally-averaged Einstein coefficients for the fundamental and first overtone vibrational transitions of nitric oxide.

	Δv = 1			Δv = 2		
	$A_{v'v''}$[a]	BAND ORIGIN[b]		$A_{v'v''}$[a]	BAND ORIGIN[b]	
v'	(s^{-1})	(cm^{-1})	(μm)	(s^{-1})	(cm^{-1})	(μm)
1	10.8	1876	5.33	—	—	—
2	20.4	1848	5.41	0.460	3724	2.68
3	29.1	1820	5.50	1.51	3668	2.73
4	36.5	1792	5.58	3.10	3612	2.77
5	42.9	1764	5.67	4.90	3556	2.81
6	48.5	1736	5.76	7.28	3500	2.86
7	53.5	1708	5.86	9.63	3444	2.90
8	57.5	1680	5.95	12.5	3388	2.95
9	60.6	1652	6.05	15.7	3332	3.00
10	62.8	1624	6.16	19.1	3275	3.05
11	64.2	1596	6.27	22.9	3219	3.11
12	64.6	1567	6.38	27.2	3163	3.16
13	64.2	1539	6.50	31.9	3107	3.22
14	63.2	1511	6.62	36.9	3050	3.28
15	61.7	1483	6.74	41.9	2994	3.34
16	59.8	1455	6.87	46.4	2938	3.40
17	57.6	1427	7.01	50.5	2882	3.47
18	55.0	1399	7.15	54.3	2825	3.54
19	51.9	1370	7.30	58.1	2769	3.61
20	48.5	1342	7.45	61.8	2713	3.69

[a]Billingsley [1976]
[b]Calculated from vibrational constants of Goldman and Schmidt [1975].

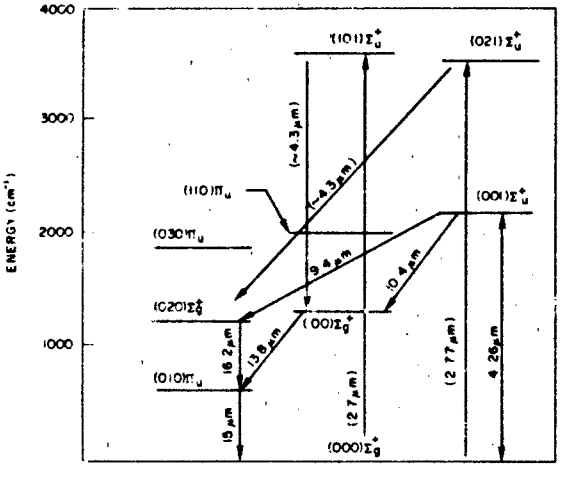


Figure 12-43. Lowest vibrational energy levels of CO_2 molecule showing some transitions of aeronomic interest.

The situation is somewhat complicated by the large variability of NO and by its transport. Since the rates of reaction of NO [Equations (12.16) and (12.17)] are proportional to the NO density, the nighttime quiescent component of the fundamental band should be stronger at high latitudes, especially in aurorally predosed regions where the NO density is expected to be large [Swider and Narcisi, 1977]. Nitric oxide is a long-lived species in the thermosphere, and model calculations have pointed out the importance of thermospheric winds in redistributing aurorally produced NO [Roble and Gary, 1979].

The NO 5.3 μm band radiance and the spectral shape of the emission depend on whether the NO(v) is produced by the quiescent mechanisms [Equations (12.16) and (12.17)] or the auroral mechanism [Equation (12.18)]. As alluded to previously, the quiescent processes can only excite the 1-0 transition, while the auroral process should yield in addition appreciable emission in the hot bands 2-1, 3-2, . . . , 12-11. Evidence of these components of the NO Δv = 1 emission in an aurora of only moderate intensity (IBC II) is clearly seen in the spectra resulting from the launch of a high-resolution (1.2 cm^{-1}) rocketborne field-widened interferometer (FWI) from Poker Flat, Alaska, in March 1983 [Steed et al., 1983; Murphy et al., 1983]. Preliminary analysis of the FWI flight data indicates that the hot band component increases with auroral energy deposition in a manner consistent with the mechanism of Equation (12.18).

In contrast to the 5.3 μm NO fundamental, the overtone band near 2.7 μm arises only when levels v ≥ 2 are excited. Thermal collisions and earthshine do not populate these levels significantly, and thus the 2.7 μm emission is negligible during quiet nighttime conditions. However, the Δv = 2 sequence is observed even for moderate IBC II auroras [Caledonia and Kennealy, 1982], the emission occurring with an efficiency of 0.4–1.0% [R.J. Huppi and Stair, 1979; Reidy et al., 1982].

The altitude profiles of zenith spectral radiance measured during several earlier auroral rocket experiments in the ICECAP series involving lower-resolution spectrometers than the FWI, are shown in Figure 12-44 [Stair et al., 1975]. The 5.3 μm radiance is seen to be enhanced for the IBC III⁺ aurora, but essentially at quiescent levels for the weaker IBC II and II⁺ auroras. On the other hand, the 2.7 μm radiance is enhanced for both the IBC II and III⁺ auroras recorded.

The failure to observe any enhancement of NO 5.3 μm radiance in moderate IBC II-II⁺ auroras during the ICECAP series of flights contrasts sharply with the enhancement of the Δv = 1 hot band component seen in the FWI experiment during a moderate aurora. The lack of enhancement in the ICECAP observations is due to the existence of two conditions: (1) The ICECAP detector's long-wavelength cutoff, shown at 5.4 μm in the data of Figure 12-41, passes all of the Q and R branches of the NO (1-0) band, but misses more than half of the hot-band emission. (2) Quiescent-atmosphere emission dominates auroral emission in the re-

Table 12-7. Data on infrared bands of atmospheric interest.[a]

Species	Rotational constant B (cm^{-1})	Transition	Band Origin		Band strength S ($cm^{-2}atm^{-1}$ at 273 K)	Einstein coefficient A (s^{-1})	Reference
			Wave number ν_0 (cm^{-1})	Wavelength λ_0 (μm)			
CO_2	0.390	$10^00–00^00$	1388	7.20	Inactive	Inactive	c
		$01^10–00^00$	667.4	15.0	205	1.28	c
		$00^01–00^00$	2349	4.26	2380	368	c
		$10^01–00^00$	3715	2.69	37.2	14.4	c
		$20^01–00^00$	4978	2.01	0.868	0.603	c
		$00^01–10^00$	961.0	10.4	0.0170[b]	0.411	c
H_2O	d	100–000	3657	2.73	12.0	4.52	e
		010–000	1595	6.27	257	18.30	f
		001–000	3756	2.66	172	67.90	e
		011–000	5331	1.88	19.9	15.90	e
N_2O	0.419	$00^01–00^00$	2224	4.50	1420	196	g
		$01^10–00^00$	588.8	17.0	24.4	0.119	e
		$10^00–00^00$	1285	7.78	218.0	10.1	e
O_3	d	100–000	1103	9.07	16.6	0.568	f
		010–000	700.9	14.3	15.6	0.214	f
		001–000	1042	9.60	345	10.5	f
NO_2	d	100–000	1320	7.58	—	—	h
		010–000	749.6	13.3	13.1	0.207	i
		001–000	1617	6.18	1520	111	i
		101–000	2906	3.44	64.0	15.2	i
CO	1.931	1–0	2143	4.67	243	31.3	e
		2–0	4260	2.35	1.86	1.12	e
NO^+	1.997	1–0	2344	4.27	70.8	10.9	j
		2–0	4656	2.15	1.15	0.697	j
		2–1	2312	4.32	135[b]	20.2	j
NO	1.705	1–0	1876	5.33	109	10.8	k
		2–0	3724	2.68	1.18	0.460	k
OH	18.871	1–0	3570	2.80	24.0	8.56	i
		2–0	6974	1.43	8.98	12.2	i

[a]Table adapted and updated from Kennealy and Del Greco [1972]
[b]Band strength is strongly temperature-dependent, since the lower level is not the ground state.
[c]Rothman and Young [1981].
[d]Molecule is an asymmetric rotor with three distinct rotational constants
[e]Rothman [1981].
[f]Rothman et al. [1983b].
[g]McClatchey et al. [1973].
[h]Herzberg [1945].
[i]Rothman et al. [1983a].
[j]Werner and Rosmus [1982].
[k]Billingsley [1976].

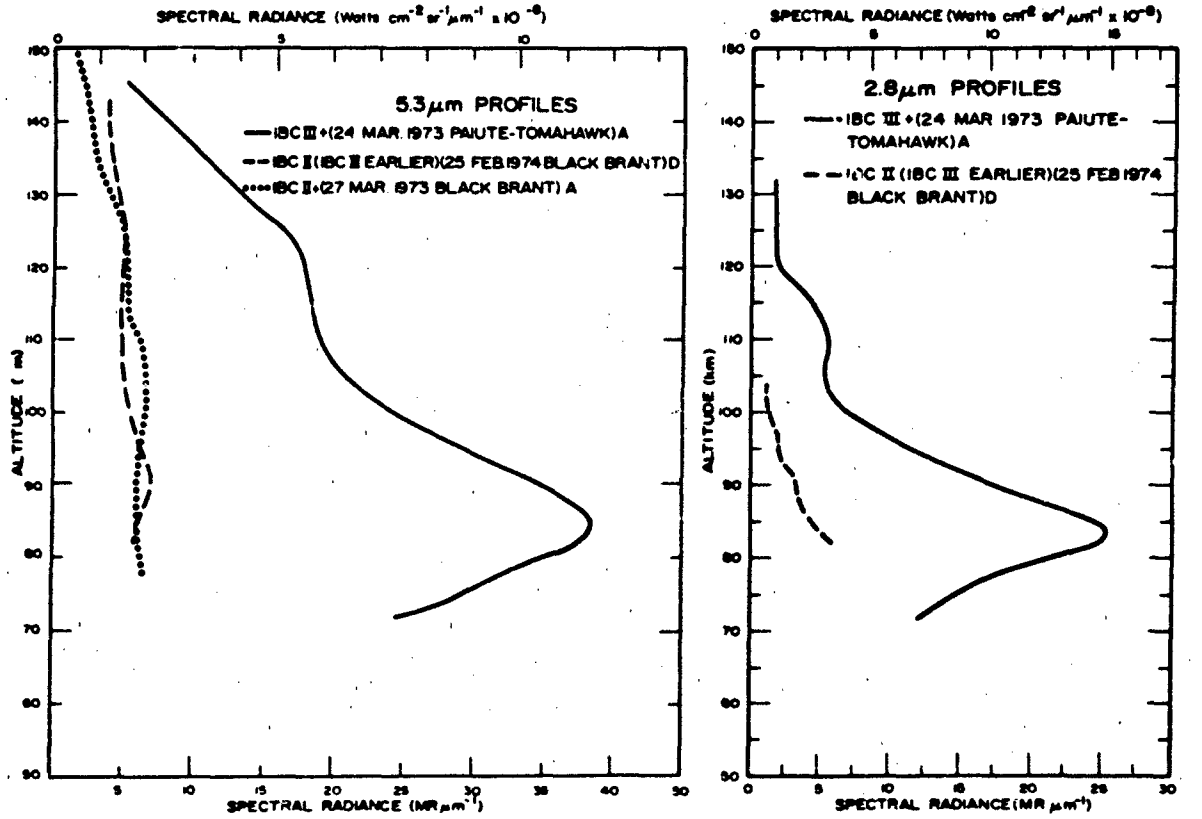


Figure 12-44. Altitude profiles of zenith peak spectral radiance observed in the NO fundamental 5.3 μm and overtone 2.8 μm bands for auroras of different brightness [Stair et al., 1975].

gion of the (1-0) band for moderate auroras and masks any auroral contribution. On the other hand, the long-wavelength cutoff for the recent FWI launch was sufficiently high to show clearly resolved lines of the NO hot bands on the low-frequency side of the NO (1-0) band, whose strength varied in altitude with the energy deposition [Steed et al., 1983; Murphy et al., 1983].

In summary, the intensity in the NO $\Delta v = 1$ system for $v' \geq 2$ responds to energy input from moderately strong auroras in the same way that the NO overtone does, in accord with model predictions. By contrast, for moderately strong auroras the NO (1-0) emission is dominated by airglow from the background atmosphere, as described by Equations (12.16) and (12.17).

As a final note on auroral NO(v), we mention that the HIRIS experiment previously referred to found the NO fundamental emission to be rather strongly enhanced for the very strong IBC III$^+$ aurora which it observed. This is consistent with models of NO emission [Caledonia and Kennealy, 1982], which indicate that chemiluminescent contributions to the NO $\Delta v = 1$ bands become comparable to emission from the background atmosphere under such auroral conditions.

We turn now to the 4.3 μm radiation from the ν_3 asymmetric-stretch mode of CO_2, which results from the emission process

$$CO_2(k\ell\ m) \rightleftarrows CO_2(k,\ell,m-1) + h\nu \quad (12.20)$$

and, in particular, from the fundamental band emission

$$CO_2(001) \rightleftarrows CO_2(000) + h\nu \quad (12.21)$$

at 4.26 μm. The altitude profile of this emission in the zenith is indicated in Figure 12-45, showing results from four ICECAP rocket launches [Stair et al., 1975], one launched under quiet conditions and three under various levels of electron precipitation in discrete arcs. The increase of spectral radiance between 75 and 90 km and the peak at 92 km are due to the fact that the ν_3 band becomes optically thick in the zenith at about 90 km. There is a general correlation between the peak radiance level and the electron flux, or the 5577 Å radiance, except that the IBC II$^+$ aurora is considerably less bright at 4.3 μm than the IBC II aurora. The explanation for this apparent anomaly is that a very

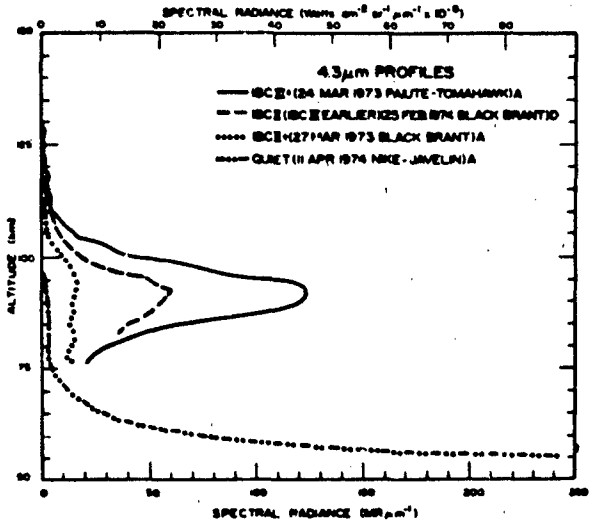


Figure 12-45. Altitude profiles of peak spectral radiance observed in the zenith in the CO_2 4.3 μm band for several classes of aurora [Stair et al., 1975].

bright IBC III arc had occupied the observation region in the latter case before the rocket arrived.

There are two possible sources of aurorally enhanced CO_2 excited vibrational states: (1) direct excitation of CO_2 by precipitating secondary electrons,

$$CO_2\,(000) + e \rightarrow CO_2\,(k\ell m) + e. \qquad (12.22)$$

and (2) vibrational pumping of N_2 by secondary electrons followed by vibration-to-vibration (VV) energy transfer from N_2 to CO_2 (Figure 12-29),

$$N_2 + e \rightarrow N_2^{\ddagger} + e \qquad (12.23a)$$

$$N_2^{\ddagger} + CO_2\,(k,\ell,m-1) \rightleftarrows N_2 + CO_2\,(k\ell m). \qquad (12.23b)$$

Due to the low CO_2 mixing ratio, the rate for the direct process [Equation (12.22)] is too small by several orders of magnitude to explain the observed radiance. Moreover, since the radiative lifetime of the 001 state is $A^{-1} = 2.3 \times 10^{-3}$ s (Table 12-7), the 4.3 μm emission due to the direct reaction of Equation (12.22) should be prompt, following the energy deposition closely in space and time. However, the actual altitude profiles and time histories for the 3914 Å N_2^+ first negative (0-0) radiance and for the 4.3 μm radiance are quite different with the CO_2 emission having a time delay ≳ 5 min [Kumer, 1975]. The delay arising from the N_2-CO_2 transfer is responsible for the apparently anomalous brightness at 4.3 μm of the IBC II auroral profile in Figure 12-45.

Another potential mechanism can also be ruled out by the delay as a major factor in accounting for enhanced 4.3 μm emission, namely $NO^+(\Delta v = 1)$ chemiluminescence (Table 12-7). The NO^+ 4.3 μm emission results from vibrationally excited NO^+ produced (Figure 12-29) in reactions such as

$$O_2^+ + NO \rightarrow O_2 + NO^+ \qquad (12.24)$$

and the fourth reaction in Equation (12.19). However, it cannot account for the observed time delay since the time constant for NO^+ formation is ≲10 s [Kumer, 1975].

On the other hand, the reaction sequence of Equation (12,23) has the slow step (b). Assuming a nitrogen vibrational temperature $T_v = 750$ K, one obtains a time constant of 200 s at an altitude of 115 km for the V-V transfer. The effective time constant for response of the ν_3 emission to electron precipitation is lengthened further by repeated transfer of excitation back and forth between N_2 and CO_2 before photon emission occurs and radiative trapping of the emitted photon. The net effect is an appreciable delay in the 4.3 μm emission. To model the process [Kumer, 1975, 1977], one must add quenching of both $N_2^{\ddagger}$ and the CO_2 excited states to the reaction set of Equations (12.20) and (12.23). The reverse reactions, indicated by backward arrows in the reactions of Equations (12.20) and (12.23b), account for repeated collisional transfer of the excitation between CO_2 and N_2 and for radiative transfer of the ν_3 photons. Kumer [1977] has shown that up to 11 combination bands and isotopic bands must be considered in Equation (12.20) and that an error of a factor of three can result if only the fundamental 001-000 band is used.

12.2.6 Summary

Our current understanding of infrared auroral processes suffers from a lack of sufficient high-resolution, high-sensitivity data on infrared emissions. This deficiency should be partially alleviated as analysis of data from the March 1983 rocketborne FWI experiment proceeds. In addition, several planned rocket-probe and shuttle-based experiments will carry aloft telescoped high-resolution interferometers to view the infrared aurora in earthlimb geometry. The combination of improved sensor sensitivity and the preferred earthlimb viewing aspect with greater auroral brightness will provide a comprehensive spectral survey of auroral infrared emissions. The observation of additional auroral emissions is anticipated as detection thresholds are improved and as infrared measurements are extended to higher spectral resolution and longer wavelengths. These experiments should contribute significantly to our understanding of the production and loss processes for infrared auroral emitters.

There is also a great need for data on the variability of the infrared aurora in order to identify seasonal, diurnal, and latitudinal trends, as well as dependences on solar and

geomagnetic activity. Such a data base may be obtained from a satellite platform or, in part, from ground-based observatories and series of aircraft or balloon flights. None of these experiments are without difficulty; the satellite experiments require solving the problem of maintaining detectors at cryogenic temperatures for extended periods, while ground-based and atmospheric platform experiments suffer from the masking and/or absorption of the auroral emission by the lower atmosphere and from the impossibility of using any limb enhancement technique. Nevertheless, the existence of the space shuttle and the success of unmanned infrared satellite experiments such as IRAS (InfraRed Astronomical Satellite) [Neugebauer et al., 1984] portend well for the future of infrared auroral studies.

ACKNOWLEDGEMENTS

The authors are indebted to A.J. Ratkowski, J.R. Winick, and R.D. Sharma for reading Section 12.2 and making useful recommendations.

REFERENCES

Akasofu, S.-I., "Energy Coupling Between the Solar Wind and the Magnetosphere," *Space Sci. Rev.*, **28:** 121, 1981.

Andriyenko, D.A., "Some Vertical Distribution Patterns," *Geomag. and Aeron.*, **5:** 838, 1965.

Baker, D., W. Pendleton, Jr., A. Steed, R. Huppi, and A.T. Stair, Jr., "Near-Infrared Spectrum of an Aurora," *J. Geophys. Res.*, **82:** 1601, 1977.

Baker, D.J., W.R. Pendleton, Jr., G.J. Romick, V.B. Wickwar, and M.J. Baron, "An Auroral Enhancement of O_2 λ1.27-μm Emission," *Planet. Space Sci.*, **26:** 619, 1978.

Baker, K.D., D.J. Baker, J.C. Ulwick, and A.T. Stair, Jr., "Measurements of 1.5- to 5.3-μm Infrared Enhancements Associated with a Bright Auroral Breakup," *J. Geophys. Res.*, **82:** 3518, 1977.

Banks, P.M., C.R. Chappell, and A.F. Nagy, "A New Model for the Interaction of Auroral Electrons with the Atmosphere: Spectral Degradation, Backscatter, Optical Emission, and Ionization," *J. Geophys. Res.*, **79:** 1459, 1974.

Bates, D.R., "Airglow and Auroras," in *Applied Atomic Collision Physics*, **1**, Academic Press, New York, 1982.

Berger, M.J., S.M. Seltzer, and K. Maeda, "Energy Deposition by Auroral Electrons in the Atmosphere," *J. Atm. Terr. Phys.*, **32:** 1015, 1970.

Berkey, F.T., V.M. Driatskiy, K. Henriksen, B. Hultqvist, D.H. Jelly, T.I. Shchuka, A. Theander, and J. Yliniemi, "A Synoptic Investigation of Particle Precipitation Dynamics for 60 Substorms in IQSY (1964–1965) and IASY (1969)," *Planet. Space Sci.*, **22:** 255, 1974.

Besprozvonnaya, A.S. and T.I. Shchuka, "Distribution of Anomalous Ionization in the High Latitude E Region According to Ground-Based Sounding," *Geomag. Aeron.*, **16:** 430, 1976.

Billingsley, F.P., II, "Calculated Vibration-Rotation Intensities for NO($X^2\Pi$)," *J. Mol. Spectrosc.*, **61:** 53, 1976.

Borst, W.L. and E.C. Zipf, "Lifetimes of Metastable CO and N_2 Molecules," *Phys. Rev. A*, **3:** 979, 1971.

Boyd, J.S., "Rocket-Borne Measurements of Auroral Electrons," *Rev. Geophys. Space Phys.*, **13:** 735, 1975.

Boyd, J.S., A.E. Belon, and G.C. Romick, "Latitude and Time Variations in Precipitated Electron Energy Inferred from Measurements of Auroral Heights," *J. Geophys. Res.*, **76:** 7694–7700, 1971.

Buchau, J., G.J. Gassmann, C.P. Pike, R.A. Wagner, and J.A. Whalen, "Precipitation Patterns in the Arctic Ionosphere Determined from Airborne Observations," *Ann. Geophys.*, **28:** 443, 1972.

Buchau, J., C.P. Pike and M. Wong, "Detailed Specification of the Arctic Ionosphere and an Application to Three-Dimensional Raytracing," AFCRL-TR-73-0726, AD775084, 1973.

Caledonia, G.E. and J.P. Kennealy, "NO Infrared Radiation in the Upper Atmosphere," *Planet. Space Sci.*, **30:** 1043, 1982.

Chamberlain, J.W., *Physics of the Aurora and Airglow*, Academic Press, New York, 1961.

Cohn, A. and G. Caledonia, "Spatial Distribution of the Fluorescent Radiation Emission Caused by an Electron Beam," *J. Appl. Phys.*, **41:** 3767, 1970.

Dalgarno, A., I.D. Latimer, and J.W. McConkey, "Corpuscular Bombardment and N_2^+ Radiation," *Planet. Space Sci.*, **13:** 1008, 1965.

Eather, R.H. and K.M. Burrows, "Excitation and Ionization by Auroral Protons," *Aust. J. Phys.*, **19:** 309, 1966.

Eyler, E.E. and F.M. Pipkin, "Lifetime Measurements of the B $^1\Pi_g$ State of N_2 Using Laser Excitation," *J. Chem. Phys.*, **79:** 3654, 1983.

Feldman, P.D. and E.P. Gentieu, "The Ultraviolet Spectrum of an Aurora 530-1520 Å," *J. Geophys. Res.*, **87:** 2453, 1982.

Feldstein, Y.I., "Pecularities in the Auroral Distribution and Magnetic Disturbance in High Latitudes Caused by the Asymmetrical Form of the Magnetosphere," *Planet. Space Sci.*, **14:** 121, 1966.

Feldstein, Y.I. and G.V. Starkov, "Dynamics of Auroral Belt and Polar Geomagnetic Disturbances," *Planet. Space Sci.*, **15:** 209, 1967.

Frank, L.A., J.D. Craven, J.L. Burch, and J.D. Winningham, "Polar Views of the Earth's Aurora with Dynamics Explorer," *Geophys. Res. Lett.*, **9:** 1001, 1982.

Freund, R.S., "Molecular-Beam Measurements of the Emission Spectrum and Radiative Lifetime of N_2 in the Metastable E $^3\Sigma_g^+$ State," *J. Chem. Phys.*, **50:** 3734, 1969.

Freund, R.S., "Radiative Lifetime of N_2 (a $^1\Pi_g$) and the Formation of Metastable N_2 (a' $^1\Sigma_u^-$)," *J. Chem. Phys.*, **56:** 4344, 1972.

Garriott, O.K., "Visual Observations from Space," *J. Opt. Soc. Am.*, **69:** 1064, 1979.

Gassmann, G.J., "Analog Model of the Arctic Ionosphere," AFCRL-TR-73-0151, AD762280, 1973.

Gassman, G.J., J. Buchau, R.A. Wagner, C.P. Pike, and M.G. Hurwitz, "Arctic Ionospheric Modelling—Five Related Papers," AFCRL-72-0305, AD775084, 1972.

Gattinger, R.L. and A. Vallance Jones, "Quantitative Spectroscopy of the Aurora. II. The Spectrum of Medium Intensity Aurora Between 4500 and 8900 Å," *Can. J. Phys.*, **52:** 2343, 1974.

Gattinger, R.L. and A. Vallance Jones, "Quantitative Spectroscopy of the Aurora. V. The Spectrum of Strong Aurora between 10,000 and 16,000 Å," *Can. J. Phys.*, **59:** 480, 1981.

Gilmore, F.R., "Potential Energy Curves for N_2, NO, O_2 and Corresponding Ions," *J. Quant. Spectrosc. Radiat. Transfer*, **5:** 369, 1965.

Goldman, A. and S.C. Schmidt, "Infrared Spectral Line Parameters and the Absorptance Calculations of NO at Atmospheric and Elevated Temperatures for the $\Delta v = 1$ Bands Region," *J. Quant. Spectrosc. Radiat. Transfer*, **15:** 127, 1975.

Goody, R.M., *Atmospheric Radiation*, **1**, Clarendon Press, Oxford, U.K., 1964.

Gordiets, B.F., M.N. Markov, and L.A. Shelepin, "I.R. Radiation of the Upper Atmosphere," *Planet. Space Sci.*, **26:** 933, 1978.

Gussenhoven, M.S., D.A. Hardy, and N. Heinemann,

"Systematics of the Equatorward Diffuse Auroral Boundary," *J. Geophys. Res.*, **88**: 5692, 1983.

Hanson, W.B. (ed.), "Special Issue on Atmospheric Explorer Mission," *Radio Sci.*, **8**(No. 4): 1973.

Harang, L., *The Aurora*, Wiley, New York, 1951.

Hartz, T.R. and N.M. Brice, "The General Pattern of Auroral Particle Precipitation," *Planet. Space Sci.*, **15**: 301, 1967.

Heelis, R.A., J.D. Winningham, W.B. Hanson, and J.L. Burch, "The Relationships Between High-Latitude Convection Reversals and the Energetic Particle Morphology Observed by Atmosphere Explorer," *J. Geophys. Res.*, **85**: 3315, 1980.

Herzberg, G., *Molecular Spectra and Molecular Structure*, **2**, Van Nostrand, Princeton, N.J., 1945.

Herzberg, G., *Molecular Spectra and Molecular Structure*, **1**, second edition, Van Nostrand, New York, 1950.

Herzberg, G., *Molecular Spectra and Molecular Structure*, **3**, Van Nostrand, Princeton, N.J., 1967.

Hilliard, R.L. and G.G. Shepherd, "Upper Atmospheric Temperatures from Doppler Line Widths—IV," *Planet. Space Sci.*, **14**: 383, 1966.

Holzworth, R.H. and C.-I Meng, "Mathematical Representation of the Auroral Oval," *Geophys. Res. Lett.*, **2**: 377, 1975.

Hook, J.L., "Morphology of Auroral Zone Radiowave Absorption in the Alaska Sector," *J. Atmos. Terr. Phys.*, **30**: 1341, 1968.

Huber, K.P. and G. Herzberg, *Molecular Spectra and Molecular Structure*, **4**, Van Nostrand Reinhold, New York, 1979.

Huffman, R.E., F.J. LeBlanc, J.C. Larrabee, and D.E. Paulsen, "Satellite Vacuum Ultraviolet Airglow and Auroral Observations," *J. Geophys. Res.*, **85**: 2201, 1980.

Hunten, D.M., "Some Photometric Observations of Auroral Spectra," *J. Atmos. Terr. Phys.*, **7**: 141, 1955.

Hunten, D.M., E.G. Rawson, and J.K. Walker, "Rapid Measurement of N_2^+ Rotational Temperatures in Aurora," *Can. J. Phys.*, **41**: 258, 1963.

Hunten, D.M., F.E. Roach, and J.W. Chamberlain, "A Photometric Unit for the Airglow and Aurora," *J. Atmos. Terr. Phys.*, **8**: 345, 1956.

Huppi, R.J., and A.T. Stair, Jr., "Aurorally Enhanced Infrared Emissions," *Appl. Opt.*, **18**: 3394, 1979.

Huppi, E.R., J.W. Rogers, and A.T. Stair, Jr., "Aircraft Observations of the Infrared Emission of the Atmosphere in the 700–2800 cm^{-1} Region," *Appl. Opt.*, **13**: 1466, 1974.

International Union of Geodesy and Geophysics, *International Auroral Atlas*, Edinburgh University Press, 1963.

Jasperse, J.R. and B. Basu, "Transport Theoretic Solutions for Auroral Proton and H Atom Fluxes and Related Quantities," *J. Geophys. Res.*, **87**: 811, 1982.

Jones, R.A. and M.H. Rees, "Time Dependent Studies of the Aurora. I. Ion Density and Composition," *Planet. Space Sci.*, **21**: 537, 1973.

Kennealy, J.P. and F.P. Del Greco, "The Kinetics of Atmospheric Radiative Processes in the Infrared," Chapter 11, *Defense Nuclear Agency Reaction Rate Handbook*, second edition, edited by M.H. Bortner and T. Baurer, DNA 1948H, 1972.

Kennealy, J.P., F.P. DelGreco, G.E. Caledonia, and B.D. Green, "Nitric Oxide Chemiexcitation Occurring in the Reaction Between Metastable Nitrogen Atoms and Oxygen Molecules," *J. Chem. Phys.*, **69**: 1574, 1978.

Kockarts, G., "Nitric Oxide Cooling in the Terrestrial Atmosphere," *Geophys. Res. Lett.*, **7**: 137, 1980.

Krukonis, A.P. and J.A. Whalen, "Occurrence and Lifetimes of Discrete Auroras Near Midnight," *J. Geophys. Res.*, **85**: 119, 1980.

Krupenie, P.H., "The Spectrum of Molecular Oxygen," *J. Phys. Chem. Ref. Data*, **1**: 423, 1972.

Kumer, J.B., "Summary Analysis of 4.3μm Data," in *Atmospheres of Earth and the Planets*, edited by B.M. McCormac, D. Reidel, Dordrecht, Holland, 347–358, 1975.

Kumer, J.B., "Theory of the CO_2 4.3-μm Aurora and Related Phenomena," *J. Geophys. Res.*, **82**: 2203, 1977.

Loftus, A. and P.H. Krupenie, "The Spectrum of Molecular Nitrogen," *J. Phys. Chem. Ref. Data*, **6**: 113, 1977.

Lui, A.T.Y. and C.D. Anger, "A Uniform Belt of Diffuse Aurora and Airglow from ISIS 2 Spacecraft," *Planet. Space Sci.*, **21**: 819, 1973.

Lui, A.T.Y., P. Perreault, S.-I. Akasofu, and C.D. Anger, "The Diffuse Aurora," *Planet. Space Sci.*, **21**: 857, 1973.

Lui, A.T.Y., D. Venkatesan, C.D. Anger, S.-I. Akasofu, W.J. Heikkila, J.D. Winningham, and J.R. Burrows, "Simultaneous Observations of Particle Precipitations and Auroral Emissions by the ISIS 2 Satellite in the 19–24 MLT Sector," *J. Geophys. Res.*, **82**: 2210, 1977.

McClatchey, R.A., W.S. Benedict, S.A. Clough, D.E. Burch, R.F. Calfee, K. Fox, L.S. Rothman, and J.S. Garing, "AFCRL Atmospheric Absorption Line Parameters Compilation," AFCRL TR-73-0096, AD762904, 1973.

Mende, S.B. and R.H. Eather, "Monochromatic All Sky Observations and Auroral Precipitation Patterns," *J. Geophys. Res.*, **81**: 3771, 1976.

Meng, C.-I., "Diurnal Variation of the Auroral Oval Size," *J. Geophys. Res.*, **84**: 5319, 1979.

Meng, C.-I., R.H. Holzworth, and S.-I. Akasofu, "Auroral Circle—Delineating the Poleward Boundary of the Quiet Auroral Belt," *J. Geophys. Res.*, **82**: 164, 1977.

Muldrew, D.B., "F-Layer Ionization Troughs Deduced from Alouette Data," *J. Geophys. Res.*, **70**: 2635, 1965.

Murphy, R., R. Sharma, R. Picard, J. Winick, K.D. Baker, and J.C. Ulwick, "Preliminary Interpretation of Infrared Interferometer Rocket Measurements," *Eos*, **64**: 785, 1983.

Neugebauer, G., et al., "The Infrared Astronomical Satellite (IRAS) Mission," *Astrophys. J.*, **278**: L1, 1984.

Omholt, A., "The Auroral E-Layer Ionization and the Auroral Luminosity," *J. Atmos. Terr. Phys.*, **7**: 73, 1955.

O'Neil, R.R., F. Bien, D. Burt, J.A. Sandock, and A.T. Stair, Jr., "Summarized Results of the Artificial Auroral Experiment, Precede," *J. Geophys. Res.* **83**: 3273, 1978a.

O'Neil, R.R., O. Shepherd, W.P. Reidy, J.W. Carpenter,

T.N. Davis, D. Newell, J.C. Ulwick, and A.T. Stair, Jr., "Excede 2 Test, an Artificial Auroral Experiment: Ground-Based Optical Measurements," *J. Geophys. Res.*, **83:** 3281, 1978b.

O'Neil, R.R., E.T.P. Lee, and E.R. Huppi, "Auroral O(^{1}S) Production and Loss Processes: Ground-Based Measurements of the Artificial Auroral Experiment Precede," *J. Geophys. Res.*, **84:** 823, 1979.

O'Neil, R.R., A.T. Stair, Jr., W.R. Pendleton, Jr., and D.A. Burt, "The EXCEDE SPECTRAL Artificial Auroral Experiment: An Overview," in *Artificial Particle Beams in Space Plasma Studies*, edited by B. Grandal, Plenum, New York, pp. 207–215, 1982.

Peterson, L.R., T. Sawada, J.N. Bass, and A.E.S. Green, "Electron Energy Deposition in a Gaseous Mixture," *Computer Phys. Commun.*, **5:** 239, 1973.

Pike, C.P., "An Analytical Model of the Main F-Layer Trough," AFGL-TR-76-0098, ADA026031, 1976.

Pike, C.P. and I.A. Whalen, "Satellite Observations of Auroral Substorms," *J. Geophys. Res.*, **79:** 985, 1974.

Piper, L.G., "The Excitation of O(^{1}S) in the Reaction Between $N_2(A\ ^3\Sigma_u^+)$ and O(^{3}P)," *J. Chem. Phys.*, **77:** 2373, 1982.

Piper, L.G., G.E. Caledonia, and J.P. Kennealy, "Rate Constants for Deactivation of $N_2(A)$ v′ = 0,1 by O_2," *J. Chem. Phys.*, **74:** 2888, 1981a.

Piper, L.G., G.E. Caledonia, and J.P. Kennealy, "Rate Constants for Deactivation of $N_2(A\ ^3\Sigma_u^+, v' = 0,1)$ by O," *J. Chem. Phys.*, **75:** 2847, 1981b.

Rawlins, W.T., G.E. Caledonia, J.J. Gibson, and A.T. Stair, Jr., "Infrared Emission from NO ($\Delta v = 1$) in an Aurora: Spectral Analysis and Kinetic Interpretation of HIRIS Measurements," *J. Geophys. Res.*, **86:** 1313, 1981a.

Rawlins, W.T., G.E. Caledonia, and J.P. Kennealy, "Observation of Spectrally Resolved Infrared Chemiluminescence from Vibrationally Excited $O_3(\nu_3)$," *J. Geophys. Res.*, **86:** 5247, 1981b.

Rearwin, S. and E.W. Hones, "Near Simultaneous Measurement of Low-Energy Electrons by Sounding Rocket and Satellite," *J. Geophys. Res.*, **79:** 4322, 1974.

Rees, M.H., "Auroral Ionization and Excitation by Incident Energetic Electrons," *Planet. Space Sci.*, **11:** 1209, 1963.

Rees, M.H., "Note on the Penetration of Energetic Electrons into the Earth's Atmosphere," *Planet. Space Sci.*, **12:** 722, 1964.

Rees, M.H. and D. Luckey, "Auroral Electron Energy Derived from Ratio of Spectroscopic Emissions I. Model Computations," *J. Geophys. Res.*, **79:** 5181, 1974.

Rees, M.H. and R.G. Roble, "Effect of a Horizontal Wind on 1.27-μm Auroral Emission from O_2 ($^1\Delta_g$) Molecules," *J. Geophys. Res.*, **85:** 4295, 1980.

Rees, M.H., A.I. Stewart, and J.C.G. Walker, "Secondary Electrons in Aurora," *Planet. Space Sci.*, **17:** 1997, 1969.

Reidy, W.P., T.C. Degges, A.G. Hurd, A.T. Stair, Jr., and J.C. Ulwick, "Auroral Nitric Oxide Concentration and Infrared Emission," *J. Geophys. Res.*, **87:** 3591, 1982.

Robinson, R.M., R.R. Vondrak, and T.A. Potemra, "Electrodynamic Properties of the Evening Sector Ionosphere Within the Region 2 Field-Aligned Current Sheet," *J. Geophys. Res.*, **87:** 731, 1982.

Roble, R.G. and J.M. Gary, "The Effect of Horizontal Transport on Auroral NO Densities," *Geophys. Res. Lett.*, **6:** 703, 1979.

Roble, R.G. and M.H. Rees, "Time-Dependent Studies of the Aurora: Effects of Particle Precipitation on the Dynamic Morphology of Ionospheric and Atmospheric Properties," *Planet. Space Sci.*, **25:** 991, 1977.

Rothman, L.S., "AFGL Atmospheric Absorption Line Parameters Compilation: 1980 Version," *Appl. Opt.*, **20:** 791, 1981.

Rothman, L.S. and L.D.G. Young, "Infrared Energy Levels and Intensities of Carbon Dioxide—II," *J. Quant. Spectrosc. Radiat. Transfer*, **25:** 505, 1981.

Rothman, L.S., A. Goldman, J.R. Gillis, R.R. Gamache, H.M. Pickett, R.L. Poynter, N. Husson, and A. Chedin, "AFGL Trace Gas Compilation: 1982 Version," *Appl. Opt.*, **22:** 1616, 1983a.

Rothman, L.S., R.R. Gamache, A. Barbe, A. Goldman, J.R. Gillis, L.R. Brown, R.A. Toth, J.-M. Flaud, and C. Camy-Peyret, "AFGL Atmospheric Absorption Line Parameters Compilation: 1982 Edition," *Appl. Opt.*, **22:** 2247, 1983b.

Sandford, B.P., "Variations of Auroral Emissions with Time, Magnetic Activity and the Solar Cycle," *J. Atmos. Terr. Phys.*, **30:** 1921, 1968.

Seaton, M.J., "Excitation Processes in the Aurora and Airglow. I. Absolute Intensities, Relative Ultraviolet Intensities, and Electron Densities in High Latitude Aurorae," *J. Atmos. Terr. Phys.*, **4:** 285, 1954.

Sharber, J.R., "The Continuous (Diffuse) Aurora and Auroral-E Ionization," in *Physics of Space Plasmas*, edited by T.S. Chang, B. Coppi, and J.R. Jasperse, *SPI Conference Proceedings and Reprint Series, 4*, Scientific Publishers, Cambridge, Mass., 1981.

Sharp, W.E., "The Ultraviolet Aurora: The Spectrum Between 2100 Å and 2300 Å," *Geophys. Res., Lett.*, **5:** 703, 1978.

Sharp, W.E. and M.H. Rees, "Auroral Spectrum between 1200 and 4000 Angstroms," *J. Geophys. Res.*, **77:** 1810, 1972.

Sheehan, R.E. and R.L. Carovillano, "Characteristics of the Equatorward Auroral Boundary Near Midnight Determined from DMSP Images," *J. Geophys. Res.*, **83:** 4749, 1978.

Sheehan, R.E., R.L. Carovillano, and M.S. Gussenhoven, "Occurrence and Recurrence in Auroral Activity in DMSP Images," *J. Geophys. Res.*, **87:** 3581, 1982.

Shepherd, M.M. and R.H. Eather, "On the Determination of Auroral Electron Energies and Fluxes from Optical Spectral Measurements," *J. Geophys. Res.*, **81:** 1407, 1976.

Slater, D.W., L.L. Smith, and E.W. Kleckner, "Correlated Observations of the Equatorward Diffuse Auroral Boundary," *J. Geophys. Res.*, **85:** 531, 1980.

Spiro, R.W., P.F. Reiff, and L.J. Maher, "Precipitating Electron Energy Flux and Auroral Zone Conductances—An Empirical Model," *J. Geophys. Res.*, **87:** 8215, 1982.

Stair, A.T., Jr., J.C. Ulwick, K.D. Baker, and D.J. Baker, "Rocketborne Observations of Atmospheric Infrared Emissions in the Auroral Region," in *Atmospheres of Earth and the Planets*, edited by B.M. McCormac, D. Reidel, Dordrecht, Holland, pp. 335–346, 1975.

Stair, A.T., Jr., R. Nadile, J.C. Ulwick, K.D. Baker, and D.J. Baker, "Infrared Measurements of Aurora, Airglow and the Upper Atmosphere," in *Proceedings of the 8th Annual Meeting on Upper Atmospheric Studies by Optical Methods*, edited by E. O'Mongain, University College, Dublin, pp. 27–35, 1980.

Stair, A.T., Jr., J. Pritchard, I. Coleman, C. Bohne, W. Williamson, J. Rogers, and W.T. Rawlins, "Rocketborne Cryogenic (10 K) High-Resolution Interferometer Spectrometer Flight HIRIS: Auroral and Atmospheric IR Emission Spectra," *Appl. Opt.*, **22:** 1056, 1983.

Starkov, G.V., "Auroral Heights in the Polar Cap," *Geomag. Aeron.*, **8:** 28, 1968.

Starkov, G.V., "Analytical Representation of the Equatorial Boundary of the Oval Auroral Zone," *Geomag. Aeron.*, **9:** 614, 1969.

Starkov, G.V. and Y.I. Feldstein, "Position of the Auroral Belt at Night During the IGY and IQSY" in *Complex Investigations of the Polar Ionosphere*, edited by Isaev, Nauka Press, Leningrad, 1970.

Steed, A., J.C. Ulwick, C. Harris, F. Cook, and R. Straka, "Rocketborne Interferometer Measurements of SWIR/MWIR Spectra," *Eos*, **64:** 784, 1983.

Stolarski, R.S. and A.E.S. Green, "Calculations of Auroral Intensities from Electron Impact," *J. Geophys. Res.*, **72:** 3967, 1967.

Strickland, D.J., D.L. Book, T.P. Coffey, and J.A. Fedder, "Transport Equation Techniques for the Deposition of Auroral Electrons," *J. Geophys. Res.*, **81:** 2755, 1976.

Strickland, D.J., J.R. Jasperse, and J.A. Whalen, "Dependence of Auroral FUV Emissions on the Incident Electron Spectrum and Neutral Atmosphere," *J. Geophys Res.*, **88:** 8051, 1983.

Swider, W. and R.S. Narcisi, "Auroral E-Region: Ion Composition and Nitric Oxide," *Planet. Space Sci.*, **25:** 103, 1977.

Torr, M.R. and D.G. Torr, "The Role of Metastable Species in the Thermosphere," *Rev. Geophys. Space Phys.*, **20:** 91, 1982.

Vallance Jones, A., *Aurora*, D. Reidel, Dordrecht, Holland, 1974.

Vallance Jones, A. and R.L. Gattinger, "Quantitative Spectroscopy of the Aurora. III. The Spectrum of Medium Intensity Aurora Between 3100 Å and 4700 Å," *Can. J. Phys.*, **53:** 1806, 1975.

Vallance Jones, A. and R.L. Gattinger, "Quantitative Spectroscopy of the Aurora. IV. The Spectrum of Medium Intensity Aurora between 8800 Å and 11 400 Å," *Can. J. Phys.*, **54:** 2128, 1976.

Wadzinski, H.T. and J.R. Jasperse, "Low Energy Electron and Photon Cross Sections for O, N_2, and O_2, and Related Data," AFGL-TR-82-0008, ADA118921, 1982.

Wagner, R.A. and C.P. Pike, "A Discussion of Arctic Ionograms" in *Radar Propagation in the Arctic*, Chapter 4, edited by I. Frihagen, Technical Editing and Reproduction, London, 1972.

Wagner, R.A., A.L. Snyder, and S.-I. Akasofu, "The Structure of the Polar Ionosphere During Exceptionally Quiet Periods," *Planet. Space Sci.*, **21:** 1911, 1973.

Werner, H.-J. and P. Rosmus, "Ab Initio Calculations of Radiative Transition Probabilities in the $X\ ^1\Sigma^+$ Ground State of the NO^+ Ion," *J. Mol. Spectrosc.*, **96:** 362, 1982.

Whalen, J.A., "Auroral Oval Plotter and Nomograph for Determining Corrected Geomagnetic Local Time, Latitude and Longitude for High Latitudes in Northern Hemisphere," AFCRL 70-0422, AD713170, 1970.

Whalen, J.A. "Characteristics of the High Latitude Ionosphere Produced by Auroral Particle Precipitation," in AGARD Conference Proceedings, No. 263, *Special Topics in HF Propagation*, 1979.

Whalen, J.A. "General Characteristics of the Auroral Ionosphere" in *Physics of Space Plasmas*, edited by T.S. Chang, B. Coppi, and J.R. Jasperse, SPI Conference Proceedings and Reprint Series 4, Scientific Publishers, Cambridge, Mass., 1981.

Whalen, J.A., "A Quantitative Description of the Spatial Distribution and Dynamics of the Energy Flux in the Continuous Aurora," *J. Geophys. Res.*, **88:** 7155, 1983.

Whalen, J.A. and J.R. Sharber, "TheNighttime Auroral-E Layer: Partical Production, Latitudinal and Longitudinal Structure and Dynamics" in *The Effects of the Ionosphere on Radiowave Systems*, edited by J. Goodman and F. Clarke, U.S. Government Printing Office, Washington, D.C., 1981.

Whalen, J.A., J. Buchau, and R.A. Wagner, "Airborne Ionospheric and Optical Measurements of Noontime Aurora," *J. Atmos. Terr. Phys.*, **33:** 661, 1971.

Whalen, J.A., R.A. Wagner, and J. Buchau, "A 12-Hour Case Study of Auroral Phenomena in the Midnight Sector: Oval, Polar Cap and Continuous Auroras," *J. Geophys. Res.*, **82:** 3529, 1977.

Winckler, J.R., "The Use of Artificial Electron Beams as Probes of the Distant Magnetosphere," in *Artificial Particle Beams in Space Plasma Studies*, edited by B. Grandal, Plenum, New York, pp. 3–33, 1982.

Winningham, J.D., F. Yasuhara, S.-I. Akasofu, and W.J. Heikkila, "The Latitudinal Morphology of 10 eV to 10 keV Electron Fluxes during Magnetically Quiet and Disturbed Times in the 2100 to 0300 MLT Sector," *J. Geophys. Res.*, **80:** 3148, 1975.

Chapter 13

INFRARED AIRGLOW

R.D. Sharma

Infrared radiance observed by a detector looking at the earth's undisturbed atmosphere is called infrared airglow. This radiation may arise from rotational transitions, as is the case for radiation from water vapor around 17 μm, or from electronic transitions, for example, the O_2 ($a^1\Delta_g \rightarrow X^3\Sigma_g^-$) transition causing the airglow at 1.27 μm (0-0 band) and at 1.58 μm (0-1 band). However, the most common sources of infrared airglow are the vibrational-rotational (V-R) transitions within the ground electronic state. This class of transitions includes emission from OH, CO_2, O_3, NO, and H_2O. Figure 13-1 gives the energy level diagram of CO_2 and points out the more commonly encountered transitions. Table 13-1 gives the band origins and Table 13-2 the Einstein A coefficients for some OH V-R bands. For additional information on the energy levels and the radiative lifetimes of the infrared active species see Chapter 12. The extent of the wavelength range covered by the term infrared is not well defined. We will limit this discussion to the spectral region from 1.2 to 17 μm. This rather arbitrarily restricted wavelength range leaves out several important airglow emissions, including the 63 μm emission from oxygen atoms [Offermann and Grossmann, 1978]. This latter process together with the 5.3 μm ($1 \rightarrow 0$) emission from NO and the 15 μm radiation from the bending mode of CO_2 are the important processes in the cooling of the thermosphere [Craig and Gille, 1969; Kockarts, 1980; Gordiets et al., 1983; Zachor, Sharma, Nadile and Stair, 1981].

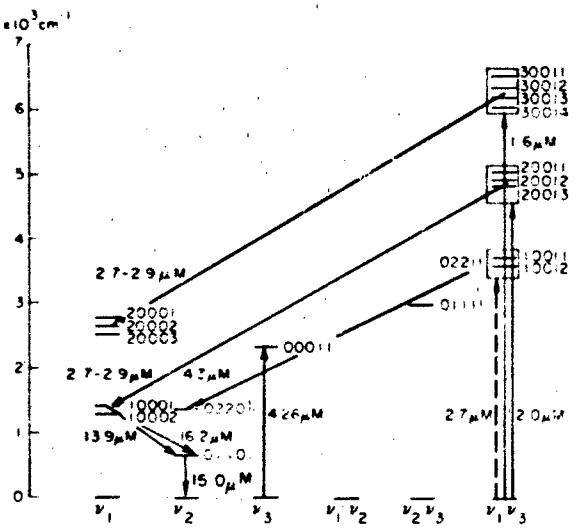


Figure 13-1. The energy levels of carbon dioxide. The lowest levels of bending mode and asymmetric stretch, which lead to transitions at 15.0 and 4.26 μm, are shown. Also shown are the upper and lower levels for some of the important transitions observed in the infrared airglow of the earth's atmosphere. The notation in the figure is the same as in AFGL line tape [Rothman and Young, 1981]. (Reprinted with permission from Pergamon Press Ltd. © 1981.)

Table 13-1. OH Vibrational Transition Band Origins [L.S. Rothman et al., 1983]

	Δ v = 1		Δ v = 2		Δ v = 3	
v'	(cm^{-1})	(μm)	(cm^{-1})	(μm)	(cm^{-1})	(μm)
1	3570	2.80				
2	3404	2.94	6974	1.43		
3	3240	3.09	6644	1.50	10214	0.98
4	3078	3.25	6318	1.58	9722	1.03
5	2915	3.43	5993	1.67	9233	1.08
6	2752	3.63	5667	1.76	8745	1.14
7	2585	3.87	5337	1.87	8252	1.21
8	2415	4.14	5000	2.00	7752	1.29
9	2237	4.47	4652	2.15	7237	1.38

Infrared airglow has been the subject of a number of studies [Stair et al., 1981; Gordiets et al., 1978; Ogawa, 1976]. It is a quickly changing field because of the rapidly improving measurement and analysis technology and it is not surprising, therefore, that the information in the literature is a few years behind that available to workers in the field. For this reason, the emphasis in this chapter has been placed upon results obtained at AFGL.

13.1 SOURCES OF THE INFRARED AIRGLOW

13.1.1 Resonant Scattering of the Earthshine

Radiation upwelling through the atmosphere from the surface of the earth is called earthshine. Although the average temperature of the earth's surface is 296 K, the tem-

Table 13-2. Thermally-averaged OH ($X^2\pi_i$) vibrational Einstein coefficients $A_{v'v''}$ (s^{-1}) for 200 K [Mies, 1974] (Reprinted with permission from Academic Press © 1974)

V"=	0	1	2	3	4	5	6	7	8	9
V' = 1	20.15									
2	14.07	25.24								
3	0.92	39.93	20.90							
4	0.079	4.29	72.61	12.25						
5	0.050	0.392	10.62	108.4	4.47					
6		0.053	1.274	20.99	141.5	2.347				
7			0.183	2.907	37.32	163.4	9.142			
8			0.030	0.569	5.67	60.84	166.8	25.84		
9				0.130	1.189	10.81	90.28	146.6	50.65	

perature of the earthshine depends upon altitude and wavelength. For example in the 12-13 μm region, where the absorption due to the earth's atmosphere is minimal, the upwelling radiation can be described by 296 K blackbody radiation independent of altitude. On the other hand, radiation in the 15 μm region up to the lower mesosphere (about 70 km altitude) can be described as radiation from a blackbody with temperature given by the local translational temperature. At higher altitudes, because of the lack of sufficient collisions, the upwelling 15 μm radiation can be described by about 200 K blackbody temperature.

Whatever the variations in the temperature of the earthshine, it is clear that this radiation has appreciable intensity only at longer wavelengths ($\lambda > 5$ μm). It can, therefore, excite only the lower vibrational levels of the infrared active molecules. The excited molecules can reradiate isotropically without change of frequency causing infrared airglow. Resonant scattering of earthshine by the ν_3 mode of ozone is an important contributor to the airglow at 9.6 μm. The 15 μm radiation from the bending mode of carbon dioxide also contains an important contribution from the earthshine.

13.1.2 Scattering of Sunshine

In contrast to the earthshine which populates only the lowest lying vibrational levels, solar radiation (6000 K blackbody temperature) plays an important role in exciting higher vibrational levels corresponding to $\lambda \leq 5$ μm. The vibrational levels thus produced can lose their energy by optical transitions to the ground state (resonant scattering) or by radiation to other lower excited vibrational levels (nonresonant scattering or fluorescence). The effect of sunshine is shown in the dramatic increase in the airglow near 4.3 μm during the day due to scattering by the ν_3 mode of CO_2 [James and Kumer, 1973]. Scattered solar radiation, resonant and nonresonant, also contributes to the CO_2 airglow at 2.7 μm, the important absorption bands being located near 2.7 μm [Kumer, 1981], 2.0, and 1.6 μm [Sharma et al., 1981].

A process capable of producing large amounts of infrared airglow radiation is the electronic excitation of molecules by sunshine followed by emission of a longer wavelength photon. The molecule is thus left in an excited vibrational state, the final vibrational level being determined by the Franck-Condon integrals. Several quanta of infrared photons may be emitted before the molecule relaxes to the ground state. This process, however, has not been identified as the source of any airglow from the quiescent atmosphere.

13.1.3 Photodissociation

Dissociation of molecular species by solar radiation contributes to the airglow by producing fragments in excited states. An example is the airglow of O_2 ($a^1\Delta_g \rightarrow X^3\Sigma_g^-$) at 1.27 μm (0-0 transition) and at 1.58 μm (0-1 transition). The $O_2(a^1\Delta_g)$ state is mainly produced by the photodissociation of O_3 by solar ultraviolet radiation around 255 nm.

Photodissociation of molecular species can also modify the infrared airglow due to reduced density of the photodissociating molecules and/or increased density of the product species. Photodissociation of O_3 also leads to reduced daytime airglow of OH not only because of the reduced production of OH by the reaction $H + O_3 \rightarrow OH + O_2$ but also due to the more rapid $OH + O \rightarrow O_2 + H$ depletion reaction. The latter reaction becomes faster because of the enhanced oxygen atom concentrations produced by photodissociation of O_3. The net result is that daytime OH emissions around 1.5 μm ($\Delta v = 2$) and around 3.0 μm ($\Delta v = 1$) are much weaker than their nighttime intensities.

13.1.4 Chemical Reactions

Creation of molecular species in the excited state by chemical reactions and their subsequent decay contributes to airglow. The airglow produced by this mechanism is also called chemiluminescence. A well-known example of this process is the OH radiation around 1.5 and 3.0 μm and at

shorter wavelengths, even visible regions, OH being produced by the chemical reaction

$$H + O_3 \rightarrow OH^* + O_2,$$

the asterisk denoting the vibrational excitation. Another example of this process [Stair et al., 1981, 1985] is the radiation around 9.6 μm from the ν_3 mode of O_3 produced by the three body recombination reaction

$$O + O_2 + M \rightarrow O_3^* + M .$$

13.1.5 Atomic and Molecular Collisions

Collisions of infrared radiating species with other molecules lead to transfer of vibrational energy from one collision partner to the other. This transfer is much more efficient if the vibrational frequencies of the transitions involved are not too far apart (near-resonant), that is $\omega\tau \sim 1$, where ω is the frequency separation between the transitions involved and τ the time duration of the collision. Because of the near-resonance, N_2 transfers its vibrational energy efficiently to the ν_3 mode of CO_2 [Sharma and Brau, 1969]. This energy transfer process plays an important role in the enhanced nighttime 4.3 μm airglow of CO_2 around 70 km [Kumer et al., 1978].

Another important type of collision producing infrared airglow is that in which the relative translational energy of atom-molecule or molecule-molecule is converted into vibrational excitation of the molecule. These collisions lead to establishment of local thermodynamic equilibrium (LTE) between various degrees of freedom, the only requirement being that enough of these take place during the effective radiative lifetime of the emitter (taking into account selfabsorption). Even when the frequency of collisions is not sufficient to establish LTE, this type of collision may be an important source of vibrational excitation. Examples are collisions above 90 km of oxygen atoms with CO_2 and NO which play a very important role in the airglow at 15 μm and 5.3 μm, respectively [Degges, 1971; Zachor et al., 1981; Sharma and Nadile, 1981].

Another important source of vibrational energy is the collisional excitation of molecules by electronically excited atoms. For example, oxygen atoms produced by photodissociation of O_3 are created in the excited ^1D state with unit efficiency. These atoms are very efficiently quenched by molecular nitrogen. During the quenching process, a fraction of the electronic excitation of oxygen atoms is converted into vibrational energy of N_2 [McEwan and Phillips, 1975] and can be collisionally transferred to CO_2 and emitted as 4.3 μm airglow.

The processes described in the above sections are the most important but by no means the only sources of infrared radiation from the undisturbed atmosphere. Raman scattering of the solar radiation is a less important source of airglow because of the very small cross sections involved. In addition, there is Rayleigh scattering of solar radiation important only below about 40 km altitude.

13.2 NON-LOCAL THERMODYNAMIC EQUILIBRIUM (NLTE)

We note again that except for 1.27 μm and 1.58 μm $O_2(a\,^1\Delta_g \rightarrow X^3\Sigma_g^-)$ radiation the infrared airglow from 1.2 to 15 μm involves molecular vibrational rotational transitions. In the lower part of the atmosphere (below about 40 km altitude) these vibrational transitions can generally be described by local thermodynamic equilibrium (LTE), that is, vibrational excitation of the molecules can be described by a Boltzmann distribution with a vibrational temperature equaling the translational temperature. Even at these altitudes the $O_2(a^1\Delta_g)$ electronically excited state population cannot be described by a distribution with a temperature equaling the translational temperature and is much larger than predicted on the basis of LTE. The reason for this lies in the metastable nature of the $O_2(a^1\Delta_g)$ state (its radiative lifetime is about 65 min) and small rates of collisional quenching by major atmospheric species [McEwan and Phillips, 1975]. This metastable nature and the fact that collisions are not a rapid process for depopulating this state, account for the great abundance of $O_2(a^1\Delta_g)$ in the stratosphere and mesosphere.

At higher altitudes (starting about 70 km) the vibrational population of the radiating species, even those that are not metastable and are neither produced by chemical reaction nor pumped by sunshine or earthshine, becomes different from what would be predicted on the basis of LTE. Situations deviating from LTE behavior are called NLTE cases. Sometimes it is useful to introduce the concept of vibrational temperature. The concept is based on the fact that near-resonant vibration-to-vibration transfer processes are much faster than nonresonant vibration-to-translation processes. The time needed for the vibrational energy to be partitioned according to a Boltzmann distribution in a collision between identical molecules is usually several orders of magnitude shorter than the time needed to establish LTE in gases [Sharma, 1969, Taylor, 1974]. The vibrational temperature T_v is defined by the relation

$$\ln\left(\frac{g_e\, n_l}{g_l\, n_e}\right) = \frac{C_2\, E}{T_v}$$

where n_e and n_l are the excited state and low state number densities, g_e and g_l are the respective statistical weights, $C_2 = 1.4399$ K/cm^{-1} is the second radiation constant and E is the energy difference between the upper and lower states in cm^{-1}.

The concept of vibrational temperature has been used to calculate [Sharma and Nadile, 1981] the contribution to

the 15 μm radiation by the transition 02201 → 01101 of CO_2. Only when this contribution is added to that of the 01101 → 00001 transition from all the isotopes is a reasonable agreement obtained between the 15 μm radiation observed by the rocket experiment SPIRE in the 70 and 90 km tangent height and the calculated CO_2 radiances [Stair et al., 1981, 1985].

The concept of vibrational temperature cannot be used in all NLTE situations. For example, the vibrational population of OH around 80 km altitude cannot be described by a Boltzmann distribution. This concept, however, is used where possible because it also gives an idea of the extent of deviation from LTE.

It should be pointed out that while fewer collisions lead to NLTE situations, more frequent collisions do not guarantee LTE behavior. This is illustrated by the 2.7 μm daytime CO_2 emission that can still be characterized around 50 km altitude by vibrational temperature about 100 K higher than the translational temperature. This emission, in part, is shown to arise from the absorption of sunlight at 1.5 μm by very weak 00001 → 30011 (x = 1,2,3) bands (Figure 13-1) [Sharma et al., 1981]. Although each absorbed photon has a rather small probability of being emitted in the 2.7 μm region, the large number of absorbed photons yields a detectable signal.

Having stated the differences between LTE and NLTE situations in some detail, we will focus for the remainder of this discussion on NLTE airglows, specifically 1–18 μm emissions originating from altitudes higher than 40 km. The LTE situation has been discussed in great detail by other workers [see, for example, Clough et al., 1981].

13.3 METHODS OF MEASUREMENT

Infrared airglow has been studied with rockets [Stair et al., 1974; Nadile et al., 1977; Stair et al., 1981; Markov, 1969; Offermann and Grossmann, 1978] and satellites [Markov et al., 1976]. In this chapter the primary emphasis will be, for reasons cited earlier, on the work at AFGL. Many of the AFGL rockets were launched from Poker Flats, Alaska [Stair et al., 1975]. The rocket payloads contained either limb-viewing or zenith-viewing light collecting and analyzing instruments. In these experiments the instrument package included a protective cover that was opened at a specified altitude during the free ballistic flight phase. The altitude for cover opening, determined by the requirement that the infrared emission due to shock heating by the rocket be sufficiently small, was 50 km or higher.

The simplest infrared sounding rockets measure the infrared radiance reaching the payload from the atmosphere above it. If the payload is not provided with an attitude control system, the rocket usually cones around an axis that makes a small angle with the vertical direction. When measuring the infrared fluorescence, care is taken that the sun is never directly in the line-of-sight (LOS). As the rocket gains altitude, less and less of the gas is being looked at so that the signal during ascent usually decreases with time (except during atmospheric disturbances); during the descent the reverse is true. However, both in ascent and descent, the magnitude of the derivative of the signal with altitude is a measure of the local radiance—provided we are dealing with an optically thin and steady state one-dimensional atmosphere. Figure 13-2 shows a typical spectral profile of

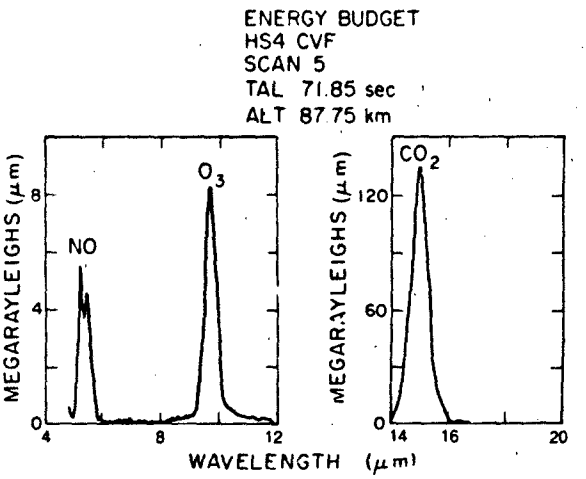


Figure 13-2. Plot of radiance vs wavelength from the zenith looking CVF "Energy Budget" rocket at 87.75 km altitude. The signal from CO_2 at 15 μm is at least an order of magnitude larger than that from NO at 5.3 μm or ozone at 9.6 μm.

zenith emission from the NO vibrational-rotational (V-R) fundamental transition ($\Delta v = 1$) at 5.3 μm, the 001 → 000 V-R band of O_3 at 9.6 μm, and the 01101 → 00001 band of CO_2 at 15 μm obtained by a rocket launched from Andoya, Norway during the Energy Budget Campaign [Ulwick et al., 1983]. This rocket was launched in November 1980 at 0314 UT. Since the solar shadow height was 400 km above the launch site, the experiment was essentially conducted in a night atmosphere.

Another approach useful for viewing weak infrared atmospheric emissions is to add a telescope to the optical system and observe the infrared radiation from the earth-limb. The telescope with a small fraction of degree full viewing angle (usually about 1/4 degree) and excellent off-axis rejection keeps the footprint of the LOS a few kilometers wide. The limb-looking SPIRE (SPectral Infrared Rocket Experiment) has obtained very useful results yielding many new insights into the infrared airglow [Nadile et al., 1977; Stair et al., 1981]. In the SPIRE experiment with the rocket in the exosphere, the earthlimb was scanned vertically through a telescoped circularly variable filter (CVF) spectrometer. The resolution of the instrument was 3% to 4% depending on the wavelength. The tangent height (the distance of closest approach to the earth's surface) is varied

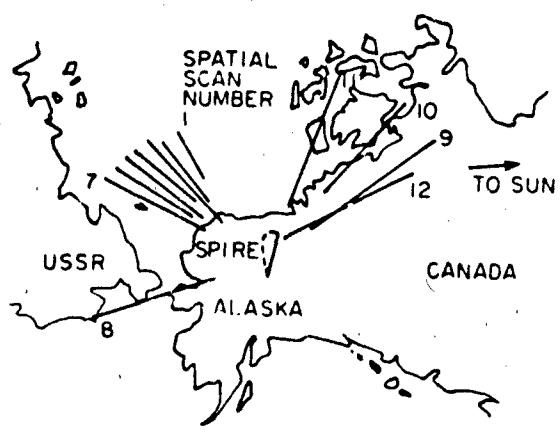


Figure 13-3. Geographical location of projection on earth's surface of tangent heights for various SPIRE scans. Tangent height increases going outwards from the center of figure. Farthest tangent point is 1500 km away from launch site.

by changing the telescope zenith angle. As the tangent height is varied, the location of the LOS and the tangent point (point closest to the earth's surface) also change (Figure 13-3). One result of this spatial scanning is the ability to obtain a day, twilight or night LOS (Figure 13-4). This is in contrast to the zenith-looking rocket where in ascent or descent the region sampled is less than 100 km over the surface of the earth. While the possibility of investigating day-night variations in a zenith-looking experiment is rather limited, it does offer the possibilities of determining the morphology on a small scale.

There are numerous other differences between the zenith and limb experiments. Two important ones are the amount of the gas sampled and the altitude dependence of the volume probed. In the zenith looking experiment, the length of the LOS between any 1 km altitude interval is sec θ km, where θ is the zenith angle of the LOS as long as $\theta \leq 75°$ so that the curvature of the earth is not important. The total volume of the gas sampled is relatively small. Moreover, the zenith experiment weights equal increments of altitude in the LOS equally. Such is not the case in a limb experiment. For example, a detector positioned at an altitude of 285 km looking at 100 km tangent height gathers photons from a path length in excess of 3000 km. Further, for density decreasing monotonically with altitude the volume of the gas sampled drops sharply as we move away from the tangent height. The length of the gas column with altitude between 100 and 101 km is about 227 km. The length of the path sampled with each km increase in altitude decreases in proportion to $[\sqrt{n} - \sqrt{(n-1)}]$ where n and (n − 1) are the upper and lower altitudes in km measured from the tangent height. In the example cited above, the length of the LOS between 110 and 111 km is 0.154 times that between 100 and 101 km. Thus, equal increments of altitude in the LOS are not equally weighted in the limb experiment and the radiation from tangent height usually predominates. The above remarks are valid only for an optically thin gas.

A further characteristic of the limb experiment is that because the volume of the gas sampled is so large, the local variations in the atmospheric structure are averaged out. On the other hand, the large column sampled also permits the limb experiment to see some emissions missed by the zenith experiment when performed at the same sensitivity. For example, Figures 13-5 and 13-6 show 15 μm radiance from ν_2 band of CO_2 seen by the SPIRE and Energy Budget rockets as a function of altitude. It can be seen in Figure 13-5 that between 95 and 110 km tangent height the band

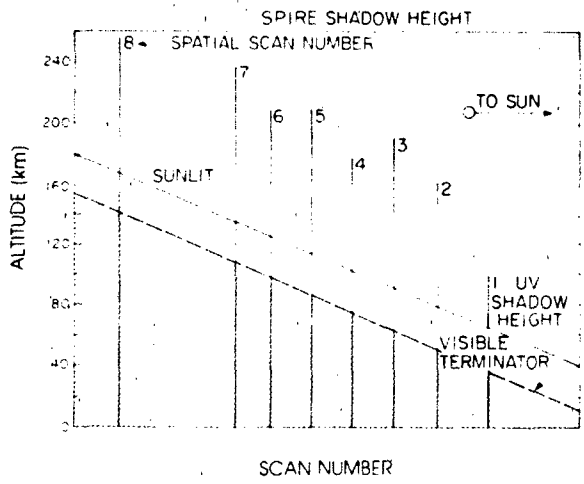


Figure 13-4. Plot showing visible ultraviolet terminators for eight of twelve SPIRE scans. Ultraviolet shadow height is 23 km higher than visible shadow height due to thickness of ozone layer. Scans 9–12 are completely sunlit.

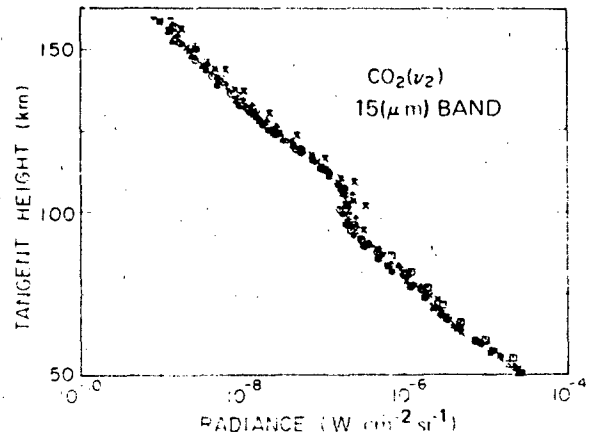


Figure 13-5. Integrated radiance observed by SPIRE in the 14–16 μm interval as function of tangent height. Night scan 8 has 30% higher radiance above 100 km.

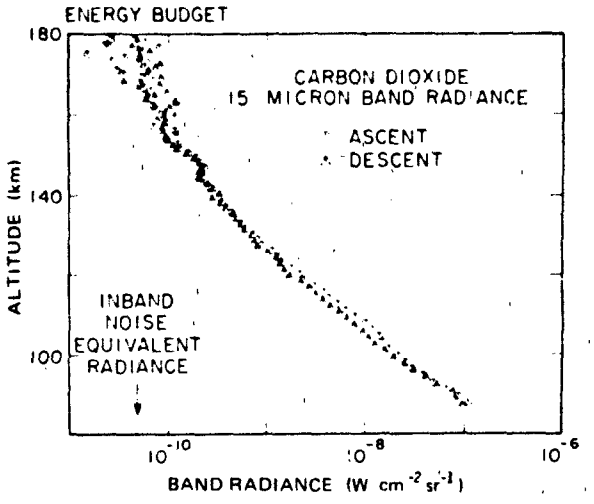


Figure 13-6. Integrated band radiance around 15 μm observed by zenith-looking Energy Budget rocket during ascent and descent as function of altitude.

radiance stays constant although the CO_2 concentration decreases in this altitude range by a factor of 30 to 40. Detailed modeling has shown this phenomenon to be due to oxygen atoms, which are much more efficient at collisionally pumping the CO_2 (01101) level than N_2 or O_2 [Sharma and Nadile, 1981]. The zenith experiment (Figure 13-6) does not give any hint of this behavior in the measured radiances.

Finally, the tangent height in the earthlimb experiment can be varied from zero (earth's surface) to 250 km, whereas the range of altitudes attainable in the zenith experiment is more limited. The two types of experiments emphasize different aspects of infrared airglow and each have their own advantages and disadvantages.

13.4 RESULTS AND DISCUSSION

In this section we introduce the prominent features of the measured airglow and compare them with model calculations. Because of the nature and scope of this chapter, we omit the details of the calculations and approach the subject as an introduction to infrared airglow from the undisturbed atmosphere. Where possible, the published work is referenced; however, a great deal of work is available so far only in the form of non-refereed reports and unpublished work.

We will use SPIRE to illustrate the broad features of the data. SPIRE was launched on September 28, 1977 [Nadile et al., 1977; Stair et al., 1981, 1983]. This payload had three coaligned primary sensors: (1) a telescoped two-channel photometer, (2) a cryogenic (77 K) telescoped short wavelength infrared (SWIR) CVF spectrometer, and (3) a cryogenic (10 K) telescoped long wavelength infrared (LWIR) CVF spectrometer. The spectrometers were coaligned and covered the 1.4 to 16 μm region. The telescopes had 1/4° fields of view with high off-axis rejection (point source rejection of 10^{-5} at 1° off-axis) and were coaligned to better than 0.1°. The payload reached an altitude of 285 km, and the earthlimb was spatially scanned from hard earth to local horizontal at 0.5 deg/s while taking spectra at 2 scans/s. The spectrometers were built at Utah State University and the telescopes at the Honeywell Radiation Center. With a 1/4° field of view, the footprint of the LOS varied from 4 to 8 km, depending upon the tangent height and the payload altitude.

Before we give the profiles of the emitted band radiance as a function of tangent height (for the limb-looking experiment) or altitude (for the zenith-looking experiment), we would like to discuss airglow in general terms as a function of altitude. For this purpose we would like to divide the atmosphere above 40 km into several regions. This division is created to facilitate discussion of the SPIRE data. Taking into account the fact that it was an experiment of about 5 min duration at dawn, it is quite possible that the boundaries of the regions created here may have to be modified in context of other data bases. Still we believe it is a useful exercise to point out underlying sources common to the airglow from different bands.

13.4.1 Thermosphere (above 100 km)

This region is characterized by a rapid increase of temperature with altitude. Important emitters in this region are restricted to NO and CO_2. The airglow, at 5.3 and 15 μm, in this region can be explained primarily in terms of oxygen atom collisions with the emitters. The oxygen atom concentration in this region shows no diurnal variation. The result is that infrared radiation from the excited states produced during the collision of the radiating species with the

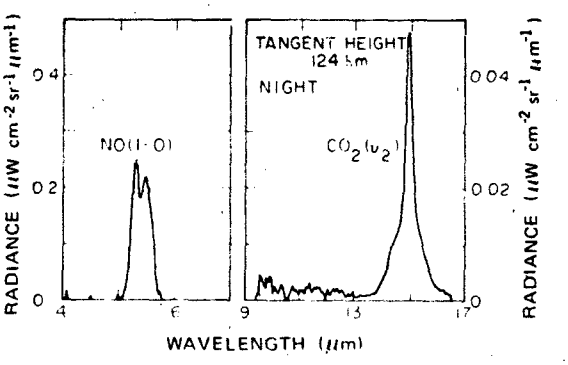


Figure 13-7. Nighttime radiance seen by the long wavelength SPIRE spectrometer as function of wavelength at 124 km tangent height. NO ($1 \rightarrow 0$) band at 5.3 μm and CO_2 (01101 $\rightarrow$ 00001) band at 15 μm are the only reproducible spectra seen. Scale for the NO band is 10 times larger than that for CO_2 band. (Daytime results were essentially the same.)

oxygen atoms shows no diurnal variation. A spectrum obtained with the SPIRE liquid-helium-cooled spectrometer at 124 km tangent height is shown in Figure 13-7. Although this spectrum is obtained at night, the post-dawn results are essentially the same. Emission from the 1-0 vibrational band of NO and the 15 μm (01101 → 00001) bending mode of CO_2 are the only two features seen at this altitude. Detailed modeling of these emissions indicates that (1) the emissions originate from the NLTE atmosphere with molecular vibrational temperatures lower than the local translational temperatures, (2) the dominant mechanism for vibrational pumping is the collisional excitation by oxygen atoms, and (3) radiative cooling due to NO emissions is the most important atmospheric cooling mechanism (~500 K/day) at these altitudes.

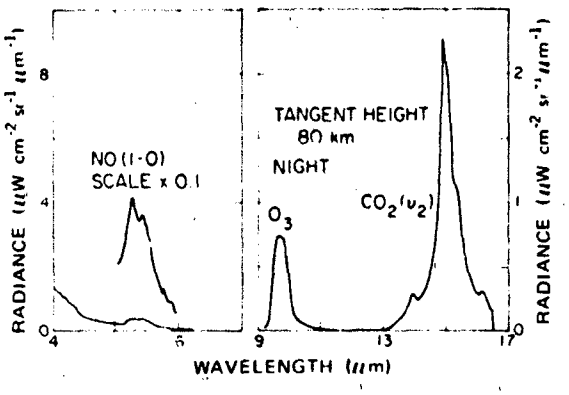


Figure 13-9. Nighttime signal for long wavelength SPIRE spectrometer as function of wavelength at tangent height of 80 km. Notice appearance of O_3 (ν_3) signal. The apparent signal in the 4–6 μm region is due to dielectric relaxation response of the detector to the 15 μm radiation. The 5–6 μm region has been magnified 10-fold to highlight the NO spectral features.

13.4.2 Lower Thermosphere (between 100 km and 85 km)

In this region the oxygen atom concentration begins to decrease without showing any diurnal variation while the temperature shows little change. In this region infrared airglow is characterized by the effect of the solar radiation and oxygen atoms on the radiating species. In addition to the NO and CO_2 bands mentioned above, we see emission from ozone at 9.6 μm (ν_3 band) and the 4.26 μm ν_3 band of CO_2 (Figures 13-8 and 13-9). Because of the large absorption coefficient of 00011 → 00001 transition at 4.26 μm, this transition is severely self-absorbed. The ν_3 photons reaching the detector most likely originate from the transition between hot bands and have slightly longer wavelength (around 4.3 μm). The diurnal airglow variations also begin to manifest themselves in the 9.6 μm emission from ozone. At about 100 km tangent height at night, the ν_3 transition of O_3 begins to appear in the spectrum between 9 and 11 μm (Figure 13-10). At a slightly lower tangent height, this transition appears during the dawn observations as well. While the 001-000 transition of ozone is centered around 9.6 μm, the profile of the emitted radiation appears, in addition, to consist of some hot-band transitions. These transitions are not unambiguously identified as yet and are referred to in the following discussion as the long-wavelength tail of the

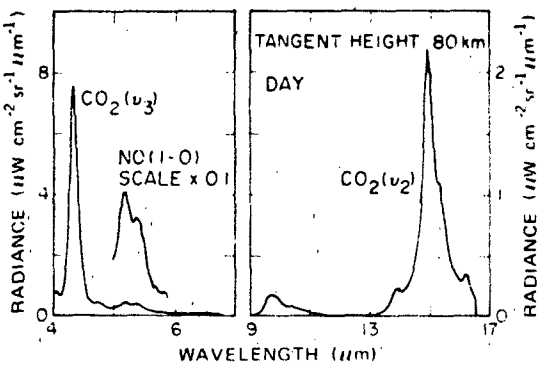


Figure 13-8. Daytime signal from the long wavelength SPIRE spectrometer as function of wavelength at tangent height of 80 km. The 15 μm CO_2 and 5.3 μm NO signals are about the same as in Figure 13-9 (nighttime). The 4.3 μm CO_2 signal is enhanced and 9.6 μm O_3 signal is reduced over nighttime values (Figure 13-9).

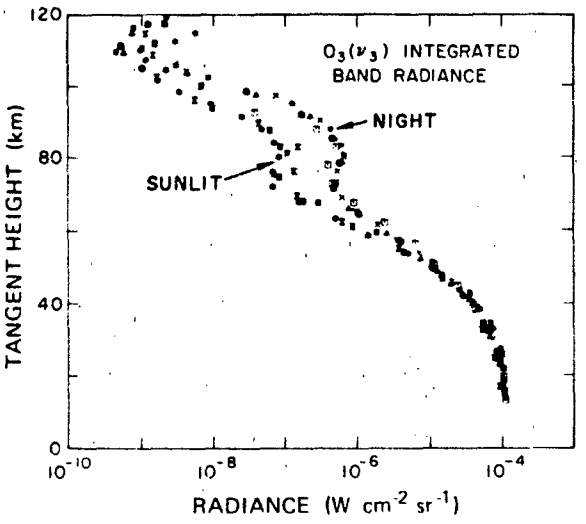


Figure 13-10. 9.6 μm (ν_3) ozone (19.20–10.50 μm) band radiance as function of tangent height. Note stronger nighttime radiance and peak around 80 km. A terminator crossing is evident in scan 1 (□) where the radiation decreases by an order of magnitude from 85 to 90 km.

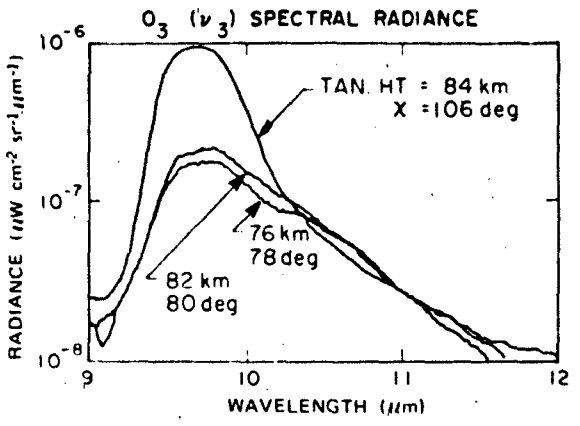


Figure 13-11. 9.6 μm (ν_3) ozone spectrum at 80 km tangent height for different solar elevation angles. The signal for wavelengths longer than 10.2 μm shows no diurnal variation.

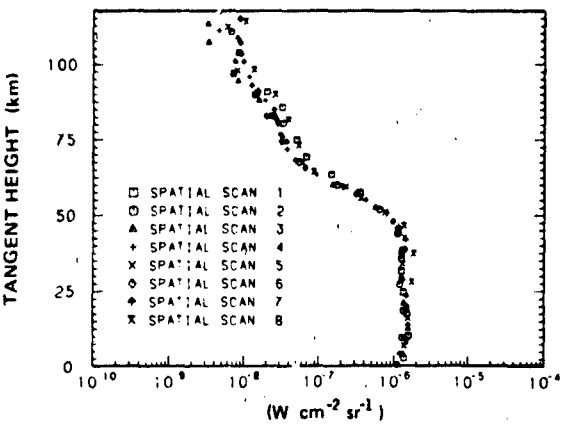


Figure 13-13. Nighttime integrated 4.3 μm band radiance observed by short wavelength SPIRE spectrometer. Notice much faster decrease of the signal above 70 km tangent height than for the daytime case.

9.6 μm emission. The night and day long wavelength tails are essentially the same (Figure 13-11) even though the total O_3 emission decreases dramatically in the daytime. This appears quite reasonable since the day and night oxygen atom concentrations at these altitudes are the same and the vibrationally excited O_3 formed by the $O + O_2 + M \rightarrow O_3 + M$ reaction is the suggested source of the long wavelength ν_3 photon emission. The peak at 9.6 μm results mainly from resonant scattering of earthshine by the ν_3 mode of O_3 and is a measure of O_3 concentration; the smaller daytime value of this scattered radiation reflects the lower daytime O_3 concentration due to photodissociation. The CO_2 4.3 μm ν_3 band daytime radiance seen clearly in Figure 13-12 is over ten times more intense during the day than during the night (Figure 13-13) in agreement with the earlier theoretical prediction of James and Kumer [1973]. As pointed out earlier, the daytime 4.3 μm signal is due to resonance fluorescence as well as nonresonant fluorescence excited by a shorter wavelength, for example, solar radiation (2.7 μm). Near the top of this layer, the collisions of oxygen atoms with CO_2 and NO also play an important role in 15 and 5.3 μm emissions.

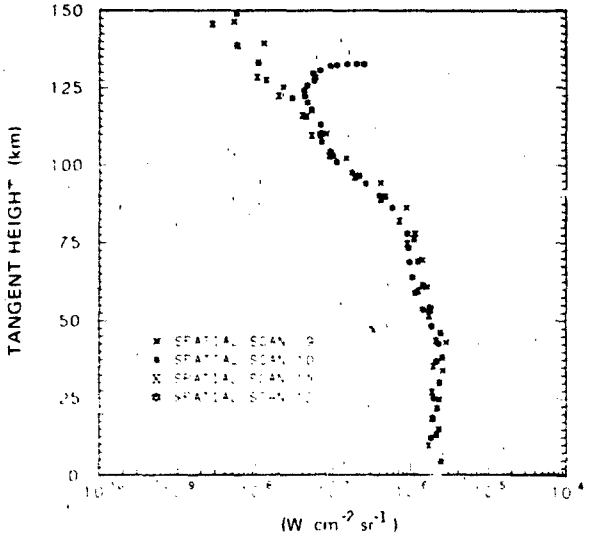


Figure 13-12. Daytime integrated 4.3 μm band radiance observed by short wavelength SPIRE spectrometer as function of tangent height. Larger signal in scan 12 is due to the fact that a deliberate effort was made in this last scan to test the telescope.

13.4.3 Upper Mesophere (between 85 km and 70 km)

Infrared airglow is primarily determined by processes involving solar radiation during the day (Figure 13-14) and the airglow reaction, $H + O_3 \rightarrow OH + O_2$ and $HO_2 + O \rightarrow OH + O_2$ during the night (Figure 13-15). Large

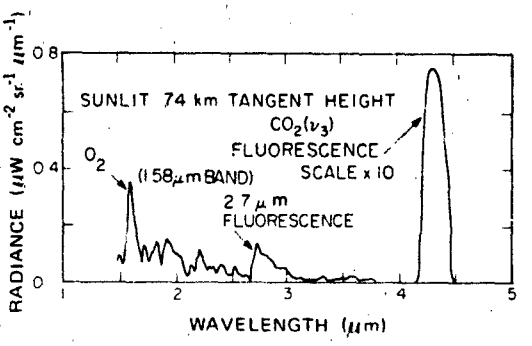


Figure 13-14. The daytime signal from near infrared spectrometer at 74 km tangent height. The nighttime OH radiation is reduced. $O_2(a\Delta_g, v = 0) \rightarrow {}^3\Sigma_g^-, v = 1)$ radiation appears at 1.58 μm. The 2.7 μm radiation is fluorescence signal from $CO_2(\nu_1 + \nu_3$ combination bands) and H_2O (ν_1 and ν_3 bands).

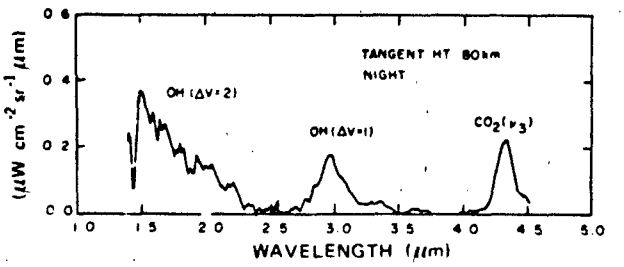


Figure 13-15. The nighttime signal from near infrared spectrometer as function of wavelength. Notice the progression of OH overtone bands ($\Delta v = 2$) between 1.5-2.4 μm and the OH fundamental bands ($\Delta v = 1$) peaked at about 3 μm. The fluorescent emission at 4.3 μm is shown here without the effects of the dielectric relaxation problem encountered in the long wavelength spectrometer.

diurnal differences in the density of oxygen atoms are present. Collisions of the radiating species with oxygen atoms become an unimportant mechanism for depopulating the radiating states because of the small oxygen atom concentration even in daytime. The collisions of the radiating species with N_2 are believed important for OH and CO_2 airglows at night [Kumer et al., 1978].

At a tangent height of about 80 km, the emission around 4.3 μm from CO_2 emission is about 30 times stronger during the day than at night due to absorption of solar radiation (Figures 13-12 and 13-13). Nighttime 4.26 μm fluorescence is primarily due to energy transfer from vibrationally excited N_2 [Kumer et al., 1978], in turn produced by energy transfer from vibrationally hot OH produced by chemical reactions, for example, $H + O_3 \rightarrow OH + O_2$.

In contrast to the strong daytime emission from CO_2 around 4.3 μm, a strong nighttime airglow is seen from OH (Figure 13-15). This airglow is composed of many vibration-rotation bands extending over a wide spectral region from visible to 4 μm. Fundamental vibration-rotation bands ($\Delta v = 1$) and first overtone ($\Delta v = 2$) bands are observed in the infrared [Murphy, 1971] and are shown in Figure 13-11. The overtone hydroxyl emission is broad and appears to be a composite of several vibrational transitions. The fundamental ($\Delta v = 1$) band is less broad and is peaked around 2.9 μm indicating strong contributions from 1-0 and 2-1 vibrational transitions (Table 13-1). It should be pointed out that the Einstein A coefficients for OH overtone emission are larger than those for OH fundamental emission for all upper vibrational levels except $v = 2$ (Table 13-2). The observations cited above appear reasonable if the collisional deactivation rates are closer to the larger of the two sets of Einstein A coefficients. The daytime hydroxyl (Figure 13-14) emissions appear weak. The 1.6 μm region is dominated by $O_2(a\,^1\Delta_g)$ emission and the 2.7 μm region by CO_2 and H_2O emission below 70 km.

The emission around 15 μm continues to show no diurnal variations (Figure 13-15). Also, emission at 13.8 and 16.2 μm in the wings of CO_2 15 μm band is observed (Figures 13-8 and 13-9). This emission is due to the transition from 10001 and 10002 levels to the 01101 level of CO_2; the transition from the third level, that is, the 02201 → 01101 transition at 668.3 cm^{-1} (compared to 667.3 cm^{-1} for the $01^10 \rightarrow 000$ transition) makes a substantial contribution to the band radiance but is not resolved from the 01101 → 00001 emission. Also not separated are the trace isotope emissions from that of the main isotope. Modeling studies indicate that the contribution of the trace isotopes to the radiance at 15 μm is much larger than expected on grounds of abundance because of the fact that the emission from the $01^10 \rightarrow 000$ transition of the main isotope is self absorbed at tangent height below 110 km.

Around 80 km, the rates for collisional relaxation of ozone, $O_3(001) + M \rightarrow O_3(000) + M$, become comparable to the Einstein A coefficient for the radiative relaxation. Rates of collisional relaxation from higher vibrational levels of ozone are not known. A clear description of the underlying processes, therefore, cannot be formulated.

Around 75 km tangent height, the daytime scans show the $a^1\Delta_g \rightarrow X^3\Sigma_g^-$ (0-1) O_2 transition at 1.58 μm; the 0-0 band of the same electronic transition at 1.27 μm is outside the range of detection and, therefore, not observed. No signal at 1.58 μm is observed at night. The metastable a state has a radiative lifetime of 3.88×10^3 s. The Franck-Condon integral for the 0-1 transition is 0.0131 times that of the 0-0 transition, making the Einstein coefficient of the 0-1 transition about $1.76 \times 10^{-6}\ s^{-1}$ [Krupenie, 1972]. That a finite signal is observed is due to the large length over which light is collected and the large densities of $O_2(^1\Delta_g)$ (about $5 \times 10^9\ cm^{-3}$). Total ground state O_2 density at 74 km is 1.7×10^{14}/cc and that in vibrational level 1 of the ground electronic state is about $3 \times 10^9\ cm^{-3}$. We thus see that at 74 km tangent height there is complete optical inversion between the a(v = 0) and X(v = 1) states. The gain in the electronic transition is about 10^{-8} per km and assuming a path length of 10^3 km we get signal amplifications of $e^{(xxxx)}$. Although the laser action on this transition does not appear likely, the signal shows a large amount of energy stored [$O_2(a^1\Delta_g) \rightarrow O_2(X^3\Sigma_g^-)$, 0-1]. The lack of 1.58 μm emission at night indicates that the main source of $O_2(a^1\Delta_g)$ for the region of the atmosphere discussed here is the photodissociation of O_3.

13.4.4 Lower Mesosphere (between 70 km and 40 km)

Solar radiation in the strong absorption bands starts becoming depleted as the sunlight traverses this region. Infrared airglow in this region is therefore very dependent on the LOS and three-dimensional atmospheric calculations are needed to model the radiance in this region. Figure 13-16 shows the 0-1 band radiance of the electronic transition ($a^1\Sigma_g^+ \rightarrow X^3\Sigma_g^-$) of O_2 as a function of tangent height. It is seen that the larger the solar elevation angle, the greater the

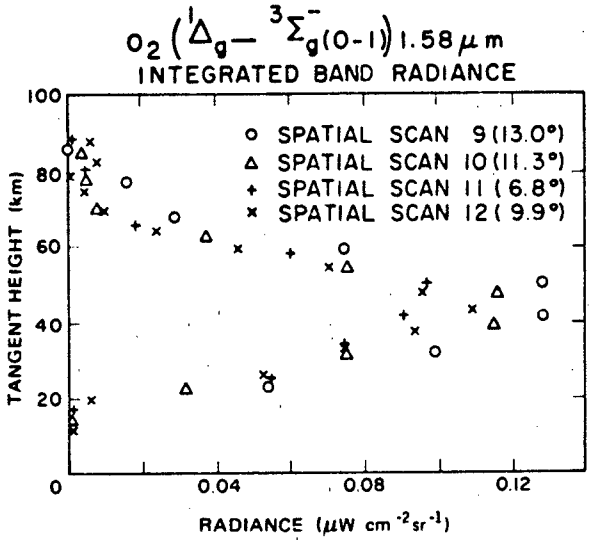


Figure 13-16. Integrated (1.51–1.65 μm) O_2 ($a^1\Delta_g$) radiance at 1.58 μm as function of solar elevation angle. Peak signal around 40 km shows inverse correlation with solar elevation (χ) angle.

1.58 μm radiance. Also in contrast to the higher altitudes where nighttime 9.6 μm ν_3 radiance from ozone is larger than the daytime value, the diurnal variation vanishes in the lower part of this altitude range (below about 50 km).

We again point out that this division of the atmosphere into four parts for the discussion of the airglow is rather arbitrary and with a different data base the boundaries between the regions may have to be slightly shifted. However we believe that basic qualitative features of infrared airglow are well described by this division. This method of presenting the data also stresses the common sources that give rise to airglow in different bands and species.

13.4.5 Infrared Radiators

We now present a brief discussion of the airglow classified according to the radiating species in order to summarize the previous discussion as well as to bring out several new points.

13.4.5.1 CO_2. The CO_2 radiation is seen in the limb-scanning SPIRE experiment at wavelengths of 15, 4.3, and 2.7 μm. 15μm radiation (Figure 13-5) shows no diurnal variation and is excited mainly by collisions with oxygen atoms above 90 km. Below 90 km altitude the main mechanisms of excitation are the collisions with N_2 and O_2 with the earthshine also playing a role. Above 70 km the 15 μm vibrational temperature lags behind the translational temperature, whereas below 70 km LTE forms a good basis for describing the 15 μm radiation (Figure 13-17).

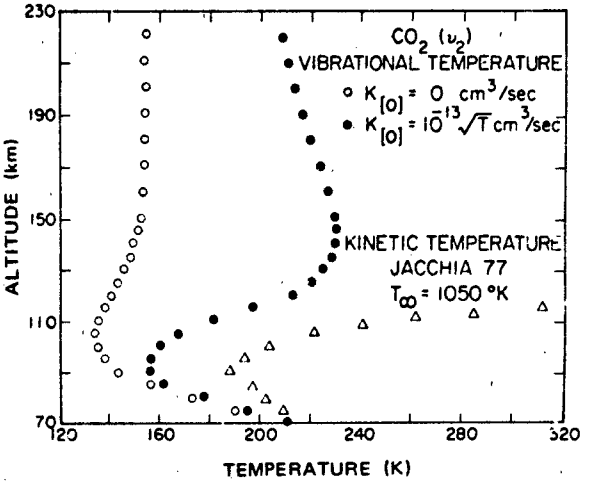


Figure 13-17. Vibrational temperature of bending mode of CO_2 for two rate constants for the deactivation by oxygen above. Kinetic temperature is denoted by triangles.

Solar radiation plays an important role in the fundamental ($\Delta v = 1$) emission from the ν_3 mode at 4.3 μm. Figure 13-18 gives a plot of the vibrational temperature and kinetic temperatures [Degges and Smith, 1977] as functions of altitude. Figures 13-12 and 13-13 give a plot of the 4.3 μm emission observed by SPIRE as a function of tangent height during day and night, respectively. The daytime ra-

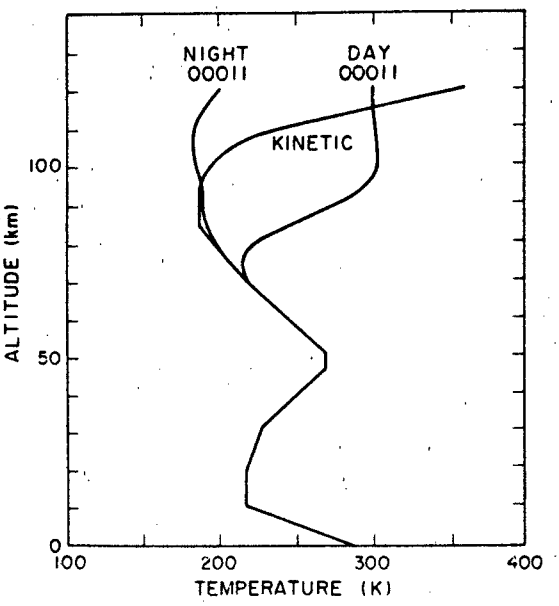


Figure 13-18. Daytime and nighttime vibrational temperatures of asymmetric stretch of CO_2 as function of altitude. Kinetic temperature is also plotted for comparison [Degges and Smith, 1977].

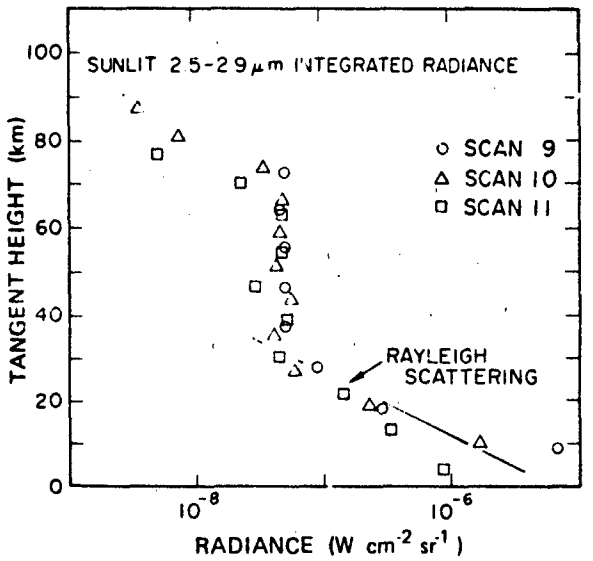


Figure 13-19. Daytime integrated 2.5–2.9 μm band CO_2 and H_2O radiance as function of tangent height. Signal below 30 km is due to Rayleigh scattering.

diance at around 80 km tangent height is about a factor of 30 larger than in nighttime value. Also, Scan 12 was programmed to look almost at the sun at higher altitude. The increased signal around 125 km tangent height is due to telescope leakage.

Radiation at 2.7 μm (Figure 13-19) is observed from CO_2 during daytime only. In addition to CO_2, there is contribution from H_2O in this band [Kumer, 1981; Sharma et al., 1981].

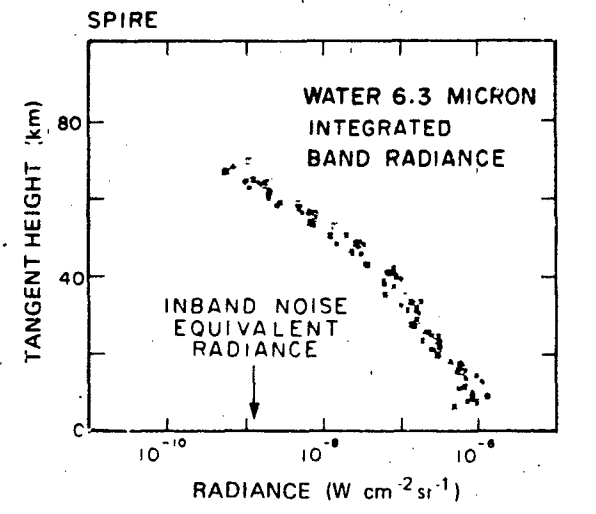


Figure 13-20. 6.3 μm (v_2) water band radiance (integrated 5.50–6.74 μm) as function of tangent height.

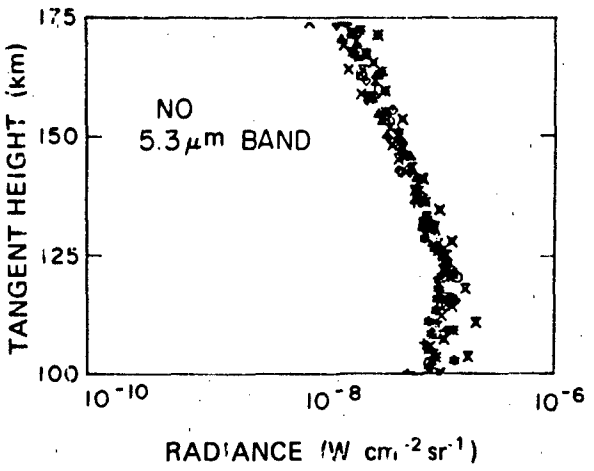


Figure 13-21. Integrated 5.00–5.70 μm NO band radiance as function of tangent height. Night scan 8 gives systematically larger signals around 100 km.

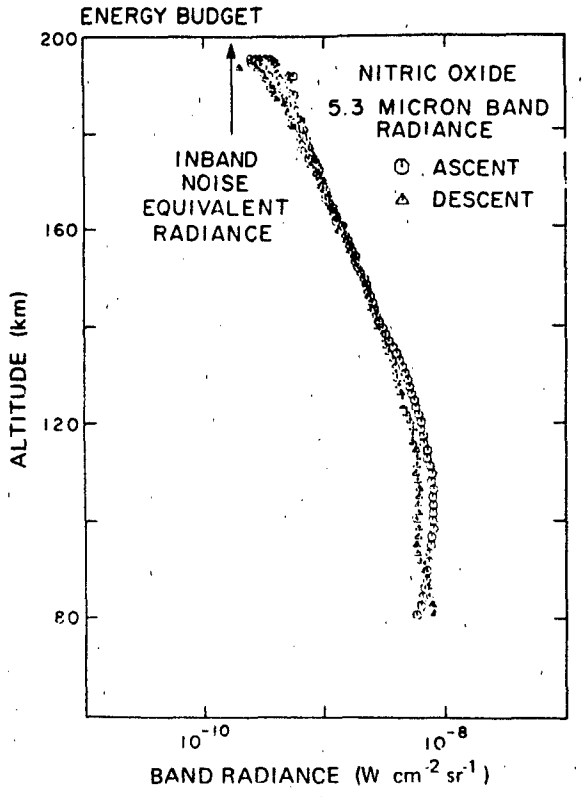


Figure 13-22. Integrated band radiance for the 5.3 μm NO signal observed in the Energy Budget campaign during ascent and descent as a function of tangent height.

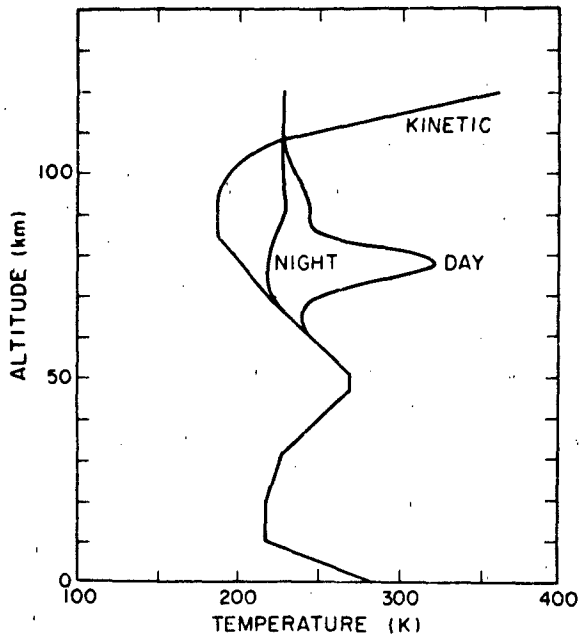


Figure 13-23. Daytime and nighttime vibrational temperatures of ν_3 mode of ozone as function of altitude [Degges and Smith, 1977].

13.4.5.2 H_2O. Radiation from H_2O is observed at 6.3 μm (ν_2 bending mode) and during the daytime around 2.7 μm (ν_1 and ν_3 modes). The radiation at 6.3 μm as a function of tangent height is shown in Figure 13-20. The LOWTRAN 4 radiance model [Selby et al., 1978] using 5 ppm H_2O adequately represents the data between 60 and 70 km tangent heights. Above 70 km the daytime 2.7 m emission is partly explained at 1.5 ppm of water in a non-LTE model by Degges (private communication, 1984).

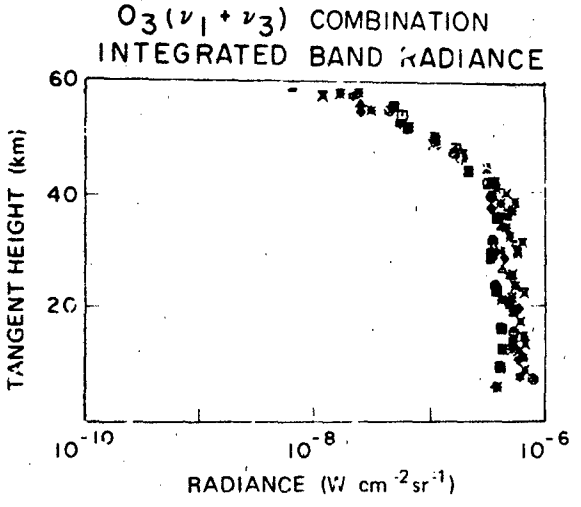


Figure 13-24. 4.8 μm ($\nu_1 + \nu_3$) combination band signal for ozone.

13.4.5.3 NO. The 5.3 μm radiation from NO in the SPIRE data (Figure 13-21) shows no diurnal valuation. Figure 13-22 shows NO radiation from the Energy Budget zenith-looking rocket as a function of altitude. Above 90 km, the oxygen atoms play a dominant role in the 5.3 μm radiation process because of the large rate constant [Degges, 1971; Fernando and Smith, 1979] for the collisional excitation of NO. The 5.3 μm radiation from NO is the major cooling mechanism above 100 km [Kockarts, 1980; Zachor et al., 1981].

13.4.5.4 O_3. Ozone emission is observed around 9.6 μm (ν_3 band) and around 4.8 μm ($\nu_1 + \nu_3$ combination band). Ozone emission at 9.6 μm has been one of the better studied

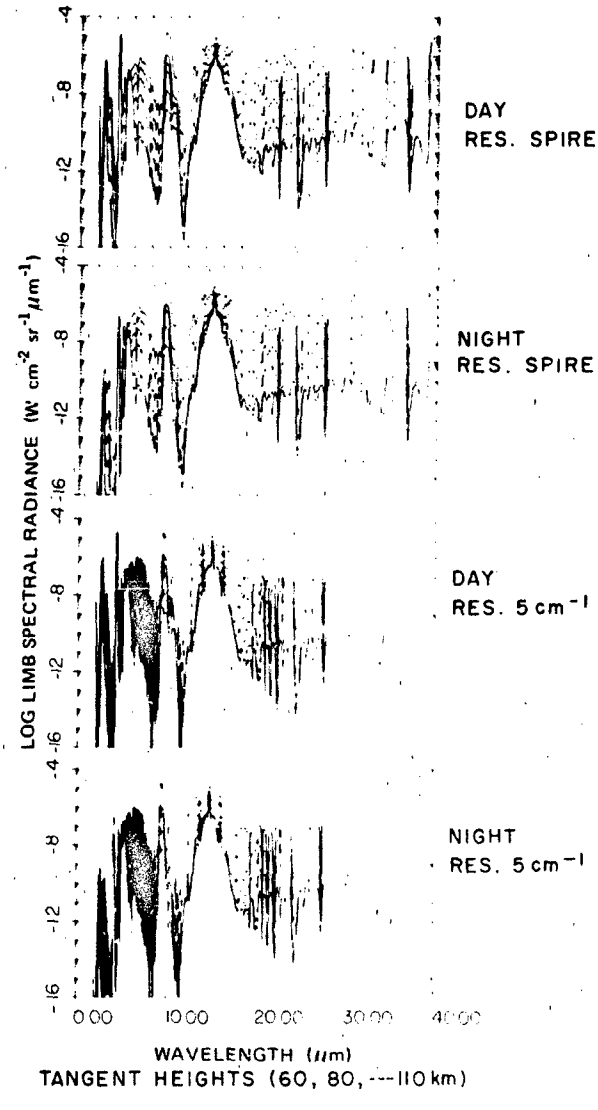


Figure 13-25. Day and night signals predicted using model calculations of T. Degges at SPIRE and 5 cm^{-1} resolutions.

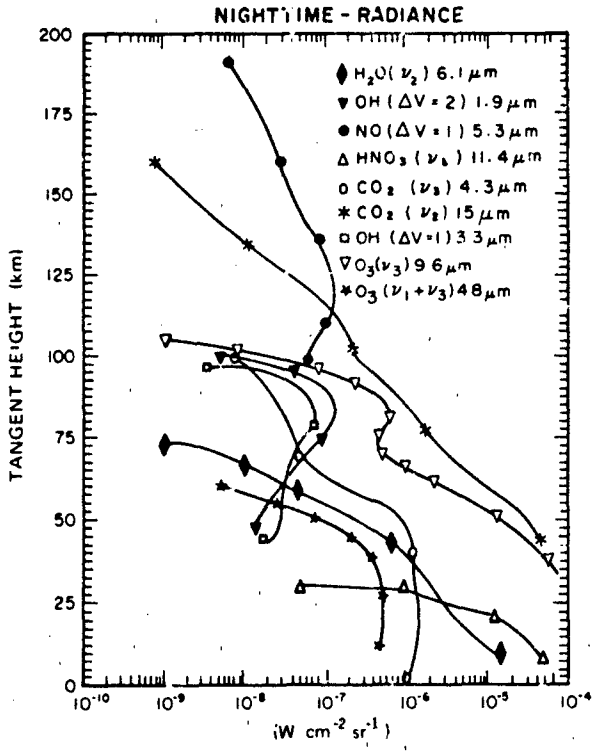


Figure 13-26. Summary of the nighttime signals observed by SPIRE.

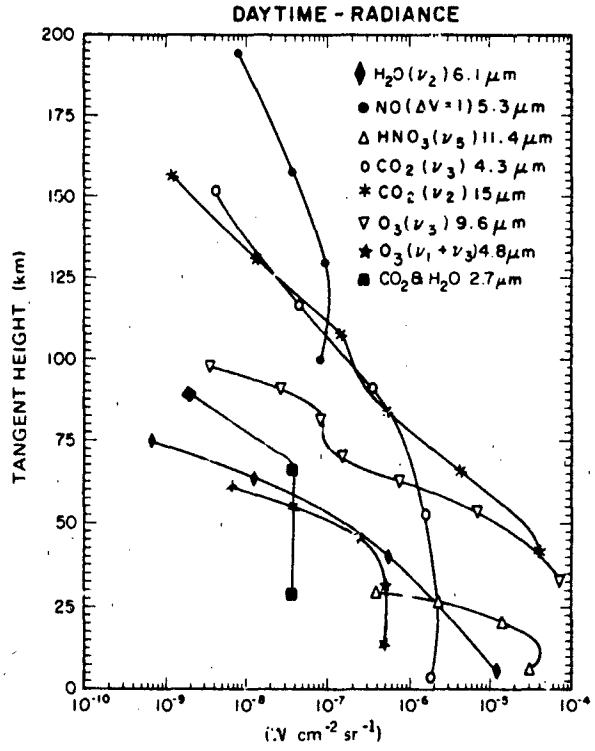


Figure 13-27. Summary of the daytime signals observed by SPIRE.

infrared airglow processes [Rawlins et al., 1981]. Figure 13-23 shows the ν_3 vibrational temperature [Degges and Smith, 1977] as a function of altitude. Figures 13-10 and 13-24 show the radiance from the 9.6 and 4.8 μm bands as functions of tangent height.

13.4.5.5 $O_2(^1\Delta_g)$. The 0-1 band emission at 1.58 μm has been already described in detail. Most of the earlier rocket work has studied the 0-0 emission [Paulsen et al., 1972]. The conclusion of the studies is that the concentration of $O_2(^1\Delta_g)$ in the atmosphere is $10^9 - 10^{11}$ cm^{-3} depending on the altitude.

13.4.5.6 OH. The hydroxyl airglow emission occurs over a relatively narrow altitude region of the mesosphere [Baker et al., 1977]. The peak emission occurs around 85 km and the thickness of the emitting layer is about 10km. The origin of the hydroxyl emission is believed to be chemiluminescence from two chemical reactions,

$$H + O_3 \rightarrow OH(v \leq 9) + O_2$$

and

$$O + HO_2 \rightarrow OH(v \leq 6) + O_2.$$

The model for hydroxyl emission is complex and involves a set of several aeronomic reactions of ozone in an atmosphere containing hydrogen (Moreels et al., 1977).

In addition to the principal radiators mentioned above, there are several minor radiators in the altitude range covered in this article, for example, NO_2, N_2O and CH_4.

Figure 13-25 gives a summary of the emission in the 2-40 μm regions as computed by the AFGL Limb Model [Degges and Smith, 1977] at the resolution of the SPIRE experiment ($\Delta\lambda \sim 0.04\lambda$). Principal features discussed above are seen, that is, enhanced daytime emission at 2.7 and 4.3 μm, enhanced nighttime emission at 9.6 μm, as well as diurnally independent emission at 5.3 and 15 μm, and long wavelengths. Also clearly seen is the region (11 to 13 μm) characterized by the absence of infrared airglow. Figures 13-26 and 13-27 give a summary of the nighttime and daytime radiances observed by SPIRE. At high altitudes, NO (5.3 μm) is the most intense emission observed while at low altitudes CO_2 (15 μm) and O_3 (9.6 μm) are the dominant radiators. It should be noted that in the absence of a data base in the 17–40 μm, the radiances in Figure 13-25 can be regarded as theoretical predictions. In the light of future work these predictions may have to be modified.

CHAPTER 13

REFERENCES

Baker, D.J., T. Conley, and A.T. Stair, "On the Altitude of OH Airglow," (Abstract), *Eos Trans. AGU*, **58:** 460, 1977.

Clough, S.A., F.X. Kneizys, L.S. Rothman, and W.O. Gallery, "Atmospheric Spectral Transmittance and Radiance: FASCODE1B." *SPIE*, **277:** 152–166, 1981.

Craig, R.A., and J.C. Gille, "Cooling of the Thermosphere by Atomic Oxygen," *J. Atmos. Sci.*, **26:** 205–209, 1969.

Degges, T.C., "Vibrationally Excited Nitric Oxide in the Upper Atmosphere," *Appl. Opt.*, **10:** 1856, 1971.

Degges, T.C. and H.J.P. Smith, "A High Altitude Infrared Radiance Model," AFCRL-TR-77-0271, ADA059242, 1977.

Fernando, R.P. and I.W.M. Smith, "Vibrational Relaxation of NO by Atomic Oxygen," *Chem. Phys. Lett.*, **66:** 218, 1979.

Gordiets, B.F., Y.N. Kulikov, M.N. Markov, and M.Ya. Marov, "The Influence of Radiative Cooling and Turbulence on the Heat Budget of the Thermosphere," (To be published in *J. Geophys. Res.*).

Gordiets, B.F., M.N. Markov, and L.A. Shelpen, "IR Radiation of the Upper Atmosphere," *Planet. Space Sci.*, **26:** 933–948, 1978.

James, T.C. and J.B. Kumer, "Fluorescence of CO_2 near 4.3 μm: Application to Daytime Limb Radiance Calculations," *J. Geophys. Res.*, **78:** 8320–8329, 1973.

Kockarts, G., "Nitric Oxide Cooling in the Terrestrial Thermosphere," *Geophys. Res. Lett.*, **7:** 137, 1980.

Krupenie, P.H., "Spectrum of Molecular Oxygen," *J. Phys. Chem. Ref. Data*, **1:** 423–534, 1972.

Kumer, J.B., "Analysis of Non-LTE CO_2 SPIRE Earthlimb Data," AIAA Reprint 81-0425, Aerospace Sciences Meeting, Jan 1981.

Kumer, J.B., A.T. Stair Jr., N. Wheeler, K.D. Baker, and D.J. Baker, "Evidence for an $OH^{\ddagger\nu v} \rightarrow N_2^{\nu v} \rightarrow CO_2^{\ddagger} \rightarrow CO_2 + h\nu(4.3\ \mu m)$ Mechanism for 4.3 μm Airglow," *J. Geophys. Res.*, **83:** 4743–4747, 1978.

Markov, M.N., "Molecular Radiation of the Upper Atmosphere in the 3–8 μm Spectral Region," *Appl. Opt.*, **8:** 887, 1969.

Markov, M.N., G.M. Grechko, A.A. Gubarev, Yu.S. Ivanov, and V.S. Petrov, "Infrared Spectrum of Nitric Oxide Obtained in the Upper Atmosphere at Middle Latitudes by the Orbiting Scientific Station Salyut 4," *Cosmic Res.*, **15:** 102, 1977.

McClatchey, R.A., W.S. Benedict, S.A. Clough, D.E. Burch, R.E. Calfee, K. Fox, L.S. Rothman, and J.S. Garing, "AFCRL Atmospheric Absorption Line Parameters Compilation," AFCRL-TR-73-0096, AD762904, 1973.

McEwan, M.J. and L.F. Phillips, *Chemistry of the Atmosphere*, The rate constants cited in the book for the deactivation of $O_2(^1\Delta_g)$ state by N_2, O_2 and O are $\leq 3 \times 10^{-21}$, 2×10^{-18} and $\leq 3 \times 10^{-16}$ cm molecule^{-1}s^{-1}, respectively, Wiley, New York, p. 140, 1975.

Mies, F.M., "Calculated Vibrational Transition Probabilities of OH ($\chi^2\pi$)," *J. Mol. Spec.*, **53:** 150, 1974.

Moreels, G., G. Megie, A. Vallance Jones, and R.L. Gattinger, "An Oxygen-Hydrogen Atmospheric Model and Its Application to the OH Emission Problem," *J. Atmos. Terr. Phys.*, **39:** 551, 1977.

Murphy, R.E., "Infrared emission of OH in the Fundamental and First Overtone Bands," *J. Chem. Phys.*, **54:** 4852–4859, 1971.

Nadile, R.M., N.B. Wheeler, A.T. Stair Jr., D.C. Frodsham, and L.C. Wyatt, "SPIRE-Spectral Infrared Experiment" *SPIE*, **24:** 18–24, 1977.

Offerman, D., and K.V. Grossman, "Spectrometric Measurement of Atomic Oxygen 63 μm Emission in the Thermosphere," *Geophys. Res. Lett.*, **5:** 381, 1978.

Ogawa, T., "Excitation Processes of Infrared Atmospheric Emissions," *Planet. Space Sci.*, **24:** 749–756, 1976.

Paulsen, D.E., R.E. Huffman, and J.C. Larabee, "Improved Photoionization Rates of $O_2(^1\Delta_g)$ in the D Region," *Radio Sci.*, **7:** 51, 1972.

Rawlins, W.T., G.E. Caledonia, and J.P. Kennealy, "Observation of Spectrally Resolved Infrared Chemiluminescence from Vibrationally Excited $O_3(\nu_3)$," *J. Geophys. Res.*, **86:** 5297, 1981.

Rothman, L.S. and L.D.G. Young, "Infrared Energy Levels and Intensities of Carbon Dioxide," *J. Quant. Spectrosc. Radiat. Transfer*, **25:** 505–524, 1981.

Rothman, L.S., A. Goldman, J.R. Gillis, R.R. Gamache, H.M. Pickett, R.L. Poynter, N. Husson, and A. Chedin, "AFGL Trace Gas Compilation: 1982 Version," *Appl. Opt.*, **22:** 1616, 1983a.

Rothman, L.S., R.R. Gamache, A. Barbe, A. Goldman, J.R. Gillis, L.R. Brown, R.A. Toth, J.-M. Flaud, and C. Camy-Peyret, "AFGL Atmospheric Absorption Line Parameters Compilation: 1982 Edition," *Appl. Opt.*, **22:** 2247, 1983b.

Selby, J.E.A., F.X. Kneizys, J.H. Chetwynd, Jr., and R.A. McClatchey, "Atmospheric Transmittance/Radiance: Computer Code LOWTRAN 4," AFGL TR-78-0053, ADA058643, 1978.

Sharma, R.D., "Near-Resonant Vibrational Energy Transfer Among Isotopes of CO_2," *Phys. Res.*, **177:** 102, 1969.

Sharma, R.D., and C.A. Brau, "Energy Transfer in Near-Resonant Molecular Collisions due to Long-Range Forces with Application to Transfer of Vibrational Energy from ν_3 Mode of CO_2 to N_2," *J. Chem. Phys.*, **50:** 924, 1969.

Sharma, R.D. and R.M. Nadile, "Carbon Dioxide (ν_2) Radiance Results Using a New Nonequilibrium Model," AIAA Reprint 81-0426, Aerospace Sciences Meeting, 1981.

Sharma, R.D., R.M. Nadile, A.T. Stair Jr., and W. Gallery, "Earthlimb Emission Analysis of Spectral Infrared Rocket Experiment (SPIRE) Data at 2.7 μm Micrometers," *SPIE*, **304:** 139, 1981.

Stair, A.T., Jr., J.C. Ulwick, D.J. Baker, C.L. Wyatt, and K.D. Baker, "Altitude Profiles of Infrared Radiance of O_3 (9.6 μm) and CO_2 (15 μm)," *Geophys. Res. Lett.*, **1:** 117, 1974.

Stair, A.T. Jr., J.C. Ulwick, K.D. Baker, and D.J. Baker, "Rocketborne Observations of Atmospheric Infrared Emissions in the Auroral Region," in *Atmosphere of Earth and the Planets*, edited by B.M. McCormac, D. Reidel, Dordrecht, Holland, p. 335–346, 1975.

Stair, A.T. Jr., R.M. Nadile, J.C. Ulwick, K.D. Baker, and D.J. Baker. "Infrared Measurements of Auroral, Airglow and the Upper Atmosphere," AIAA Reprint 81-0421, Aerospace Sciences Meeting, Jan 1981.

Stair, A.T. Jr., R.D. Sharma, R.M. Nadile, D.J. Baker, and W.F. Grieder, "Observations of Limb Radiance with Cryogenic Spectral Infrared Rocket Experiment (SPIRE)," *J. Geophys. Res.*, In Press, 1985.

Taylor, R.L. "Energy Transfer Processes in the Stratosphere," *Can. J. Chem.*, **52:** 1436–1451, 1974.

Ulwick, J.C., K.D. Baker, A.T. Stair Jr., W. Frings, R. Hennig, K.V. Grossman, and E.R. Hegblom, "Rocketborne Measurements of Atmospheric Infrared Fluxes," *J. Atmos. Terr. Phys.*, In Press, 1985.

Zacho . A.S., R.D. Sharma, R.M. Nadile, and A.T. Stair Jr., 'Inversion of SPIRE NO Data," AFGL TR-81-0325, ADA108866, 1981.

Chapter 14

Standard and Reference Atmospheres

K.S.W. Champion, A.E. Cole, and A.J. Kantor

14.1 STANDARD ATMOSPHERE

The World Meterological Organization (WMO) has defined the Standard Atmosphere as follows:

"A hypothetical vertical distribution of atmospheric temperature, pressure and density which by international agreement and for historical reasons, is roughly representative of year-round, midlatitude conditions. Typical usages are as a basis for pressure altimeter calibrations, aircraft performance calculations, aircraft and rocket design, ballistic tables and meteorological diagrams. The air is assumed to obey the perfect gas law and the hydrostatic equation which, taken together, relate temperature, pressure, and density with geopotential. Only one standard atmosphere should be specified at a particular time and this standard atmosphere must not be subjected to amendment except at intervals of many years."

Because of the interest of aerospace engineers and atmospheric scientists in conditions at much higher altitudes than those currently being considered by the WMO, members of the U.S. Committee on Extension to the Standard Atmosphere (COESA) agreed to add the following paragraph to the above definition in describing the U.S. Standard Atmosphere, 1976, which extends to 1000 km:

"The atmosphere shall also be considered to rotate with the earth, and be an average over the diurnal cycle, semiannual variation, and the range of conditions from active to quiet geomagnetic, and active to quiet sunspot conditions. Above the turbopause (about 110 km) generalized forms of the hydrostatic equations apply."

14.1.1 U.S. Standard Atmosphere, 1976

The U.S. Standard Atmosphere, 1976 is an idealized, steady-state representation of mean annual conditions of the earth's atmosphere from the surface to 1000 km at latitude 45°N, as it is assumed to exist during a period with moderate solar activity. The defining meteorological elements are sea-level temperature and pressure, and a temperature-height profile to 1000 km. The latter includes a linearly segmented temperature-height profile in geopotential heights between the surface and 86 km. The first 32 km portion of this profile is identical to the International Civil Aviation Organization (ICAO) Standard Atmosphere. The mean annual midlatitude vertical temperature profile for the portion between 32 and 55 km was developed from both routine radiosonde observations and meterological rocketsonde observations. The portion between 55 and 86 km is based primarily on measurements derived from grenade, pitot static tube and falling sphere experiments.

The temperature-height profile between 86 and 1000 km is expressed in geometric altitude. It is not a series of linear functions as it is at altitudes below 86 km. Rather, it is defined in terms of four successive functions chosen not only to provide a reasonable approximation to observations, but also to yield a continuous first derivative with respect to altitude over the entire regime. Observational data of various kinds provided the basis for determining the four segments of the temperature-height profile between 86 and 1000 km. The observed temperatures at altitudes between 110 and 120 km were particularly important in imposing limits on the temperature-height structure for that region, while the observed density at 150 km and above strongly influenced the selection of both the temperature and the extent of the low-temperature isothermal layer immediately above 86 km. Values of the thermodynamic properties for the region between 140 and 1000 km were determined almost exclusively from satellite related observations and radar incoherent scatter techniques. For altitudes between 90 and 140 km, however, there is only a very limited amount of data from rocket soundings and incoherent scatter observations, and almost none from satellite observations.

The 1976 Standard, like its predecessors, uses the following sea-level values of temperature, pressure, and density which have been standard for decades:

Temperature	288.15 K, 15°C or 59°F
Pressure	1013.25 mb, 760 mm of Hg, or 29.92 in. of Hg
Density	1225.00 g/cm^3 or 0.076474 lbs/ft^3.

In spite of the various independent data sets upon which the several temperature-height segments are based, it is desirable, for purposes of mathematical reproducibility of the tables of the 1976 Standard, to express the temperature in

a series of consecutive height functions from the surface to 1000 km, with the expression for each successive function depending upon the end-point value of the preceding function as well as upon certain terms and coefficients peculiar to the related altitude interval. This total temperature-height profile applied to the fundamental continuity models (hydrostatic equation and the equation of motion) together with all the ancillary required constants, coefficients, and functions, defines the U.S. Standard Atmosphere, 1976.

14.1.1.1 Equilibrium Assumptions. The air is assumed to be dry, and at altitudes below 86 km the atmosphere is assumed to be homogeneously mixed with a relative volume composition leading to a constant mean molecular weight M_0. The air is treated as if it were a perfect gas, and total pressure P, temperature T, and total density ρ at any point in the atmosphere are related by the equation of state, that is, the perfect gas law, one form of which is

$$P = \frac{\rho \cdot R^* \cdot T}{M_0}, \tag{14.1}$$

where R^* is the universal gas constant. An alternate form of the equation of state in terms of the total number density N and the Avogadro constant N_A is

$$P = \frac{N \cdot R^* \cdot T}{N_A}. \tag{14.2}$$

This form represents the summation of P_i, the partial pressures of the individual gas species, where P_i is related to n_i the number density of the ith gas species in the following expression:

$$P_i = n_i \cdot k \cdot T, \tag{14.3}$$

where k is the Boltzmann constant.

Within the altitude region of complete mixing, the atmosphere is assumed to be in hydrostatic equilibrium and to be horizontally stratified so that dP, the differential of pressure, is related to dZ, the differential of geometric altitude, by the relationship

$$dP = -g \cdot \rho \cdot dZ, \tag{14.4}$$

where g is the altitude-dependent acceleration of gravity. The elimination of ρ between Equations (14.1) and (14.4) yields another form of the hydrostatic equation, which serves as the basis for the low-altitude pressure calculation:

$$d \ln P = \frac{dP}{P} = \frac{-g \cdot M_i}{R^* \cdot T} \cdot dZ. \tag{14.5}$$

Above 86 km the hydrostatic equilibrium of the atmosphere gradually breaks down as diffusion and vertical transport of individual gas species lead to the need for a dynamically oriented model including diffusive separation. Under these conditions, Colegrove et al. [1965] found it convenient to express the variations with altitude of the atmospheric number density in terms of the vertical component of the flux of the molecules of individual gas species. In terms of the ith gas species, this expression is

$$n_i \cdot v_i + D_i \cdot \left(\frac{dn_i}{dZ} + \frac{n_i \cdot (1 + \alpha_i)}{T} \cdot \frac{dT}{dZ} + \frac{g \cdot n_i \cdot M_i}{R^* \cdot T} \right) + K \cdot \left(\frac{dn_i}{dZ} + \frac{n_i}{T} \cdot \frac{dT}{dZ} + \frac{g \cdot n_i \cdot M}{R^* \cdot T} \right) = 0 \tag{14.6}$$

where

n_i = number density of the ith gas species,

v_i = the vertical transport velocity of the ith species,

D_i = the altitude-dependent, molecular-diffusion coefficient of the ith species diffusing through N_2,

α_i = the thermal-diffusion coefficient of the ith species,

M_i = the molecular weight of the ith species,

M = the molecular weight of the gas through which the the ith species is diffusing,

and

K = the altitude-dependent, eddy-diffusion coefficient.

14.1.1.2 Atmospheric Properties. Because of limited space, the 1976 Standard Atmosphere Tables are not included in this handbook. For detailed tables consult *U.S. Standard Atmosphere, 1976* [COESA]. The temperature-height profile of the 1976 Standard for altitudes up to 86 km is shown in Figure 14-1. The profile is the same as the 1962 Standard for altitudes up to 51 km. From 51 to 70 km it is colder, and from 70 to 86 km it is warmer than the 1962 Standard. Temperature profiles to 500 km for the 1962 and 1976 standard and pressure and density profiles from the surface to 1000 km for the 1976 Standard are presented in Figures 14-2 and 14-3. The equations used for computing other atmospheric properties tabulated in the 1976 Standard are discussed below.

Mean Particle Speed

The mean particle speed V is the arithmetic average of the speeds of all air particles in the volume element being considered. All particles are considered to be neutral. For

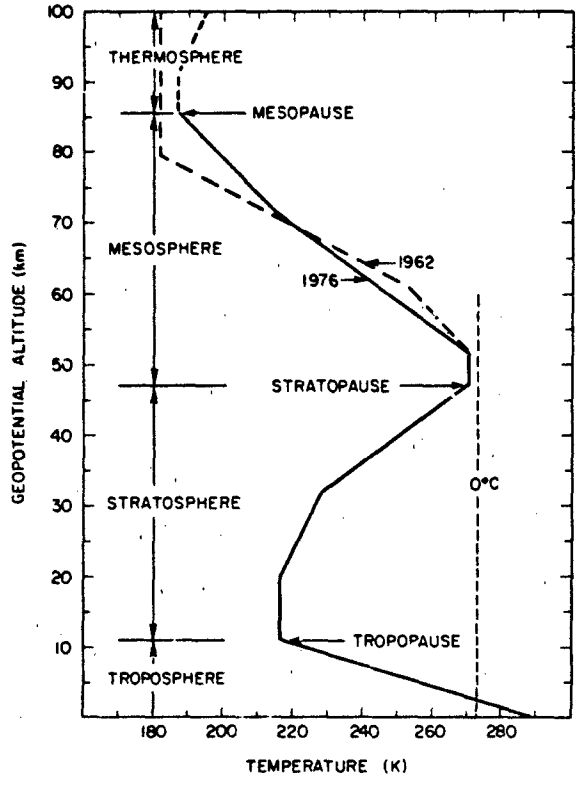


Figure 14-1. Temperature-height profile for U.S. Standard Atmosphere 0-86 km.

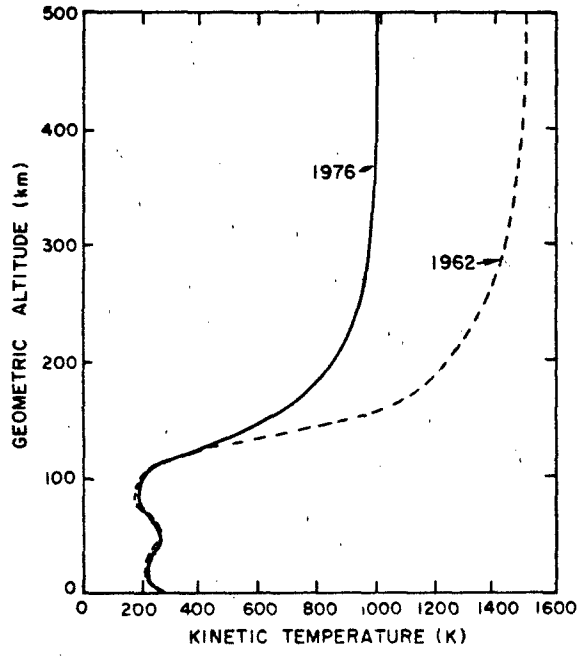


Figure 14-2. Kinetic temperature as a function of geometric altitude.

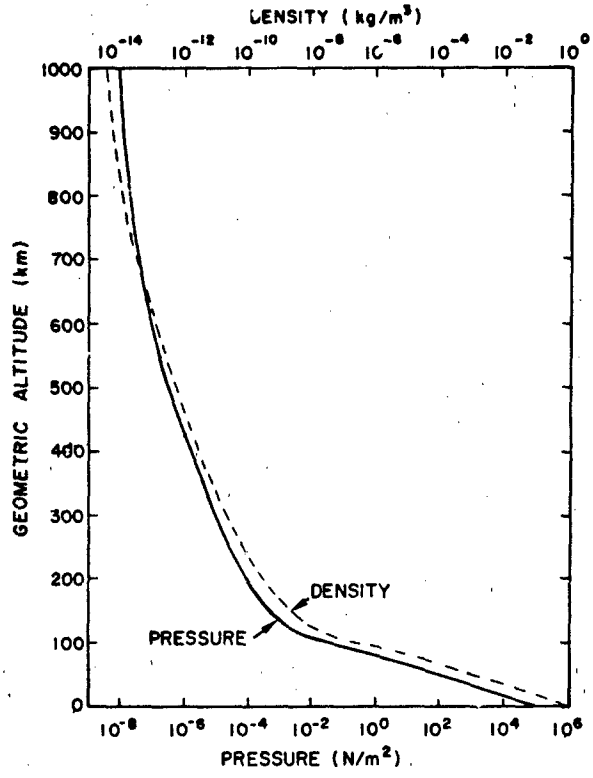


Figure 14-3. Total pressure and mass density as a function of geometric altitude.

a valid average to occur, there must be a sufficient number of particles involved to represent mean conditions. Pressure and temperature gradients within the volume must be negligible. The analytical expression for V is closely related to that for the speed of sound and is proportional to the ratio of T/M.

$$V = \left[\frac{8 \cdot R^* \cdot T}{\pi \cdot M}\right]^{1/2} \qquad (14.7)$$

where M is the mean molecular weight at a specific altitude. A graph of V as a function of geometric altitude is given in Figure 14-4.

Mean Collision Frequency

The mean collision frequency, $\nu(s^{-1})$, between the hypothetical mean neutral particles is a function of pressure P, temperature T, and molecular weight M, as well as of an effective mean collision diameter σ. The expression used for computing this property is

$$\nu = 4\,\sigma^2\,N_A\,P\left(\frac{\pi}{M \cdot R^* \cdot T}\right)^{1/2}, \qquad (14.8)$$

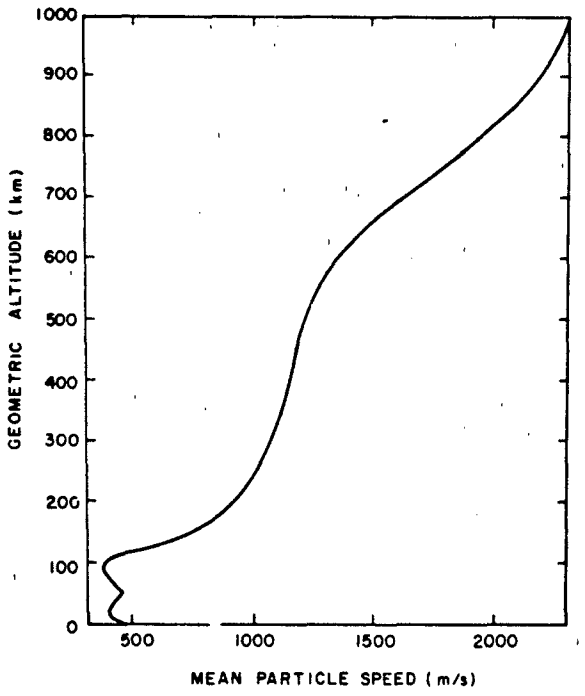


Figure 14-4. Mean particle speed as a function of geometric altitude.

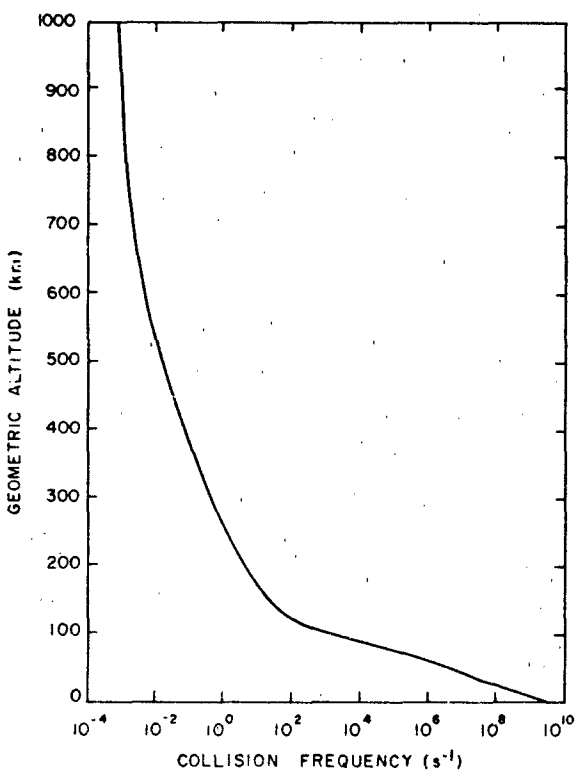


Figure 14-5. Collision frequency as a function of geometric altitude.

where π = the ratio of the circumference to the diameter of a circle, that is 3.14159 (dimensionless) and σ = the effective collision diameter, 3.65×10^{-10}m. Figure 14-5 shows the variation in mean collision frequency with altitude.

Mean Free Path

The mean free path L is equal to the ratio of mean particle speed to mean collision frequency,

$$L = \frac{V}{\nu}. \qquad (14.9)$$

This quantity represents the average travel distance between collisions of the hypothetical mean neutral particles in any direction along which the total change of number density is small over the mean free path length. Thus at the altitude of 130 km where L is less than 10 m, the concept is applicable in all directions. At altitudes above 175 km, where L is the order of 100 m or greater, the concept of mean free path, at best, becomes blurred and must be used with caution. The values of L versus altitude are plotted in Figure 14-6.

Mean Molecular Weight

The mean molecular weight M of a mixture of gases is by definition

$$M = \frac{\sum(n_i M_i)}{\sum n_i} \qquad (14.10)$$

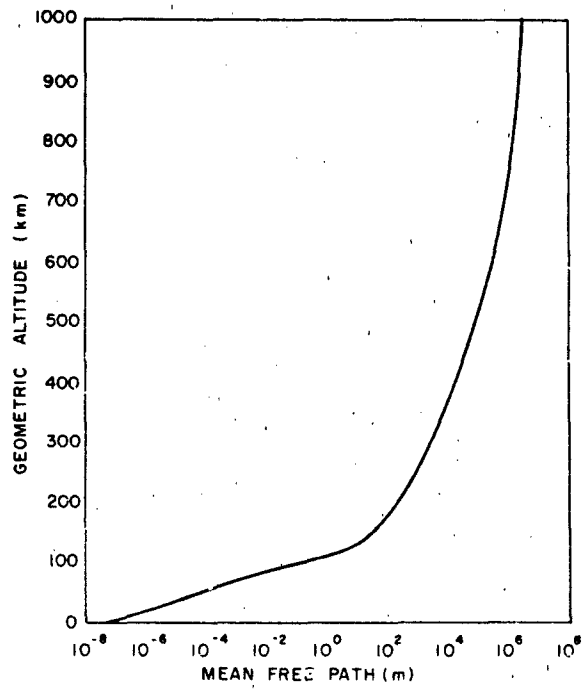


Figure 14-6. Mean free path as a function of geometric altitude.

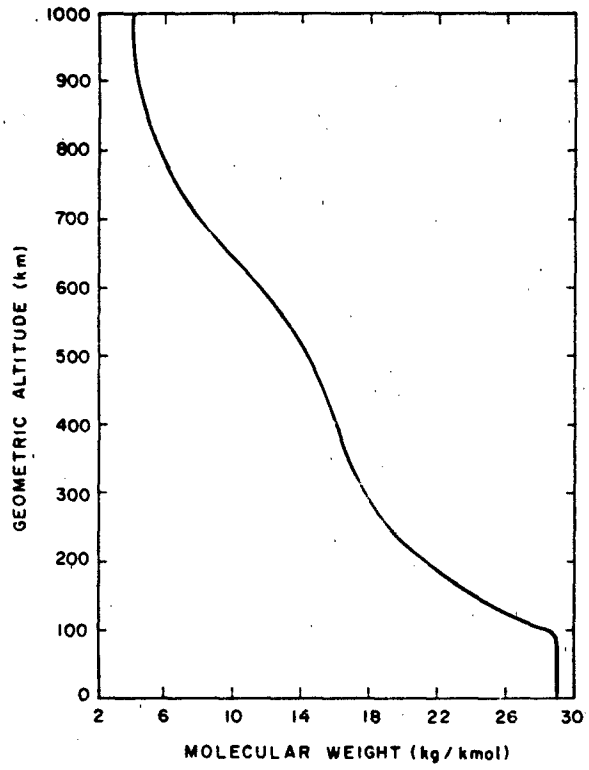


Figure 14-7. Mean molecular weight as a function of geometric altitude.

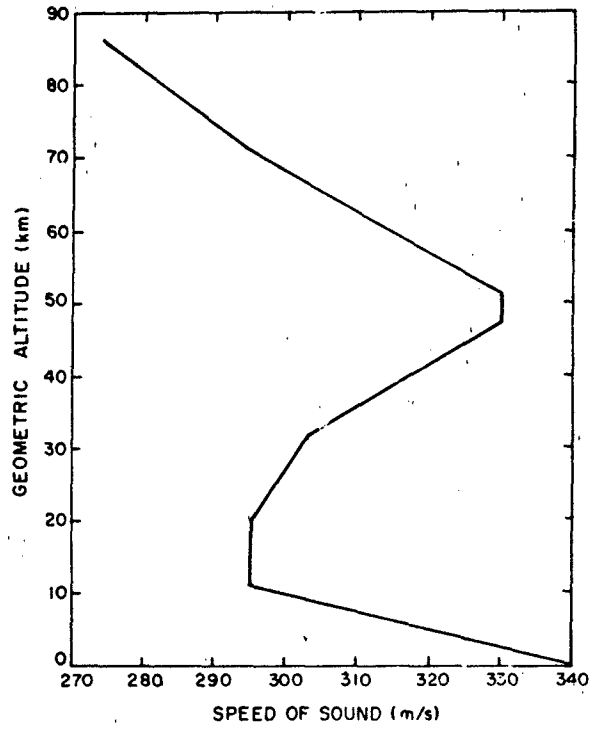


Figure 14-8. Speed of sound as a function of geometric altitude.

where n_i and M_i are the number density and molecular weight, respectively, of the ith gas species. In that part of the atmosphere between the surface and 86 km altitude, mixing is dominant, and the effect of diffusion and photochemical processes upon M is negligible. Consequently, M remains constant at its sea level value, $M_0 = 28.9644$ kg/kmol. The variation of the mean molecular weight with altitude is shown in Figure 14-7.

Speed of Sound

The expression adopted for the speed of sound C_s is

$$C_s = \left(\frac{\gamma \cdot R^* \cdot T}{M}\right)^{1/2}, \qquad (14.11)$$

where γ is the ratio of specific heat of air at constant pressure to that at constant volume and is taken to be 1.40 exact (dimensionless). The above equation for speed of sound applies only when the sound wave is a small perturbation on the ambient condition. Calculated values for C_s have been found to vary slightly from experimentally determined values. The limitation of the concept of speed of sound due to extreme attenuation is also of concern. The attenuation that exists at sea level for high frequencies applies to successively lower frequencies as atmospheric pressure decreases or as the mean free path increases. For this reason, the concept of speed of sound (except for frequencies approaching zero) progressively loses its range of applicability at high altitudes. Hence the main tables listing the values of speed of sound are terminated at 86 km. The variation of the speed of sound with altitude is shown in Figure 14-8.

Dynamic Viscosity

The coefficient of dynamic viscosity μ(N s/m^2) is defined as a coefficient of internal friction developed where gas regions move adjacent to each other at different velocities:

$$\mu = \frac{\beta \cdot T^{3/2}}{T + S} \qquad (14.12)$$

where β is a constant equal to 1.458×10^{-6} kgs^{-1}m^{-1}K$^{-1/2}$ and S is Sutherland's constant, equal to 110.4 K. The equation fails for conditions of very high and very low temperatures and under conditions occurring at great altitudes. Variation of dynamic viscosity with altitude is shown graphically in Figure 14-9.

Kinematic Viscosity

Kinematic viscosity η is defined as the ratio of the dynamic viscosity of a gas to the density of that gas, ρ.

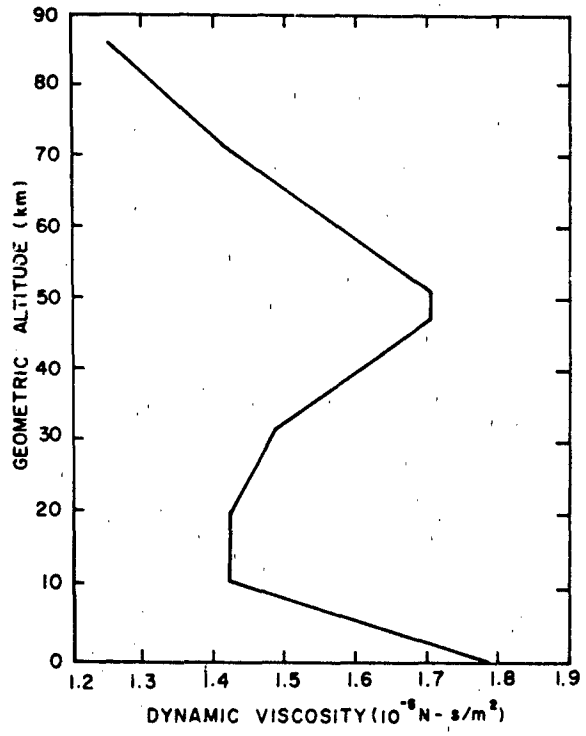


Figure 14-9. Dynamic viscosity as a function of geometric altitude.

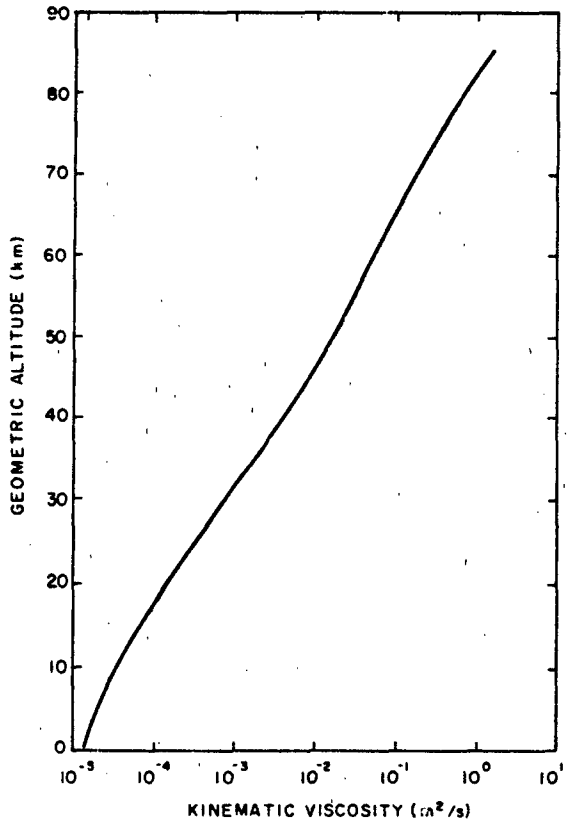


Figure 14-10. Kinematic viscosity as a function of geometric altitude.

$$\eta = \frac{\mu}{\rho} \qquad (14.13)$$

Limitations of this equation are comparable to those associated with dynamic viscosity. A graphical representation of the variation of kinematic viscosity with altitude is shown in Figure 14-10.

Coefficient of Thermal Conductivity

The empirical expression adopted for developing tabular values of the coefficient of thermal conductivity k_t for altitudes up to 86 km is as follows:

$$k_t = \frac{2.65019 \times 10^{-3} T^{3/2}}{T + 245.4 \times 10^{-(12/T)}} \qquad (14.14)$$

This expression differs from that used in the U.S. Standard Atmosphere, 1962 in that the numerical constant has been adjusted to accommodate a conversion of the related energy unit from the temperature-dependent kilogram calorie to the invariant joule. Thus, the values of k_t in units of J m^{-1}s^{-1}K^{-1} or W m^{-1}K^{-1} are greater than the values of k_t in units of kcal m^{-1}s^{-1}K^{-1} by a factor of exactly 4.19002×10^3. Kinetic theory determinations of thermal conductivity of some monatomic gases agree well with observation. The variation of the coefficient of thermal conductivity with altitude is shown in Figure 14-11.

14.1.1.3 Atmospheric Structure. Temperature variation with altitude is the most pronounced feature of the earth's atmosphere and the various layers and shells can be described more easily by temperature than by any other element. This feature is illustrated in Figure 14-1 with the temperature height profile for the 1976 Standard up to 86 km. Usually, temperature decreases with increasing altitude in the lowermost several kilometers. Except for the first kilometer above the earth's surface, the normal lapse rate is about 6.5 K/km in the troposphere, which extends from the surface to the tropopause.

The tropopause is the atmospheric surface at which the temperature decrease with altitude either stops abruptly or the lapse rate becomes less than some arbitrary low value. The height of the tropopause varies with latitude, season, and weather situation; in general it is lowest (8 to 10 km) in arctic regions in winter, and highest (16 to 18 km) in the tropical and equatorial regions. Above the tropopause, temperatures increase with altitude, slowly at first, then more

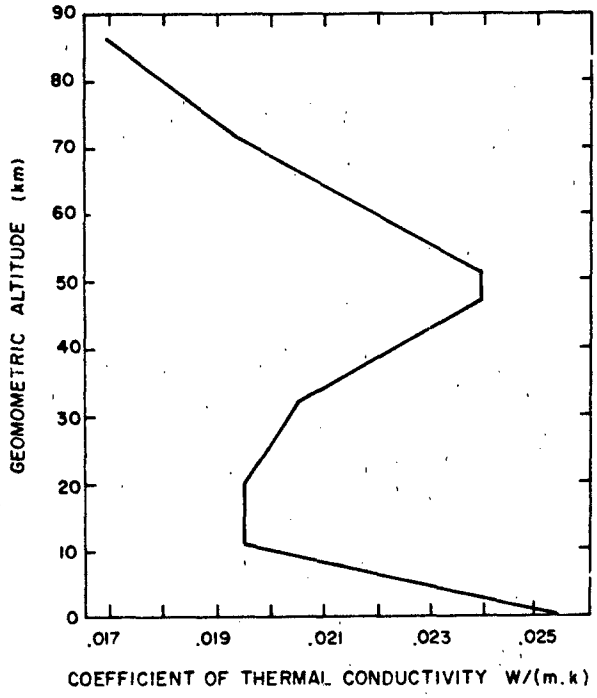


Figure 14-11. Coefficient of thermal conductivity as a function of geometric altitude.

rapidly up to about 50 km, where the average temperature is within a few degrees of the freezing point. This surface of maximum temperature is designated the stratopause, and the entire layer below it down to the tropopause is the stratosphere.

The layer above the stratopause in which the temperature normally decreases with increasing altitude is called the mesosphere. It ends at the mesopause at 80 to 90 km, where the lowest temperatures of the atmosphere are found (135 to 190 K). Above the mesopause lies the thermosphere; its upper limit varies with solar and geophysical conditions. The temperature in the thermosphere increases rapidly with altitude up to about 200 km, then, more slowly, becoming nearly isothermal above 300 to 500 km. The temperature above 200 km varies widely according to the degree of solar activity; it is about 600 K when the sun is quiet and possibly 2000 K during sunspot maxima. The exosphere, sometimes called region of escape, is the outermost, or topmost portion of the atmosphere. Its lower boundary is the critical level of escape, estimated to be 500 to 1000 km above the earth's surface. The air density is so low that the mean free path of individual particles depends upon their direction with respect to the local vertical, being greatest for upward moving particles. It is only from the exosphere that atmospheric gases can escape into outer space. In this region temperature can no longer be defined in the usual way.

14.2 REFERENCE ATMOSPHERES

14.2.1 Air Force Reference Atmospheres for Altitudes Below 90 Km

The Reference Atmospheres presented in this section [Cole and Kantor, 1978] were developed to provide Air Force engineers and designers of aerospace systems with up to date information on the seasonal, latitudinal, and longitudinal variations in the vertical distributions of the thermodynamic properties (pressure, temperature, and density) of the atmosphere in the Northern Hemisphere for altitudes up to 90 km. Special models are included to illustrate the magnitude of the changes in the vertical distributions of these atmospheric properties during extreme winter warmings and coolings of the stratosphere and mesophere in arctic and subarctic regions.

14.2.1.1 Basic Assumptions and Formulas. The atmospheric models are defined by temperature-altitude profiles in which temperatures change linearly with respect to geopotential altitude. It is assumed that the air is dry, is in hydrostatic equilibrium, and behaves as a perfect gas. The molecular weight of air at sea level, 28.9644 kg/kmol, is assumed to be constant to 90 km. Actually, dissociation of molecular oxygen begins to take place near 80 km and molecular weight starts decreasing slowly with height. Consequently, the temperatures given in the tables for altitudes above 80 km are slightly but not significantly larger (less than one degree K) than the ambient kinetic temperature.

Numerical values for the various thermodynamic and physical constants used in computing the tables of atmospheric properties for these Reference Atmospheres are identical to those used in the preparation of the *U.S. Standard Atmosphere, 1976*, with two exceptions. Surface conditions for the atmospheres are based on mean monthly sea-level values of pressure and temperature for the appropriate latitude rather than on standard conditions. The accelerations due to gravity at sea level for the latitudes were obtained from the following expression by Lambert [List, 1968] in which gravity, g, varies with latitude ϕ:

$$g_\phi = 9.780356\,(1. + 0.0052885 \sin^2 \phi - 0.0000059 \sin^2 2\phi). \qquad (14.15)$$

The relationship between geopotential altitude and geometric altitude is the same as that used for the U.S. Standard Atmosphere, 1976:

$$H = \left(\frac{r_\phi Z}{r_\phi + Z}\right)\left(\frac{g_\phi}{G}\right), \qquad (14.16)$$

where H is the geopotential altitude in geopotential meters (m'), Z is the geometric altitude, r_ϕ is the effective earth

Table 14-1. Sea level acceleration of gravity and the effective radius, r_ϕ, of the earth for each latitude.

Latitude ϕ	Sea Level Gravity g_ϕ (m/s^2)	Effective Earth Radius r_ϕ, (km)
0	9.78036	6334.984
15	9.78381	6337.838
30	9.79324	6345.653
45	9.80665	6356.766
60	9.81911	6367.103
75	9.82860	6374.972
90	9.83208	6377.862

radius, g_ϕ is the sea-level value for acceleration of gravity at a specific latitude ϕ as given by Lambert's equation, and G is the unit geopotential set equal to 9.80665 $m^2 s^{-2} (m')^{-1}$. Values of r_ϕ and g_ϕ are given in Table 14-1.

14.2.1.2 Data. Initial sea-level pressures for each atmosphere were taken from mean monthly sea-level charts for the northern hemisphere. Temperature-height profiles for altitudes up to 30 km were obtained for specific latitudes and months by giving equal weight to radiosonde temperatures interpolated for each 10° of longitude.

Temperature distributions between 30 and 55 km are based on Meterological Rocket Network (MRN) observations taken during the period 1964 through 1976 and temperatures from sets of 5, 2, and 0.4 mb constant pressure maps prepared by the Upper Air Branch of the National Weather Service for the years 1964 through 1968 and from January 1972 through June 1974. The temperature distributions between 55 and 90 km are based primarily on temperatures derived from grenade, falling sphere, and pressure gauge experiments. A detailed description of the analysis of these data and the development of the monthly models is contained in Air Force Surveys in Geophysics No. 382, Air Force Reference Atmospheres [Cole, 1978].

14.2.1.3 Seasonal and Latitudinal Models. Sets of mean monthly Reference Atmospheres for altitudes up to 90 km, [Cole, 1978] have been developed for 15° intervals of latitude, including the equator and north pole, to provide information on the seasonal and latitudinal variations in the thermodynamic properties of the atmosphere. Due to limitations on the length of individual presentations in the handbook, only the properties of the January, April, July, and October models are presented by geometric altitudes in Table 14-2.

The distribution of atmospheric temperature between the equator and north pole, based on the January and July Reference Atmospheres for latitudes 0°, 15°, 30°, 45°, 60°, 75° and 90°N, is shown in Figure 14-12 for geometric altitudes

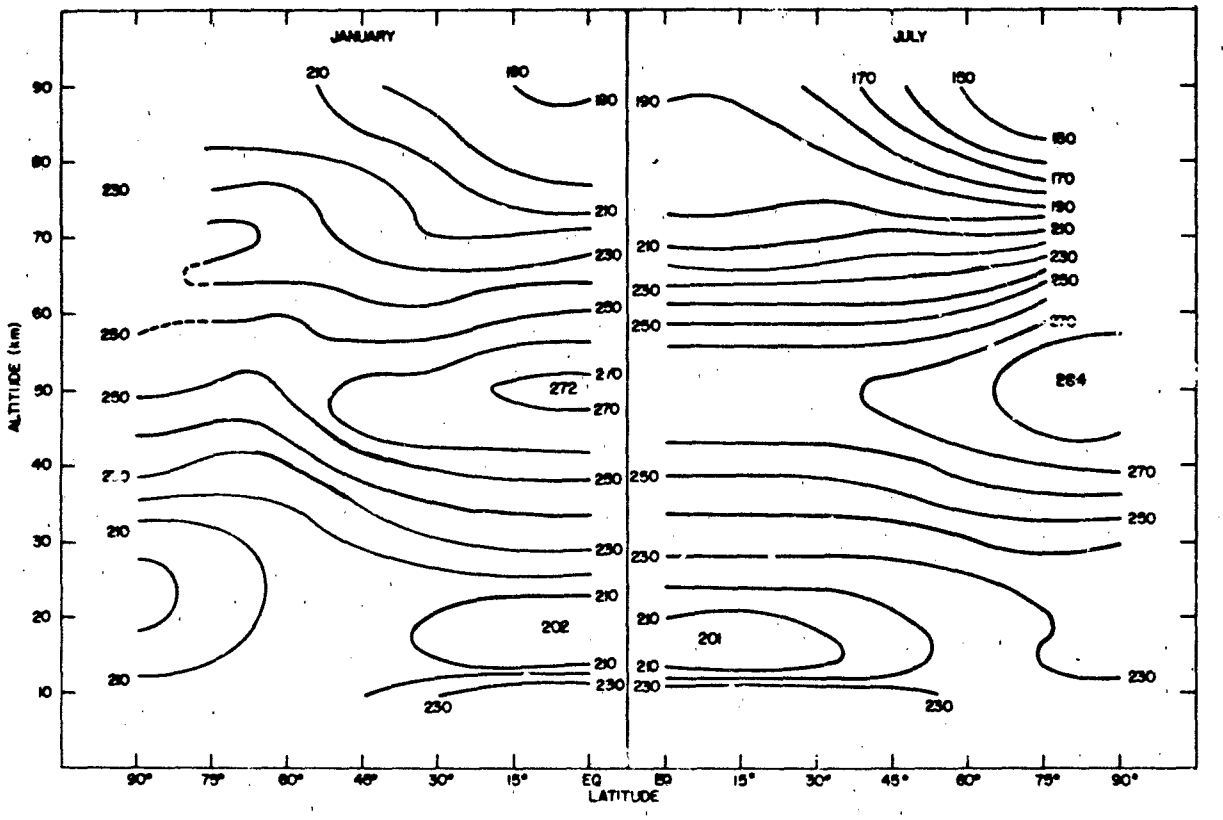


Figure 14-12. Latitudinal temperature-height cross-sections of monthly temperatures for January and July.

Table 14-2a. Reference Atmosphere Temp (K)—January

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N
0.000	299.15	296.65	287.15	272.15	257.15	248.15	237.15
2.000	288.78	287.27	281.16	265.15	255.94	251.37	245.88
4.000	278.41	277.90	268.20	255.66	247.73	240.36	236.61
6.000	268.06	266.28	255.24	243.68	234.13	229.35	226.60
8.000	256.40	252.35	242.28	231.71	220.54	218.35	216.59
10.000	240.49	238.42	229.34	219.74	217.15	214.14	212.64
12.000	224.58	224.50	216.40	218.65	217.15	212.14	210.64
14.000	208.69	210.59	211.08	217.66	217.15	210.14	208.63
16.000	197.34	199.46	205.91	216.67	216.66	208.15	204.64
18.000	195.03	195.15	203.15	215.67	215.66	206.16	200.65
20.000	201.37	203.65	207.73	215.15	214.66	204.16	196.67
22.000	207.71	212.57	212.49	215.15	213.67	202.93	196.65
24.000	214.05	217.22	217.25	215.15	212.68	204.12	196.65
26.000	220.38	221.58	221.79	215.15	212.98	205.32	196.65
28.000	226.71	225.93	225 36	217.08	214.77	206.51	196.65
30.000	232.70	230.28	228.92	221.44	216.55	209.00	203.04
32.000	236.65	234.63	233.13	225.79	218.34	212.97	209.59
34.000	240.60	238.97	237.67	230.15	220.12	216.94	216.14
36.000	244.55	244.29	242.24	235.54	222.07	220.90	222.69
38.000	249.66	250.20	246.98	241.47	224.25	224.86	229.23
40.000	254.98	256.12	251.71	247.40	226.43	228.82	232.98
42.000	260.30	262.02	256.45	253.32	232.87	232.78	236.55
44.000	265.61	265.70	261.17	259.24	239.78	236.74	240.11
46.000	270.92	268.84	265.65	262.82	245.16	240.69	243.67
48.000	272.15	271.15	265.65	264.65	247.93	244.64	247.22
50.000	272.15	271.15	265.65	264.65	250.69	248.59	250.78
52.000	269.12	268.95	261.73	261.41	251.15	252.53	252.15
54.000	265.01	264.43	256.62	255.51	251.15	255.15	252.15
56.000	260.71	259.92	251.65	250.83	250.93	254.06	
58.000	255.61	254.49	247.13	247.29	250.54	250.72	
60.000	250.52	249.01	242.62	243.76	250.11	247.37	
62.000	245.44	243.53	238.12	240.22	245.59	244.03	
64.000	240.35	238.05	233.61	236.69	241.08	240.69	
66.000	235.27	232.58	229.11	233.17	238.65	238.52	
68.000	230.19	226.62	224.62	229.64	238.65	240.09	
70.000	225.12	228.57	220.12	226.12	238.65	241.66	
72.000	217.25	214.88	218.15	225.65	238.07	239.43	
74.000	206.72	210.20	218.15	225.65	234.16	235.52	
76.000	196.20	205.53	216.35	223.44	230.25	231.60	
78.000	195.65	200.85	212.84	219.54	226.34	227.69	
80.000	195.65	198.15	209.33	215.63	222.43	223.77	
82.000	195.65	198.15	205.83	212.50	218.53	219.87	
84.000	195.65	198.15	202.32	210.16	214.65	215.96	
86.000	194.59	196.97	198.82	207.82	214.65	214.15	
88.000	191.29	193.48	195.33	205.49	214.65	214.15	
90.000	187.99	189.99	191.83	203.15	214.65	214.15	

Table 14-2b. Reference Atmosphere Temp (K)—April

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N
0.000	300.15	297.15	292.15	279.15	269.15	255.15	248.65
2.000	288.98	286.97	282.16	273.15	263.14	254.88	252.12
4.000	277.82	276.81	272.19	263.66	257.13	244.86	242.11
6.000	266.67	264.78	260.24	250.68	243.13	233.85	232.10
8.000	254.63	250.85	246.29	237.71	229.14	222.85	222.09
10.000	239.71	236.92	232.35	224.75	222.15	222.56	221.91
12.000	224.81	223.00	218.42	218.15	222.15	224.15	224.15
14.000	209.91	209.09	213.47	218.15	222.15	224.15	224.15
16.000	198.56	200.02	208.69	218.15	222.15	224.15	224.15
18.000	196.12	196.15	206.15	218.15	222.15	224.15	224.15
20.000	202.66	204.84	210.92	218.15	222.15	224.15	224.15
22.000	209.20	213.96	215.88	218.15	222.15	224.15	224.15
24.000	215.74	221.02	220.46	220.26	222.15	224.15	224.66
26.000	222.27	225.38	225.02	223.24	223.17	224.15	226.85
28.000	228.39	229.73	229.57	226.21	225.35	224.15	229.04
30.000	233.93	234.08	234.13	229.54	227.53	227.61	231.23
32.000	239.46	238.43	238.68	234.49	231.97	231.97	233.41
34.000	244.98	242.77	243 05	239.44	236.53	234.76	236.16
36.000	250.51	247.24	247.39	244.54	241.59	238.33	241.12
38.000	256.03	252.77	251.73	250.47	247.33	241.89	246.08
40.000	261.54	258.28	256.07	256.40	253.07	248.03	251.03
42.000	267.06	263.80	260.41	261.09	258.80	254.36	255.98
44.000	269.72	268.24	262.77	265.04	264.53	260.69	260.93
46.000	271.69	270.21	266.92	268.08	269.24	267.02	265,87
48.000	272.15	272.15	269.15	271.65	272.20	273.34	270.15
50.000	272.15	272.15	269.15	271.65	274.15	274.15	270.15
52.000	269.27	269.52	267.98	269.06	274.15	274.15	270.15
54.000	265.35	265.99	263.46	264.34	270.80	272.01	270.15
56.000	261.42	261.78	258.95	259.62	264.90	265.71	
58.000	256 93	254.72	253.74	254.31	259.00	259.41	
60.000	249.49	247.67	248.25	249.01	253.10	253.11	
62.000	242.05	240.62	242.76	243.71	247.21	246.82	
64.000	234.62	233.58	237.28	238.42	241.31	240.53	
66.000	227.20	226.54	231.80	233.13	236.24	234.24	
68.000	219.76	219.51	226.20	229.09	232.12	230.47	
70.000	211.95	212.48	220.53	225.53	228.01	226.75	
72.000	204.15	205.86	214.87	221.26	223.90	223.02	
74.000	196.35	199.62	209.21	217.35	219.79	219.30	
76.000	188.55	193.39	203.55	213.44	215.33	215.58	
78.000	190.73	193.15	200.17	209.54	210.25	209.96	
80.000	193.45	193.15	199.20	205.63	205.17	204.09	
82.000	196.40	195.44	198.22	201.11	200.09	198.23	
84.000	199.70	197.97	197.25	196.13	195.02	192.37	
86.000	203 00	200.49	196.28	191.27	189.95	186.51	
88.000	206.30	203.02	195.31	189.65	184.88	180.66	
90.000	209.60	205.54	194.34	189.65	179.81	176.15	

Table 14-2c. Reference Atmosphere Temp (K)—July

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N
0.000	298.65	299.65	301.15	294.15	287.15	275.65	273.15
2.000	287.68	288.87	288.16	285.15	276.34	273.62	271.12
4.000	276.72	278.11	277.19	273.16	265.53	260.60	260.09
6.000	265.77	267.35	266.22	261.18	253.13	247.59	246.88
8.000	254.82	255.38	252.29	248.21	239.14	234.59	233.67
10.000	238.99	238.46	238.35	235.25	225.17	228.55	228.76
12.000	223.08	223.55	224.42	222.29	225.15	229.35	231.15
14.000	207.19	207.65	210.49	215.65	225.15	230.15	231.15
16.000	198.59	200.30	203.15	215.65	225.15	230.15	231.15
18.000	201.05	202.87	207.38	216.78	225.15	230.15	231.15
20.000	207.99	208.82	211.75	219.17	225.15	230.15	231.15
22.000	214.92	214.76	216.11	221.55	225.15	230.15	231.15
24.000	219.97	220.71	220.47	223.75	227.03	230.15	231.15
26.000	224.13	224.81	224.83	225.93	231.00	232.62	233.48
28.000	228.28	228.77	229.19	229.16	234.97	237.79	236.67
30.000	232.43	232.72	233.55	233.72	239.18	242.95	239.84
32.000	236.58	236.68	237.90	238.28	244.33	248.12	243.02
34.000	240.72	240.63	237.90	238.28	244.33	248.12	249.82
36.000	244.87	244.57	246.59	247.38	254.64	258.64	256.76
38.000	249.01	248.59	251.45	251.93	259.78	263.29	263.70
40.000	253.14	252.78	256.38	256.47	264.83	268.24	270.63
42.000	257.28	256.86	261.31	261.01	270.07	273 19	275.90
44.000	261.84	260.99	264.44	265.55	275.20	278.14	278.86
46.000	266.76	265.13	267.19	270.09	278.13	283.08	281.83
48.000	269.15	268.15	269.15	273.15	279.15	283.65	283.65
50.000	269.15	268.15	269.15	273.15	279.15	283.65	283.65
52.000	266.56	266.71	267.34	272.91	277.60	283.65	283.65
54 000	263.03	263.77	264.98	267.01	272.87	283.23	283.65
56.000	258.78	258.71	261.50	260.61	268.15	278.30	
58.000	251.33	251.07	254.04	253.73	263.43	273.38	
60.000	243.89	243.43	246.59	246.86	258.71	268.46	
62.000	236.45	235.80	239.14	239.99	252.79	263.41	
64.000	229.02	228.17	231.70	233.13	243.17	252.01	
66.000	221.57	220.54	224.27	226.27	233.56	240.62	
68.000	213.76	212.92	216.83	219.41	223.95	229.23	
70.000	205.95	205.31	209.41	212.56	214.34	217.86	
72.000	198.15	201.81	204.50	205.72	204.75	206.49	
74.000	192.59	199.47	201.58	198.88	195.16	195.12	
76.000	194.54	197.65	198.65	192.70	185 57	183.76	
78.000	196.49	197.65	195.72	188.01	175.99	172.41	
80.000	198.44	197.65	192.80	183.33	166.42	161.07	
82.000	199.65	197.65	189.88	178.65	157.09	149.73	
84.000	199.65	197.65	186.96	173.97	155.13	147.01	
86.000	197.85	193.52	184.04	169.30	153.18	146.23	
88.000	192.22	188.66	181.13	164.63	151.23	145.45	
90.000	186.59	183.81	178.22	164.15	149.29	144.67	

Table 14-2c. Reference Atmosphere Temp (K)—October

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N
0.000	299.15	298.15	293.65	284.15	275.15	262.65	252.65
2.000	288.38	288.57	283.66	274.15	267.14	259.44	254.37
4.000	277.62	277.81	273.69	263.16	259.13	248.85	243.35
6.000	265.69	266.66	261.74	251.18	246.13	235.84	232.34
8.000	252.55	254.56	247.79	239.21	233.14	222.84	221.34
10.000	239.43	239.44	233.85	227.24	220.16	219.65	220.16
12.000	223.58	224.33	219.92	215.28	220.15	219.65	222.14
14.000	207.69	209.23	211.85	215.15	220.15	219.65	220.14
16.000	196.77	197.92	203.89	215.15	220.15	219.65	218.14
18.000	199.19	200.82	203.65	215.15	219.86	217.66	216.15
20.000	206.33	207.76	205.96	215.15	218.67	215.66	214.16
22.000	213.09	214.70	210.92	217.07	217.47	213.67	212.16
24.000	218.24	219.59	215.88	219.06	218.30	211.68	210.17
26.000	223.39	224.35	219.52	221.04	219.89	211.79	210.15
28.000	228.53	229.09	223.09	223.02	221.48	214.17	210.15
30.000	233.67	223.84	226.66	225.00	223.06	216.56	210.15
32.000	238.78	238.58	230.22	227.91	224.65	218.94	210.15
34.000	243.72	243.32	233.78	232.07	227.87	221.32	213.99
36.000	248.65	248.30	237.34	236.22	231.83	225.02	217.96
38.000	253.58	253.43	241.77	240.37	235.79	230.37	221.93
40.000	258.50	258.55	246.70	244.52	339.75	235.71	225.89
42.000	263.32	263.67	251.63	249.68	244.59	241.06	231.71
44.000	266.47	266.72	256.56	256.59	249.53	246.40	237.65
46.000	269.62	269.28	261.49	261.82	254.47	251.73	243.58
48.000	271.15	271.15	266.25	266.75	259.40	257.06	249.51
50.000	271.15	271.15	267.65	267.65	260.15	262.39	253.15
52.000	267.84	267.07	267.65	266.08	260.15	263.15	253.15
54.000	263.33	261.57	263.43	260.77	257.14	263.15	253.15
56.000	258.60	255.93	257.93	255.47	251.82	259.05	
58.000	252.92	249.66	252.52	249.72	246.51	251.96	
60.000	247.24	243.39	247.42	243.83	241.20	244.88	
62.000	240.21	236.37	242.32	237.94	235.90	237.80	
64.000	232.39	228.94	232.13	233.42	232.20	230.72	
66.000	224.57	221.51	227.04	229.50	230.24	227.15	
68.000	216.76	214.08	221.96	225.59	228.28	227.15	
70.000	208.95	206.66	216.88	221.67	226.32	227.15	
72.000	201.15	201.09	211.80	217.76	224.36	227.15	
74.000	195.15	197.19	206.72	213.85	222.40	225.52	
76.000	195.15	195.15	201.65	209.94	218.75	221.60	
78.000	195.15	195.15	200.67	206.04	214.84	217.69	
80.000	195.15	195.15	199.70	202.15	210.93	213.77	
82.000	195.15	195.15	199.65	202.15	207.03	209.87	
84.000	195.15	195.15	199.65	202.15	204.15	205.96	
86.000	195.15	195.15	199.65	202.15	204.15	205.15	
88.000	195.15	195.15	199.65	202.15	204.15	205.15	
90.000	195.15	195.15	199.65	202.15	204.15	205.15	

Table 14-2e. Reference Atmosphere Density (kg/m^3)—January

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N	
0.000	1.1761	1.1891	1.2363	1.3013	1.3742	1.4244	1.4910	+0*
2.000	9.6636	9.7230	9.9321	1.0357	1.0590	1.0711	1.0854	
4.000	7.8839	7.8982	8.1234	8.2700	8.3466	8.4810	8.4962	−1
6.000	6.3834	6.4213	6.5793	6.6018	6.6507	6.6428	6.6028	
8.000	5.1518	5.2127	5.2712	5.2114	5.2286	5.1417	5.0740	
10.000	4.1779	4.1820	4.172	4.0631	3.8811	3.8185	3.7553	
12.000	3.3404	3.3114	3.2591	2.9928	2.8353	2.7984	2.7457	
14.000	2.6276	2.5835	2.4310	3.2009	2.0717	2.0452	2.0020	
16.000	1.9873	1.9567	1.7992	1.6165	1.5172	1.4905	1.4674	
18.000	1.4185	1.4158	1.3073	1.1859	1.1126	1.0832	1.0692	
20.000	0.9761	9.6547	9.1885	8.6716	8.1486	7.8497	7.7432	−2
22.000	6.7966	6.6802	6.5057	6.3254	5.9605	5.6510	5.4786	
24.000	4.7853	4.7721	4.6425	4.6148	4.3545	4.0226	3.8768	
26.000	3.4046	3.4372	3.3411	3.3675	3.1592	2.8697	2.7438	
28.000	2.4462	2.4921	2.4296	2.4375	2.2814	2.0517	1.9256	
30.000	1.7766	1.8183	1.7760	1.7546	1.6522	1.4615	1.3392	
32.000	1.3103	1.3348	1.3015	1.2715	1.2000	1.0400	9.3378	
34.000	0.9715	0.9857	0.9582	0.9272	0.8740	0.7448	0.6584	
36.000	7.2401	7.2933	7.0967	6.7763	6.3791	5.3685	4.6932	−3
38.000	5.3993	5.4221	5.2832	4.9797	4.6657	3.8925	3.3786	
40.000	4.0487	4.0598	3.9559	3.6876	3.4235	2.8388	2.4805	
42.000	3.0546	3.0605	2.9787	2.7508	2.4799	2.0819	1.8316	
44.000	2.3182	2.3396	2.2549	2.0662	1.8100	1.5352	1.3589	
46.000	1.7692	1.7981	1.7174	1.5752	1.3411	1.1379	1.0123	
48.000	1.3760	1.3906	1.3334	1.2119	1.0088	0.8477	0.7582	
50.000	1.0753	1.0856	1.0354	0.9399	0.7615	0.6347	0.5700	
52.000	8.4906	8.5422	8.1503	7.3740	5.8134	4.7744	0.4338	−4
54.000	6.7090	6.7568	6.4162	5.8169	4.4471	3.6259	3.3209	
56.000	5.2860	5.3239	5.0251	4.5435	3.4051	2.7964		
58.000	4.1600	4.1913	3.9109	3.5198	2.6086	2.1707		
60.000	3.2586	3.2837	3.0303	2.7172	1.9082	1.6796		
62.000	2.5402	2.5592	2.3371	2.0900	1.5522	1.2952		
64.000	1.9702	1.9835	1.7939	1.6016	1.2002	0.9954		
66.000	1.5200	1.5285	1.3701	1.2227	0.9162	0.7587		
68.000	1.1663	1.1730	1.0410	0.9297	0.6920	0.5695		
70.000	8.8979	8.9430	7.8672	7.0413	5.2271	0.4284		
72.000	6.8278	6.7570	5.8499	5.2477	3.9583	0.3275		
74.000	5.2406	5.0487	4.2[illegible]9	3.9033	3.0319	2.5119		−5
76.000	3.9678	3.7483	3.1[illegible]9	2.9301	2.3122	1.9182		
78.000	2.8323	2.7644	2.3840	2.2067	1.7555	1.4583		
80.000	2.0164	2.0059	1.7683	1.6538	1.3267	1.1037		
82.000	1.4358	1.4343	1.3053	1.2290	0.9979	0.8313		
84.000	1.0226	1.0258	0.9587	0.9069	7.4677	0.6231		−6
86.000	7.3229	7.3799	7.0051	6.6705	5.4751	4.6052		
88.000	5.2827	5.3488	5.0910	4.8902	4.0150	3.3735		
90.000	3.7900	3.8549	3.6793	3.5731	2.9448	2.4717		

*Power of 10 by which preceding numbers should be multiplied.

Table 14-2f. Reference Atmosphere Density (kg/m³)—April

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N	
0.000	1.1716	1.1857	1.2113	1.2678	1.3116	1.3884	1.4297	+0*
2.000	9.6565	9.7223	9.8903	1.0117	1.0375	1.0635	1.0745	
4.000	7.8992	7.9159	8.0172	8.1340	8.1648	8.4225	8.4831	−1
6.000	6.4096	6.4388	6.4947	6.5615	6.5709	6.6274	6.6313	
8.000	5.1728	5.2202	5.2446	5.2336	5.2213	5.1557	5.1285	
10.000	4.1737	4.1823	4.1832	4.1224	3.9711	3.7924	3.7669	
12.000	3.3219	3.3066	3.2908	3.1158	2.9216	2.7756	2.7470	
14.000	2.6032	2.5754	2.4578	2.2808	2.1499	2.0475	2.0262	
16.000	1.9726	1.9299	1.8223	1.6700	1.5823	1.5106	1.4947	
18.000	1.4115	1.3954	1.3288	1.2229	1.1648	1.1147	1.1029	
20.000	0.9725	9.5265	9.3798	8.9578	8.5768	8.2281	8.1399	−2
22.000	6.7816	6.6002	6.6703	6.5625	6.3162	6.0742	6.0084	
24.000	4.7828	4.6845	4.7883	4.7676	4.6523	4.4850	4.4260	
26.000	3.4092	3.3930	3.4617	3.4649	3.4129	3.3122	3.2438	
28.000	2.4585	2.4733	2.5195	2.5294	2.4976	2.4465	2.3849	
30.000	1.7923	1.8140	1.8456	1.8516	1.8336	1.7839	1.7589	
32.000	1.3165	1.3383	1.3603	1.3540	1.3391	1.3068	1.3011	
34.000	0.9741	0.9929	1.0092	0.9968	0.9836	0.9621	0.9631	−3
36.000	7.2575	7.4043	7.5300	7.3817	7.2550	7.1172	7.1019	
38.000	5.4428	5.5315	5.6476	5.4858	5.3740	5.2896	5.2700	
40.000	4.1077	4.1592	4.2575	4.1059	4.0089	3.9134	3.9347	
42.000	3.1189	3.1468	3.2253	3.1075	3.0108	2.9155	2.9550	
44.000	2.4044	2.4043	2.4632	2.3692	2.2758	2.1882	2.2319	
46.000	1.8624	1.8595	1.8890	1.8139	1.7371	1.6540	1.6949	
48.000	1.4527	1.4412	1.4582	1.4007	1.7371	1.6540	1.6949	
50.000	1.1352	1.1262	1.1360	1.0933	1.0396	0.9812	1.010[illegible]	
52.000	8.9597	8.8794	8.8895	8.6128	8.1322	7.6739	0.7872	−4
54.000	7.0759	7.0044	7.0239	6.8140	6.4353	6.0473	6.1347	
56.000	5.5694	5.5228	5.5281	5.3690	5.1181	4.8196		
58.000	4.3731	4.3794	4.3441	4.2210	4.0501	3.8208		
60.000	3.4519	3.4507	3.4001	3.3023	3.1882	3.0122		
62.000	2.7184	2.7007	2.6471	2.5703	2.4960	2.3609		
64.000	2.1189	2.0987	2.0495	1.9898	1.9429	1.8391		
66.000	1.6387	1.6186	1.5777	1.5319	1.4984	1.4234		
68.000	1.2568	1.2384	1.2075	1.1670	1.1456	1.0840		
70.000	9.5682	9.3944	9.1854	8.8452	8.7179	0.8218		−5
72.000	7.2109	7.0488	6.9386	6.6727	6.6024	0.6203		
74.000	5.3759	5.2331	5.2032	5.0094	4.9755	4.6613		
76.000	3.9611	3.8493	3.8718	3.7419	3.7359	3.4859		
78.000	2.7561	2.7309	2.8279	2.7806	2.7954	2.6145		
80.000	1.9224	1.9355	2.0362	2.0551	2.0773	1.9477		
82.000	1.3465	1.3584	1.4640	1.5152	1.5325	1.4389		−6
84.000	0.9471	0.9568	1.0512	1.1105	1.1220	1.0535		
86.000	6.7014	6.7706	7.5368	8.0774	8.1494	7.6418		
88.000	4.7692	4.8129	5.3958	5.7400	5.8691	5.4874		
90.000	3.4132	3.4364	3.8574	4.0433	4.1894	3.8701		

*Power of 10 by which preceding numbers should be multiplied.

Table 14-2g. Reference Atmosphere Density (kg/m^3)—July

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N	
0.000	1.1797	1.1751	1.1717	1.2003	1.2252	1.2778	1.2919	+0*
2.000	9.7077	9.6694	9.7103	0.9780	0.9987	1.0032	1.0120	
4.000	7.9287	7.8986	7.9314	7.9946	8.0757	8.1533	8.1528	−1
6.000	6.4239	6.4017	6.4266	6.4769	6.5120	6.5568	6.5576	
8.000	5.1594	5.1685	5.2140	5.2145	5.2231	5.2120	5.2124	
10.000	4.1775	4.1875	4.1827	4.1502	4.1345	3.9747	3.9519	
12.000	3.3352	3.3444	3.3109	3.2612	3.0546	2.9399	2.9076	
14.000	2.6191	2.6275	2.5823	2.4580	2.2570	2.1773	2.1645	
16.000	1.9556	1.9493	1.9214	1.7932	1.6679	1.6192	1.6116	
18.000	1.3723	1.3730	1.3521	1.3021	1.2329	1.2043	1.2001	
20.000	0.9525	9.5976	9.5807	9.4313	9.1149	8.9601	8.9390	−2
22.000	6.6925	6.7780	6.8380	6.8564	6.7399	6.6671	6.6592	
24.000	4.7926	4.8334	4.9144	5.0062	4.9463	4.9618	4.9518	
26.000	3.4691	3.5029	3.5556	3.6673	3.6143	3.6569	3.6647	
28.000	2.5265	2.5552	2.5891	2.6835	2.6556	2.6803	2.7081	
30.000	1.8509	1.8743	1.8969	1.9638	1.9598	1.9781	2.0097	
32.000	1.3637	1.3823	1.3981	1.4460	1.4495	1.4694	1.4975	
34.000	1.0103	1.0248	1.0363	1.0712	1.0791	1.0985	1.1062	
36.000	7.5248	7.6366	7.7248	7.9815	8.0839	8.2646	8.2372	−3
38.000	5.6332	5.7168	5.7779	5.9797	5.0918	6.2554	6.1828	
40.000	4.2381	4.2984	4.3456	4.5040	4.6169	4.7601	4.6763	
42.000	3.2037	3.2475	3.2867	3.4100	3.5184	3.6410	3.5820	
44.000	2.4293	2.4650	2.5144	2.5946	2.6955	2.7989	2.7782	
46.000	1.8493	1.8795	1.9321	1.9836	2.0908	2.1619	2.1604	
48.000	1.4275	1.4450	1.4931	1.5315	1.6358	1.7007	1.6911	
50.000	1.1125	1.1250	1.1632	1.1971	1.2850	1.3408	1.3332	
52.000	8.7481	8.8047	9.1197	9.3670	1.0150	1.0572	1.0511	−4
54.000	6.8831	6.9145	7.1504	7.4642	8.0867	8.3501	8.2891	
56.000	5.4137	5.4579	5.6185	5.9204	6.4179	6.6864		
58.000	4.2874	4.3247	4.4591	4.6911	5.0734	5.3337		
60.000	3.3722	3.4028	3.5152	3.6881	3.9941	4.2378		
62.000	2.6331	2.6574	2.7514	2.8804	3.1454	3.3550		
64.000	2.0402	2.0588	2.1373	2.2339	2.4948	2.7023		
66.000	1.5679	1.5815	1.6469	1.7196	1.9607	2.1552		
68.000	1.1961	1.2038	1.2581	1.3134	1.5256	1.7003		−5
70.000	9.0350	9.0750	9.5234	9.9476	1.1742	1.3256		
72.000	6.7522	6.6490	7.0578	7.4673	8.9324	1.0199		
74.000	4.9325	4.8263	5.1557	5.5522	6.7071	7.7331		
76.000	3.4623	3.4821	3.7497	4.0732	4.9649	5.7676		
78.000	2.4394	2.4867	2.7148	2.9407	3.6178	4.2228		
80.000	1.7251	1.7763	1.9564	2.1061	2.5905	3.0274		
82.000	1.2285	1.2690	1.4031	1.4957	1.8162	2.1197		
84.000	0.8809	0.9068	1.0013	1.0528	1.1999	1.3730		−6
86.000	6.3733	6.6002	7.1100	7.3426	7.8878	8.7593		
88.000	4.6691	4.7844	5.0219	5.0703	5.1588	5.5760		
90.000	3.3898	3.4399	3.5279	3.3925	3.3563	3.5420		

*Power of 10 by which preceding numbers should be multiplied

Table 14-2h. Reference Atmosphere Density (kg/m³)—October

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N	
0.000	1.1769	1.1816	1.2063	1.2474	1.2795	1.3385	1.3981	+0*
2.000	9.6821	9.6775	9.8598	1.0122	1.0240	1.0425	1.0617	
4.000	7.9068	7.9044	8.0015	8.1810	8.1419	8.3111	8.4294	−1
6.000	6.4329	6.4133	6.4897	6.5738	6.5407	6.6134	6.6221	
8.000	5.2053	5.1764	5.2471	5.2270	5.1931	5.1955	5.1425	
10.000	4.1652	4.1790	4.1912	4.1081	4.0696	3.8648	3.7877	
12.000	3.3261	3.3274	3.3023	3.1874	2.9860	2.8327	2.7569	
14.000	2.6135	2.6080	2.5021	2.3246	2.1912	2.0766	2.0434	
16.000	1.9701	1.9722	1.8579	1.6946	1.6082	1.5226	1.5107	
18.000	1.3779	1.3810	1.3313	1.2356	1.1821	1.1252	1.1140	
20.000	0.9524	9.5802	9.3893	9.0114	8.7162	8.2945	8.1936	−2
22.000	6.6793	6.7278	6.6771	6.5238	6.4173	6.0980	6.0101	
24.000	4.7657	4.8169	4.8106	4.7364	4.6790	4.4711	4.3964	
26.000	3.4279	3.4765	3.4862	3.4493	3.4076	3.2395	3.1819	
28.000	2.4848	2.5268	2.5398	2.5196	2.4879	2.3289	2.3031	
30.000	1.8144	1.8489	1.8599	1.8460	1.8208	1.6807	1.6673	
32.000	1.3343	1.3617	1.3687	1.3512	1.3358	1.2175	1.1889	
34.000	0.9885	1.0091	1.0122	0.9890	0.9760	0.8852	0.8528	
36.000	7.3687	7.5165	7.4956	7.2803	7.1469	6.4248	6.1550	−3
38.000	5.5256	5.6298	5.5731	5.3888	5.2620	4.6616	4.4688	
40.000	4.1672	4.2419	4.1689	4.0101	3.8948	3.4080	3.2409	
42.000	3.1613	3.2145	3.1366	2.9879	2.8883	2.5095	2.3687	
44.000	2.4237	2.4664	2.3732	2.2277	2.1543	1.8607	1.7450	
46.000	1.8643	1.9011	1.8063	1.6837	1.6164	1.3888	1.2952	
48.000	1.4467	1.4728	1.3971	1.2809	1.2197	1.0431	0.9767	
50.000	1.1295	1.1498	1.0869	0.9928	0.9386	0.7882	0.7482	
52.000	8.9199	9.1029	8.5800	7.7661	7.2462	6.0836	0.5732	−4
54.000	7.0493	7.2119	6.7740	6.1400	5.6554	4.7098	4.3930	
56.000	5.5537	5.6898	5.3200	4.8317	4.4337	3.7001		
58.000	4.3709	4.4754	4.1526	3.7897	3.4585	2.9235		
60.000	3.4219	3.4993	3.2253	2.9575	2.6836	2.2949		
62.000	2.6779	2.7270	2.4920	2.2944	2.0710	1.7889		
64.000	2.0863	2.1125	1.9149	1.7597	1.5788	1.3842		
66.000	1.6118	1.6230	1.4632	1.3402	1.1915	1.9476		
68.000	1.2341	1.2360	1.1114	1.0161	8.9728	0.7798		
70.000	9.3589	9.3244	8.3906	7.6683	6.7413	0.5805		−5
72.000	7.0241	6.9050	6.2932	5.7589	5.0531	0.4323		
74.000	5.1671	5.0394	4.6880	4.3034	3.7788	3.2416		
76.000	3.6738	3.6236	3.4675	3.1991	2.8382	2.4457		
78.000	2.6127	2.5767	2.5022	2.3654	2.1236	1.8363		
80.000	1.8585	1.8326	1.8032	1.7391	1.5808	1.3719		
82.000	1.3222	1.3036	1.2925	1.2508	1.1705	1.0196		
84.000	0.9409	0.9276	0.9264	0.8998	8.5777	0.7537		−6
86.000	6.6974	6.6017	6.6420	6.4750	6.1893	5.4678		
88.000	4.7680	4.6993	4.7628	4.6600	4.4668	3.9511		
90.000	3.3951	3.3458	3.4160	3.3544	3.2244	2.8557		

*Power of 10 by which preceding numbers should be multiplied.

Table 14-2i. Reference Atmosphere Pressure (mb)—January

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N	
0.000	1.0100	1.0126	1.0191	1.0166	1.0144	1.0147	1.0150	+3*
2.000	8.0106	8.0179	8.0160	7.8835	7.7806	7.7293	7.6613	+2
4.000	6.3009	6.3007	6.2540	6.0693	5.9354	5.8516	5.7706	
6.000	4.9119	4.9087	4.8205	4.6180	4.4699	4.3734	4.2948	
8.000	3.7918	3.7760	3.6661	3.4662	3.3102	3.2228	3.1547	
10.000	2.8842	2.8621	2.7470	2.5629	2.4192	2.3472	2.2922	
12.000	2.1535	2.1340	2.0245	1.8785	1.7673	1.7041	1.6602	
14.000	1.5741	1.5618	1.4729	1.3751	1.2914	1.2337	1.1990	
16.000	1.1258	1.1203	1.0634	1.0054	1.9436	1.8906	0.8620	+1
18.000	7.9414	7.9315	7.6236	7.3420	6.8878	6.4106	6.1588	
20.000	5.6424	5.6442	5.4791	5.3555	5.0213	4.6004	4.3714	
22.000	4.0525	4.0763	3.9683	3.9065	3.6559	3.2919	3.0926	
24.000	2.9403	2.9757	2.8952	2.8501	2.6584	2.3570	2.1884	
26.000	2.1538	2.1863	2.1271	2.0797	1.9314	1.6913	1.5488	
28.000	1.5920	1.6163	1.5717	1.5189	1.4065	1.2162	1.0869	
30.000	1.1867	1.2020	1.1670	1.1153	1.0271	1.8768	1.7805	
32.000	8.9018	8.9907	8.7098	8.2413	0.7521	0.6358	0.5618	+0
34.000	6.7103	6.7618	6.5377	6.1259	5.5230	4.6386	4.0855	
36.000	5.0825	5.1143	4.9348	4.5816	4.0665	3.4042	3.0000	
38.000	3.8695	3.8943	3.7456	3.4517	3.0035	2.5126	2.2231	
40.000	2.9634	2.9848	2.8584	2.6188	2.2252	1.8647	1.6590	
42.000	2.2824	2.3019	2.1927	2.0003	1.6577	1.3912	1.2437	
44.000	1.7675	1.7844	1.6905	1.5376	1.2458	1.0432	0.9366	
46.000	1.3759	1.3876	1.3096	1.1884	0.9438	0.7862	0.7084	−1
48.000	1.0749	1.0823	1.0168	0.9207	7.1802	0.5953	0.5380	
50.000	8.4007	8.4501	7.8958	7.1403	5.4799	0.4529	0.4103	
52.000	6.5594	6.5948	6.1234	5.5334	4.1911	3.4610	3.1400	
54.000	5.1036	5.1289	4.7265	4.2665	3.2060	2.6556	2.4037	
56.000	3.9560	3.9723	3.6300	3.2713	2.4528	2.0394		
58.000	3.0525	3.0619	2.7745	2.4986	1.8761	1.5623		
60.000	2.3434	2.3472	2.1105	1.9012	1.4346	1.1926		
62.000	1.7897	1.7890	1.5975	1.4412	1.0942	1.9073		
64.000	1.3593	1.3554	1.2030	1.0882	1.8306	0.6877		
66.000	1.0265	1.0204	0.9011	0.8184	0.6276	0.5195		−2
68.000	7.7069	7.6307	6.7122	6.1291	4.7405	0.3925		
70.000	5.7500	5.6623	4.9711	4.5704	3.5808	2.9720		
72.000	4.2580	4.1680	3.6632	3.3991	2.7051	2.2512		
74.000	3.1098	3.0464	2.6983	2.5283	2.0379	1.6982		
76.000	2.2347	2.2114	1.9867	1.8794	1.5282	1.2753		
78.000	1.5906	1.5938	1.4565	1.3907	1.1406	0.9531		
80.000	1.1324	1.1409	1.0626	1.0237	0.8471	0.7089		−3
82.000	8.0642	8.1586	7.7126	7.4973	6.2598	5.2469		
84.000	5.7436	5.8351	5.5682	5.4712	4.6013	3.8628		
86.000	4.0905	4.1728	3.9981	3.9795	3.3735	2.8309		
88.000	2.9008	2.9707	2.8545	2.8846	2.4738	2.0738		
90.000	2.0453	2.1023	2.0260	2.0837	1.8144	1.5194		

*Power of 10 by which preceding numbers should be multiplied

Table 14-2j. Reference Atmosphere Pressure (mb)—April

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N	
0.000	1.0095	1.0114	1.0159	1.0159	1.0134	1.0169	1.0205	+3*
2.000	8.0104	8.0090	8.0108	7.9327	7.8374	7.7818	7.7772	+2
4.000	6.2997	6.2899	6.2640	6.1568	6.0265	5.9201	5.8956	
6.000	4.9065	4.8940	4.8518	4.7216	4.5860	4.4489	4.4181	
8.000	3.7810	3.7589	3.7079	3.5712	3.4344	3.2982	3.2696	
10.000	2.8720	2.8444	2.7901	2.6596	2.5323	2.4228	2.3996	
12.000	2.1437	2.1167	2.0633	1.9511	1.8631	1.7859	1.7675	
14.000	1.5686	1.5457	1.5060	1.4283	1.3709	1.3174	1.3037	
16.000	1.1243	1.1081	1.0917	1.0457	1.0090	0.9720	0.9617	+1
18.000	7.9466	7.8573	7.8633	7.6583	7.4282	7.1729	7.0967	
20.000	5.6576	5.6017	5.6791	5.6094	5.4693	5.2942	5.2374	
22.000	4.0725	4.0537	4.1336	4.1095	4.0277	3.9083	3.8660	
24.000	2.9619	2.9721	3.0303	3.0144	2.9667	2.8857	2.8544	
26.000	2.1751	2.1952	2.2361	2.2204	2.1863	2.1311	2.1123	
28.000	1.6118	1.6311	1.6604	1.6425	1.6156	1.5741	1.5680	
30.000	1.2035	1.2189	1.2404	1.2200	1.1976	1.1656	1.1674	
32.000	9.0499	9.1598	9.3203	9.1144	0.8917	0.8672	0.8718	+0
34.000	6.8505	6.9201	7.0416	6.8516	6.6787	6.4835	6.5297	
36.000	5.2189	5.2550	5.3475	5.1817	5.0313	4.8691	4.9156	
38.000	4.0001	4.0135	4.0811	3.9442	3.8154	3.6729	3.7226	
40.000	3.0840	3.0837	3.1296	3.0220	2.9122	2.7863	2.8353	
42.000	2.3909	2.3829	2.4110	2.3290	2.2367	2.1288	2.1714	
44.000	1.8616	1.8513	1.8650	1.8025	1.7281	1.6375	1.6717	−1
46.000	1.4525	1.4423	1.4473	1.4006	1.3426	1.2678	1.2935	
48.000	1.1349	1.1259	1.1266	1.0922	1.0466	0.9876	1.0058	
50.000	8.8690	8.7980	8.7773	8.5260	8.1813	0.7722	0.7835	
52.000	6.9254	6.8698	6.8383	6.6521	6.3997	6.0390	6.1049	
54.000	5.3897	5.3482	5.3121	5.1704	5.0026	4.7219	4.7573	
56.000	4.1795	4.1501	4.1092	4.0012	3.8918	3.6761		
58.000	3.2290	3.2022	3.1641	3.0814	3.0111	2.8452		
60.000	2.4763	2.4533	2.4229	2.3604	2.3164	2.1886		
62.000	1.8888	1.8654	1.8447	1.7981	1.7712	1.6727		
64.000	1.4271	1.4072	1.3959	1.3618	1.3459	1.2698		
66.000	1.0687	1.0526	1.0497	1.0252	1.0161	0.9571		−2
68.000	7.9288	7.8036	7.8411	7.6745	7.6335	0.7171		
70.000	5.8214	5.7300	5.8149	5.7172	5.7060	5.3494		
72.000	4.2257	4.1655	4.2797	4.2381	4.2435	3.9715		
74.000	3.0300	2.9988	3.1248	3.1254	3.1391	2.9343		
76.000	2.1440	2.1369	2.2623	2.2926	2.3092	2.1572		
78.000	1.5090	1.5141	1.6249	1.6725	1.6871	1.5758		
80.000	1.0675	1.0731	1.1643	1.2131	1.2234	1.1411		
82.000	7.5916	7.6212	7.3308	8.7433	8.8026	8.1878		−3
84.000	5.4293	5.4373	5.9522	6.2526	6.2813	5.8179		
86.000	3.9051	3.8967	4.2465	4.4348	4.4435	4.0914		
88.000	2.8243	2.8048	3.0251	3.1248	3.1148	2.8457		
90.000	2.0536	2.0275	2.1519	2.2011	2.1624	1.9569		

*Power of 10 by which preceding numbers should be multiplied.

Table 14-2k. Reference Atmosphere Pressure (mb)—July

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N	
0.000	1.0114	1.0108	1.0129	1.0135	1.0099	1.0111	1.0130	+3*
2.000	8.0167	8.0182	8.0323	8.0057	7.9221	7.8803	7.8763	+2
4.000	6.2981	6.3057	6.3110	6.2688	6.1556	6.0994	6.0870	
6.000	4.9007	4.9130	4.9112	4.8560	4.7319	4.6602	4.6473	
8.000	3.7740	3.7889	3.7768	3.7154	3.5855	3.5098	3.4964	
10.000	2.8659	2.8784	2.8618	2.8026	2.6724	2.6077	2.5951	
12.000	2.1358	2.1462	2.1329	2.0810	1.9742	1.9355	1.9293	
14.000	1.5577	1.5662	1.5603	1.5215	1.4587	1.4384	1.4362	
16.000	1.1148	1.1208	1.1204	1.1100	1.0780	1.0697	1.0693	+1
18.000	7.9201	7.9959	8.0491	8.1030	7.9683	7.9568	7.9632	
20.000	5.6870	5.7531	5.8235	5.9336	5.8909	5.9195	5.9312	
22.000	4.1290	4.1786	4.2420	4.3606	4.3560	4.4046	4.4185	
24.000	3.0263	3.0622	3.1103	3.2154	3.2235	3.2780	3.2922	
26.000	2.2319	2.2606	2.2948	2.3784	2.3966	2.4419	2.4562	
28.000	1.6556	1.6780	1.7034	1.7653	1.7912	1.8295	1.8398	
30.000	1.2349	1.2521	1.2717	1.2175	1.3455	1.3795	1.3836	
32.000	9.2614	9.3917	9.5480	9.8912	1.0167	1.0466	1.0447	+0
34.000	6.9815	7.0790	7.2069	7.4672	7.7286	7.9867	7.9335	
36.000	5.2893	5.3614	5.4680	5.6678	5.9090	6.1289	6.0713	
38.000	4.0266	4.0794	4.1705	4.3244	4.5428	4.7279	4.6802	
40.000	3.0797	3.1183	3.1982	3.3159	3.5111	3.6654	3.6329	
42.000	2.3661	2.3945	2.4654	2.5549	2.7277	2.8553	2.8376	
44.000	1.8260	1.8468	1.9087	1.9778	2.1294	2.2346	2.2240	−1
46.000	1.4161	1.4304	1.4819	1.5379	1.6693	1.7567	1.7478	
48.000	1.1028	1.1123	1.1536	1.2008	1.3108	1.3847	1.3770	
50.000	8.5953	8.6599	8.9876	9.3866	1.0297	1.0917	1.0855	
52.000	6.6938	6.7410	6.9986	7.3382	8.0887	8.6086	8.5588	
54.000	5.1970	5.2354	5.4389	5.7211	6.3343	6.7889	6.7492	
56.000	4.0215	4.0533	4.2174	4.4357	4.9401	5.3417		
58.000	3.0932	3.1169	3.2518	3.4168	3.8364	4.1856		
60.000	2.3609	2.3778	2.4883	2.6135	2.9661	3.2658		
62.000	1.7873	1.7987	1.8888	1.9844	2.2825	2.5368		
64.000	1.3412	1.3484	1.4216	1.4949	1.7415	1.9549		
66.000	0.9972	1.0012	1.0602	1.1169	1.3145	1.4886		−2
68.000	7.3396	7.3582	7.8313	8.2726	9.8076	1.1189		
70.000	5.3414	5.3484	5.7247	6.0699	7.2253	8.2902		
72.000	3.8406	3.8519	4.1432	4.4096	5.2499	6.0456		
74.000	2.7270	2.7636	2.9833	3.1697	2.6448	3.0425		
76.000	1.9335	1.9756	2.1382	2.2531	2.6448	3.0425		
78.000	1.3759	1.4108	1.5253	1.5871	1.8277	2.0899		
80.000	0.9826	1.0078	1.0827	1.1083	1.2375	1.3997		−3
82.000	7.0408	7.2002	7.6480	7.6707	8.1900	9.1071		
84.000	5.0489	5.1453	5.3742	5.2582	5.3435	5.7944		
86.000	3.6196	3.6665	3.7563	3.5684	2.2396	2.3281		
88.000	2.5763	2.5911	2.6111	2.3961	2.2396	2.3281		
90.000	1.8156	1.8150	1.8048	1.5985	1.4383	1.4709		

*Power of 10 by which preceding numbers should be multiplied

Table 14-21. Reference Atmosphere Pressure (mb)—October

Alt (km)	0°	15°N	30°N	45°N	60°N	75°N	90°N	
0.000	1.0107	1.0113	1.0169	1.0175	1.0106	1.0092	1.0140	+3*
2.000	8.0149	8.0165	8.0286	7.9662	7.8528	7.7639	7.7527	
4.000	6.3010	6.6036	6.2863	6.1801	6.0564	5.9371	5.8884	+2
6.000	4.9062	4.9091	4.8760	4.7399	4.6213	4.4773	4.4166	
8.000	3.7737	3.7826	3.7323	3.5891	3.4755	3.3235	3.2674	
10.000	2.8627	2.8723	2.8135	2.6797	2.5720	2.4368	2.3937	
12.000	2.1347	2.1427	2.0847	1.9697	1.8870	1.7860	1.7580	
14.000	1.5581	1.5664	1.5216	1.4356	1.3847	1.3093	1.2913	
16.000	1.1128	1.1205	1.0974	1.0456	1.0163	0.9600	0.9460	+1
18.000	7.8791	7.9611	7.8713	7.6313	7.4607	7.0305	6.9124	
20.000	5.6412	5.7136	5.6849	5.5654	5.4712	5.1350	5.0370	
22.000	4.0858	4.1464	4.1378	4.0651	4.0062	3.7403	3.6603	
24.000	2.9857	3.0364	3.0315	2.9783	2.9321	2.7169	2.6524	
26.000	2.1982	2.2389	2.2325	2.1886	2.1509	1.9695	1.9194	
28.000	1.6300	1.6617	1.6525	1.6131	1.5817	1.4318	1.3893	
30.000	1.2170	1.2411	1.2291	1.1923	1.1659	1.0448	1.0058	
32.000	9.1464	9.3262	9.1856	8.8406	0.8614	0.7652	0.7303	+0
34.000	6.9158	7.0486	6.8961	6.5885	6.3843	5.6244	5.3358	
36.000	5.2593	5.3575	5.2020	4.9367	4.7562	4.1500	3.9211	
38.000	4.0221	4.0956	3.9467	3.7183	3.5616	3.0827	2.8978	
40.000	3.0923	3.1487	3.0113	2.8147	2.6804	2.3059	2.1556	
42.000	2.3896	2.4330	2.3100	2.1415	2.0279	1.7365	1.6158	
44.000	1.8540	1.8884	1.7813	1.6408	1.5431	1.3161	1.2201	−1
46.000	1.4429	1.4695	1.3805	1.2654	1.1807	1.0035	0.9277	
48.000	1.1260	1.1463	1.0734	0.9808	9.0823	0.7697	0.7097	
50.000	8.7921	8.9499	8.3510	7.6277	7.0097	0.5937	0.5427	
52.000	6.8580	6.9786	6.4881	5.9319	5.4112	4.5954	4.1658	
54.000	5.3285	5.4152	5.0155	4.5962	4.1744	3.5576	3.1923	
56.000	4.1227	4.1800	3.8563	3.5432	2.4473	2.7515		
58.000	3.1734	3.2073	2.9493	2.7166	2.4473	2.1145		
60.000	2.4286	2.4449	2.2435	2.0700	1.9581	1.6132		
62.000	1.8465	1.8503	1.6969	1.5671	1.4024	1.2211		
64.000	1.3917	1.3883	1.2760	1.1791	1.0524	0.9168		−2
66.000	1.0390	1.0320	0.9536	0.8829	0.7875	0.6831		
68.000	7.6789	7.5958	7.0818	6.5802	5.8798	0.5084		
70.000	5.6135	5.5316	5.2237	4.8795	4.3796	4.7856		
72.000	4.0557	3.9860	3.8261	3.5999	3.2544	2.8189		
74.000	2.8945	2.8526	2.7819	2.6417	2.4125	2.0985		
76.000	2.0580	2.0299	2.0072	1.9279	1.7822	1.5558		
78.000	1.4636	1.4434	1.4414	1.3990	1.3096	1.1475		
80.000	1.0411	1.0266	1.0336	1.0091	0.9571	0.8419		−3
82.000	7.4071	7.3031	7.4076	7.2585	6.9561	6.1426		
84.000	5.2710	5.1963	5.3096	5.2217	5.0267	4.4562		
86.000	3.7517	3.6981	3.8065	3.7572	3.6270	3.2199		
88.000	2.6709	2.6324	2.7295	2.7040	2.6176	2.3267		
90.000	1.9019	1.8742	1.9577	1.9465	1.8895	1.6817		

*Power of 10 by which preceding numbers should be multiplied

up to 90 km. The largest seasonal changes in mean monthly temperatures occur north of 60° latitude in arctic and subarctic regions. At altitudes up to 60 km, coldest temperatures occur in January and warmest in July. Conditions are reversed at altitudes between 60 and 90 km with coldest temperatures occurring in July and warmest in January.

Cross sections of the January and July densities between the equator and north pole are shown in Figure 14-13. Densities at the various altitudes are expressed as percentages of the densities in the U.S. Standard Atmosphere. The largest departures from standard in January and July densities occur north of 65° latitude near 70 km, where mean January densities, 51% of standard, are a third of the July densities which are 168% of standard.

14.2.1.4 Longitudinal Variations. In summer, longitudinal variations in the structure of the atmosphere are relatively small at all latitudes and at all altitudes above 20 km. Isotherms and contour lines on constant pressure-height charts in the stratosphere and mesosphere parallel the latitude circles, and the associated circulation pattern is symmetrical about the pole. During the winter season, changes with longitude remain relatively small at low latitudes but become as important as changes with latitude and season in arctic and subarctic regions. The magnitude of the longitudinal variations in arctic and subarctic regions during winter is illustrated by a set of atmospheric models depicting January conditions between the surface and 54 km at 10°, 100°, and 140°W for 60°N, and 10° and 140°W for 75°N. The models are based on radiosonde observations, rocketsonde observations, constant pressure maps between 5 and 0.4 mb, hydrostatic build-up techniques from the 5- and 10-mb levels, and the thermal wind equation. The atmospheric properties for these January models at 60° and 75°N are given in Table 14-3. Temperature-height profiles for the individual models are shown in Figures 14-14 and 14-15.

The density-height profiles for the 60°N January models developed for 10°, 100° and 140°W (Figure 14-16) indicate that the longitudinal variation in mean monthly densities at 40 km in winter ranges from 5% less than standard at 140°W to 20% less at 10°W. Density profiles for longitudes 10° and 140°W at 75°N (Figure 14-17), indicate that the lon-

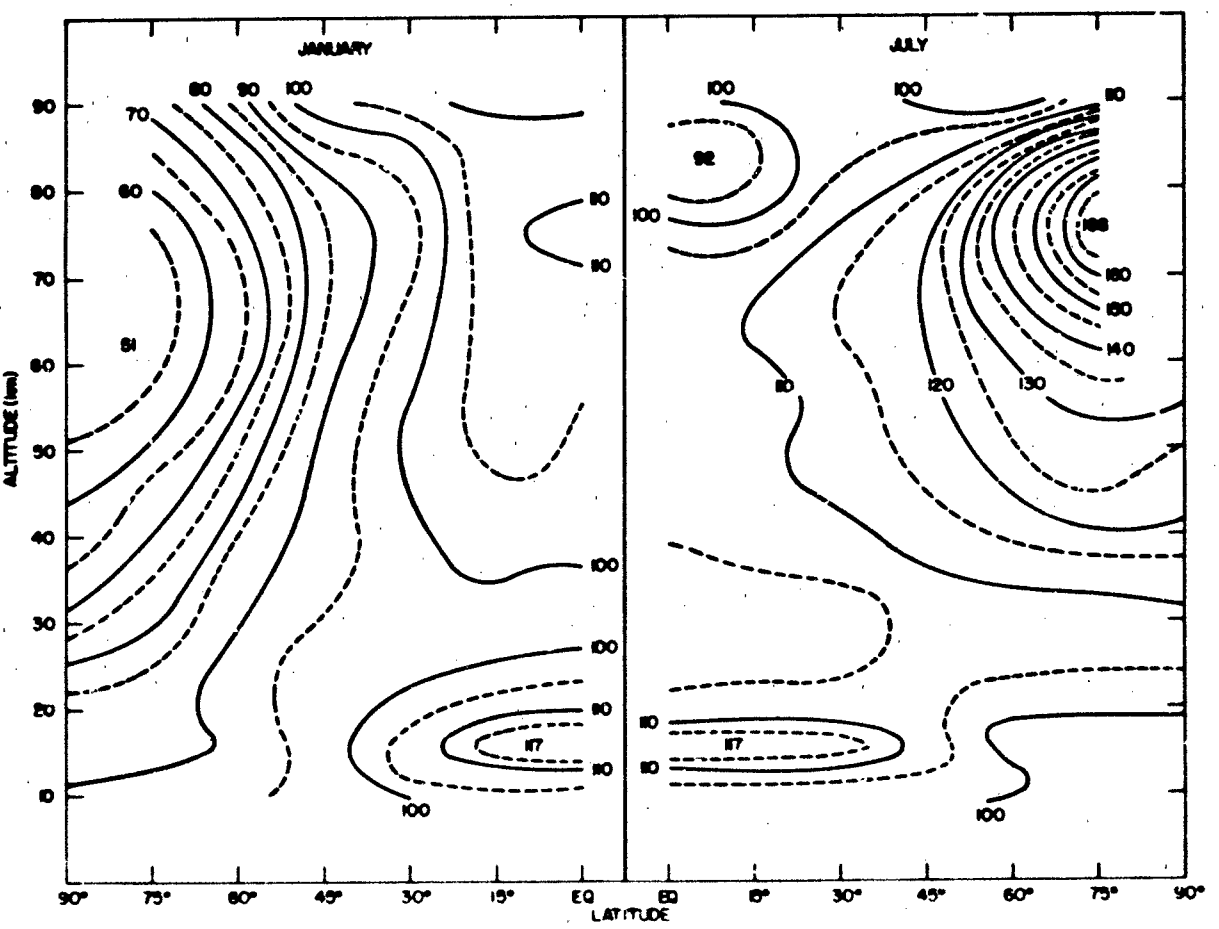


Figure 14-13. Latitudinal density-height cross sections for January and July

Table 14-3a. Temperature (K) in January for 60°N and 75°N at specified longitudes.

Altitude	60°N			75°N	
(km)	10°W	100°W	140°W	10°W	140°W
0.000	278.15	246.13	269.15	257.65	242.15
2.000	265.73	248.64	261.14	252.64	246.95
4.000	253.32	242.13	251.73	244.11	242.12
6.000	240.93	232.14	238.13	232.10	232.34
8.000	228.54	222.14	224.54	220.10	217.34
10.000	216.16	217.15	221.15	213.64	213.65
12.000	216.15	217.15	221.15	212.64	213.65
14.000	216.15	217.15	221.15	211.65	213.65
16.000	216.15	215.96	221.15	210.65	213.65
18.000	214.17	214.76	221.15	209.65	213.65
20.000	212.18	213.57	221.15	206.40	212.16
22.000	207.27	212.37	218.22	200.82	210.17
24.000	202.29	211.18	215.24	195.25	208.18
26.000	204.84	213.07	217.46	195.15	211.27
28.000	207.62	215.06	219.84	199.22	214.45
30.000	210.40	217.04	222.22	203.39	217.63
32.000	213.18	219.03	224.60	207.55	220.80
34.000	215.95	221.94	226.98	211.73	223.98
36.000	220.38	224.91	229.36	215.89	227.15
38.000	224.94	227.88	231.73	221.17	230.32
40.000	229.49	232.65	234.03	226.51	233.49
42.000	234.03	237.59	236.20	231.86	236.66
44.000	238.58	242.53	238.37	237.20	239.82
46.000	243.12	247.47	240.55	242.53	242.80
48.000	247.66	249.85	242.72	244.89	245.77
50.000	252.19	251.82	245.83	246.87	248.73
52.000	256.73	252.15	250.76	248.84	251.69
54.000	257.55	252.15	254.15	250.15	252.15

Table 14.3b Density (kg/m^3) in January for 60°N and 75°N at Specified Longitudes

Altitude	60°N			75°N		
(km)	10°W	100°W	140°W	10°W	140°W	
0.000	1.2555	1.4400	1.3075	1.3608	1.4681	+0*
2.000	1.0219	1.0829	1.0412	1.0612	1.0880	
4.000	8.2374	8.4250	8.2788	8.3467	8.4007	−1
6.000	6.5693	6.5882	6.6212	6.5870	6.5674	
8.000	5.1775	5.0972	5.2271	5.1341	5.1802	
10.000	4.0275	3.8155	3.9010	3.8527	3.8306	
12.000	2.9382	2.7874	2.8660	2.8103	2.7832	
14.000	2.1438	2.0367	2.1061	2.0473	2.0226	
16.000	1.5645	1.4954	1.5479	1.4895	1.4701	
18.000	1.1508	1.0963	1.1379	1.0822	1.0688	
20.000	8.4435	8.0254	8.3672	7.9325	7.8198	−2
22.000	6.2514	5.8655	6.2233	5.8370	5.7198	
24.000	4.5969	4.2802	4.6120	4.2589	4.1721	
26.000	3.2519	3.0803	3.3352	3.0083	2.9732	
28.000	2.3090	2.2230	2.4190	2.0881	2.1291	
30.000	1.6474	1.6094	1.7610	1.4600	1.5325	
32.000	1.1808	1.1689	1.2865	1.0284	1.1085	
34.000	8.5018	8.4851	9.4326	7.2973	8.0574	−3
36.000	6.1087	6.1851	6.9392	5.2134	5.8840	
38.000	4.4174	4.5282	5.1221	3.7332	4.3165	
40.000	3.2158	3.3069	3.7944	2.6943	3.1806	
42.000	2.3561	2.4296	2.8208	1.9597	2.3537	
44.000	1.7370	1.7967	2.1031	1.4361	1.7491	
46.000	1.2881	1.3370	1.5724	1.0599	1.3059	
48.000	0.9607	1.0100	1.1790	0.7959	0.9788	
50.000	7.2057	7.6608	8.8318	6.0021	7.3628	−4
52.000	5.4328	5.8578	6.6017	4.5369	5.5581	
54.000	4.1700	4.4858	4.9924	3.4462	4.2473	

*Power of 10 by which preceding numbers should be multiplied

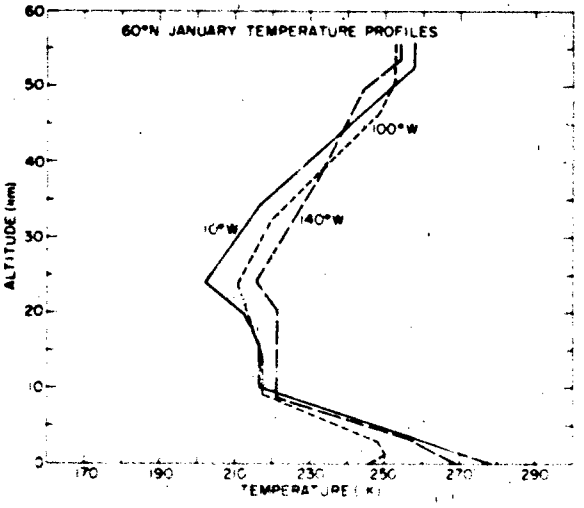


Figure 14-14. Mean monthly temperature height profiles for the 60°N models at longitudes 10°, 100°, and 140°W.

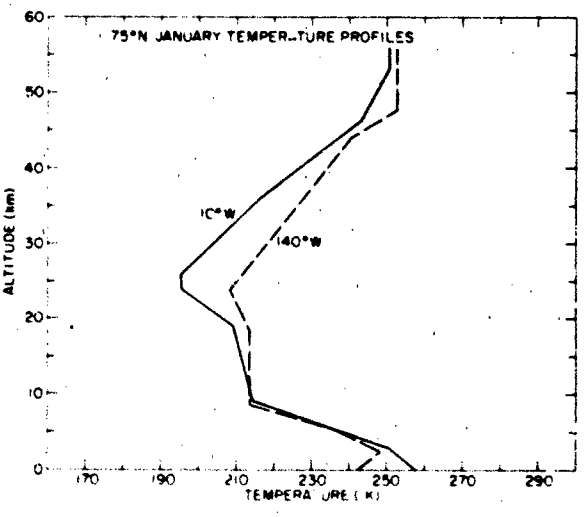


Figure 14-15. Mean monthly temperature-height profiles for the 75°N models at 10° and 140°W.

Table 14-3c. Pressure (mb) in January for 60°N and 75°N at Specified Longitudes

Altitude (km)	60°N 10°W	60°N 100°W	60°N 140°W	75°N 10°W	75°N 140°W	
0.000	1.0025	1.0175	1.0102	1.0065	1.0205	+3*
2.000	7.7954	7.7292	7.6054	7.6963	7.7132	+2
4.000	5.9903	5.8559	5.9823	5.8488	5.8387	
6.000	4.5435	4.3901	4.5261	4.3886	4.3801	
8.000	3.3967	3.2504	3.3693	3.2438	3.2318	
10.000	2.4991	2.3783	2.4761	2.3628	2.3493	
12.000	1.8230	1.7375	1.8194	1.7154	1.7069	
14.000	1.3301	1.2695	1.3370	1.2438	1.2404	
16.000	9.7072	9.2705	9.8268	9.0069	9.0164	+1
18.000	7.0753	6.7589	7.2240	6.5134	6.5550	
20.000	5.1428	4.9201	5.3116	4.6998	4.7625	
22.000	3.7194	3.5758	3.8983	3.3648	3.4508	
24.000	2.6695	2.5947	2.8495	2.3870	2.4933	
26.000	1.9122	1.8840	2.0819	1.6852	1.8032	
28.000	1.3762	1.3723	1.5266	1.1941	1.3106	
30.000	0.9949	1.0027	1.1233	0.8524	0.9573	
32.000	7.2258	7.3493	8.2951	6.1279	7.0263	+0
34.000	5.2703	5.4058	6.1459	4.4351	5.1805	
36.000	3.8645	3.9933	4.5687	3.2309	3.8367	
38.000	2.8523	2.9621	3.4072	2.3701	2.8539	
40.000	2.1184	2.2084	2.5490	1.7519	2.1318	
42.000	1.5829	1.6570	1.9126	1.3043	1.5990	
44.000	1.1896	1.2508	1.4391	0.9778	1.2041	
46.000	0.8990	0.9498	1.0858	0.7379	0.9102	
48.000	6.8305	7.2441	8.2146	5.5957	6.9055	−1
50.000	5.2165	5.5378	6.2323	4.2534	5.2570	
52.000	4.0037	4.2399	4.7520	3.2407	4.0156	
54.000	3.0829	3.2468	3.6422	2.4746	3.0742	

*Power of 10 by which preceding numbers should be multiplied.

gitudinal variability is slightly smaller at 75°N than at 60°N. The lowest mean monthly densities between 35 and 55 km occur at 10°W at both 60° and 75°N.

14.2.1.5 Cold and Warm Winter Stratosphere/Mesosphere. In arctic and subarctic regions, sudden warmings and coolings of the winter stratosphere and mesosphere produce large changes in the vertical and horizontal structure of the atmosphere. Both the magnitude and altitude of maximum temperature and density fluctuations during major warmings and coolings vary considerably. Some of the largest changes in the vertical temperature profiles have been observed in the upper stratosphere between 35 and 45 km. The observed 35- to 45-km temperatures have a range of roughly 85 K in winter compared with 20 K in summer. As a result, mean monthly atmospheric models for the winter months are of limited value for specifying temperatures at

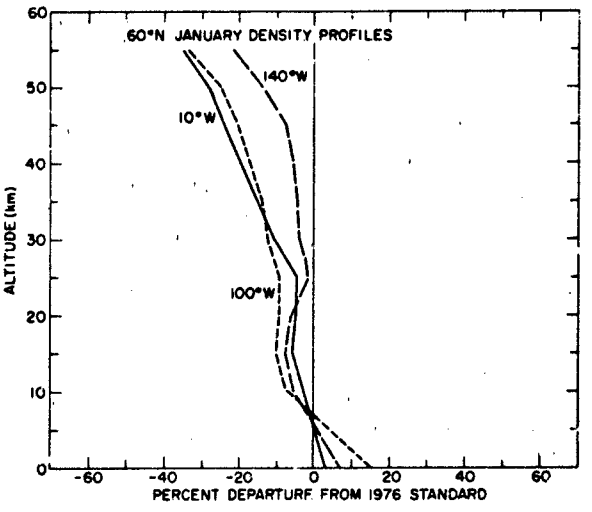


Figure 14-16. Density-height profiles for the 60°N models at 10°, 100°, and 140°W.

these altitudes since the day-to-day variations in temperature are in some cases as great or greater than the seasonal or latitudinal changes. Although these warmings and coolings occur throughout the arctic and subarctic region, the largest changes generally occur between latitude 60° and 75°N; they have been observed much more frequently at some longitudes than at others.

A family of warm and cold atmospheric models typical of the region between 60° and 75°N has been prepared to provide an indication of the magnitude of the variations that can occur in the vertical distributions of temperature, density, and pressure in winter for altitudes up to 90 km. The

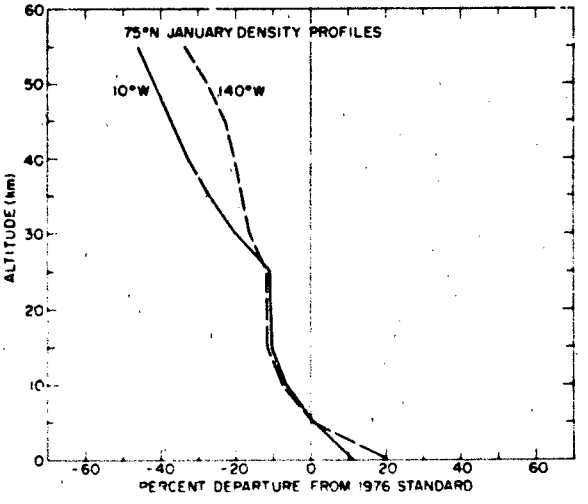


Figure 14-17. Density height profiles for the 75°N models at 10° and 140°W

atmospheric properties representative of one cold and three warm stratospheric regimes that occur at these latitudes are given in Table 14-4. The temperature-height profiles are shown in Figure 14-18. They are based on radiosonde, meteorological rocket, and experimental observations (grenades, falling spheres, and pressure gauges) taken at Poker Flat, 64°N, Ft. Churchill, 59°N, Pt. Barrow, 71°N and West Geirinish 57°N during the period 1965 to 1976. Mean January conditions at 60°N are assumed below 9 km since the temperature-height profiles for this region during the various

Table 14-4 High-latitude thermodynamic properties of warm and cold winter stratosphere/mesosphere.

Altitude (km)	Model A (Warm)	Model B (Warm)	Model C (Warm)	Model D (Cold)	Model A (Warm)	Model B (Warm)	Model C (Warm)	Model D (Cold)		Model A (Warm)	Model B (Warm)	Model C (Warm)	Model D (Cold)	
	Temperature (K)				Density (kg/m^3)					Pressure (mb)				
0 000	257 15	257 15	257.15	257 15	1.3742	1.3742	1.3742	1.3742	+0	1.0144	1.0144	1.0144	1.0144	+3*
2 000	255 94	255 94	255.94	255.94	1.0590	1.0590	1.0590	1.0590		7.7806	7.7806	7.7806	7.7806	+2
4 000	247 73	247 73	247 73	247.73	8.3466	8.3466	8.3466	8.3466	−1	5.9354	5.9354	5.9354	5.9354	
6 000	234 13	234 13	234 13	234.13	6.6507	6.6507	6.6507	6.6507		4.4699	4.4699	4.4699	4.4699	
8 000	220 54	220 54	220.54	220 54	5.2286	5.2286	5.2286	5.2286		3.3102	3.3102	3.3102	3.3102	
10 000	217 15	217 15	217 15	219.39	3.8811	3.8811	3.8811	3.8460		2.4192	2.4192	2.4192	2.4221	
12 000	215 16	215 66	214.17	222.38	2.8495	2.8534	2.8716	2.7866		1.7661	1.7664	1.7654	1.7789	
14 000	211 17	212 66	208 18	223 15	2.1163	2.1050	2.1394	2.0461		1.2829	1.2850	1.2785	1.3106	
16 000	207 19	209 68	202.21	223 15	1.5577	1.5465	1.5806	1.5080		9.2646	9.3087	9.1750	9.6600	+1
18 000	203 20	206 69	196.23	223 15	1.1399	1.1314	1.1574	1.1116		6.6496	6.7132	6.5200	7.1210	
20 000	199 22	203 70	196.15	223 15	8.2929	8.2419	8.1870	8.1965	−2	4.7425	4.8194	4.6097	5.2503	
22 000	195 24	200 72	204 93	222 17	5.9953	5.9767	5.5825	6.0670		3.3601	3.4436	3.2840	3.8693	
24 000	194 56	201 50	213.88	221 18	4.2395	4.2401	3.8671	4.4895		2.3677	2 4525	2.3742	2.8481	
26 000	197 54	206 46	222 82	220 18	2.9535	2.9667	2.7200	3.3130		1.6747	1.7583	1.7397	2.0940	
28 000	200 51	211 43	231.75	219.19	2.0692	2 0939	1.9402	2.4439		1.1910	1.2708	1.2908	1.5377	
30 000	203 49	216 39	240 68	218 20	1.4576	1 4901	1.4021	1.8006		0.8514	0.9256	0.9687	1.1278	
32 000	215 87	227 36	253 37	217.21	0 9942	1 0947	1.0122	1.3251		6.1609	6.8189	7.3624	8.4621	+0
34 000	228 75	238 26	266 25	214 35	6 9184	7.4429	7.4226	9.8119	−3	4.5429	5.0991	5.6730	6.0374	
36 000	241 63	248 52	279 13	211 38	4 9118	5.4164	5.5241	7.2403		3.4068	3.8640	4.4261	4.3933	
38 000	254 49	256 44	275.59	208 41	3.5505	4 0156	4.3868	5.3208		2.5938	2.9560	3.4704	3.1832	
40 000	267 35	264 35	270 64	211 39	2.6084	3.0049	3.4878	3.8002		2.0018	2.2802	2.7096	2.3060	
42 000	282 87	272 26	265 70	214 95	1.9283	2 2683	2.7617	2.7224		1.5657	1.7727	2.1063	1.6798	
44 000	298 67	278 41	260 76	219 78	1.4478	1.7360	2.1775	1.9506		1.2413	1.3874	1.6299	1.2306	
46 000	296 64	280 65	255 83	226 69	1.1640	1.3537	1.7093	1.3979		0.9912	1.0906	1.2553	0.9097	
48 000	285 79	276 24	250 89	233 60	0 9585	1.0732	1.3357	1 0121		7.8635	8.5719	9.6204	6.7872	−1
50 000	274 94	274 30	246 43	240 50	0.7834	0 8530	1.0370	0.7398		6 1836	6.7169	7.3358	5.1080	
52 000	264 10	270 36	242 89	245 76	6 3533	6 7591	7.9903	5 4900	−4	4.8166	5.2456	5 5710	3.8730	
54 000	253 27	266 42	239 34	250 68	5 1074	5 3380	6.1340	4 1041		3.7132	4.0823	4.2143	2.9533	
56 000	244 48	260 58	235 80	255 60	4 0350	4.2290	4 6913	3.0861		2.8317	3.1633	3.1754	2 2643	
58 000	240 55	254 29	232 26	255 87	3.1084	3 3379	3 5739	2.3716		2.1464	2.4365	2.3828	1.7419	
60 000	236 61	248 00	228 72	254 89	2.3847	2.6194	2.7118	1 8302		1.6197	1.8647	1.7804	1.3391	
62 000	232 69	241 71	225 18	253 91	1 8217	2 0431	2.0492	1.3112		1.2168	1.4176	1.3246	1.0285	
64 000	228 76	235 43	221 65	252 92	1 3855	1 5835	1 5419	1 0872		0.9098	1.0701	0.9810	0.7893	
66 000	224 83	229 15	218 11	251 94	1.0489	1.2190	1 1551	0 8369		6 7699	8 0188	7.2325	6 0527	−2
68 000	220 91	222 87	214 59	250 28	7 9039	9 3184	8 6146	6.4526	−5	5 0122	5 9616	5.3064	4 6359	
70 000	216 99	216 60	211 32	248 32	5.9267	7 0697	6 3867	4 9718		3 6917	4 3957	3.8742	3 5440	
72 000	213 07	210 33	209 36	246 36	4 4216	5 3212	4 6899	3 8235		2 7045	3 2128	2 8 86	2 7040	
74 000	209 16	204 07	207 40	244 40	3 2815	3 9716	3 4346	2 9348		1 9702	2.3266	2.0446	2 0590	
76 000	205 35	199 74	205 45	242 45	2 4210	2 9126	2 5084	2 2482		1 4271	1 6699	1 4793	1 5646	
78 000	202 41	196 60	203 49	239 89	1 7700	2.1123	1.8268	1 7234		1.0284	1 1921	1.0671	1 1863	
80 000	199 48	193 47	201 54	236 67	1 2884	1 5244	1.3266	1 3195		7 3779	8 4666	7 6749	8 9645	−3
82 000	198 15	190 35	199 59	233 55	0 9268	1 0945	0 9605	1 0068		5 2718	5 9806	5 5034	6 7501	
84 000	198 15	187 23	197 63	230 43	6 6205	7 8172	6 9347	7 6561	−6	3 7657	4 2013	3 9342	5 0642	
86 000	198 15	184 68	196 65	227 31	4 7301	5 5370	4 9689	5 8011		2 6904	2 9354	2 8049	3 7852	
88 000	198 15	182 73	196 65	223 82	3 3802	3 8948	3 5417	4 3861		1 9226	2 0430	1 9992	2 8181	
90 000	198 15	180 79	196 65	219 93	2.4160	2 7299	2.5250	3 3072		1 3742	1 4167	1 4253	2 0879	

*Power of 10 by which preceding numbers should be multiplied

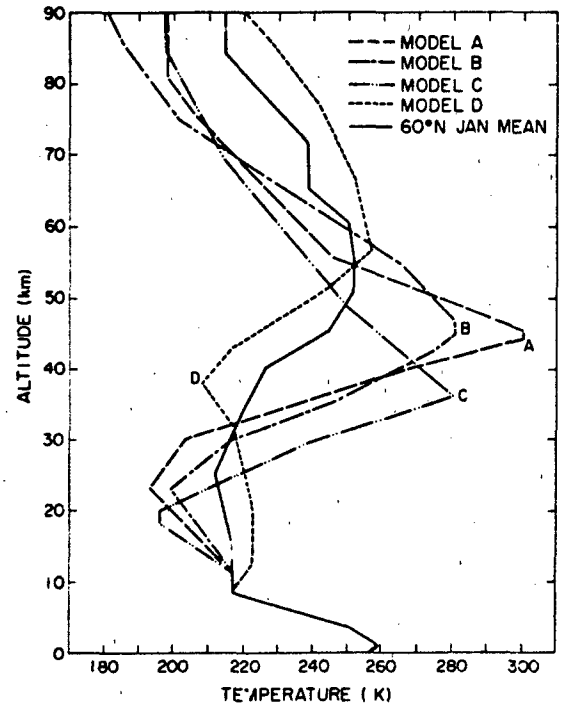


Figure 14-18. Temperature-height profiles associated with extreme warm and cold regimes in the winter stratosphere and mesosphere near 60°N.

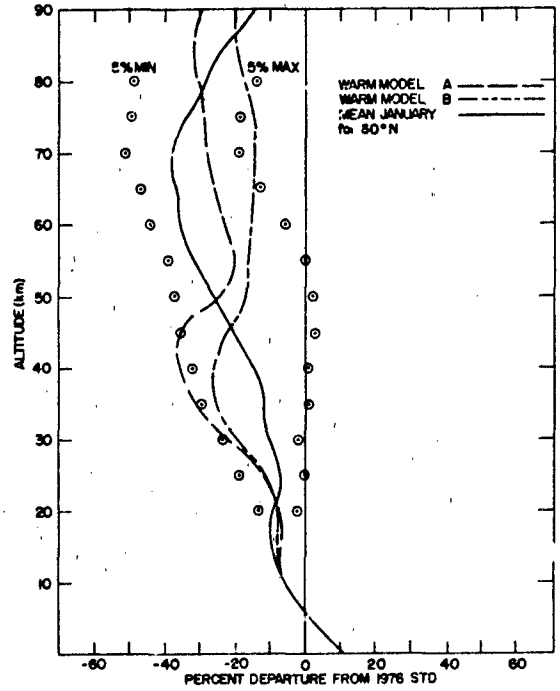


Figure 14-19. Density profiles (warm) associated with extreme temperatures in the upper stratosphere and mean January conditions at 60°N. Circled points form an envelope of high and low densities equaled or surpassed 5% of time.

warmings and coolings are not significantly different from the mean January 60°N atmosphere.

The three warm models could all occur during various stages of one large-scale warming. However, available observations indicate that a temperature of 300 K at 45 km is equaled or exceeded 2% of the time at West Geirinish and 0.4% of the time at Ft. Churchill during January, whereas a temperature of 280 K at 45 km is equaled or exceeded 10% of the time at West Geirinish and 4% of the time at Ft. Churchill. A temperature of 280 K near 36 km is equaled or exceeded 0.6% of the time at West Geirinish and 0.1% of the time at Ft. Churchill. Frequencies of occurrence were obtained by plotting the observed temperature distributions on probability paper.

The cold profile, Model D, is based on an average of five observations in which the temperature at 45 km was within 2° of 223 K. Observed data indicate that a temperature of 223 K or colder occurs at Ft. Churchill 6% of the time, at West Geirinish 4% of the time, and at Poker Flat 9% of the time in January.

The portions of the temperature-height profiles between 55 and 85 km are based on estimates obtained by using interlevel temperature correlations with the temperatures adopted at 40, 45, and 50 km.

The density profiles associated with both the warm and cold models are provided along with the mean January 60°N profile in Figures 14-19 and 14-20. The densities are portrayed as percent departures from the 1976 Standard Atmosphere. Envelopes of the high and low values of density which are equaled or surpassed 5% of the time at 60°N in January are also shown. They are envelopes rather than realistic profiles since 5% values do not occur simultaneously at all altitudes. The density profiles for the warm and cold models illustrate the negative correlations that exist between the densities at various levels in the atmosphere [Quiroz, 1971; Labitzke, 1971; Cole, 1972]. For example, when the density is much less than the mean monthly value at altitudes between 25 and 40 km (Figure 14-20), it is greater than the mean value between 45 and 75 km. In most cases the departures of density from the monthly mean fall within the 5% envelope. However, as shown in Figure 14-20, density profiles associated with an extreme winter warming or cooling will approach both the 5% maximum and 5% minimum values at different altitudes.

The altitudes of the maximum density departures from the monthly mean are related to the altitudes of maximum temperature deviations in that the maximum density departures are roughly 10 to 20 km above the maximum temperature deviations. For example, the largest positive density departure for profile C (Figure 14-20) occurs near 49 km, whereas the maximum stratospheric temperature, 280 K for profile C (Figure 14-18), is at 36 km. The largest

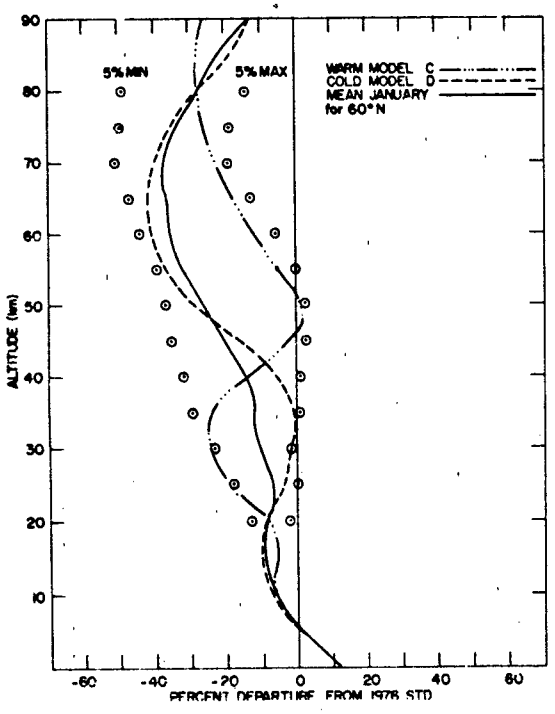


Figure 14-20. Density profiles associated with extreme temperatures in the upper stratosphere and mean January conditions at 60°N. Circled points form an envelope of high and low densities equaled or surpassed 5% of time.

negative density departure for the same profile occurs near 33 km and its minimum stratospheric temperature, 196 K is at 18-20 km.

14.2.2 High Altitude Reference Atmospheres

Although there were early high altitude model atmospheres based on upward extrapolation of rocket data, they were not accurate and the first official U.S. models of the satellite era, including density values derived from satellite drag, were the *ARDC Model Atmosphere, 1959* [Minzner, 1959] and the *U.S. Standard Atmosphere, 1962* [COESA]. These models accurately represented thermospheric density at the time the observations were made, but since the observations were made at the peak of solar cycle 19 which had unusually high activity, they represente ' conditions that are not frequently observed.

The next generation of reference atmospheres started with the COSPAR International Reference Atmosphere (CIRA) 1965 (COSPAR). These atmospheres included a Mean CIRA between 30 and 300 km prepared by Champion and a set of models for 120 to 800 km prepared by Harris and Priester. The latter are semi-theoretical models but with some free parameters whose values were chosen so that the models reproduced observed densities. The following year the *U.S. Standard Atmosphere Supplements, 1966* [COESA] were published. This publication included models up to 1000 km. The lower thermospheric models were based on work by Champion and the upper thermospheric models on work by Jacchia. An updated set of COSPAR models was published in 1972 as *CIRA 1972* (COSPAR). In these models Champion and Schweinfurth [1972] prepared the Mean CIRA for altitudes between 25 and 500 km and Jacchia [1971] prepared a set of models for the altitude range 110 to 2000 km. More recent models include the U.S. Standard Atmosphere, 1976 (COESA), and the Jacchia [1977] model.

The preceding models have all been based principally on density data from orbital drag and *in situ* measurements of density and composition. Recently, models have been developed which are also based on airglow temperatures (DTM) [Barlier et al., 1978] or incoherent radar scatter temperatures (MSIS) [Hedin et al., 1977]. A number of theoretical models are being developed, including the NCAR thermospheric general circulation model by Roble and colleagues [Dickinson et al., 1981]. This is a three-dimensional model of the global neutral gas temperatures and circulation of the thermosphere. Rees and colleagues [Fuller-Rowell and Rees, 1980] are also developing a three-dimensional global model of the thermosphere. They start with a realistic steady-state model and then progressively modify it in response to the sources and sinks of energy and momentum and to the winds that result from the various driving forces redistributing mass, momentum, and energy. Important components of theoretical models are the amplitudes and phases of the tidal effects. The principal tides are the solar diurnal, semidiurnal and terdiurnal. The magnitudes of these tides and their effects on density and composition are reviewed by Champion [1981].

14.2.2.1 Development of Reference Atmospheres. Since the earliest upper atmosphere models, it has been customary to define them in part by means of temperature profiles which, by means of the appropriate physical relations, yield density profiles, and in the more modern models, composition and other properties.

At this point we will review the temperature profiles of some models of particular interest. Figure 14-21 shows the temperature profiles of the U.S. Standard 1962, Mean CIRA 1965, Mean CIRA 1972 and U.S. Standard 1976. It can be seen that as time progressed the mean model temperatures in the thermosphere have become lower. There is a good reason for this which can be understood by referring to Figure 14-22. This figure shows a plot of the mean annual number of sunspots from 1820 to 1976. The 11-yr cycle of activity can be seen and also the wide variation in the number of sunspots at the cycle peaks. Table 14-5 shows that the succeeding models, which are based on satellite orbital drag and *in situ* measurements, have sampled the atmosphere starting from very non-typical conditions in 1957 to increasingly representative conditions with time. In addition, Slowey [1979] has done a study using both sunspot numbers

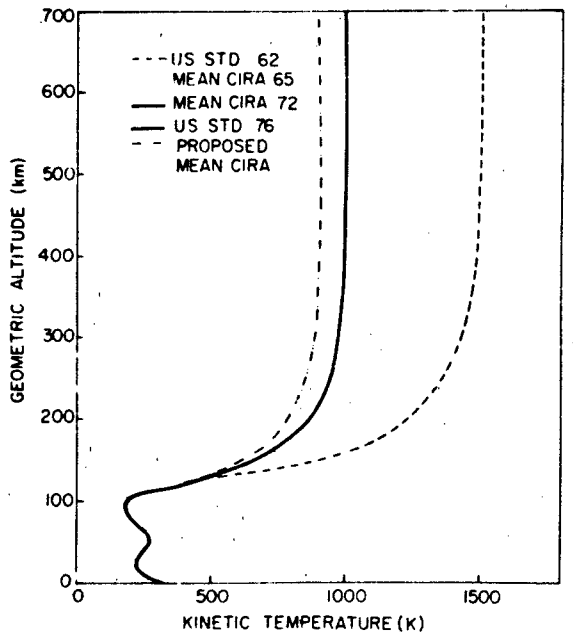


Figure 14-21. Kinetic temperature profiles for several standard and reference atmospheres, including a proposed new mean CIRA.

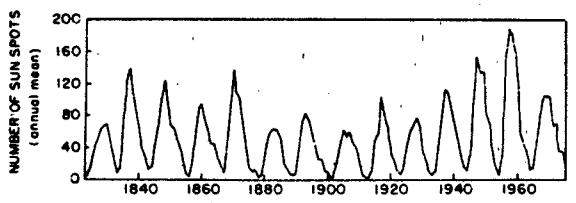


Figure 14-22. The mean annual number of sunspots from 1820 to 1976.

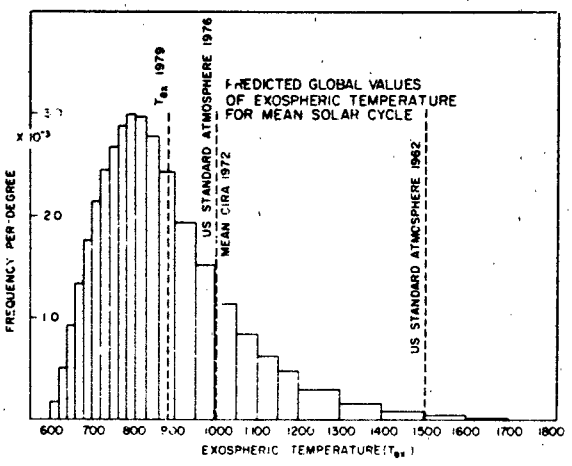


Figure 14-23. Histogram of frequency per degree of predicted exospheric temperatures for a mean solar cycle.

Table 14-5. Average exospheric temperature.

	AVERAGE
US STANDARD ATMOSPHERE 1962 (3 YRS)	1500 K
MEAN CIRA 65	1200 K
MEAN CIRA 72	1000 K
US STANDARD ATMOSPHERE 1976 (1½ SOLAR CYCLES)	1000 K
SPECIAL STUDY (12 SOLAR CYCLES)	882 K

back to 1847 (12 cycles) and the Jacchia 1971 model to determine relative values and has found that the expected mean exospheric temperature averaged over all conditions for this period would be 882 K. A histogram showing the predicted exospheric temperature distribution is given in Figure 14-23. The exact mean temperature will depend on the time period chosen (for example, including the Maunder minimum would have a profound effect) and the model used to determine relative values. However, it is recommended that 900 K be used for the exospheric temperature for the next Mean CIRA as being appropriate to this phase of solar activity. The corresponding density profiles of the four previous standard or mean models and that of the proposed new model are shown in Figure 14-24.

The most up-to-date set of approved reference atmospheres for altitudes above 90 km is that contained in *CIRA 72*. There are plans to prepare revised versions of both CIRA and the U.S. Standard Atmosphere Supplements. The latest published version of the latter is the *U.S. Standard Atmosphere Supplements, 1966*.

The CIRA 1972 models up to 120 km were prepared by Groves of University College, London, England and the Air Force Cambridge Research Laboratory (AFCRL). An independent set of models for the region 110 to 2000 km were prepared by Jacchia of the Smithsonian Astrophysical

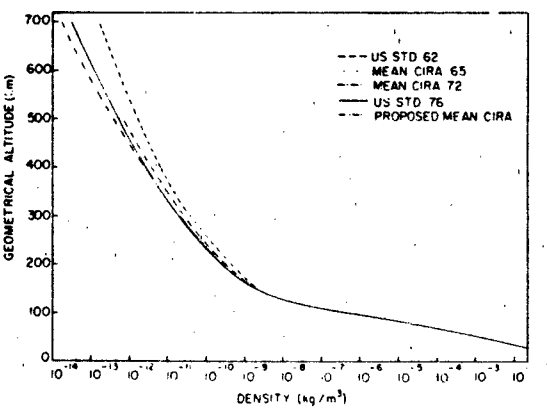


Figure 14-24. Density profiles of several standard and reference atmospheres, including a proposed new mean CIRA.

Observatory. Since these two sets of models do not, in general, match at 110 km, Champion and Schweinfurth of AFCRL developed the Mean CIRA which is continuous from 25 to 500 km and applies under specified conditions.

14.2.2.2 Mean CIRA Reference Atmosphere. The basis of the mean reference atmosphere is as follows: Between 25 and 75 km the model represents annual mean conditions for latitudes near 30°. Between 120 and 500 km the model corresponds to diurnal, seasonal, and semi-annual variation average conditions for a latitude near 30° and a solar flux F of 145×10^{-22} W m^{-2} Hz^{-1}. Between 75 and 120 km a model has been developed which provides a smooth connection between the lower and upper sections of the mean atmosphere. This atmosphere contains kinetic temperature, molecular-scale temperature, density, pressure, density and pressure scale heights, mean molecular weight, densities of major constituents, and total number densities.

The data used to develop the model between 25 and 75 km were the values at 30° latitude of the annual mean pressure at 25 km and the annual mean temperature between 25 and 75 km. The pressure equation was integrated numerically.

$$p = p_1 \exp\left[-\frac{M_0}{R} \int_{z_1}^{z} g\,dz/T_M \right] \qquad (14.17)$$

where

p = pressure at altitude z

p_1 = pressure at reference altitude z_1

M_0 = sea level value of mean molecular weight = 28.96

$\mathcal{R}$ = 8.31432×10^7 ergs K^{-1} g mol^{-1}

g = acceleration due to gravity at 30° latitude

T_M = molecular-scale temperature

$$= \frac{M_0 T}{M}$$

where

M = mean molecular weight

T = kinetic temperature.

The model between 75 and 120 km was constrained to provide a transition between the low altitude model and the high altitude model. These two models are not only different but are functions of different parameters. Thus a compromise was devised. A temperature profile was developed with a constraint that it yield a specified density value at 120 km. The equations and method used to determine the composition are given in Champion and Schweinfurth [1972].

Table 14-6. Mean Reference Atmosphere Structure Parameters 25 to 120 km.

Height km	Mol Temp K	Density kg/m^3	Pressure N/m^2	Number Density m^{-3}	Pressure Scale Ht(km)	g m/s^2
25	221.7	3.899E−02	2.483E+03	8.111E+23	6.55	9.716
30	230.7	1.774E−02	1.175E+03	3.690E+23	6.83	9.701
35	241.5	8.279E−03	5.741E+02	1.722E+23	7.16	9.686
40	255.3	3.972E−03	2.911E+02	8.265E+22	7.58	9.671
45	267.7	1.995E−03	1.535E+02	4.148E+22	7.96	9.656
50	271.6	1.057E−03	8.241E+01	2.198E+22	8.09	9.641
55	263.9	5.821E−04	4.406E+01	1.210E+22	7.87	9.626
60	249.3	3.206E−04	2.296E+01	6.669E+21	7.45	9.611
65	232.7	1.718E−04	1.146E+01	3.568E+21	6.96	9.596
70	216.2	8.770E−05	5.445E+00	1.822E+21	6.48	9.581
75	205.0	4.178E−05	2.460E+00	8.696E+20	6.15	9.566
80	195.0	1.905E−05	1.067E+00	3.964E+20	5.86	9.551
85	185.1	8.337E−06	4.426E−01	1.736E+20	5.57	9.536
90	183.8	3.396E−06	1.795E−01	7.087E+19	5.54	9.521
95	190.3	1.343E−96	7.345E−02	2.808E+19	5.75	9.506
100	203.5	5.297E−07	3.090E−02	1.125E+19	6.16	9.492
105	228.0	2.173E−07	1.422E−02	4.768E+18	6.91	9.477
110	265.5	9.061E−08	7.362E−03	2.182E+18	8.06	9.462
115	317.1	4.645E−08	4.236E−03	1.076E+18	9.64	9.448
120	380.6	2.438E−08	2.667E−03	5.772E+17	11.58	9.433

Table 14-7 Kinetic Temperature and Composition of the Mean Reference Atmosphere, 75 to 120 km

Height km	Temp K	Mean Mol Wt	Log $N(N_2)$ m^{-3}	Log $N(O_2)$ m^{-3}	Log N(O) m^{-3}	Log N(Ar) m^{-3}	Log N(HE) m^{-3}	Log $N(O_3)$ m^{-3}
75	205.0	28.96	20.832	20.261		18.910	15.659	
80	194.9	28.95	20.492	19.914	16.794	18.563	15.315	14.495
85	184.9	28.93	20.135	19.550	17.144	18.199	14.954	14.098
90	183.4	28.89	19.747	19.153	17.220	17.802	14.565	13.425
95	189.3	28.81	19.346	18.739	17.280	17.389	14.178	12.665
100	199.4	28.37	18.940	18.299	17.618	16.945	13.928	12.099
105	216.6	27.51	18.556	17.823	17.647	16.435	13.811	11.178
110	245.1	26.73	18.200	17.398	17.509	15.961	13.711	10.132
115	285.2	26.05	17.872	17.036	17.332	15.539	13.620	9.298
120	334.5	25.45	17.579	16.734	17.153	15.173	13.538	8.167

The exospheric temperature for the conditions specified above is 1000 K. The Jacchia models were recomputed using the acceleration due to gravity for 30° latitude. The original models use a value valid only near latitude 45°. The model values were then changed at all altitudes so that they matched the values at 120 km for the intermediate altitude model.

The properties of the Mean Reference Atmosphere are presented in abbreviated form in Tables 14-6 to 14-9. (The complete tables are provided in CIRA, 1972 [COSPAR, 1972]. Table 14-6 contains values of molecular-scale temperature, density, pressure, number density, pressure scale height, and acceleration due to gravity over the altitude range 25–120 km. Table 14-7 contains values of kinetic temperature, mean molecular weight and log number densities of N_2, O_2, O, Ar, He, and O_3 over the altitude range 75–120 km. In Table 14-8 molecular-scale temperature, density, density scale height, pressure, pressure scale height, and acceleration due to gravity for the altitude range 120–500 km are given. Table 14-9 contains the corresponding values of kinetic temperature, mean molecular weight, number density, and log number density of N_2, O_2, O, Ar, and He for the altitudes 120–500 km.

Some of the properties are illustrated in Figures 14-25 to 14-30. Figure 14-25 shows the kinetic temperature of the mean atmosphere plus curves indicating low extreme and high extreme temperatures whose frequency of occurrence is 1% or less. Figure 14-26 shows the pressure scale heights

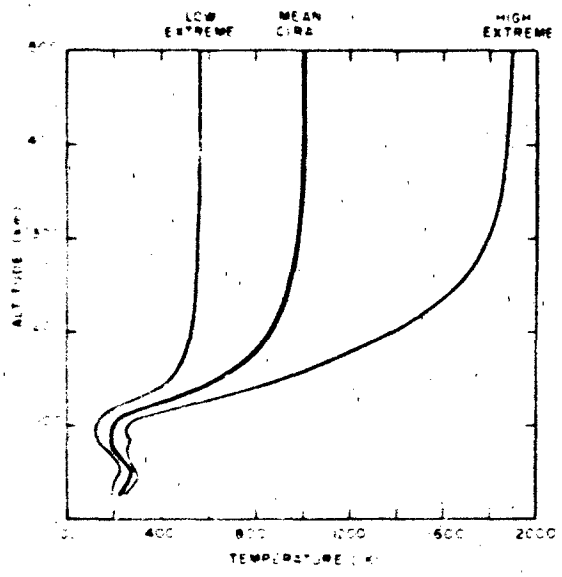


Figure 14-25 Mean CIRA temperatures and low, extreme and high extreme temperatures

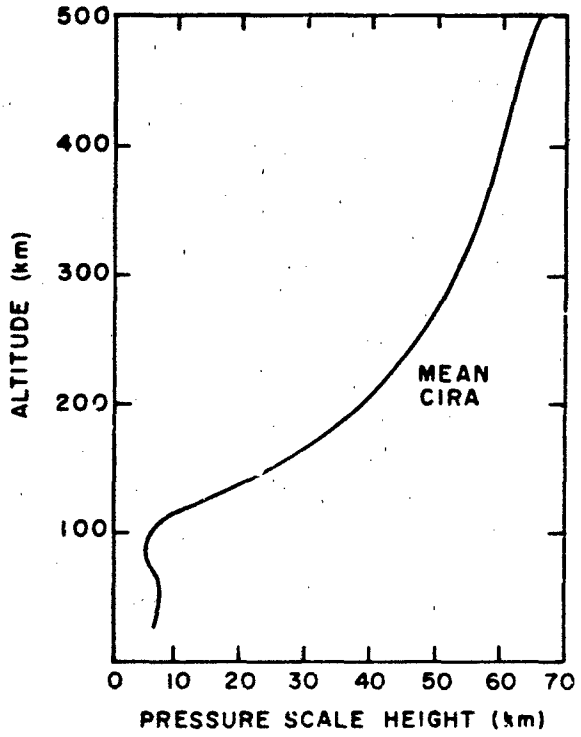


Figure 14-26 Pressure scale heights of the mean CIRA

Table 14-8. Mean Reference Atmosphere Structure Parameters 120 to 500 km.

Height km	Molec Temp K	Density kg/m^3	Density Scale Ht (km)	Pressure N/m^2	Pressure Scale Ht(km)	g m/s^2
120	380.6	2.440E−08	8.17	2.666E−03	11.58	9.433
125	452.3	1.382E−08	9.51	1.795E−03	13.79	9.418
130	526.9	8.484E−09	11.05	1.283E−03	16.08	9.404
135	600.9	5.563E−09	12.70	9.597E−04	18.37	9.389
140	672.4	3.845E−09	14.44	7.423E−04	20.59	9.375
145	739.8	2774E−09	16.21	5.891E−04	22.69	9.360
150	802.7	2.070E−09	17.98	4.770E−04	24.66	9.346
155	861.0	1.587E−09	19.70	3.923E−04	26.49	9.332
160	914.8	1.244E−09	21.37	3.267E−04	28.19	9.317
165	964.7	9.927E−10	22.96	2.749E−04	29.77	9.303
170	1011.1	8.040E−10	24.48	2.334E−04	31.25	9.289
175	1054.4	6.593E−10	25.92	1.996E−04	32.64	9.274
180	1095.0	5.464E−10	27.29	1.718E−04	33.95	9.260
185	1133.2	4.568E−10	28.60	1.468E−04	35.19	9.246
190	1169.3	3.850E−10	29.85	1.292E−04	36.36	9.232
195	1203.4	3.267E−10	31.05	1.129E−04	37.48	9.218
200	1235.8	2.789E−10	32.21	9.896E−05	38.55	9.204
210	1295.9	2.066E−10	34.40	7.686E−05	40.55	9.176
220	1350.3	1.558E−10	36.45	6.039E−05	42.38	9.148
230	1399.6	1.192E−10	38.39	4.792E−05	44.06	9.120
240	1444.4	9.246E−11	40.21	3.834E−05	45.61	9.092
250	1484.9	7.248E−11	41.92	3.090E−05	47.03	9.065
260	1521.5	5.735E−11	43.52	2.505E−05	48.33	9.037
270	1554.6	4.576E−11	45.03	2.042E−05	49.54	9.010
280	1584.5	3.677E−11	46.43	1.673E−05	50.64	8.983
290	1611.5	2.974E−11	47.73	1.376E−05	51.66	8.955
300	1636.0	2.418E−11	48.95	1.136E−05	52.60	8.928
310	1658.1	1.976E−11	50.07	9.405E−06	53.48	8.902
320	1678.2	1.621E−11	51.12	7.812E−06	54.29	8.875
330	1696.5	1.336E−11	52.09	6.507E−06	55.05	8.848
340	1713.4	1.104E−11	52.98	5.432E−06	55.76	8.822
350	1728.8	9.158E−12	53.81	4.545E−06	56.43	8.795
360	1743.2	7.615E−12	54.59	3.811E−06	57.07	8.769
370	1756.7	6.348E−12	55.31	3.202E−06	57.69	8.743
380	1769.5	5.304E−12	55.98	2.694E−06	58.28	8.717
390	1781.8	4.441E−12	56.60	2.272E−06	58.86	8.691
400	1793.7	3.725E−12	57.20	1.918E−06	59.43	8.665
410	1805.4	3.130E−12	57.75	1.622E−06	60.00	8.639
420	1817.2	2.635E−12	58.28	1.374E−06	60.57	8.614
430	1829.1	2.221E−12	58.79	1.166E−06	61.14	8.588
440	1841.3	1.875E−12	59.26	9.910E−07	61.74	8.563
450	1854.0	1.585E−12	59.73	8.435E−07	62.35	8.537
460	1867.4	1.341E−12	60.18	7.191E−07	62.98	8.512
470	1881.6	1.137E−12	60.63	6.140E−07	63.65	8.487
480	1896.9	9.644E−13	61.08	5.252E−07	64.36	8.462
490	1913.4	8.192E−13	61.52	4.500E−07	65.11	8.437
500	1931.4	6.967E−13	61.95	3.863E−07	65.91	8.412

Table 14-9. Kinetic Temperature and Composition of the Mean Reference Atmosphere, 120 to 500 km.

Height km	Temp K	Mean Mol Wt	Number Density m^{-3}	Log $N(N_2)$ m^{-3}	Log $N(O_2)$ m^{-3}	Log N(O) m^{-3}	Log N(Ar) m^{-3}	Log N(HE) m^{-3}
120	334.5	25.45	5.772E + 17	17.579	16.734	17.153	15.173	13.538
125	389.7	24.95	3.336E + 17	17.322	16.449	16.978	14.835	13.469
130	445.4	24.48	2.087E + 17	17.098	16.203	16.826	14.541	13.410
135	499.0	24.05	1.393E + 17	16.903	15.987	16.693	14 284	13.358
140	549.0	23.65	9.793E + 16	16.731	15.796	16.577	14.056	13.314
145	594.5	23.27	7.178E + 16	16.577	15.624	16.474	13.850	13.275
150	635.2	22.92	5.439E + 16	16.437	15.469	16.381	13.663	13.241
155	671.3	22.58	4.233E + 16	16.308	15.325	16.298	13.490	13.212
160	703.1	22.26	3.366E + 16	16.189	15.192	16.221	13.328	13.185
165	731.3	21.95	2.723E + 16	16.077	15.066	16.150	13.176	13.161
170	756.2	21.66	2.236E + 16	15.971	14.947	16.083	13.031	13.139
175	778.4	21.38	1.857E + 16	15.870	14.834	16.020	12.892	13.118
180	798.1	21.11	1.559E + 16	15.773	14 724	15.960	12.758	13.099
185	815.9	20.85	1.319E + 16	15.679	14.619	15.902	12.629	13.081
190	831.9	20.60	1.125E + 16	15.589	14.517	15.847	12.504	13.064
195	846.3	20.37	9.661E + 15	15.501	14.418	15.794	12.382	13.048
200	859.3	20.14	8.342E + 15	15.415	14.321	15.742	12.262	13.033
210	882.0	19.71	6.312E + 15	15.250	14.133	15.642	12.031	13.004
220	900.7	19.32	4.856E + 15	15.090	13 952	15.547	11.807	12.977
230	916.4	18.96	3.788E + 15	14.936	13.777	15.456	11.590	12.951
240	929.4	18.63	2.988E + 15	14.785	13.606	15.367	11.378	12.927
250	940.2	18.34	2.380E + 15	14.638	13.438	15.281	11.171	12.903
260	949.3	18.07	1.912E + 15	14.494	13.274	15.197	10.967	12.881
270	956 8	17.82	1.546E + 15	14.352	13.112	15.114	10.766	12.859
280	963.1	17.60	1.258E + 15	14.212	12.953	15.033	10.567	12.837
290	968.4	17.40	1.029E + 15	14.073	12.795	14.953	10.371	12.816
300	972.8	17.22	8.456E + 14	13.937	12.639	14.874	10.177	12.796
310	976.5	17.06	6.977E + 14	13.801	12.485	14.796	9.984	12.776
320	979.7	16.91	5.776E + 14	13.667	12.332	14.719	9.793	12.756
330	982.3	16.77	4.798E + 14	13.533	12.179	14.642	9.604	12.736
340	984.6	16.64	3.997E + 14	13.401	12.028	14.566	9.415	12.717
350	986.5	16.52	3.338E + 14	13.269	11.878	14.490	9.228	12.698
360	988.1	16.42	2.794E + 14	13.139	11.729	14.415	9.041	12.679
370	989.5	16.31	2.344E + 14	13.008	11.580	14.341	8.856	12.660
380	990.7	16.21	1.970E + 14	12.879	11.432	14.267	8.672	12.641
390	991.7	16.12	1.659E + 14	12.750	11.285	14.193	8.488	12.622
400	992.6	16.03	1.400E + 14	12.621	11.138	14.119	8.305	12.604
410	993.4	15.93	1.183E + 14	12.494	10.992	14.046	8.123	12.585
420	994.1	15.84	1.001E + 14	12.365	10.847	13.973	7.941	12.567
430	994.7	15.75	8.493E + 13	12.239	10.702	13.901	7.761	12.549
440	995.2	15.65	7.213E + 13	12.113	10.558	13.829	7.581	12.531
450	995.7	15.55	6.136E + 13	11.987	10.414	13.757	7.401	12.513
460	996.1	15.45	5.229E + 13	11.862	10.271	13.685	7.222	12.495
470	996.4	15.34	4.464E + 13	11.737	10.128	13.613	7.044	12.477
480	996.8	15.22	3.817E + 13	11.612	9.986	13.542	6.867	12.459
490	997.0	15.09	3.269E + 13	11.488	9.844	13.471	6.690	12.441
500	997.3	14.95	2.806E + 12	11.364	9.703	13.400	6.513	12.423

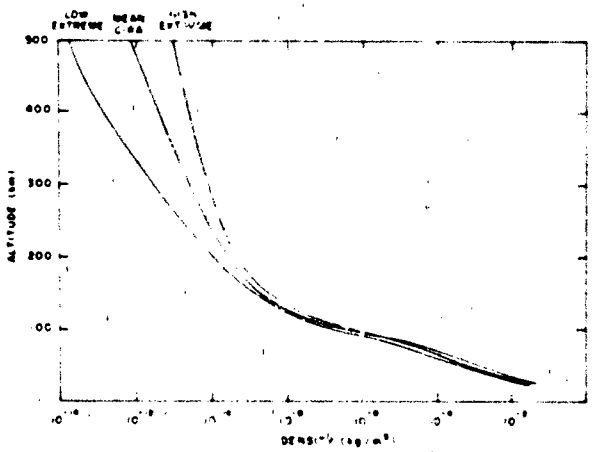


Figure 14-27 Mean CIRA densities and curves of extreme densities

of the mean atmosphere as a function of altitude. Figure 14-27 contains low extreme, high extreme, and mean density values. Total number densities and densities of N_2, O_2, O, O_3, Ar, He, and H are given in Figure 14-28. Figure 14-29 contains the Mean CIRA temperature profile, median warm temperatures and those exceeded 10% and 1% of the time and, similarly, median cold temperatures and those above which 90% or 99%, respectively, of the temperatures lie. The corresponding density curves are shown in Figure 14-30.

14.2.2.3 Reference Atmospheres 90 to 120 km. These empirical atmospheres are based on a report entitled, "Atmospheric Structure and its Variations in the region from 25 to 120 km" by Groves [1971]. In this section only the properties above 90 km are presented. The upper altitude of the models varies with latitude and time of year, de-

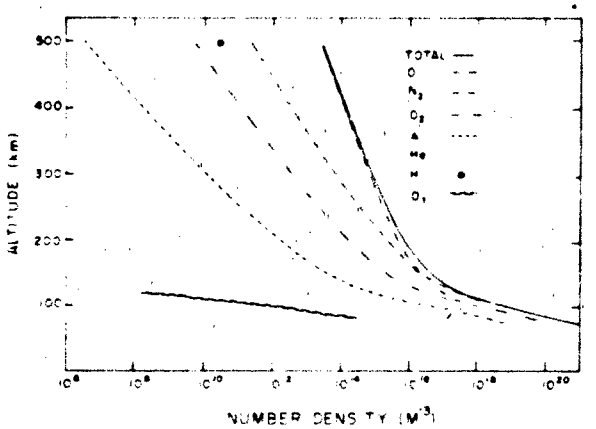


Figure 14-28 Total number densities and densities of N_2, O_2, O, O_3, A, He, and H

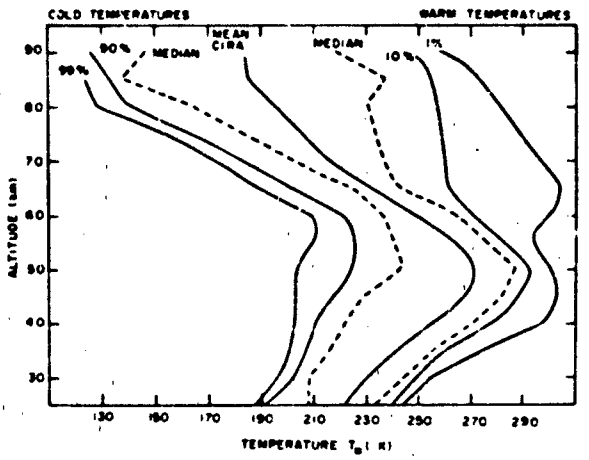


Figure 14-29 Mean CIRA temperatures and temperatures which are exceeded 50%, 10%, 1% of the time during warmest months, and temperatures exceeded 50%, 90%, and 99% of the time during coldest months, at latitudes between 0° and 80°N.

pending on the availability of data. Above 90 km the data are primarily from rocketborne falling sphere, grenade and chemical release techniques. Table 14-10 contains average values of zonal (W - E) winds at latitudes from the equator to 50° for the first day of each month. Data used are from all longitudes, with southern hemisphere data shifted by six months. Mean temperature values at latitudes from the equator to 70° are given in Table 14-11. Less accurate values (based on few data points) are indicated by an asterisk. Table 14-12 gives the corresponding pressure values and Table 14-13 the density values.

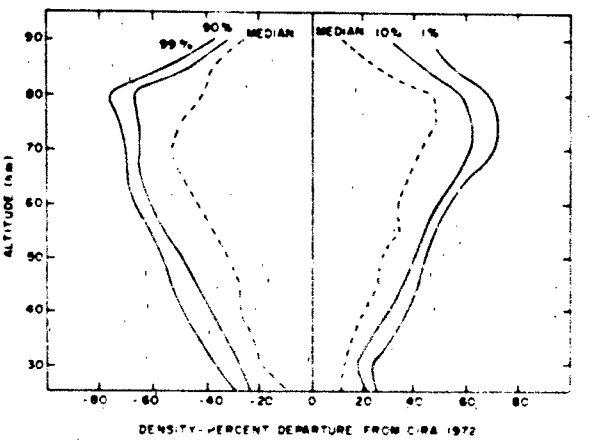


Figure 14-30 Densities relative to Mean CIRA exceeded 50%, 10%, and 1% of the time during months with highest densities, and densities exceeded 50%, 90%, and 99% of the time during months with lowest densities, at latitudes between 0° and 80°N.

Table 14-10. W-E Winds 90 to 130 km based on data from all longitudes with S. Hemisphere data shifted six months in time. Winds to the east are positive in m/s. Values apply to the first day of each month

km	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0 Degrees N												
90	−9	−31	−22	−34	−15	−23	−9	−31	−22	−34	−15	−23
95	−28	−23	−13	−21	−22	−32	−28	−23	−13	−21	−22	−32
100	−28	−3	7	−3	−22	−29	−28	−3	7	−3	−22	−29
105	−7	22	26	3	−11	−11	−7	22	26	3	−11	−11
110	17	37	30	−13	10	14	17	37	30	−13	10	14
115	31	37	20	−29	32	33	31	37	20	−29	32	33
10 Degrees N												
90	−14	−30	−36	−54	−5	−7	0	−21	4	3	−2	−22
95	−39	−41	−43	−52	−15	−8	−14	−5	15	14	−5	−36
100	−40	−29	−27	−25	−19	1	−21	19	28	15	−10	−43
105	−20	8	7	−3	−10	5	1	36	31	2	−10	−21
110	1	41	27	−20	9	3	22	40	19	−17	2	20
115	18	49	27	−44	30	13	30	32	1	−27	19	41
20 Degrees N												
90	−8	−29	−23	−32	16	17	12	5	19	22	19	−6
95	−23	−36	−34	−40	6	21	−6	23	25	28	18	−18
100	−28	−16	−48	−36	−2	22	−12	45	29	15	9	−32
105	−18	17	−41	−28	−3	11	−3	55	20	−11	−3	−24
110	0	34	−21	−31	5	2	7	46	−1	−27	−8	5
115	20	27	−7	−36	19	8	12	26	−19	−25	−1	31
30 Degrees N												
90	5	−18	15	0	29	34	22	14	24	27	45	11
95	0	−19	17	−16	25	42	1	31	12	21	38	−3
100	−8	5	0	−30	18	40	−1	50	6	−5	21	−19
105	−9	39	−8	−31	4	15	0	63	−6	−39	−2	−25
110	2	36	1	−26	−4	−12	0	52	−28	−43	−19	−11
115	11	0	12	−22	−1	−18	−5	26	−44	−21	−30	15
120	7	−27	10	−21	−1	−7	−7	−1	−42	−5	−28	34
125	−12	−39	−1	−15	−12	−4	−13	−17	−25	−16	−17	37
130	−33	−50	−13	−8	−24	−4	−19	−32	−9	−32	−2	38
40 Degrees N												
90	31	32	20	−6	−10	28	34	15	29	24	29	37
95	17	15	17	−3	20	54	43	27	22	23	20	37
100	7	14	32	6	57	60	58	33	19	28	−4	31
105	1	17	20	12	73	36	65	28	9	29	−19	17
110	1	6	−33	3	56	−7	49	16	−21	5	−15	9
115	1	−18	7	−16	19	−43	13	−4	−55	−39	−9	4
50 Degrees N												
90	28	21	21	−18	−18	−13	8	22	27	8	22	20
95	14	21	19	−20	−6	12	18	18	20	−2	7	10
100	8	8	13	−3	27	18	43	11	12	−5	−13	11

The preceding tables contain mean climatological values. One variation of these values that is important and can be readily modeled is that due to solar tides. Figure 14-31 contains the amplitude and phase of the solar diurnal tide in density as a function of latitude and Figure 14-32 contains the corresponding amplitude and phase of the solar semidiurnal tide. Note that the tide amplitudes are quite small at low altitudes, but rapidly increase at higher altitudes. These values were calculated theoretically by Forbes [Private Communication, 1982].

Table 14-11 Temperatures (K) 90 to 110 km

km	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0 Degrees N												
90	185	193	198	200	193	189	185	193	198	200	193	189
95	187	200*	204*	199*	187*	184*	187	200*	204*	199*	187*	184*
100	204	219*	220*	209*	193*	191*	204	219*	220*	209*	193*	191*
105	231	251*	255*	239*	219*	215*	231	251*	255*	239*	219*	215*
110	273	296*	306*	293*	269*	261*	273	296*	306*	293*	269*	261*
10 Degrees N												
90	185	191*	193	197	194	189	186	194	199	200	192	188
95	187*	197*	199	193	187*	186	189	199	207*	203*	191	184
100	203*	215*	214*	201*	191*	193*	204	216*	225*	217*	200	193
105	232*	246*	245*	229*	215*	214*	229	247*	257*	247*	227*	220*
110	276*	290*	291*	278*	259*	253*	265	291*	307*	301*	282*	270*
20 Degrees N												
90	186*	189	190	194	192	188	186	191	197	198	191	189
95	190*	197	195	190*	187*	189	193	197	206*	206*	196	187
100	204*	213	210	195*	194*	198*	206	211*	222*	222*	210	197*
105	228*	237*	236*	220*	216*	221*	229	236*	248*	250*	239*	225*
110	271*	278*	275*	264*	255*	253*	260	276*	294*	299*	286*	271*
30 Degrees N												
90	189*	190	189	190	185	179	179	187	193	195	194	194
95	195*	197	195	189*	187	189*	195	197	205	205	202	196
100	204*	210	208	196*	197*	206*	214	208	216	221	219	204*
105	222*	230*	227	216*	220*	232*	237	226	236*	249*	245*	225*
110	259*	266*	265	256*	259*	264*	265	263	278*	290*	284*	263*
40 Degrees N												
90	199*	195	191	185	172	162	167	180	191	197	200	203
95	203*	197	196	191	184	182	192	199	203	202	208	207*
100	206*	207	207	201	204	213	222	212	211	217	222	212*
105	217*	225*	226	218	230	247	246	225*	226*	241	241*	222*
110	248*	258*	263	262	274	285	276	256*	261*	276*	272*	250*
50 Degrees N												
90	208*	202	192*	179*	161*	147	153	170	185	198	207	213
95	210*	200*	197*	191*	180*	173	184	194	196*	201	214	219*
100	210*	206*	209*	210*	212*	219	221	211*	204*	209	220*	220*
105	215*	224*	231*	237*	253*	265	255	224*	216*	228*	231*	219*
110	240*	259*	274*	284*	304*	313	293	254*	244*	256*	253*	238*
60 Degrees N												
90	214	207	196*	181*	159*	144*	145	159	177	197	211*	217
95	214*	206*	199*	193*	180*	171*	174*	181	188*	199	215*	222
100	214*	210*	215*	221*	223*	222*	216*	203*	195*	201	215*	221
105	217*	228*	245*	263*	279*	282*	258*	220*	202*	209*	217*	215*
110	235*	261*	291*	317*	343*	343*	306*	250*	223*	227*	231*	227*
70 Degrees N												
90	214	208	199*	183*	162*	145*	141*	150	173*	196*	212*	217
95	216*	209*	205*	197*	184*	173*	168*	170*	179*	195*	212*	219*
100	215*	214*	223*	232*	234*	226*	212*	193*	184*	193*	209*	217*
105	215*	230*	256*	285*	302*	295*	260*	213*	187*	192*	206*	211*
110	229*	261*	303*	344*	373*	366*	314*	241*	198*	199*	212*	216*

*Temperature data lacking or less than two data points within about one month and 10 deg lat/lon.
Values apply to the first day of each month.

Table 14-12 Pressure (N m^{-2}) 90 to 110 km. Insert decimal point on the right of the three digits and multiply by 10^n

km	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	N
0 Degrees N													
90	197	200	206	212	213	208	197	200	206	212	213	208	−3
95	803	855*	903*	920*	887*	853*	803	855*	903*	920*	887*	853*	−4
100	350	395*	419*	416*	379*	360*	350	395*	419*	416*	379*	360*	−4
105	168	199*	213*	202*	172*	162*	166	199*	213*	202*	172*	162*	−4
110	090	113*	122*	112*	091*	084*	090	113*	122*	112*	091*	084*	−4
10 Degrees N													
90	194	188*	191	206	214	206	198	214	230	228	216	204	−3
95	079*	080*	081	088	090*	085	081	093	101*	099*	090	083	−3
100	345*	364*	374*	386*	379*	362*	357	421*	480*	462*	395	353	−4
105	164*	182*	183*	182*	172*	164*	170	213*	246*	230*	184*	161*	−4
110	089*	101*	104*	098*	088*	084*	090	117*	141*	130*	100*	085*	−4
20 Degrees N													
90	187*	188	191	213	215	201	191	214	236	233	212	199	−3
95	077*	080	080	090*	090*	083	079	090	104*	102*	089	081	−3
100	338*	362	362	386*	382*	361*	354	412*	487*	480*	405	353*	−4
105	160*	178*	175*	178*	175*	167*	168	199*	247*	242*	195*	161*	−4
110	086*	096*	096*	093*	089*	087*	089	109*	138*	137*	109*	087*	−4
30 Degrees N													
90	184*	187	188	212	199	183	170	189	200	202	192	194	−3
95	778*	784	789	873*	819	737*	688	796	872	876	826	822	−4
100	345*	357	354	380*	351*	327*	317	359	404	413	384	368*	−4
105	163*	171*	169	172*	163*	156*	153	171	199*	207*	192*	173*	−4
110	084*	091*	089	089*	084*	084*	083	090	107*	116*	106*	091*	−4
40 Degrees N													
90	190*	191	187	198	181	156	145	158	166	172	175	191	−3
95	833*	814	785	813	700	577	569	660	721	751	774	849*	−4
100	379*	366	354	359	302	260	262	301	328	343	365	393*	−4
105	177*	174*	167	165	145	127	132	145*	158*	172	183*	188*	−4
110	899*	908*	885	858	784	719	725	748*	823*	928*	988*	962*	−5
50 Degrees N													
90	191*	192	191*	192*	169*	141	131	140	150	158	169	185	−3
95	873*	843*	804*	769*	619*	480	479	567	633*	694	772	867*	−4
100	401*	377*	365*	348*	276*	214	218	255*	280*	310	364*	412*	−4
105	190*	179*	174*	167*	135*	108	110	122*	131*	151*	180*	210*	−4
110	939*	929*	937*	923*	785*	640	626	629*	656*	778*	933*	992*	−5
60 Degrees N													
90	174	175	182*	184*	170*	145*	131	134	153	169	178*	179	−3
95	813*	792*	783*	754*	622*	489*	458*	504	619*	737	819*	841	−4
100	376*	359*	355*	341*	280*	216*	198*	216*	265*	325	386*	405	−4
105	181*	174*	177*	176*	145*	112*	101*	102*	119*	150*	184*	195*	−4
110	089*	090*	098*	102*	089*	069*	058*	052*	056*	072*	091*	095*	−4

Table 14-12 (Continued)

km	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	N
70 Degrees N													
90	138	133	146*	158*	170*	152*	135*	138	164*	170*	176*	162	– 3
95	646*	600*	645*	658*	649*	526*	456*	489*	638*	735*	811*	764*	– 4
100	301*	278*	298*	305*	292*	230*	191*	199*	261*	316*	372*	361*	– 4
105	125*	135*	153*	164*	162*	124*	098*	091*	110*	139*	174*	172*	– 4
110	070*	071*	086*	099*	100*	077*	056*	045*	048*	061*	081*	082*	– 4

*Data lacking (i.e., less than two data points within about one month and 10 deg latitude)
Values apply to the first day of each month

14.2.2.4 Reference Atmospheres Above 120 km. These atmospheres are based on a report by Jacchia [1971]. Previous thermospheric models that assumed a single, constant set of boundary conditions at 120 km did not accurately represent variations in properties in the altitude region 120 to 180 km. To attempt to remedy this problem, Jacchia lowered the boundary to 90 km where it is known that there is an approximate isopycnic (constant density) layer, although the temperature shows important variations. In his models, all temperature profiles start from a constant value, 183 K at 90 km. He did not print out the model values below 110 km in order to disregard the less accurate region.

The following are the major types of variation of properties of the thermosphere:

1. Variations with the solar cycle
2. Variations with daily changes in solar activity
3. Diurnal and semidiurnal variations
4. Atmospheric response to storms of various kinds
5. Semi-annual variation
6. Seasonal-latitudinal variations of the lower thermosphere
7. Effects of large scale circulation
8. Fluctuations associated with gravity waves.

The models described in this section include both long and short-term solar activity effects and diurnal, but not semidiurnal, variations. Atmospheric response to storms is represented in terms of Kp, the 3-h geomagnetic planetary index. This provides reasonable representation of the response to most storms, but not all. For example, the Auroral Electrojet (AE) index is better for high latitude response, both because it provides a better measure of the change in energy input in that region, and because of its much smaller time resolution. Unfortunately, values of the AE index are not always available.

The average semi-annual variation is reasonably modeled, but it must be pointed out that the amplitude of this effect, whose cause is not completely understood, varies from year to year. The seasonal-latitudinal variations of the lower thermosphere are not modeled, except for an empirical expression for density but not temperature changes. Champion [COESA, 1966] developed a limited set of models for this portion of the atmosphere, but further modeling efforts have been delayed because of a lack of information on the systematic variations of this part of the atmosphere. Attempts have been made to model empirically some of the effects of large scale circulation which primarily affect the

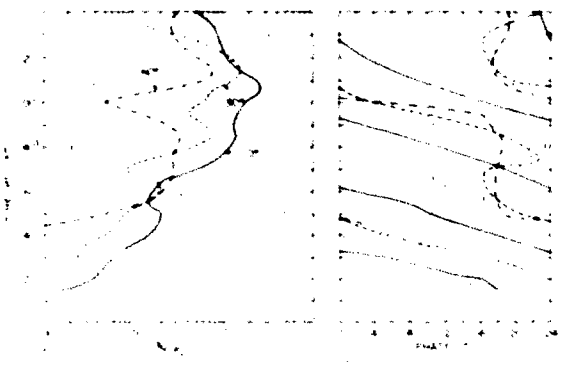


Figure 14-31. Relative amplitude and phase of density variation for the diurnal tide at 0°, 30°, and 60° latitude under equinox conditions.

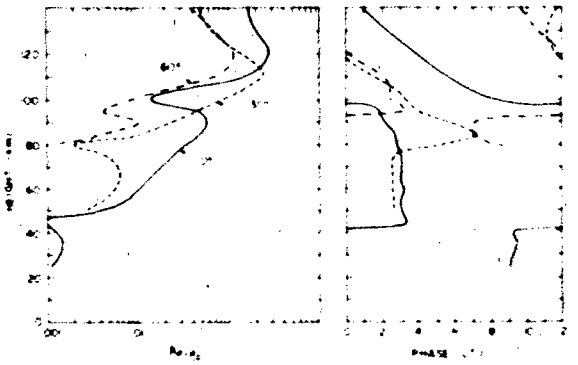


Figure 14-32. Relative amplitude and phase of density variation for the semidiurnal tide at 0°, 30°, and 60° latitude under equinox conditions.

Table 14-13. Densities (kg m^3) 90 to 110 km. Insert decimal point on the right of the three digits and multiply by 10^{-9}.

km	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	N
0 Degrees N													
90	370	360	361	368	384	384	370	360	361	368	284	384	−8
95	148	147*	153*	159*	163*	160*	148	147*	153*	159*	163*	160*	−8
100	580	610*	644*	673*	663*	636*	580	610*	644*	673*	663*	636*	−9
105	242	264*	279*	282*	262*	252*	242	264*	279*	282*	262*	252*	−9
110	108	125*	131*	126*	111*	106*	108	125*	131*	126*	111*	106*	−9
10 Degrees N													
90	365	343*	345	364	383	380	370	384	403	396	390	378	−8
95	146*	140*	140	157	166*	158	148	161	169*	169*	162	156	−8
100	574*	571*	591*	648*	670*	633*	591	658*	721*	719*	668	617	−9
105	236*	246*	249*	265*	266*	256*	247	287*	318*	310*	270*	244*	−9
110	106*	114*	117*	116*	112*	109*	112	133*	151*	142*	116*	103*	−9
20 Degrees N													
90	339*	327	350	382	390	372	358	390	417	409	386	366	−8
95	139*	139	142	163*	166*	152	142	158	174*	170*	156	149	−8
100	560*	574	583	669*	665*	616*	581	659*	741*	731*	651*	605*	−9
105	234*	249*	247*	269*	270*	252*	245	281*	332*	322*	272*	239*	−9
110	104*	114*	114*	115*	115*	113*	113	129*	154*	150*	125*	106*	−9
30 Degrees N													
90	339*	342	345	388	375	355	329	351	361	361	345	347	−8
95	138*	137	139	159*	151	134*	122	139	147	147	141	145	−8
100	571*	574	574	654*	602*	535*	500	583	633	631	593	609*	−9
105	244*	247*	248	265*	246*	224*	215	252	280*	276*	260*	256*	−9
110	107*	113*	110	114*	107*	104*	103	112	127*	132*	123*	114*	−9
40 Degrees N													
90	333*	341	340	373	366	335	302	305	302	303	304	327	−8
95	151*	142	138	147	131	109	102	114	122	128	128	141	−8
100	622*	598	578	603	510	412	399	479	526	534	555	627*	−9
105	271*	257*	247	253	210	171	178	214*	233*	237	252*	281*	−9
110	119*	116*	110	108	094	083	086	096*	104*	110*	119*	126*	−9
50 Degrees N													
90	320*	331	346*	374*	365*	333	297	287	282	277	283	302	−8
95	143*	145*	141*	139*	119*	096	090	101	111*	119	124	136*	−8
100	645*	619*	589*	560*	439*	330	333	408*	464*	502	559*	633*	−9
105	294*	266*	251*	234*	178*	136	143	182*	202*	220*	259*	305*	−9
110	128*	118*	112*	107*	085*	067	070	081*	088*	100*	121*	137*	−9
60 Degrees N													
90	283	295	323*	353*	372*	350*	314	293	300	298	294*	286	−8
95	131*	133*	136*	135*	119*	099*	091*	096	113*	128	131*	131	−8
100	593*	578*	559*	522*	424*	328*	310*	360*	459*	546	606*	619	−9
105	278*	254*	241*	223*	173*	132*	130*	154*	196*	239*	282*	302*	−9
110	124*	114*	111*	106*	085*	066*	062*	068*	083*	104*	129*	137*	−9
70 Degrees N													
90	225	222	255*	301*	365*	365*	332*	321	329*	302*	288*	260	−8
95	103*	099*	108*	115*	122*	105*	094*	099*	123*	130*	132*	120*	−8
100	474*	439*	451*	445*	422*	345*	305*	349*	479*	553*	602*	563*	−9
105	224*	195*	200*	192*	178*	140*	125*	142*	197*	241*	281*	272*	−9
110	101*	089*	094*	095*	088*	069*	058*	062*	080*	101*	125*	124*	−9

*Data lacking or less than two data points within about one month and 10 deg. latitude.
Values apply to the first day of each month.

composition of the thermosphere. For example, the strong increase in helium concentration above the winter pole (the helium winter bulge) is a result of large scale circulation, which causes preferential transport of lighter constituents away from the summer hemisphere. Density and composition fluctuations caused by gravity waves are not modeled.

The tables in this section are given with the exospheric temperature as the defining parameter. To determine the exospheric temperature for a particular time and location the following equations are used. The temperature T_c is the global night-time minimum of the exospheric temperature when Kp is zero.

$$T_c = 379 + 3.24\,\bar{F}_{10.7} + 1.3\,(F_{10.7} - \bar{F}_{10.7}) \quad (14.18)$$

where $F_{10.7}$ is in units of 10^{-22} W m^{-2} Hz^{-1}. The diurnal variation of the exospheric temperature can be represented by

$$T_D = T_c\,(1 + R\cos^m\eta)$$
$$T_N = T_c\,(1 + R\sin^m\theta), \quad (14.19)$$

where T_D is the maximum daily temperature and T_N is the minimum daily temperature. R is a coefficient with a value of approximately 0.3.

$$\eta = 1/2\,|\,\phi - \delta_o\,|$$
$$\theta = 1/2\,|\,\phi + \delta_o\,|$$

where ϕ is the latitude and δ_o is the sun's declination. A value of 2.2 can be used for m.

For Kp above zero the exospheric temperature is increased by

$$\Delta T_x = 28\,Kp + 0.08\exp Kp\,. \quad (14.20)$$

Formulas to give the exospheric temperature at any time of day, to include the effect of the semi-annual variation, and to include the winter helium bulge are given in the reference [Jacchia, 1971].

Table 14-14 contains model values of kinetic temperature, logarithms of the concentrations of N_2, O_2, O, Ar, He, and H, mean molecular weight, and density as functions of altitude for exospheric temperatures ranging from 500 to 1600 K.

Table 14-14 Atmospheric temperature, density and composition as functions of height and exospheric temperature

Height km	Temp K	LogN(N_2) m^{-3}	LogN(O_2) m^{-3}	LogN(O) m^{-3}	LogN(Ar) m^{-3}	LogN(HE) m^{-3}	LogN(H) m^{-3}	Mean Mol Wt	Density kg/m^3
Exospheric Temperature = 500 K									
120	267.1	17.536	16.672	17.169	15.073	13.579		25.10	2.25E − 08
130	328.3	16.981	16.051	16.814	14.320	13.457		23.74	6.80E − 09
140	378.8	16.531	15.545	16.530	13.704	13.363		22.51	2.67E − 09
150	411.9	16.149	15.114	16.296	13.174	13.291		21.39	1.25E − 09
160	433.1	15.804	14.723	16.089	12.691	13.231		20.40	6.51E − 10
180	457.7	15.170	14.002	15.717	11.798	13.129		18.78	2.13E − 10
200	471.4	14.576	13.325	15.372	10.956	13.038		17.62	8.13E − 11
250	488.1	13.169	11.721	14.562	8.956	12.830		15.79	1.05E − 11
300	494.6	11.823	10.184	13.791	7.039	12.635		13.30	1.72E − 12
350	497.3	10.512	8.686	13.041		12.446		8.33	3.29E − 13
400	498.5	9.226	7.218	12.306		12.262		3.82	8.10E − 14
500	499.5	6.718		10.873		11.903	12.860	1.44	1.94E − 14
600	499.8			9.484		11.555	12.772	1.19	1.24E − 14
800	499.9			6.822		10.889	12.604	1.06	7.24E − 15
1000	500.0					10.260	12.445	1.03	4.79E − 15
1500	500.0					8.827	12.084	1.01	2.04E − 15
2000	500.0					7.565	11.767	1.01	9.78E − 16
Exospheric Temperature = 600 K									
120	285.2	17.549	16.690	17.165	15.103	13.567		25.21	2.30E − 08
130	359.6	17.018	16.098	16.818	14.389	13.443		23.98	7.29E − 09
140	423.6	16.597	15.627	16.547	13.817	13.349		22.87	3.00E − 09
150	468.4	16.246	15.232	16.328	13.336	13.278		21.89	1.48E − 09
160	498.6	15.937	14.883	16.140	12.906	13.221		20.99	8.10E − 10

Table 14-14. *(Continued)*

Height km	Temp K	LogN(N_2) m^{-3}	LogN(O_2) m^{-3}	LogN(O) m^{-3}	LogN(Ar) m^{-3}	LogN(HE) m^{-3}	LogN(H) m^{-3}	Mean Mol Wt	Density kg/m^3
180	534.9	15.381	14.252	15.809	12.127	13.127		19.49	2.93E−10
200	555.8	13.869	13.670	15.510	11.404	13.046		18.35	1.23E−10
250	581.6	13.676	12.310	14.820	9.711	12.866		16.67	1.99E−11
300	591.6	12.546	11.021	14.171	8.103	12.701		15.72	4.15E−12
350	595.8	11.450	9.769	13.544	6.541	12.543		14.41	9.69E−13
400	597.7	10.377	8.543	12.931		12.389		11.79	2.46E−13
500	599.2	8.285	6.154	11.735		12.090	12.126	4.81	2.49E−14
600	599.6	6.257		10.577		11.800	12.053	2.37	7.08E−15
800	599.9			8.359		11.245	11.913	1.54	2.54E−15
1000	600.0			6.262		10.720	11.780	1.25	1.36E−15
1500	600.0					9.525	11.484	1.04	5.27E−16
2000	600.0					8.474	11.215	1.01	2.76E−16
	Exospheric Temperature = 700 K								
120	300.5	17.559	16.705	17.161	15.126	13.558		25.29	2.35E−08
130	386.2	17.046	16.134	16.821	14.440	13.433		24.15	7.67E−09
140	462.0	16.644	15.686	16.558	13.900	13.338		23.14	3.27E−09
150	518.2	16.315	15.317	16.349	13.452	13.267		22.24	1.67E−09
160	557.8	16.029	14.995	16.172	13.058	13.211		21.43	9.46E−10
180	607.5	15.526	14.426	15.869	12.357	13.122		20.03	3.67E−10
200	636.8	15.072	13.910	15.601	11.718	13.047		18.92	1.65E−10
250	673.5	14.030	12.723	14.995	10.242	12.886		17.15	3.16E−11
300	688.0	13.054	11.610	14.434	8.854	12.742		16.24	7.80E−12
350	694.0	12.111	10.533	13.894	7.511	12.606		15.53	2.17E−12
400	696.8	11.189	9.481	13.367	6.198	12.473		14.50	6.46E−13
500	698.8	9.395	7.431	12.341		12.216	11.559	10.02	6.99E−14
600	699.5	7.656		11.348		11.968	11.496	5.19	1.26E−14
800	699.8			9.446		11.492	11.376	2.77	2.54E−15
1000	699.9			7.649		11.042	11.263	2.13	1.04E−15
1500	700.0					10.018	11.005	1.29	2.39E−16
2000	700.0					9.117	10.778	1.07	1.09E−16
	Exospheric Temperature = 800 K								
120	313.5	17.567	16.716	17.158	15.145	13.550		25.35	2.39E−08
130	408.8	17.067	16.162	16.823	14.481	13.424		24.28	7.99E−09
140	495.1	16.680	15.731	16.566	13.964	13.329		23.35	3.50E−09
150	562.1	16.366	15.380	16.363	13.539	13.258		22.52	1.82E−09
160	611.3	16.096	15.078	16.194	13.171	13.202		21.77	1.06E−09
180	675.4	15.632	14.553	15.910	12.526	13.115		20.46	4.34E−10
200	714.4	15.218	14.084	15.664	11.948	13.044		19.39	2.06E−10
250	764.0	14.288	13.026	15.120	10.633	12.897		17.57	4.46E−11
300	783.7	13.427	12.043	14.623	9.410	12.769		16.59	1.25E−11
350	791.9	12.598	11.098	14.148	8.230	12.648		15.96	3.96E−12
400	795.6	11.791	10.175	13.686	7.079	12.532		15.33	1.34E−12
500	798.4	10.219	8.381	12.788		12.307	11.108	12.86	1.77E−13

Table 14-14 *(Continued)*

Height km	Temp K	LogN(N_2) m^{-3}	LogN(O_2) m^{-3}	LogN(O) m^{-3}	LogN(Ar) m^{-3}	LogN(HE) m^{-3}	LogN(H) m^{-3}	Mean Mol Wt	Density kg/m^3
	Exospheric Temperature = 800 K								
600	799.3	8.697	6.642	11.918		12.089	11.053	8.43	3.04E−14
800	799.8			10.254		11.673	10.948	3.92	3.75E−15
1000	799.9			8.681		11.279	10.849	3.22	1.39E−15
1500	800					10.383	10.623	2.10	2.31E−16
2000	800.0					9.594	10.425	1.39	7.06E−17
	Exospheric Temperature = 900 K								
120	324.7	17.574	16.726	17.156	15.160	13.544		25.41	2.41E−08
130	428.4	17.084	16.184	16.825	14.513	13.417		24.39	8.25E−09
140	523.9	16.708	15.766	16.572	14.013	13.321		23.51	3.68E−09
150	600.9	16.404	15.428	16.373	13.606	13.249		22.73	1.95E−09
160	659.6	16.147	15.140	16.209	13.257	13.193		22.03	1.16E−09
180	738.9	15.710	14.648	15.938	12.654	13.107		20.81	4.93E−10
200	788.5	15.328	14.216	15.708	12.122	13.039		19.79	2.44E−10
250	852.9	14.483	13.255	15.211	10.931	12.902		17.97	5.83E−11
300	878.6	13.710	12.374	14.764	9.835	12.785		16.91	1.80E−11
350	889.3	12.971	11.530	14.339	8.783	12.678		16.26	6.29E−12
400	894.2	12.251	10.709	13.927	7.758	12.574		15.74	2.36E−12
500	897.9	10.853	9.112	13.128		12.373	10.741	14.22	3.76E−13
600	899.0	9.500	7.566	12.355		12.179	10.692	11.08	7.04E−14
800	899.7	6.910		10.876		11.809	10.599	5.03	6.35E−15
1000	899.9			9.478		11.459	10.511	3.81	2.05E−15
1500	900.0			6.294		10.663	10.310	3.08	3.40E−16
2000	900.0					9.962	10.133	2.21	8.36E−17
	Exospheric Temperature = 1000 K								
120	334.5	17.579	16.733	17.154	15.172	13.539		25.45	2.44E−08
130	445.4	17.098	16.202	16.826	14.539	13.411		24.48	8.47E−09
140	549.0	16.730	15.794	16.577	14.053	13.315		23.64	3.84E−09
150	635.2	16.435	15.467	16.381	13.660	13.242		22.91	2.06E−09
160	703.1	16.187	15.189	16.220	13.325	13.186		22.25	1.24E−09
180	798.1	15.770	14.721	15.959	12.755	13.100		21.10	5.44E−10
200	859.3	15.413	14.317	15.741	12.258	13.034		20.13	2.78E−10
250	940.2	14.634	13.434	15.279	11.164	12.904		18.33	7.21E−11
300	972.8	13.932	12.633	14.872	10.169	12.796		17.21	2.40E−11
350	986.5	13.263	11.871	14.488	9.219	12.698		16.52	9.09E−12
400	992.6	12.614	11.130	14.116	8.294	12.604		16.02	3.69E−12
500	997.3	11.355	9.692	13.396	6.499	12.423	10.438	14.94	6.89E−13
600	998.8	10.137	8.300	12.699		12.249	10.393	12.85	1.45E−13
800	999.6	7.805		11.368		11.915	10.309	6.55	1.17E−14
1000	999.9			10.109		11.601	10.229	4.24	3.02E−15
1500	1000.0			7.244		10.884	10.049	3.62	5.28E−16
2000	1000.0					10.253	9.890	3.10	1.32E−16

Table 14-14. (*Continued*)

Height km	Temp K	LogN(N_2) m^{-3}	LogN(O_2) m^{-3}	LogN(O) m^{-3}	LogN(Ar) m^{-3}	LogN(HE) m^{-3}	LogN(H) m^{-3}	Mean Mol Wt	Density kg/m^3
	Exospheric Temperature = 1200 K								
120	350.9	17.587	16.745	17.150	15.192	13.530		25.52	2.48E – 08
130	473.8	17.119	16.229	16.827	14.580	13.401		24.61	8.82E – 09
140	591.1	16.764	15.837	16.583	14.115	13.304		23.84	4.09E - 09
150	693.3	16.481	15.524	16.391	13.741	13.231		23.17	2.24E – 09
160	778.3	16.245	15.262	16.235	13.426	13.173		22.57	1.37E – 09
180	905.0	15.857	14.828	15.985	12.901	13.087		21.54	6.28E – 10
200	991.4	15.533	14.463	15.783	12.455	13.021		20.65	3.36E – 10
250	1110.3	14.851	13.691	15.373	11.504	12.901		18.95	9.84E – 11
300	1159.1	14.253	13.011	15.023	10.659	12.806		17.79	3.70E – 11
350	1179.6	13.691	12.370	14.699	9.860	12.722		17.01	1.57E – 11
400	1188.9	13.147	11.750	14.387	9.087	12.643		16.47	7.19E – 12
500	1195.9	12.096	10.549	13.785	7.589	12.491	9.966	15.67	1.70E – 12
600	1198.2	11.080	9.388	13.205	6.140	12.346	9.929	14.62	4.46E – 13
800	1199.4	9.136	7.169	12.095		12.063	9.858	10.17	4.09E – 14
1000	1199.8	7.300		11.046		11.806	9.792	5.74	7.21E – 15
1500	1200.0			8.658		11.208	9.642	3.96	1.09E – 15
2000	1200.0			6.557		10.682	9.509	3.82	3.25E – 16
	Exospheric Temperature = 1400 K								
120	364.1	17.593	16.754	17.147	15.207	13.523		25.57	2.51E – 08
130	496.8	17.134	16.250	16.827	14.611	13.393		24.71	9.10E – 09
140	625.2	16.789	15.869	16.587	14.160	13.296		23.99	4.29E – 09
150	740.8	16.514	15.566	16.398	13.800	13.222		23.36	2.38E – 09
160	841.0	16.286	15.313	16.245	13.498	13.163		22.81	1.48E – 09
180	998.5	15.916	14.901	16.001	13.003	13.075		21.86	6.93E – 10
200	1111.9	15.613	14.562	15.809	12.591	13.009		21.05	3.82E – 10
250	1274.4	14.998	13.867	15.431	11.738	12.893		19.45	1.22E – 10
300	1342.6	14.473	13.271	15.122	10.999	12.807		18.30	5.00E – 11
350	1371.4	13.986	12.716	14.840	10.309	12.733		17.48	2.32E – 11
400	1384.4	13.518	12.182	14.571	9.643	12.664		16.89	1.15E – 11
500	1394.3	12.614	11.151	14.054	8.356	12.534	9.621	16.09	3.23E – 12
600	1397.4	11.743	10.155	13.555	7.113	12.409	9.588	15.39	9.98E – 13
800	1399.2	10.077	8.252	12.603		12.171	9.528	12.79	1.17E – 13
1000	1399.7	8.503	6.454	11.704		11.946	9.471	8.37	1.93E – 14
1500	1399.9			9.658		11.434	9.342	4.18	1.93E – 15
2000	1400.0			7.856		10.983	9.229	3.96	6.44E – 16
	Exospheric Temperature = 1600 K								
120	375.4	17.598	16.762	17.144	15.219	13.518		25.62	2.53E – 08
130	516.4	17.147	16.266	16.827	14.635	13.387		24.79	9.32E – 09
140	654.1	16.808	15.894	16.590	14.196	13.290		24.10	4.45E – 09
150	781.1	16.540	15.598	16.404	13.846	13.215		23.51	2.50E – 09
160	894.6	16.318	15.353	16.252	13.554	13.155		23.00	1.56E – 09

Table 14-14. (*Continued*)

Height km	Temp K	Log(N(N_2) /m^3)	Log(N(O_2) /m^3)	Log(N(O) /m^3)	Log(N(A) /m^3)	Log(N(HE) /m^3)	Log(N(H) /m^3)	Mean Mol Wt	Density kg/m^3
180	1081.3	15.960	14.955	16.012	13.080	13.065		22.11	[illegible]E−10
200	1222.4	15.672	14.655	15.825	12.692	12.998		21.36	4.20E−10
250	1433.1	15.103	13.995	15.470	11.910	12.884		19.87	1.43E−10
300	1523.5	14.633	13.461	15.190	11.250	12.804		18.75	6.27E−11
350	1561.9	14.201	12.970	14.939	10.640	12.737		17.91	3.10E−11
400	1579.2	13.790	12.500	14.702	10.055	12.676		17.28	1.64E−11
500	1592.4	12.997	11.596	14.248	8.926	12.561	9.361	16.43	5.21E−12
600	1596.5	12.234	10.724	13.811	7.838	12.452	9.333	15.82	1.82E−12
800	1599.0	10.776	9.059	12.978		12.243	9.280	14.21	2.67E−13
1000	1599.6	9.398	7.485	12.192		12.046	9.230	11.01	4.88E−14
1500	1599.9	6.263		10.401		11.598	9.117	4.71	3.30E−15
2000	1600.0			8.824		11.204	9.018	4.03	1.08E−15

REFERENCES

Barlier, F., C. Berger, J.L. Falin, G. Kockarts, and G. Thuillier, "A Thermospheric Model Based on Satellite Drag Data," *Ann. Geophys.*, **34:** 9, 1978.

Champion, K.S.W., "Properties of the Mesosphere and Lower Thermosphere," *Adv. Space Res.*, **1:** 133, 1981.

Champion, K.S.W., "Properties of the Mesosphere and Thermosphere and Comparison with CIRA 72," *Adv. Space Res.*, **3:** 45, 1983.

Champion, K.S.W. and R.A. Schweinfurth, "A New Mean Reference Atmosphere for 25 to 500 km," AFCRL-72-0579, AD757483, 1972.

COESA, *U.S. Standard Atmosphere, 1962*, U.S. Government Printing Office, Washington, D.C., 1962.

COESA, *U.S. Standard Atmosphere Supplements, 1966*, U.S. Government Printing Office, Washington, D.C., 1966.

COESA, *U.S. Standard Atmosphere, 1976*, U.S. Government Printing Office, Washington, D.C. 1976.

Cole, A.E., "Models of extreme arctic and subarctic atmospheres between 20 and 90 km," *Space Res. XII*, 629, Akademie-Verlag, Berlin, 1972.

Cole, A.E. and A.J. Kantor, "Air Force Reference Atmospheres," AFGL TR-78-0051, ADA058505, 1978.

Colegrove, F.D., W.B. Hanson, and F.S. Johnson, "Eddy Diffusion and Oxygen Transport in the Lower Thermosphere," *J. Geophys. Res.*, **70:** 4931–4941, 1965.

COSPAR Working Group IV, *COSPAR International Reference Atmosphere (CIRA) 1965*, North-Holland, Amsterdam, 1965.

COSPAR Working Group IV, *COSPAR International Reference Atmosphere (CIRA) 1972*, Akademie-Verlag, Berlin, 1972.

Dickinson, R.E., E.C. Ridley, and R.G. Roble, "A Three-Dimensional General Circulation Model of the Thermosphere," *J. Geophys. Res.*, **86:** 1499, 1981.

Fuller-Rowell, T.J. and D. Rees, "A Three-Dimensional Time-Dependent Global Model of the Thermosphere," *J. Atmos. Sci.*, **37:** 2545, 1980.

Groves, G.V., "Atmospheric Structure and its Variations in the Region from 25 to 120 km," AFCRL-71-0410, 1971.

Hedin, A.E., J.E. Salah, J.V. Evans, C.A. Reber, G.P. Newton, N.W. Spencer, D.C. Kayser, D. Alcayde, P. Bauer, L. Cogger, and J.P. McClure, "A Global Thermospheric Model Based on Mass Spectrometer and Incoherent Scatter Data. 1. N_2 Density and Temperature," *J. Geophys. Res.*, **82:** 2139, 1977.

ICAO (International Civil Aviation Organization), *Manual of the ICAO Standard Atmosphere to 32 km (105,000 ft)*, Second Edition, Montreal, Canada, 1964.

ISO (International Standards Organization), *International Standard Atmosphere*, ISO Standard No. 2533, ISO/TC20/SC6 Secretariat USSR, 1973.

Jacchia, L.G., "Revised Static Models of the Thermosphere and Exosphere with Empirical Temperature Profiles," *Smithson. Astrophys. Obs. Spec. Rep.*, 332, 1971.

Jacchia, L.G., "Thermospheric Temperature, Density, and Composition: New Models," *Smithson. Astrophys. Obs. Spec. Rep.*, **375,** 1977.

List, R.J., *Smithsonian Meteorological Tables*, Smithsonian Institution Press, Washington, D.C., 1968.

Labitzke, K., "Synoptic-Scale Motions Above the Stratopause," NCAR Ms. No. 71-39, 1971.

Minzner, R.A., K.S.W. Champion, and H.L. Pond, "The ARDC Model Atmosphere, 1959," AFCRC-TR-59-267, 1959.

Quiroz, R.S., "The Determination of the Amplitude and Altitude of Stratospheric Warmings from Satellite-Measured Radiance Changes," *JAM*, **10**(3), 1971.

Slowey, J.W., "Global Frequency Distribution of Exospheric Temperature," AFGL TR-79-0143, ADA077293, 1979.

Chapter 15

ATMOSPHERIC TEMPERATURES, DENSITY, AND PRESSURE

Section 15.1 I.I. Gringorten, A.J. Kantor, Y. Izumi, and P.I. Tattelman
Section 15.2 A.J. Kantor and A.E. Cole
Section 15.3 A.J. Kantor and A.E. Cole

The three physical properties of the earth's atmosphere, temperature T, density ρ, and pressure P are related by the ideal gas law $P = \rho TR$ where R is known as the gas constant for air. Except for the one thousandth of 1% of the atmosphere by mass above 80 km, various gases comprise the atmosphere in essentially constant proportions. The principal exception is water vapor discussed in Chapters 16 and 21.

15.1 THERMAL PROPERTIES, SURFACE TO 90 KM

In the following sections the units of measurement are metric. Abbreviations are used whenever quantitative measures are presented. The temperature is in degrees Kelvin (K), density in kilograms per cubic meter (kg/m^3), pressure in millibars (mb), time in seconds (s) or hours (h), length in centimeters (cm), meters (m) or kilometers (km), and speed in meters per second (m/s) or kilometers per hour (km/h). The main unit of energy is the calorie (cal):

$$1 \text{ cal} = 4.1860 \text{ joules (J)}$$
$$= 1.163 \times 10^{-3} \text{ watt-hours (Wh)}.$$

For energy per unit area an additional unit, the Langley (L), is introduced:

$$1 \text{ L} = 1 \text{ cal/cm}^2$$
$$= 11.62 \text{ Wh/m}^2$$
$$= 41.84 \text{ kJ/m}^2.$$

The main unit of power is the watt (W), but the unit of solar power per unit area is given as Langleys per hour (L/H). In terms of the British Thermal Unit (BTU) 1 L/H = 3.686 BTU · ft^{-2} · h^{-1}.

15.1.1 Energy Supply and Transformation

The prime source of energy that produces and maintains atmospheric motions and the spatial and temporal variations of meteorological elements is the solar radiation intercepted by the earth. In comparison with solar radiation, other energy sources such as heat from the interior of the earth, radiation from other celestial bodies, or the tidal forces of the moon and sun are practically negligible. Manmade sources, such as the heat island of a city, can be neglected although their by-products, such as the increasing amounts of carbon dioxide in the atmosphere, have been subjected to intense scrutiny in recent years with respect to heat balance and climatic trends.

The rate at which solar energy is received on a plane surface perpendicular to the incident radiation outside of the atmosphere at the earth's mean distance from the sun is the *solar constant;* the approximate value used in this section is 2.0 L/min, or about 1400 W/m^2. (A detailed description of the solar constant and its empirically determined value is given in Chapters 1 and 2. The rate at which direct solar energy is received on a unit horizontal plane at the earth's surface or in the atmosphere above the earth's surface is called the *insolation*. The planetary *albedo*, which is the reflected radiation divided by the total incident solar radiation, varies primarily with angle of incidence of the radiation, the type of surface, and the amount of cloudiness. On the average, 30% to 40% of the incident solar energy is reflected back to space by the cloud surfaces, the clear atmosphere, the earth/air interface, and particles such as dust and ice crystals suspended in the atmosphere. The remaining 60% to 70% of the solar radiation is available as the energy source for maintaining and driving atmospheric processes.

Less than twenty years ago we could confidently consider the earth and its surrounding atmosphere as a self-contained thermodynamic system. No major energy changes in the system within the 50 to 100 year period of our climatological records were apparent. Globally there had been

no obvious systematic short-term change in (1) heating of the earth's surface or the atmosphere, (2) the intensity of the atmospheric circulation, or (3) the balance between evaporation and precipitation. The processes affecting the internal and latent heat and the mechanical energy within the earth-atmosphere system had appeared virtually balanced.

Over the past twenty years there has been much agonizing by many experts and authors over the possibility of climatic change. Since there have been changes in the climate throughout geological history, it is inevitable that there will be long-term and large-scale changes in the future. Man-produced local changes through the use of fossil fuels, destruction of forests and desertification, irrigation on one hand and drainage of swamps on the other hand, urbanization and the introduction of pollutants in the air all have telling effects on local climate. The broader implications, however, over large regions and over decades or centuries have been the subject of many extensive and ongoing investigations by agencies worldwide with only one universally accepted conclusion. The carbon dioxide content of the atmosphere is increasing, which may lead to a global warming [WMO, 1979]. The next 5 to 10 years might produce a valid prediction.

A consensus among climatologists on heating or cooling of large regions of the earth or changes in rainfall patterns in response to natural or manmade influences is lacking. For this chapter the climate is considered to be stationary stochastic. It is stochastic because there is much variability in weather events and conditions that can be fitted into probability distributions assuming partially random processes. It is stationary because derived statistics or parameters, such as averages and standard deviations, are assumed to be unchanging. Their true values are constant and are best estimated by large samples.

The main features of the global energy transformation are summarized in a flow chart, Figure 15-1, from which the relative importance of the major energy cycles within the earth-atmosphere thermodynamic system can be determined. The numerical data presented in this figure are useful for various quantitative estimates. For example, if all energy inputs for the system ceased and rates of energy expenditure were maintained, the reservoir of mechanical energy (momentum) would be depleted in about 3 days, the reservoir of latent heat (precipitable water) in about 12 days, and the reservoir of internal energy (heat) in about 100 days.

Although an absolutely dry and motionless atmosphere is conceivable, it is difficult to imagine an atmosphere at zero degrees K. It is perhaps more appropriate to interpret the above time intervals as fictitious cycles of turnover of the atmospheric properties. The relative magnitudes of these life cycles show that, in comparison with rainfall and winds, temperature is the most conservative and will exhibit the relatively smallest, and thus most regular, temporal and spatial large-scale variation.

The solar energy input into the atmosphere at any one point varies during the earth's rotation about its axis and revolution about the sun. A consistent feature of this variation on a global scale is the driving force produced by differential latitudinal solar heating of the earth's surface. The reaction of the atmosphere to the solar driving force on an hourly, a daily, or an annual basis is observed most easily in the temperature field at low levels.

The solar energy input varies according to season, latitude, orientation of terrain to the incident energy, soil structure, all of which can change the balance between the incoming solar and sky radiation (*short wave*) and the outgoing atmosphere-terrestrial radiation (*long wave*). The difference between short-wave and long-wave radiation is the *net radiation*. Locally, net radiation is decreased primarily by atmospheric moisture (vapor and clouds). Evaporation of soil moisture diminishes by the latent heat required the portion of net radiation available for heating air and soil at the ground. The importance of moisture in establishing general climatic zones is shown by comparing desert climates with adjacent climates at roughly the same latitude. Table 15-1 gives the effect of soil moisture on the heat budget of the earth/air interface.

Slopes facing south receive maximum solar energy. Slopes facing west are usually warmer and drier than those facing east because the time of maximum insolation on a west slope is shifted to the afternoon when the general level of air temperature is higher than in the forenoon.

The energy balance of the earth/air interface requires that net radiation equals the sum of heat fluxes into the air and soil plus the heat equivalent of evaporation. In general, the maximum of heat flux into the soil precedes the maximum of heat flux into the air. The temperature maximum at standard instrument height (about 1.8 m) follows the maximum of heat flux into the air by roughly one to two hours.

15.1.2 Surface Temperature

15.1.2.1 Official Station Temperature. The standard station temperature used in meteorology [NWS, 1979] is measured by a thermometer enclosed within a white-painted, louvered instrument shelter or Stevenson Screen. The shelter has a base about 1 m (1.7 to 2.0 m in Central Europe) above the ground and is mounted in an open field (close-cropped grass surface). The standardized height of the thermometer is about 1.8 m. The shelter permits air circulation past the thermometer and is designed to exclude direct solar and terrestrial radiation. However, the shelter unavoidably absorbs and radiates some heat energy which, although minimal, causes some deviation of the thermometer reading from the "true" air temperature. On calm, sunny days the recorded daytime shelter temperature will normally be 0.5 to 1 K higher than the free air temperature outside the shelter at the same level. On calm, clear nights it will most likely

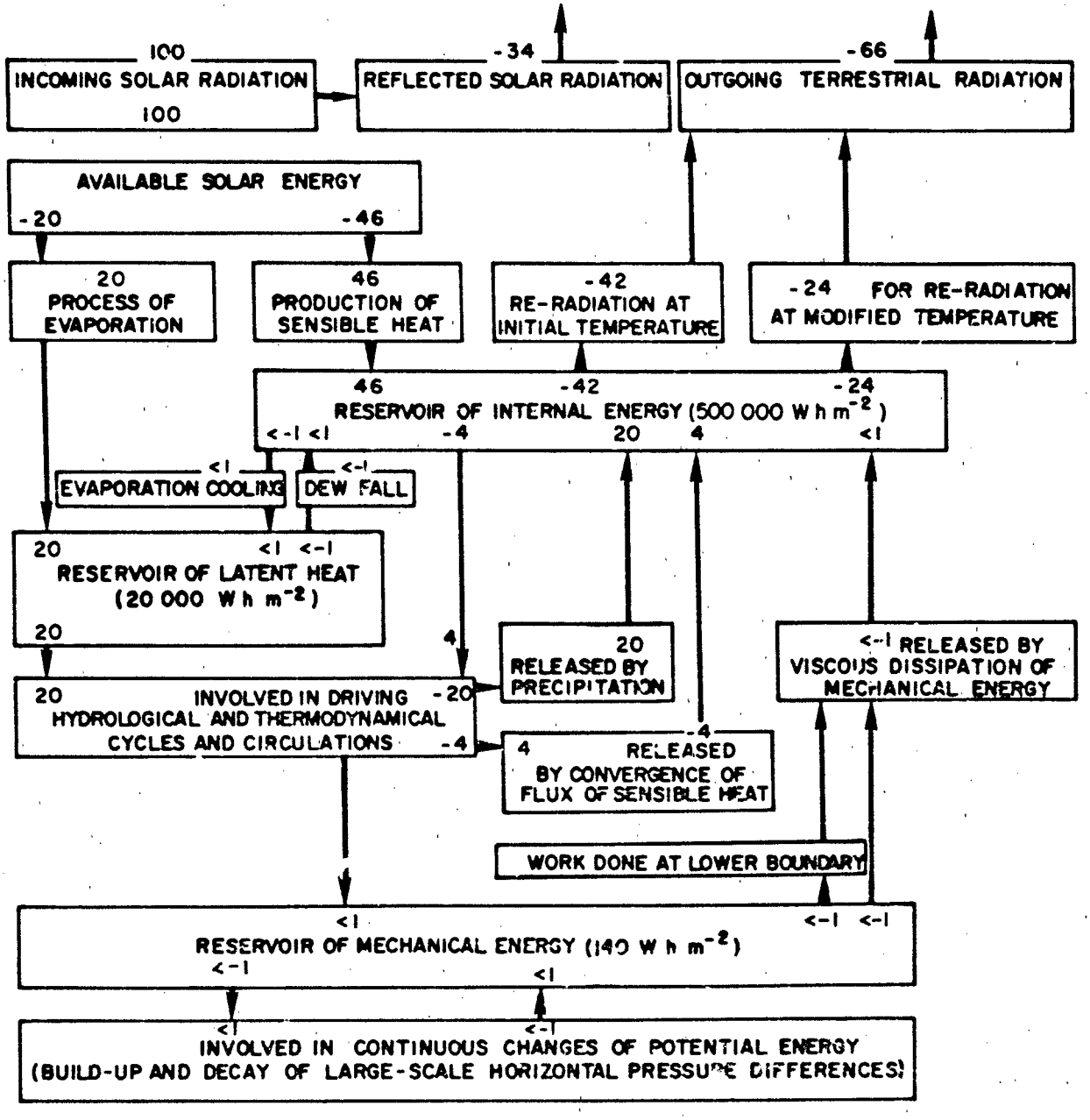


Figure 15-1 Global mean energy cycles of the atmosphere. All values are relative to 100 units of incoming solar radiation, which average 350 W/m², or 1/4 of the solar constant. [Revised from Lettau, 1954a].

be cooler by 0.5 K. Therefore, the thermometer should be artificially ventilated. Spatial variations of the ambient air temperature, especially in the first meter above the ground, are dependent upon the type of soil and ground cover. Ground effects decrease with height and for this reason the international standard heights of temperature measurement are a compromise between suppressing ground-cover effects and maintaining ease of access.

15.1.2.2 Daily Temperature. The official station temperature taken every hour on the hour reveals a fairly regular diurnal cycle. This is true despite several superimposed phenomena such as frontal passages and thunderstorms. Usually there is a temperature maximum in midafternoon and a temperature minimum near sunrise. The amplitude of the diurnal cycle varies with location and season from as little as 1 K to more than 17 K.

Table 15-1 Short-wave radiation on horizontal plane, net radiation, and estimated constituents of heat budget at the earth-air interface showing effect of difference in soil moisture caused by rains before 9 August and a dry spell before 7 September 1953 [Davidson and Lettau, 1957]

	Radiation (W/m^2)								
Mean Local Time	04h	06h	08h	10h	12h	14h	16h	18h	20h
9 August 1953*									
Short-wave	0	141	544	733	796	823	537	144	—
Net	−59	47	364	497	540	525	273	−13	—
Flux into soil	−40	29	186	63	74	73	28	−65	—
Flux into air	−11	—	81	158	176	190	64	−17	—
Heat of evap.	−8	—	97	276	290	262	181	69	—
7 September 1953**									
Short-wave	0	54	441	765	870	735	407	44	0
Net	−54	−32	181	403	488	398	154	−69	−77
Flux into soil	−44	−25	36	84	95	66	13	−29	−28
Flux into air	−6	−6	98	230	303	299	114	−30	−39
Heat of evap.	−4	−1	47	89	90	33	27	−10	−10

*Mean soil moisture in 0 to 10 cm layer, about 10% wet weight basis
**Mean soil moisture to 0 to 10 cm layer, about 4% wet weight basis

The annual cycle of daily mean temperature ranges from practically zero near the equator to as much as 40 K in the temperate zone. As an example, Figure 15-2 shows temperatures at Hanscom AFB, Mass. The middle curve reveals the annual cycle of the daily mean temperature (actually the monthly mean is plotted) and shows an annual range of 25 K. The diurnal range given here by the difference between mean daily maximum and minimum in Figure 15-2, is fairly uniform throughout the year, averaging 12 K.

Superimposed on both the diurnal and the annual cycles of temperature are many influences including cloudiness, precipitation, wind speed and direction, type of soil, ground cover, and aerodynamic roughness of the terrain. In the example of Figure 15-2, there is a range from the uppermost 1% of the daily maximum to the lowermost 1% of the daily minimum that is 3 times the mean diurnal cycle. The standard deviation of hourly temperature averages 5 K. The range from the uppermost 1% of the maximum temperature of the hottest month to the lowermost 1% of the minimum temperature of the coldest month in Figure 15-2 is about 2 1/2 times the mean annual cycle.

The pattern of surface temperature varies with geographic location. This is illustrated by the statistics of some widely scattered stations and even by the statistics of neighboring stations (Table 15-2). The annual mean temperature is generally lowest in the polar regions and highest in the equatorial belt. In addition, the mean temperature decreases generally with elevation. The diurnal range is greatest in desert climates and least in oceanic or maritime climates. The mean annual range tends to be greatest in temperate climates and least in equatorial climates.

In polar regions, where continuous darkness (daylight) endures for several months of the year, the 24-h cycle is minimal and the small diurnal variations are controlled primarily by changing winds and cloudiness. In summer, nearly all of the solar energy is expended in melting ice; hence, the maximum temperature seldom exceeds 273 K. Extra-

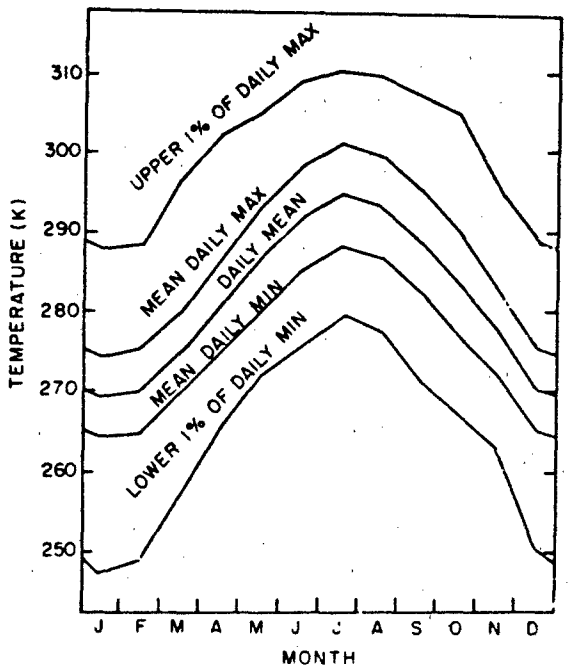


Figure 15-2. Surface temperature at Hanscom AFB, Mass. throughout the year.

Table 15-2 Temperatures at various stations around the world

Station	Lat	Long	Elev. m	Annual Mean K	Mean Diurnal Range K	Mean Annual Range K	Hottest Month 1% of Daily Max K	Coldest Month 1% of Daily Min K
Hanscom AFB, Mass.	42°28'N	71°17'W	43	282.3	11.7	25.6	311	247
Boston, Mass.	42°22'N	71°02'W	5	283.8	8.6	24.3	—	—
Blue Hill Obs., Mass.	42°13'N	71°07'W	192	282.3	9.6	24.4	—	—
Nantucket, Mass.	41°15'N	70°04'W	13	282.7	7.2	20.4	—	—
Pittsfield, Mass.	42°26'N	73°18'W	357	280.2	11.4	25.6	—	—
Worcester, Mass.	42°16'N	71°52'W	301	281.2	9.5	25.4	—	—
Thule, Greenland	76°32'N	68°42'W	59	261.8	6.4	31.9	289	233
Eielson AFB, Alaska	64°41'N	147°05'W	170	270.2	10.8	39.6	303	224
Keflavik, Iceland	63°58'N	22°36'W	50	278.1	4.4	11.2	291	258
Goose Bay, New Foundland	53°19'N	60°25'W	44	273.2	9.5	33.7	307	237
Berlin, Germany	52°28'N	13°26'E	50	282.8	7.2	20.6	307	254
Limestone, Maine	46°57'N	67°53'W	230	276.9	9.4	29.4	308	241
Bolling AFB, Wash. D.C.	38°49'N	76°51'W	20	287.0	10.2	23.0	311	259
Scott AFB, Ill.	38°33'N	89°51'W	138	286.1	6.2	26.3	311	250
Blytheville, Ark.	35°58'N	89°57'W	81	278.4	6.8	17.9	312	255
Riverside, Calif.	33°54'N	117°15'W	461	292.7	16.8	14.4	314	269
Tucson, Arizona	32°10'N	110°53'W	809	289.4	11.7	19.0	316	267
Ft. Huachuca, Arizona	31°25'N	110°20'W	1439	290.0	14.8	17.0	312	264
Dharan, Saudi Arabia	26°17'N	50°09'E	22	299.8	11.8	19.8	321	276
Wheeler, Hawaii	21°29'N	158°02'W	256	295.8	7.5	4.0	305	283
Honolulu, Hawaii	21°20'N	153°55'W	12	297.7	6.7	4.2	308	291
Guam, Phillipines	13°29'N	144°48'E	82	300.8	1.7	1.7	306	297
Diego Garcia Island	07°18'S	72°24'E	2	300.7	3.9	2.0	305	296
Canton Island	02°46'S	171°43'W	3	300.7	1.2	0.8	305	297

tropical regions characteristically have distinct diurnal and annual cycles. These cycles are superimposed over temperature variations caused by shifting air masses and frontal passages. In tropical regions, the diurnal range rarely exceeds 6 K

Depending on circumstances and ground characteristics, the surface air temperature could differ by several degrees over short distances ranging from a few meters to a few kilometers. Also, vertical temperature variations are observed from a few millimeters above the ground to the top of the instrument shelter. On windy, cloudy days or nights, the differences between thermometer readings, within short distances of one another in either the horizontal or the vertical, will be minimal. In high temperature regimes, however, with a bright sun and light winds, the ground surface, especially if dry sand, can attain temperatures 17 to 33 K higher than the free air. The temperature of air layers within a few centimeters of the surface will differ only slightly from the ground, but the decrease with height is rapid. The temperature at 0.5 to 1 m above the ground will be only slightly warmer than that observed in the instrument shelter at 1 or 1.5 m above the ground. Conversely, on calm, clear nights the ground radiation can produce a temperature inversion, as much as 4 or 5 K, in the air within several feet of the ground.

The induced temperature in military equipment exposed to the sun's heat will vary greatly with physical properties such as heat conductivity, reflectivity, capacity, and type of exposure. Surface and internal temperatures, such as are induced in a boxcar, make the reading of the shelter thermometer only the beginning of the engineering problem.

Table 15-2 is only an initial guide to the effects of various influences on station temperature. Detailed temperature information should be obtained from the climatological record of each station or of stations close by. The latter should be modified for the influences of terrain proximity to water, and elevation.

5.1.2.3 Horizontal Extent of Surface Temperature. Horizontal differences in surface temperature can arise both from large-scale weather disturbances and from local influences. Weather disturbances such as cold and warm fronts, thunderstorms, and squall lines account for unsystematic changes in the horizontal temperature gradient. Nonuniform radiational heating and cooling of the ground also contribute to turbulent mixing, cloudiness, and vertical motions in the

Table 15-3 Estimates of the horizontal scale of certain local meteorological conditions

Local Conditions	Horizontal Scale (km)	Temperature Differences (K)
Changes in Air Mass	160 to 1600	3 to 22
Weather Fronts	16 to 160	3 to 22
Squall Lines	8 to 80	3 to 17
Thunderstorms	8 to 24	3 to 17
Sea Breezes	8 to 16	1 to 11
Land Breezes	3 to 8	1 to 6

lower troposphere, resulting in constantly changing temperatures at the surface.

Horizontal transport by air currents, referred to as advection, is a key factor in surface temperature differences. Large-scale advection will bring both the relatively dry cold arctic air masses and the relatively moist warm tropical air masses alternately to the temperature zones. This can produce large changes in the day's mean and the diurnal range of temperature.

Table 15-3 gives estimates of surface temperature differences over varying horizontal distances associated with several kinds of weather phenomena. Large-scale differences are greatest in winter due to the more substantial differential heating by solar radiation from equator to pole and, consequently, the more intense large-scale motion of the atmosphere. In summer the north-south gradients in solar insolation are much less, but the general increase in the amount of insolation results in more thunderstorms and other air-mass activity.

Systematic differences in the surface temperature between neighboring stations are due to five prime factors: (1) aspect or orientation of the terrain with respect to incident solar radiation, (2) type of surface structure and of soil cover underlying the stations, (3) proximity to the moderating influences of large water bodies, (4) elevation, and (5) difference in solar time for stations that are several hundred kilometers apart. Sometimes the topography permits "pools" of cold air to drain locally at night into lower basins or valleys. Also nonuniform distribution of water vapor and cloudiness will result in uneven distributions of short-wave and long-wave radiation and, consequently, uneven cooling and heating at the surface.

A striking example of local influences on surface temperature gradients is found in the temperature contrasts between cities and the surrounding countryside. The sheltering effect of buildings, their heat storage, products of fuel combustion, smog, rain water drainage, and snow removal all act to make the city a relative heat source. Thus, the city's nightly minimum temperature might be 5 to 14 K higher than that of surrounding suburbs. As another example, in hilly or mountainous terrain the valley floor could have a diurnal temperature range 2 to 4 times as great as that over the peaks, and a temperature minimum from 5 to 17 K lower. Also, some pronounced horizontal temperature gradients occur along coastlines in temperate latitudes due to the cooling effect of coastal sea breezes.

Generally, temperatures between two stations become more independent of one another with increasing distance (Figure 15-3). One model curve [Gringorten, 1979] for fitting the correlation coefficient ρ as a function of distance s between stations is given by

$$\rho = \frac{2}{\pi}[\cos^{-1}\alpha) - \sqrt{1 - \alpha^2}], \qquad (15.1)$$

where

$$\alpha = \frac{s}{128r}, \qquad (15.2)$$

where r is the model parameter and is in the same units as the distance s. The value of r is, in fact, the distance over which the correlation coefficient is 0.99. For the curve in Figure 15-3, $r = 17.7$ km. While this curve could be fitted by other models, the given model curve has the quality that

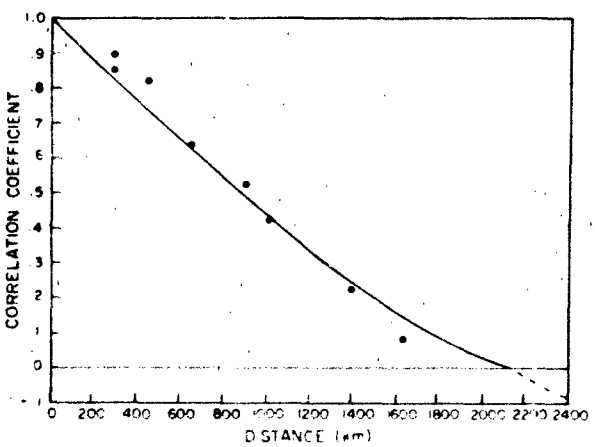


Figure 15-3 The correlation coefficient of the daily mean temperature of Columbus, Ohio with that of nine other U.S. stations at indicated distances from Columbus.

the correlation coefficient decreases exponentially with distance between stations for the first few kilometers of separation. Eventually, the correlation coefficient drops to zero at distance 128r

In the United States the separation between weather stations averages about 160 km, with the exception of the eastern states where it is 30 to 80 km. The root mean square difference of temperatures, as a function of the correlation coefficient ρ between two stations is approximated by

$$\text{rmsd} = s_t \sqrt{2(1 - \rho)}, \qquad (15.3)$$

where s_t is the standard deviation of the hourly temperature (estimated as 5 K for Hanscom AFB). For stations 150 km apart, with $\rho = 0.91$ (Equation 15.1), the rmsd should be approximately 2 K.

15.1.2.4 Runway Temperatures. At airports the desired length of the landing strip or runway is directly related to air temperature. Any discrepancy, therefore, between free air temperature over runways and shelter temperatures is important in establishing safe aircraft payloads and runway lengths. It had been thought, on days when insolation is strong, that the free air temperature over airfield landing strips is significantly higher than standard shelter temperature over the surrounding grassy areas. Results of observations, however, over four airstrips (Easterwood Airport, Hearne Air Force Satellite Field and Bryan Air Force Base in Texas, and an auxiliary naval airstrip near Houma, Louisiana) have shown that the air between 0.3 and 6 m above a landing strip is about 0.5 K cooler than indicated by the shelter thermometer over adjacent grassy areas. The relative smoothness of the runway surface is the physical cause of daytime flow of air from grass to runway. During the transition from flow over the rough grassy surface the wind speeds up and entrains the cooler air immediately above the runway. When a daytime equilibrium state is established, there will be a large lapse rate close to the ground. This is the effect over both concrete and blacktop airstrips with surrounding grass having only a slightly modifying effect.

In exceptional cases the free air temperature over the runway exceeds the shelter temperature but by no more than 0.5 K when averaged over 10 min, 1 K when averaged over 1 min with a dry soil environment, and 0.25 K (5-min mean) with a swamp environment. Thus the standard method of temperature measurement in a properly exposed shelter over grass provides a representative temperature for the estimations of runway length and aircraft payloads.

15.1.2.5 Temperature Extremes. A knowledge of the occurrence of hot and cold temperature extremes is important for the design of equipment and the selection of material that will be exposed to the natural environment. The hourly temperature observations at most locations are not normally distributed around the mean monthly values. Departures from a normal distribution are largest in the temperate and northerly latitudes during the winter months when the temperature distributions are substantially bimodal. Thus the straightforward method for determining the frequency distribution of hourly temperatures is to obtain a representative sample of observations for each location and compute the distributions. Estimates of the frequency distribution from such data can be made using the Blom formula given by

$$\hat{P}(T) = \frac{n_T - 3/8}{N + 1/4}, \qquad (15.4)$$

where $\hat{P}(T)$ is the estimated cumulative probability of the temperature T, n_T is the number of observations equal to or less than T, and N is the overall sample size. Since, representative samples of data are not easily obtained for regions outside North America, an objective method has been developed by Tattelman and Kantor [1977] so that the frequency distribution of surface temperature can be estimated at all locations from data in climatic summaries that are available for most locations throughout the world.

Because the warmest temperatures in the world are found at locations where the monthly means are high and the mean daily range is large, Tattelman et al. [1969] developed an index using these values. The index is expressed by

$$I_w = \bar{T} + (\bar{T}_{max} - \bar{T}_{min}), \qquad (15.5)$$

where I_w is the warm temperature index, $\bar{T}$ is the monthly mean, $\bar{T}_{max}$ is the mean daily maximum, and $\bar{T}_{min}$ is the mean daily minimum temperature (K − 273) for the warmest month. The index was related to temperature occurring 1%, 5% and 10% of the time during the warmest months at a number of locations; it appears in the following regression equations for estimating monthly 1%, 5% and 10% warm temperature extremes [Tattelman and Kantor, 1977]:

$$\hat{T}_{1\%} = 0.676I_w + 10.657, \qquad (15.6)$$

$$\hat{T}_{5\%} = 0.733I_w + 5.682, \qquad (15.7)$$

$$\hat{T}_{10\%} = 0.762I_w + 2.902 \qquad (15.8)$$

where $\hat{T}$ is the estimated temperature in (K − 273) occurring 1, 5, and 10% of the time, respectively. The same principle can be used to estimate cold temperature extremes. The cold temperature index is

$$I_c = \bar{T} - (\bar{T}_{max} - \bar{T}_{min}), \qquad (15.9)$$

where I_c is the cold temperature index, $\bar{T}$ is the monthly mean, $\bar{T}_{max}$ is the mean daily maximum, and $\bar{T}_{min}$ is the mean daily minimum temperature (K − 273) for the coldest months. The corresponding regression equations [Tattelman and Kantor, 1977] are

$$\hat{T}_{1\%} = 1.069I_c - 7.013, \qquad (15.10)$$

$$\dot{T}_{5\%} = 1.084\dot{T}_{\max} - 3.050. \quad (15.11)$$

$$\dot{T}_{10\%} = 1.082\dot{T}_{\max} - 0.704. \quad (15.12)$$

This technique has been used by Tattelman and Kantor [1976a,b] to map global temperature extremes using locations for which monthly climatic temperature summaries are available. Estimates of the 1% warm and cold temperature extremes for the Northern Hemisphere are shown in Figures 15-4 and 15-5.

Most extreme high temperatures have been recorded near the fringes of the deserts of northern Africa and southwestern U.S. in shallow depressions where rocks and sand reflect the sun's heat from all sides. In the Sahara, the greatest extremes have been recorded toward the Mediterranean coast, leeward of the mountains after the air has passed over the heated desert. The highest temperature on record is 331 K at Al Aziziyah, Libya (32°32'N, 13°1'E, elevation 112 m). Northern Africa and eastward throughout most of India is the hottest part of the world. Large areas

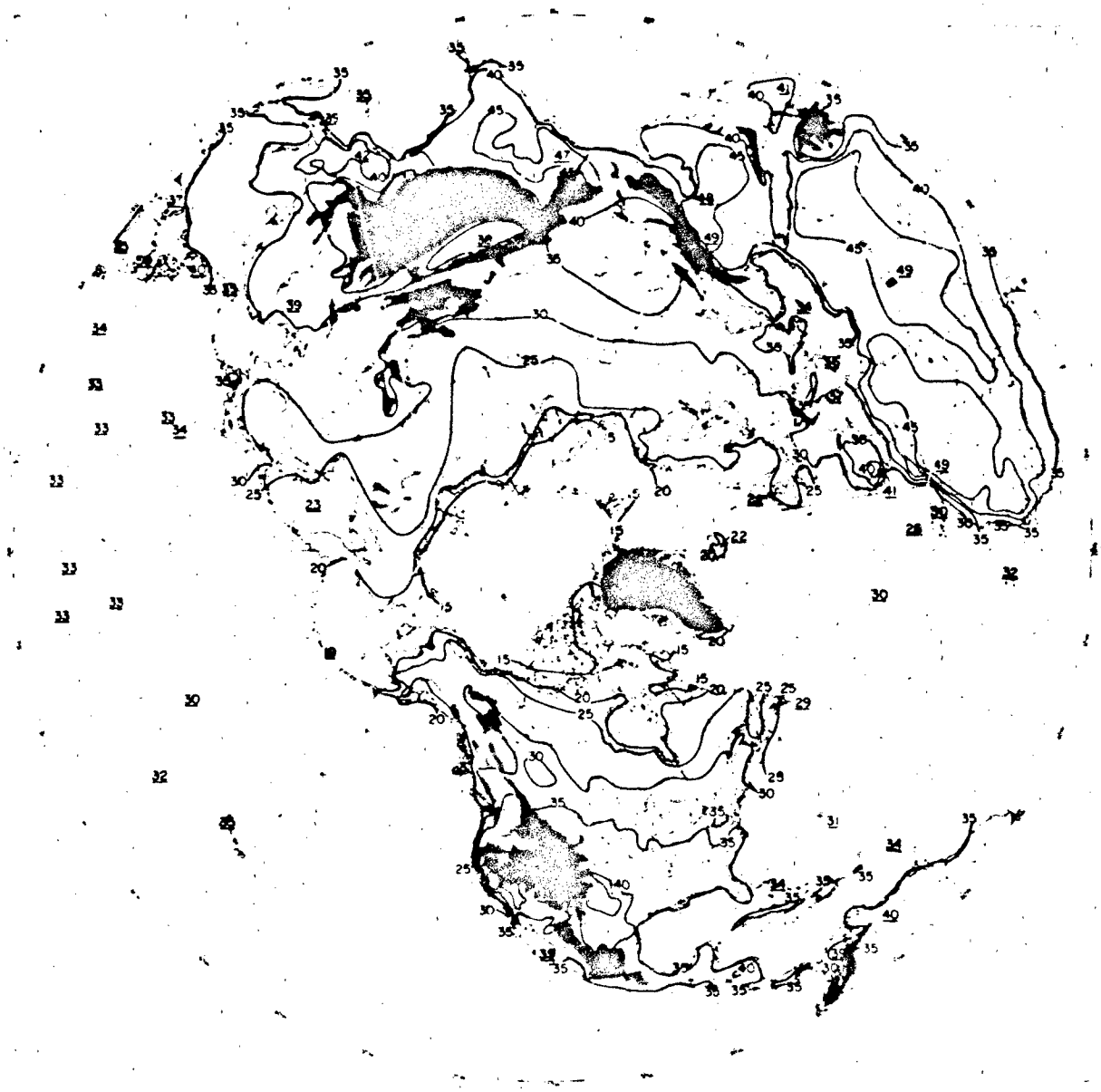


Figure 15-4. Temperature equaled or exceeded 1% of the time during the warmest month (K − 273) [Tattelman and Kantor, 1976a].

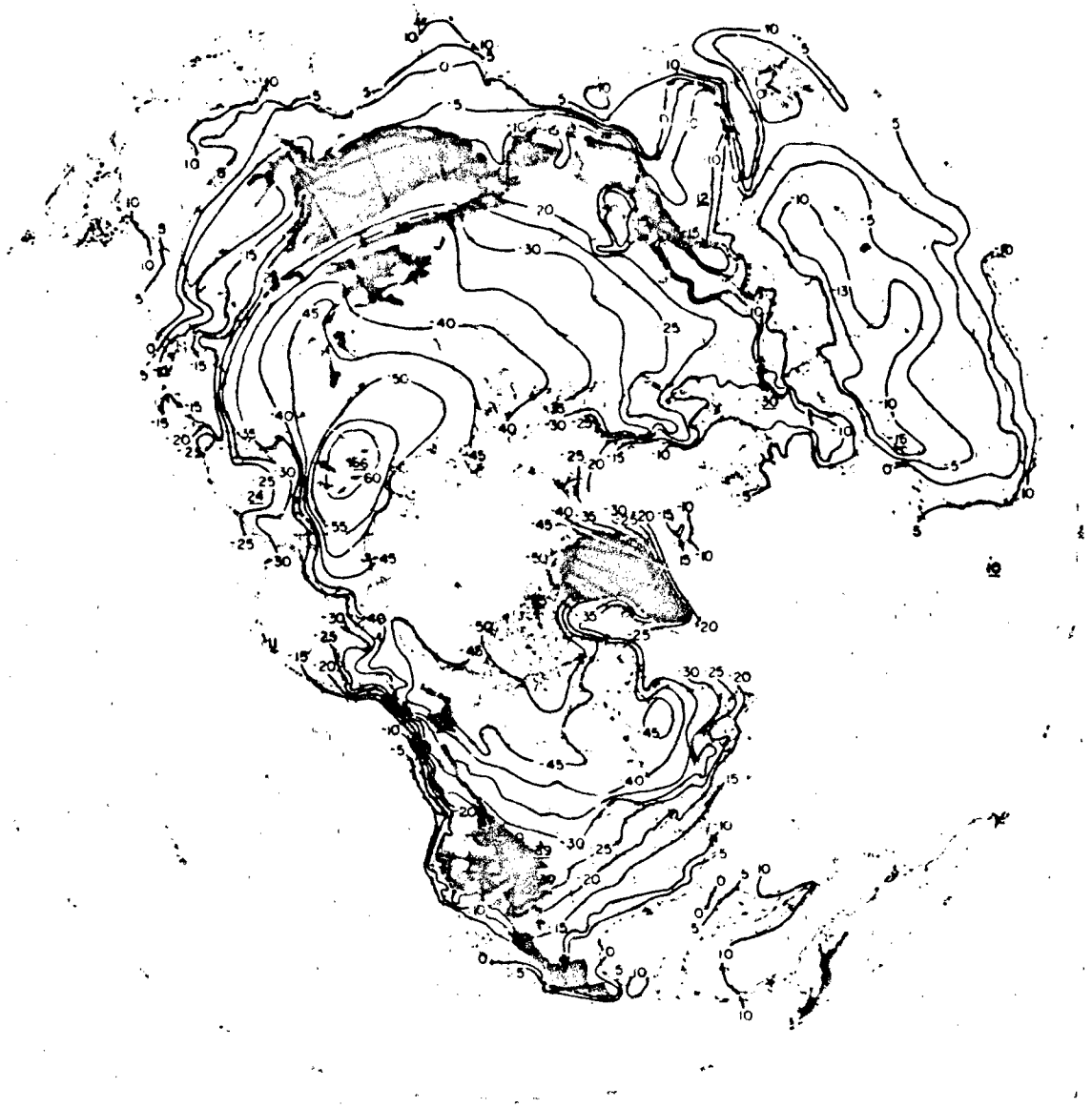


Figure 15-5. Temperature equaled or colder 1% of the time during the coldest month (K − 273) [Tattleman and Kantor, 1976a]

attain temperatures greater than 316 K more than 10% of the time during the hottest month. Sections of northwest Africa experienced temperatures greater than 322 K as much as 1% of the time during the hottest month of the year. Regions in Australia and South America have temperatures at and above 311 K much of the time, but do not experience temperatures greater than 316 K more than 1% of the time during the hottest month. The southwestern U.S. and a narrow strip of land in western Mexico are exceptionally hot. A substantial part of the area experiences temperatures equal to or greater than 316 K for 1% of the time in the hottest month. Death Valley, within this area, has temperatures equal to or greater than 322 K for 1% of the time in the hottest month and it once had a record temperature of 330 K.

Geographic areas of extreme cold include the Antarctic Plateau (2700 to 3600 m in elevation), the central part of the Greenland Icecap (2500 to 3000 m), Siberia between 63°N and 68°N, 93°E and 160°E (less than 760 m elevation) and the Yukon Basin of Northwest Canada and Alaska (less

than 760 m elevation). The generally accepted record low temperature (excluding readings in Anarctica) is 205 K in Siberia.

15.1.2.6 The Gumbel Model. For equipment that is either in continuous operation or is on standby status, thereby continuously exposed to all temperatures, the statistic of interest is the extreme temperature that is likely to occur during a full month, season, year, decade, or whatever period is considered to be the useful lifetime of the equipment.

Many extreme values have been estimated effectively by a model that has become known as the *Gumbel* distribution. Let us assume the annual highest temperature (T_i) has been recorded for each of N years (i = 1,N) with average $\bar{T}$ and standard deviation s_t. The Gumbel estimate of the cumulative probability P_T of the annual extreme high temperature (T) is then

$$P_T = \exp[-\exp(-y)], \qquad (15.13)$$

where

$$y = \bar{y} + \frac{\sigma_y}{s_t}(T - \bar{T}), \qquad (15.14)$$

where, to five decimal places

$$\bar{y} = 0.57722 \text{ and } \sigma_y = 1.28255. \qquad (15.15)$$

(There are other estimates to the Gumbel distribution. This one is preferred for its simplicity as well as degree of accuracy.) The quantity y is referred to as the *reduced variate*. If one is interested in the cold temperature, these formulas hold with the T and $\bar{T}$ reversed in Equation (15.14).

If the lifetime of a piece of equipment is intended to be n years, then the cumulative probability $P_T(n)$ that the temperature T will not be exceeded in the n years is

$$P_T(n) = \exp[-\exp(-y + \ell n\, n)] \qquad (15.16)$$

where y is given by Equation (15.14). Assuming, for example, that we want to estimate the temperature T that has only a 10% probability or risk of being exceeded over n years, we set $P_T(n)$ equal to 0.9 in Equation (15.16) and solve Equation (15.16) for y obtaining

$$y = \ell n\, n - \ell n(-\ell n\, P), \qquad (15.17)$$

which we in turn use in Equation (15.14) to obtain

$$\hat{T} = \bar{T} + \frac{s_t(y - \bar{y})}{\sigma_y} \qquad (15.18)$$

The return period is a term sometimes used in association with the extreme. In terms of the cumulative probability P_T of the annual extreme temperature, it is equal to $1/(1-P_T)$ years. The return period is not to be confused with the planned lifetime (n) of the equipment. Roughly speaking, the temperature with the 100-yr return period or the annual 1% (P_T = .99) is approximately the 10% temperature of a 10-year planned life.

The Gumbel distribution with a set of periodic extremes is the easiest model to use, but there are reservations in its application. Theoretically the basic distribution, such as the station temperature taken hourly, should be an exponential type, such as Pearson Type III or Gaussian. However, this condition may not be sufficient because the record may not be long enough to make the annual extreme fit into a Gumbel distribution. The Gumbel distribution is only the limiting form over long times and may not be adequately reached over short periods. It is advisable, therefore, to test the data to determine if the Gumbel distribution is applicable. Figure 15-6 illustrates the use of special-purpose "Extreme Probability Paper" in which the cumulative probability P_T is read on the vertical axis to correspond to T on the horizontal axis. Alongside the scale of P_T is the scale of the reduced variate y, which is uniform on this paper. A Gumbel distribution appears as a straight line.

Let us suppose a set of N extreme temperatures T_i for each of N years (i = 1,N) is ordered from lowest to highest value. The cumulative probability of the ith lowest temperature since it is an extreme is best estimated by

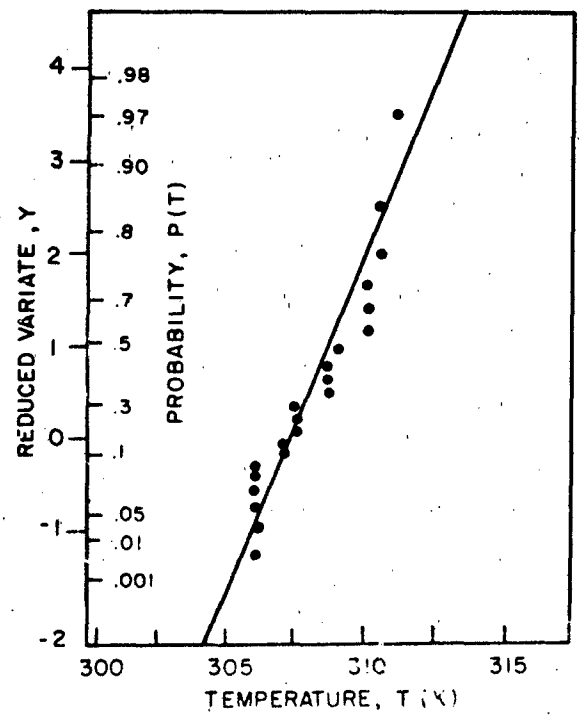


Figure 15-6 a) Annual highest temperature, Hanscom AFB, Mass., 21 ordered values (1944-1964)

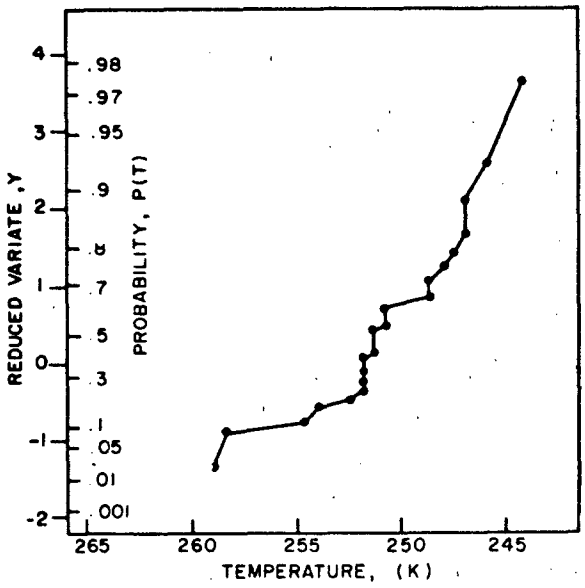


Figure 15-6. b) Lowest temperatures of 22 winter seasons (1943-1965) ordered from warmest to coldest, Hanscom AFB, Mass.

$$\hat{P}_T = \frac{i - 0.44}{N + 0.12} \qquad (15.19)$$

rather than Equation (15.4). Now in the example of Figure 15-6a, we have the plot of the annual highest temperatures of 21 years (1944-1964 at Hanscom AFB, Mass.) ordered from lowest to highest value and having cumulative probability estimates P_T given by Equation (15.19). The mean is $\bar{T}$ = 309 K and the standard deviation is s_t = 1.9 K. The solution of Equation (15.14) gives the straight line plot between y and T as shown. Whether the straight line and therefore the Gumbel distribution adequately fits the distribution is a matter of judgment. If accepted, and it should be in this example, then the 99th percentile (P_T = .99) or the 1% extreme is estimated by Equations (15.17) and (15.18) with n = 1 as 315 K. For a lifetime of 25 years (n = 25) the temperature of 10% risk (P_1 = 0.9) is given by Equations (15.17) and (15.18) as 316 K.

As another example, Figure 15-6b shows the plot of the extreme low temperatures of 22 winter seasons (1943-1965) at Hanscom AFB, Mass. The mean is $\bar{T}$ = 251 K and the standard deviation is s_t = 3.68 K. A straight line fit of these data is not satisfactory. Possibly a concave curve would be more appropriate. The Gumbel model is not acceptable in this case, and consequently another model should be tried.

5.1.2.7 Temperature Cycles and Durations. High temperature extremes are inevitably part of a well pronounced diurnal cycle, modified by wind and by moisture content. Typical of a hot climate, the record of Yuma, Arizona (32°51'N, 114°24'W) (Figure 15-7) reveals a mean diurnal temperature range of 15.3 K for the middle 20 days in July.

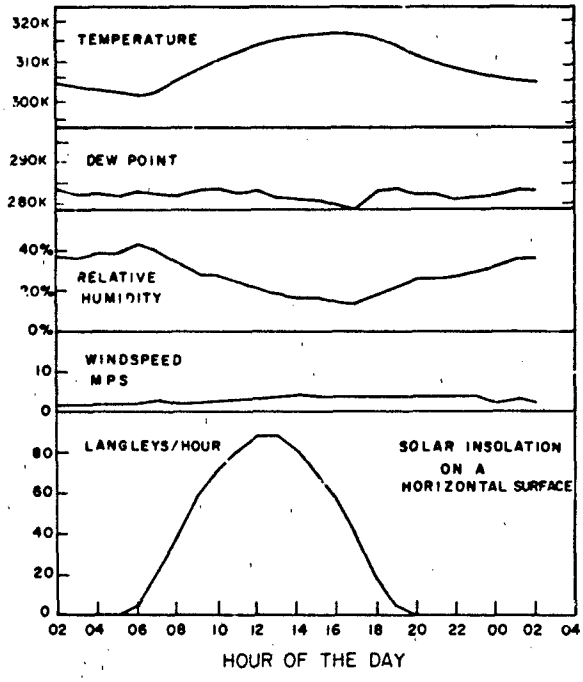


Figure 15-7 Yuma, Arizona typical July diurnal cycles when maximum daily temperature equals or exceeds 317 K (based on 1961-1968 data).

The dewpoint has a median of 287 K with a small diurnal range. Relative humidity, consequently, has a large-amplitude diurnal cycle. Wind speed at anemometer levels of 6 to 8 m above ground averages approximately 4 m/s with little diurnal range. Solar insolation, on the other hand, has a large diurnal range with a maximum clear-sky value of 88.2 L/h and a minimum value of zero from 2000 LST in the evening till 0500 LST in the morning. For the hottest areas on earth (for example, Sahara Desert) Table 15-4 presents the associated cycles of temperature, relative humidity, windspeed and solar insolation when the afternoon temperature in the middle of a 5-day period reaches 322 K which occurs about 1% of the time in the hottest month.

Death Valley, California is also one of the hottest areas but is close to 60 m below sea level resulting in extreme absorption of solar radiation before it reaches the ground. Consequently its maximum clear-sky solar insolation of 82.5 L/h is less than that shown in Figure 15-7. Solar insolation I increases with elevation roughly in accord with the exponential model given by

$$I = I_1 e^{-a(p - p_1)} \qquad (15.20)$$

where p and p_1 are, respectively, atmospheric surface pressures for a given station and another reference station at roughly the same latitude, I_1 is the solar insolation at the reference station, and the value for a is dependent on the location. For Yuma and Death Valley, where the mean

Table 15-4. Diurnal cycles of temperature and associated other elements for days when the maximum temperature equals or exceeds the operational 1% extreme temperature (322 K) in the hottest month in the hottest area.

	Time of Day (h)											
Item	1	2	3	4	5	6	7	8	9	10	11	12
	Temperature (K)											
Hottest Day	308	307	307	306	306	305	306	308	311	314	316	317
1 day before or after	309	308	307	306	306	305	306	309	311	313	315	317
2 days before or after	307	307	306	306	305	305	306	308	310	312	314	315
	Other Elements											
Relative Humidity (%) (dp = 266 K)	6	7	7	8	8	8	8	6	6	5	4	4
Windspeed (m/s)	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	4.3	4.3	4.3
Solar Radiation (L/H)	0	0	0	0	0	5	23	43	63	79	90	96
	Time of Day (h)											
Item	13	14	15	16	17	18	19	20	21	22	23	24
	Temperature (K)											
Hottest Day	320	321	321	322	321	321	319	315	314	312	311	310
1 day before or after	318	320	320	321	320	319	317	315	312	311	310	309
2 days before or after	316	317	319	320	319	318	317	314	312	311	310	309
	Other Elements											
Relative Humidity (%) (dp = 266 K)	3	3	3	3	3	3	3	4	5	6	6	6
Windspeed (m/s)	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	2.7
Solar Radiation (L/H)	96	90	79	63	43	23	5	0	0	0	0	0

atmospheric surface pressures are about 1006 mb and 1020 mb respectively, $a = 0.00461$ mb^{-1}.

The hottest locations in the Sahara Desert are relatively high (about 300 m above sea level) with atmospheric pressure about 977 mb. Thus Equation (15.20) yields an estimate for the peak solar insolation at these elevations of about 100 L/h. Most countries, however, including the U.S., Canada, and the United Kingdom, have adopted a peak figure for solar insolation for operational and design purposes of 96 L/h.

Heavy clouds and precipitation reduce the incident solar insolation. At a few stations the National Weather Service has taken records of incoming solar insolation. Table 15-5 gives the results of processing such data from Albuquerque, N.M. It presents estimates of the probabilities with which daily incoming solar insolation equals or exceeds the given amount in June. In contrast, Table 15-6 gives corresponding results for the insolation at Caribou, Maine where there is much more frequent cloudiness and precipitation.

The operability of equipment in a cold climate is very much dependent on the *duration* of extreme cold. Unlike the hot extremes, cold extremes are usually accompanied by very small diurnal ranges, if any. The direct approach for determining the duration of cold temperature is by an analysis of hourly data. Such data are available for many stations in North America but are not generally available for other regions of the world. Data from 108 stations in the U.S. and Canada have been analyzed [Tattelman, 1968] to obtain information on the longest period of time during which the temperature remained at or below eight "threshold" values (from 273 K to 220 K) during a 10-yr period. Figure 15-8 from that report shows the results for a threshold

Table 15-5. Probability of daily solar insolation equaling or exceeding given amounts for given number of consecutive days in June, at Alburquerque, N.M. Station elevation is 1620 m. Peak clear sky solar insolation was observed at 910 L/day.

Insolation L/Day	No. of Consecutive Days 1	2	4	8	15	30
850	0.03					
800	0.24	0.06				
750	0.49	0.28	0.09			
700	0.71	0.54	0.31	0.11		
650	0.31	0.70	0.52	0.29	0.09	
600	0.90	0.84	0.77	0.52	0.29	0.08
550	0.935	0.90	0.81	0.65	0.44	0.20
500	0.955	0.93	0.87	0.75	0.57	0.33
450	0.971	0.95	0.91	0.82	0.67	0.43
400	0.985	0.972	0.946	0.87	0.81	0.65
350	0.9933	—	0.973	0.945	0.902	0.82
300	0.9946	—	0.980	0.958	0.928	0.86
250	0.9975	—	—	0.980	0.963	0.932
200	0.999	—	—	—	—	0.970

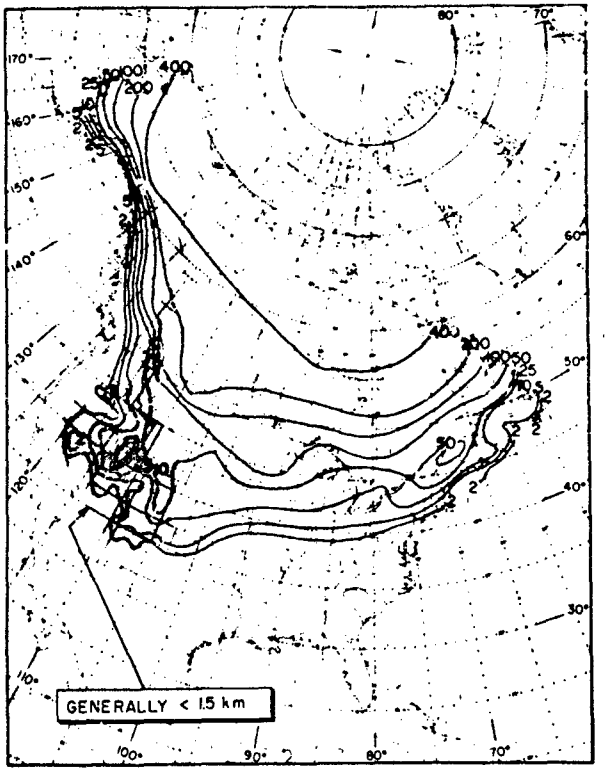


Figure 15-8. Longest duration (h) of temperature ≤ 250 K in ten winters [Tattelman, 1968].

Table 15-6. Probability of daily solar insolation equaling or exceeding given amounts for given number of consecutive days in June, at Caribou, Maine. Station elevation is 190 m. Peak clear sky solar insolation was observed at 843 L/day.

Insolation L/Day	No. of Consecutive Days 1	2	4	8	15	30
850						
800	0.019					
750	0.085	0.02				
700	0.20	0.057				
650	0.31	0.13	0.02			
600	0.40	0.20	0.05			
550	0.47	0.26	0.086			
500	0.55	0.33	0.13	0.024		
450	0.59	0.40	0.18	0.037		
400	0.66	0.50	0.26	0.075		
350	0.72	0.56	0.33	0.12	0.02	
300	0.78	0.67	0.47	0.22	0.062	
250	0.82	0.72	0.54	0.30	0.10	
200	0.88	0.80	0.66	0.42	0.22	0.05
150	0.921	0.87	0.77	0.59	0.37	0.13
100	0.965	0.943	0.90	0.79	0.63	0.40
90	0.975	0.96	0.92	0.84	0.72	0.51
80	0.980	0.962	0.93	0.86	0.75	0.56
70	0.987	0.978	0.956	0.912	0.83	0.70
60	0.9931	—	0.971	0.947	0.903	0.82
50	0.9961	—	—	0.977	0.95	0.90
40	0.99906	—	—	—	—	0.97

temperature of 250 K. The report also presents the expected (approximately 50% probability) duration of the temperature at or below six "threshold" temperatures (from 273 K to 232 K) during a single winter season. Figure 15-9 shows the single winter results for a threshold value of 250 K.

Estimates of duration have been made using data that consisted mainly of daily, monthly and annual average maximum and minimum, and monthly and annual absolute maximum and minimum, for some 35 to 50 years at Siberian, Yukon and Alaskan stations. The mean January temperature in eastern Siberia (Verkhoyansk and Oimyakon) is 225 K. Table 15-7 presents estimates of the lower 20% of the average temperature (averaged for durations ranging from one hour to 32 days), the maximum temperature for the durations shown, and the minimum temperature for the same durations.

The duration of temperature anywhere, hot or cold, is of general interest. In the midlatitude belt the temperatures of Minneapolis, Minn. are typical (Figure 15-10). The January probability distribution of all hourly temperatures has a 1% value of 244 K, and a 50% or median value of 263 K. That is, the range from the lower 1% to the median is 19 K. The 24-h averages, as expected, have a narrower range, 17 K. The range of monthly averages (768 h) is much narrower, 7 K. Similarly, the July hourly temperatures have

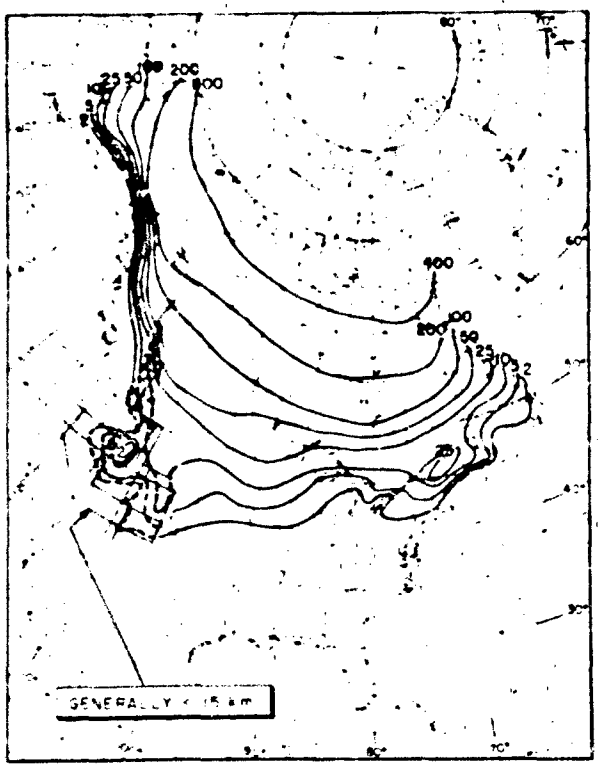


Figure 15-9. Expected longest duration (h) of temperature ≤ 250 K during a single winter season [Tattelman, 1968]

a 12 K range from the 50% value of 295 K to the upper 1% value of 307 K. The 24-h averages have an 8 K range and monthly averages have only a 2 K range. These figures imply a relatively high hour-to-hour correlation. Correlation analysis has provided estimates of 0.982 in the midwinter month of January, and 0.919 in the midsummer month of July.

Table 15-7. Durations of cold temperatures associated with the 222 K extreme. Each temperature in this table is the maximum, average, or the minimum in an operational time exposure of m hours with 20% probability of occurrence, during January, in a Siberian cold center

	Time m(h)									
	1	3	6	12	24	48	96	192	384	768
Maximum Temperature K	222	223	224	225	226	228	230	334	238	241
Average Temperature K	222	222	222	222	222	222	222	223	223	224
Minimum Temperature K	222	222	220	219	217	216	215	213	211	210

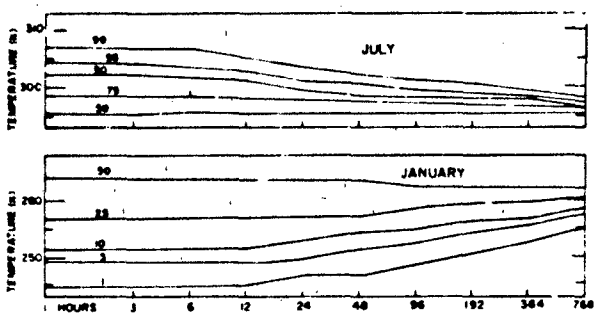


Figure 15-10. The distribution of the averages of consecutive hours of temperature at Minneapolis, Minn. (The upper half is for mid-summer month, from 1 July to 1 August; the lower half is for mid-winter month, 1 January to 1 February. Each curve is labeled with percent probability of occurrence.)

Hourly observations have been taken at Minneapolis for many years, making many useful summaries possible. Figure 15-11 shows a sample distribution (1949–1958) of hourly January temperatures alongside the left axis and the distribution of m-hour minima over the body of the graph, m from 1 hour to 768 hours (1 Jan to 1 Feb inclusive). Figure 15-12 shows the sample distributions of hourly temperatures in January of m-hour maxima. As an example of the usefulness of such a chart, freezing conditions (≤273 K) are shown as 94% frequent for 1-h durations. For 24 consecutive hours this frequency reduces to 83%, for 8 days (192 h) to 42% and for 16 days (384 h) to 10%.

15.1.3 Upper Air Temperature

The temperature data discussed in this section are from direct and indirect observations obtained from balloon-borne

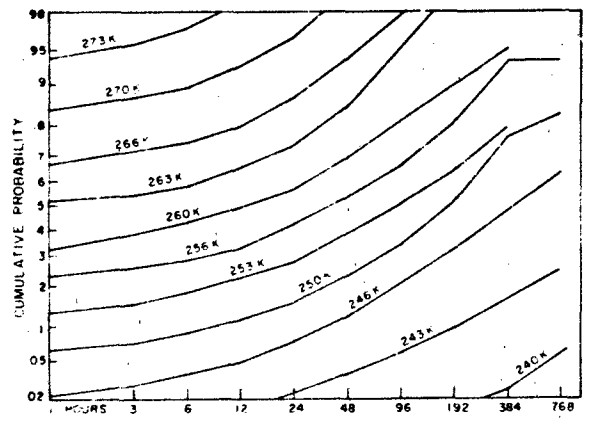


Figure 15-11. The cumulative probability of the M-hour minimum temperature (1 January to 1 February) at Minneapolis, Minn.

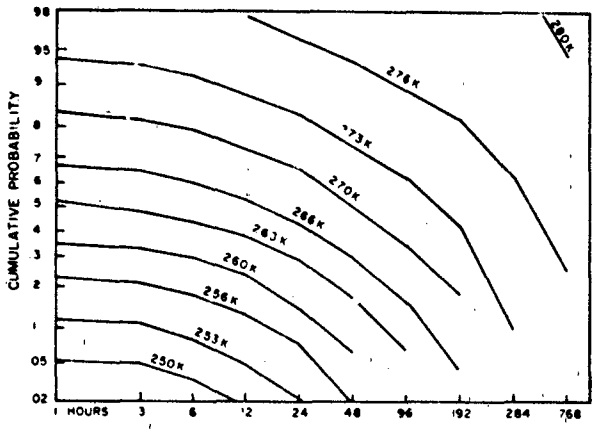
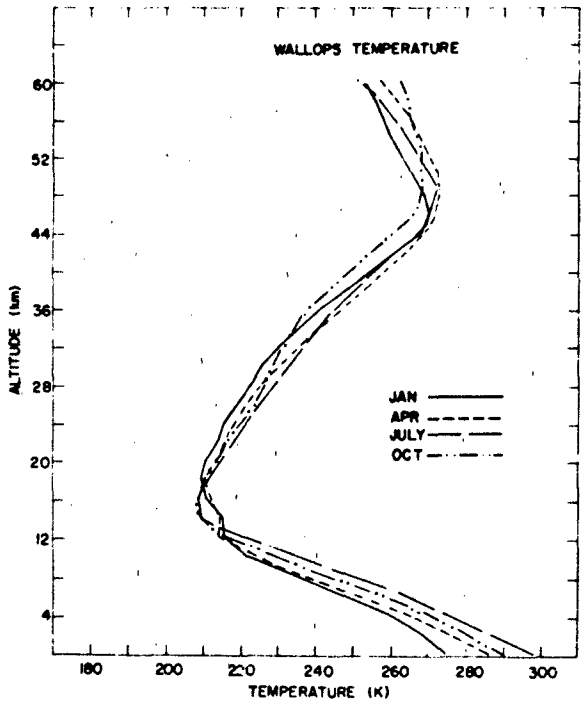
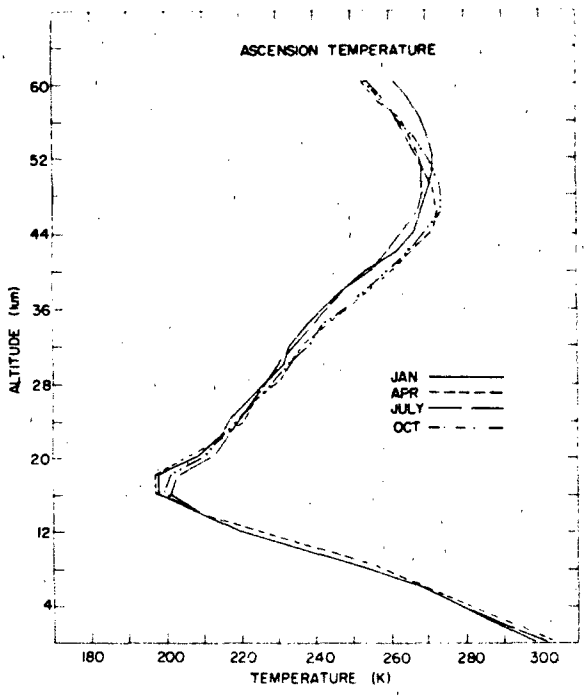


Figure 15-12 The frequency of duration (h) of the temperature (≤ T) in the mid-winter month (1 January - 1 February) at Minneapolis, Minn. (based on 1943–1952 data.)



instruments, primarily radiosondes, for altitude up to 30 km and from rockets and instruments released from rockets for altitudes between 30 and 90 km.

15.1.3.1 Seasonal and Latitudinal Variations. The Reference Atmospheres presented in Chapter 14 provide tables of mean monthly temperature-height profiles, surface to 90 km, for 15° intervals of latitude between the equator and North Pole. These profiles depict both the seasonal and



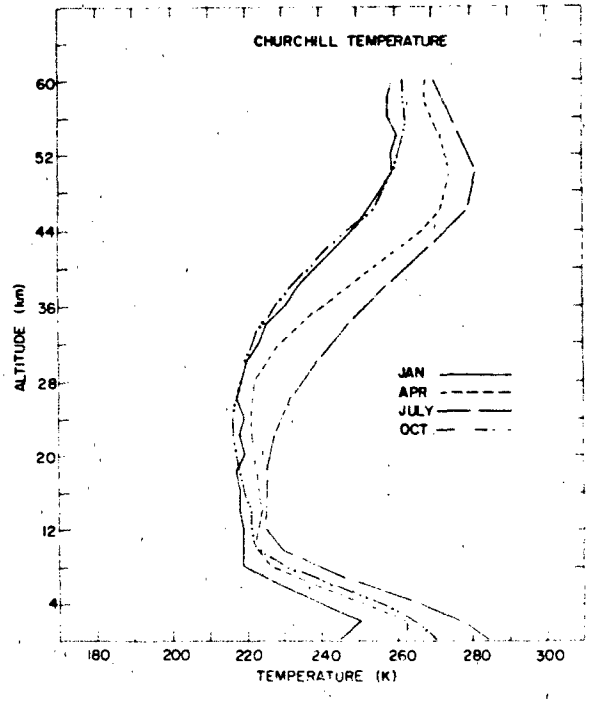


Figure 15-13 Seasonal differences in the temperature-altitude profiles at Ascension Island, Wallops Island, and Ft. Churchill.

Table 15-8a. Median, high, and low values of temperatures for January and July at 30°N

Altitude (km)	Median (K)	1% High (K)	1% Low (K)	10% High (K)	10% Low (K)	20% High (K)	20% Low (K)
January							
5	262	272	251	267	256	265	258
10	229	239	219	235	223	233	225
15	208	221	198	216	203	214	205
20	208	222	200	216	203	214	204
25	220	231	210	226	216	224	217
30	229	239	218	236	224	234	226
35	240	254	222	248	232	245	235
40	252	270	240	262	249	258	250
45	264	283	253	277	258	272	260
50	266	281	256	276	260	273	262
55	254	272	231	267	243	263	248
60	243	254	223	248	232	246	235
65	231	254	218	242	226	238	228
70	220	235	198	227	204	225	210
75	218	253	197	237	203	227	208
80	209	243	187	230	194	217	197
July							
5	272	278	262	274	266	275	268
10	238	249	227	246	232	242	234
15	204	216	196	211	200	210	200
20	212	223	203	218	206	216	206
25	223	230	216	227	218	226	219
30	234	241	226	238	229	236	231
35	244	254	237	250	240	247	242
40	256	267	247	263	251	261	253
45	266	275	259	272	264	269	265
50	269	282	258	278	262	275	264
55	264	273	247	269	253	267	256
60	247	262	231	255	240	252	243
65	228	240	215	236	219	234	222
70	209	222	186	219	194	214	200
75	200	218	178	214	192	209	196
80	193	207	182	200	189	198	191

Table 15-8b. Median, high, and low values of temperatures for January and July at 45°N.

Altitude (km)	Median (K)	1% High (K)	1% Low (K)	10% High (K)	10% Low (K)	20% High (K)	20% Low (K)
January							
5	250	263	233	257	239	254	242
10	220	233	206	227	212	225	214
15	217	231	202	225	208	222	211
20	215	227	203	222	208	220	210
25	215	233	197	226	205	224	209
30	221	240	209	230	214	226	219
35	233	258	215	251	223	243	226
40	247	272	226	264	236	257	240
45	262	288	240	283	250	271	254
50	265	282	249	274	256	270	258
55	253	275	229	267	239	263	245
60	244	266	220	263	230	257	241
65	235	255	214	246	223	243	228
70	226	246	206	238	211	234	217
75	225	261	197	245	205	235	210
80	216	248	185	237	197	228	202
July							
5	267	277	255	274	259	272	262
10	235	247	222	240	227	239	230
15	216	227	205	222	206	220	212
20	219	233	207	227	213	225	215
25	225	233	216	229	217	228	221
30	234	242	228	239	231	237	232
35	245	254	238	250	241	248	243
40	256	268	250	265	254	263	255
45	268	280	260	276	263	272	265
50	273	283	264	279	268	277	270
55	264	273	249	269	255	267	260
60	247	270	230	264	235	260	238
65	230	245	216	241	223	238	220
70	213	226	188	219	196	216	202
75	195	210	175	205	186	201	190
80	183	203	154	195	163	191	170

latitudinal variations in mean monthly temperatures. The largest seasonal variations in temperature occur at altitudes between 70 and 80 km near 75°N latitude. In this region the mean monthly temperature fluctuates from 230 K in January to 160 K in July. In the upper mesosphere, 60 to 85 km, mean monthly temperatures decrease toward the pole in summer and towards the equator in winter. In the upper stratosphere, 20 to 55 km, conditions are reversed; temperature decreases toward the pole in winter and toward the equator in summer. At altitudes between 15 and 20 km temperature decreases toward the equator in all seasons. Temperature-altitude profiles, surface to 60 km, for the mid-season months at Ascension Island, 8°S, Wallops Island, 38°N, and Ft. Churchill, 59°N, are given in Figure 15-13 and illustrate the magnitude of the seasonal and latitudinal variations in mean monthly temperatures.

5.1.3.2 Distribution Around Monthly Means and Medians. The distributions of observed temperatures around the median values for altitudes up to 80 km in January and July at 30°, 45°, 60° and 75°N are shown in Tables 15-8a to 15-8d. Median, and high and low values that are equaled

Table 15-8c. Median, high, and low values of temperatures for January and July at 60°N

Altitude	Median	1%		10%		20%	
		High (K)	Low (K)	High (K)	Low (K)	High (K)	Low (K)
January							
5	240	255	225	249	231	246	234
10	217	231	203	224	209	222	211
15	217	231	203	225	209	222	212
20	215	236	194	226	204	222	208
25	212	241	185	229	197	223	203
30	216	253	203	235	204	225	210
35	221	277	204	259	209	238	214
40	227	300	206	278	211	246	219
45	243	303	219	282	225	255	231
50	251	289	226	280	240	271	245
55	251	283	225	275	233	256	238
60	243	271	210	261	224	253	234
65	238	262	208	258	218	249	222
70	239	264	212	253	219	249	225
75	232	255	180	249	203	246	213
80	223	248	173	243	195	239	204
July							
5	260	271	250	266	254	264	256
10	225	238	214	233	219	231	221
15	225	235	217	231	221	229	223
20	225	233	219	230	222	229	223
25	229	236	222	233	225	232	226
30	239	245	232	243	234	241	235
35	252	258	243	256	247	253	248
40	265	272	259	269	263	268	262
45	277	287	271	283	274	280	275
50	279	290	273	286	277	284	279
55	271	278	257	275	264	273	266
60	259	273	212	265	250	263	253
65	228	259	225	253	230	248	233
70	214	239	202	226	208	222	211
75	190	202	178	196	182	194	186
80	166	180	142	176	153	174	155

Table 15-8d. Median, high, and low values of temperature for January and July at 75°N

Altitude (km)	Median (km)	1%		10%		20%	
		High (K)	Low (K)	High (K)	Low (K)	High (K)	Low (K)
January							
5	235	216	222	241	[illegible]	[illegible]	230
10	214	224	202	219	207	217	209
15	209	219	195	213	201	211	203
20	204	225	179	215	189	210	194
25	205	233	181	221	193	216	198
30	209	255	194	231	198	224	202
35	219	256	199	249	210	236	213
40	229	284	207	256	219	248	224
45	239	281	203	264	224	260	233
50	249	282	201	265	225	259	229
55	255	291	208	262	221	253	226
60	247	303	206	263	213	255	219
65	238	310	186	277	202	263	209
70	242	297	166	277	201	261	207
75	234	289	183	259	201	261	207
80	224	277	165	254	194	240	201
July							
5	254	264	244	259	248	257	250
10	229	238	219	234	223	232	225
15	230	237	225	235	228	233	229
20	230	237	227	235	228	234	229
25	230	240	226	238	227	237	229
30	243	262	233	247	235	246	240
35	256	262	238	260	246	258	250
40	268	275	252	271	260	270	262
45	281	292	268	287	275	284	278
50	284	296	270	291	279	288	280
55	281	288	254	284	270	283	275
60	268						
65	246	(insufficient data above 55 km in summer)					
70	218						
75	189						
80	161						

or more severe 1%, 10% and 20% of the time during these months are given for 5-km altitude increments between the surface and 80 km. Distributions below 30 km are based on radiosonde observations taken in the Northern Hemisphere, and those above 30 km are based on meteorological and experimental rocket observations taken primarily from launching sites in or near North America. The 1% values are considered to be rough estimates as they are based on the tails of the distributions of observed values plotted on probability paper. Estimates of values for altitudes above 50 km are less reliable than those below 50 km because of the paucity of data and larger observational errors at the higher altitudes. Only median temperatures are given above 55 km at 75°N for July due to the small number of observations that are available for the higher altitudes over the polar regions in summer.

In tropical regions, 0° to 15° latitude, the day-to-day variations of temperature are normally distributed about the mean at altitudes up to 50 km. Consequently, a reasonably accurate estimate of the distribution of temperature at a given altitude can be obtained from the standard deviations and

Table 15-9 Standard deviations of observed day-to-day variations in temperatures (K) at Ascension Island (8°S) at altitudes up to 50 km during the midseason months

Altitude (km)	S. D. of Temperature (K)			
	Jan	April	July	Oct
5	0.8	0.6	0.7	0.6
10	0.8	1.0	0.8	1.1
15	1.6	2.0	1.9	1.5
20	2.2	2.2	2.4	2.1
25	2.2	2.2	2.7	2.1
30	3.1	2.8	3.8	3.6
35	3.7	3.2	3.7	3.8
40	5.2	3.9	3.3	3.5
45	3.6	2.8	3.2	3.3
50	5.8	2.9	3.9	3.0

the montaly means. The standard deviations of observed temperatures around the mean monthly values for the midseason months at Ascension Island, Table 15-9, are typical of the day-to-day variations found in the tropics. Values are not given for altitudes above 50 km as there are too few daily observations on which to base the monthly temperature distributions. The observed standard deviations includes the rms instrumentation errors as well as the actual rms climatic variations. Consequently, the observed variations are somewhat larger than the actual values.

Day-to-day variations of temperature around the annual mean at levels between 50 and 90 km in tropical areas (Table 15-10) were computed from data derived from grenade and pressure-gage experiments at Natal, 6°S, and Ascension Island, 8°S. These data were not uniformly distributed with respect to season or time of day. An analysis of the relatively sparse data that are available for individual months indicates that if the seasonal and diurnal variations are removed from the data, standard deviations around monthly means due to day-to-day changes in synoptic conditions would be roughly 50% of those given in Table 15-10.

5.1.3.3 Distributions at Pressure Levels. The mean January and July temperatures over North America for standard pressure levels up to 10 mb (≈31 km) are presented in Table 15-11. Standard deviations of the daily values around these means are also shown, thereby providing information on seasonal changes in monthly mean temperatures and interdiurnal (day-to-day) variability at various pressure levels and latitudes. Standard deviations are not shown above 100 mb north of 50° latitude because a bimodal temperature distribution exists in the winter stratosphere in arctic and subarctic regions over eastern North America. As a result, the standard deviations do not provide reliable information on the temperature distributions at these levels.

15.1.3.4 Interlevel Correlation of Temperature. The manner in which the correlation between temperatures at two levels decreases (or decays) with increasing separation between the levels is an example of the general problem of correlation decay. Correlation decay is similar for most meteorological elements as the horizontal or vertical distance between the points of observations increases. As yet, no fully satisfactory description of the decay rate, based on fundamental properties or assumptions, is available. Consequently, many empirical models that are valid for specific elements over restrictive ranges have been proposed.

Profiles of correlation coefficient r, of surface temperature with temperature at other altitudes are shown in Figure 15-14 for the midseason months at Ascension Island, Kwajalein, Wallops Island, and Ft. Churchill. At most locations, the correlation between surface temperatures and temperatures at other altitudes decreases rapidly with increasing altitudes, reaching a minimum or becoming negative between 12 and 16 km and then remaining near zero, plus or minus 0.3, from 20 to 60 km. Individual arrays of the mean temperatures, standard deviations and interlevel correlation coefficients for altitudes to 60 km are given in Table 15-12a to 15-12f for the months of January and July at Ft. Churchill, Wallops Island, and Kwajalein. Additional in-

Table 15-10 Standard deviations of observed densities (%) and temperatures (K) around the mean annual values of Ascension Island (8°S) Natal (6°S)

Altitude (km)	Density S.D. (% of mean)	Temperature S.D. (K)	No. of Observations
50	4.1	6	33
55	4.3	3	33
60	4.8	6	33
65	4.7	7	33
70	6.4	9	32
75	8.6	10	31
80	7.8	10	30
85	10.2	13	29
90	12.3	21	28

Table 15-11. Mean temperature and standard deviation at standard pressure levels over North America

Pressure (mb)	20°N Mean	20°N S.D.	30°N Mean	30°N S.D.	40°N Mean	40°N S.D.	50°N Mean	50°N S.D.	60°N Mean	60°N S.D.	70°N Mean	70°N S.D.	80°N Mean	80°N S.D.
	Mean Temperature and Standard Deviation (K)													
	January													
700	280	2	275	5	267	6	256	8	251	8	247	7	245	6
500	264	3	259	4	252	6	242	7	238	7	234	5	232	5
300	236	3	232	3	227	4	221	4	220	4	217	4	214	5
200	217	3	216	5	216	6	219	7	219	7	216	6	213	6
100	198	3	204	4	212	4	218	−5	219	6	216	7	210	6
50	208	3	209	3	213	3	215	4	216	*	213	*	206	*
25	218	2	218	3	216	4	212	5	212	*	208	*	203	*
15	225	2	223	3	221	4	218	6	215	*	211	*	207	*
10	230	2	227	3	224	4	221	−6	217	*	213	*	209	*
	July													
700	283	2	283	2	282	3	275	4	270	3	268	4	265	4
500	267	2	267	2	264	3	258	4	255	4	253	4	249	4
300	240	2	240	2	237	3	232	4	229	4	228	4	227	4
200	218	2	218	2	219	3	221	5	224	5	226	5	230	4
100	200	3	203	3	210	3	220	4	226	3	228	2	231	2
50	213	2	215	2	218	3	221	3	226	3	228	3	230	3
25	222	2	222	2	225	2	227	2	229	2	232	2	233	2
15	228	2	228	2	229	2	232	2	235	3	236	3	236	3
10	232	2	233	2	234	2	237	2	239	3	240	3	241	3

*Not normally distributed

formation useful in design studies is given by Cole and Kantor [1980].

15.1.4 Speed of Sound vs Temperature

The speed of sound is primarily a function of temperature. An equation for computing the speed of sound and the limitations of such computations are presented in Chapter 14. Figure 15-15 shows the relationship between temperature and the speed of sound. It can be used with the various temperature presentations given in this section to estimate the probable speed of sound for various altitudes and geographical areas.

15.1.5 Earth/Air Interface Temperatures

The earth/air interface is either a land, snow, or water surface. At many locations, the physical structure of the interface is overwhelmingly complex. The land surface can be covered with seasonally varying vegetation of great diversity, and even without plant cover there is normally a considerable variability produced by small-scale terrain features, differences in soil moisture and cultivation. A snow surface is markedly affected by aging. The physical conditions of water in a shallow puddle are quite different from the open ocean. All these conditions reflect themselves in the micro-climatological aspects of natural or unnatural surfaces.

As discussed in Section 15.1.2, the use of ordinary thermometers to measure surface temperature, will result in meaningful values only in the rare cases of a flat, uniform, and homogeneous surface. In general, area averages of temperature obtained by an integrating method over certain defined sections will be more representative than any one of a multitude of widely varying point values. Bolometric temperature measurements from an airplane cruising at low altitude provide a more reasonable approach to the problem of surface temperature determination than a series of thermometric point measurements. Table 15-13 lists some results of bolometric measurements from an airplane. The data illustrate the great horizontal variability of surface temperature even when effects on the scale of less than 6 m linear dimension are averaged out.

The processes that determine the temperature of the earth/air interface and the surface characteristics that influ-

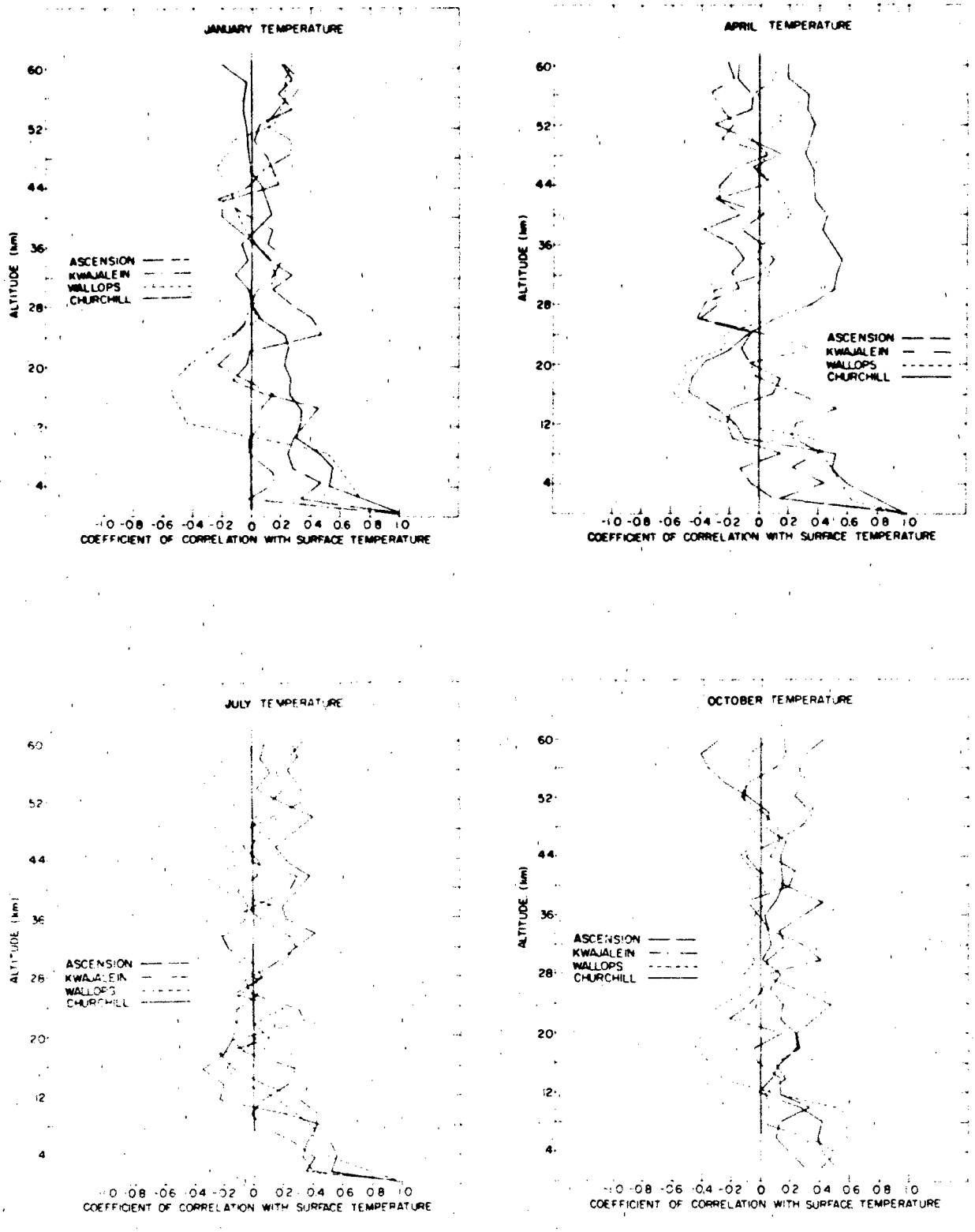


Figure Figure 15-14. Vertical profiles of interlevel coefficients of correlation of surface temperature with temperature at other altitudes up to 60 km for the mid-season months at Ascension Island, Kwajalein, Wallops Island, and Ft. Churchill

ence these processes may be separated into the following four classes:

1. radiative energy transformation (or net radiation intensity), which depends upon the albedo and selective absorption and emission;

2. turbulent heat transfer into the air (by both convective and mechanical air turbulence);

3. conduction of heat into or out of the ground, which depends upon the thermal admittance of the soil; and

4. transformation of radiant energy into latent heat by evaporation, which depends upon the dampness of the surface or available soil moisture at the ground level.

The aerodynamic roughness of a natural surface strongly influences the momentum exchange between ground and air flowing past it. The momentum exchange establishes the low-level profile of mean wind speed. The mechanical turbulence produced by surface roughness also determines to a certain degree the relative amount of heat transported into or from the air at mean ground level. Other conditions being equal, an increase in roughness and hence mechanical turbulence will cause lowering of maximum surface temperature during daytime and raising of minimum surface temperature during nighttime. For ordinary sandy soil, under average conditions of overall airflow and net radiation on summer days in temperate zones, the diurnal range of surface temperature is about 17 K if the roughness coefficient is 0.06 mm or 14 K if it is 6.35 mm (roughness coefficient, also called roughness "length," is $\varepsilon/30$ where ε is the average height of surface irregularities).

A special and rather extreme case of the influence of surface characteristics is represented by forests. The trees intercept solar radiation and the heat absorbed is given off into the air that is trapped between the stems. Although deep snow may lie on the ground, daytime temperatures in wooded areas in spring can reach 289 K.

The thermal admittance (Section 15.1.6) of most soils depends on porosity and moisture content. Because both the thermal conductivity and heat capacity of soils increase with soil moisture, the thermal admittance may be significantly affected by humidity variations during rainy or wet weather periods, whereas the normal diffusivity may remain unaltered. These effects are difficult to assess, however, because the dampness of the surface is also a major factor in the utilization of solar energy for evaporation. If soil moisture is readily available at the earth's surface, part of the net radiation that would have been used for heating air and ground is used instead for latent heat of evaporation. Table 15-14 lists observed temperatures in the air and soil at levels close to the earth/air interface.

Engineers must consider the effect of albedo and color on net radiation in artificially changing surface or ground temperature. In India, a very thin layer of white powdered lime dusted over a test surface made ground temperatures up to 15 K cooler; the effect was felt at a depth of at least 20 cm.

Another effective method of controlling surface temperature is shading. Thin roofs (metal, canvas), however, may attain a temperature so high that the under surface acts as an intense radiator of long-wavelength radiation, thus acting to warm the ground. In hot climates, multilayer shades with natural or forced ventilation in the intermediate space, or active cooling of the outer surface by water sprinkling, can be used to cool the ground with some success. Table 15-15 compares temperature measurements of various material surfaces with corresponding air and soil temperatures.

15.1.6 Subsoil Temperatures

The thermal reaction of the soil to the daily and seasonal variations due to the earth's rotation and its revolution about the sun of net radiation is governed by the *molecular thermal conductivity* of the soil, k, and by the *volumetric heat capacity* of the soil, $C = \rho c$ (where ρ is the density and c is the heat capacity per unit mass). For a cyclic forcing function of frequency n, the quotient $(nk/C)^{1/2}$ (which has the physical units of velocity) determines the downward propagation or *amplitude decrement with depth* of the soil-temperature response. The product $(nk\,C)^{1/2}$, which has the physical units of degrees divided by Langleys per unit time (1 L/s equals 4.186×10^4 W/m^2), governs the amplitude of the temperature profile in time at the soil surface. The ratio k/C is the *thermal diffusivity* (physical units of length squared per unit time). The expression $(kC)^{1/2}$ defines *thermal admittance* of the soil.

The continuous flow of heat from the earth's hot, deep interior to the surface is the order of 10^{-5} L/min. This is very small compared with a solar constant of 2 L/min, average net-radiation rates of 0.2 L/min, and induced soil-heat fluxes in the uppermost several feet of the earth's crust of 0.1 L/min. Only for depth intervals in excess of about 30 m must the heat flow from the earth's interior be considered, inasmuch as it results in vertical temperature gradients of the order of 2.5 to 25 K/km.

Table 15-16 gives experimental data on thermal admittance and theoretical values of the *half-amplitude depth interval* based on experimental thermal diffusivity data for diverse ground types. The smaller the thermal admittance, the larger the surface-temperature amplitude for a given forcing function. This latter inverse proportionality is valid only when turbulent heat transfer into the atmosphere is negligible.

In a simple theoretical model of thermal diffusion, an *effective atmospheric thermal conductivity* K is introduced. For air, K is many times larger than the molecular thermal conductivity of the air. For the same forcing function, the surface-temperature amplitudes at two different kinds of ground follow the ratio

$$\frac{(TAR)_2 + (K/k)_2^{1/2}}{(TAR)_1 + (K/k)_1^{1/2}} \qquad (15.21)$$

Table 15-12a. Ft. Churchill—Correlation of January temperatures (K) from surface to 60 km.

KM Kilometers above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	0[35]	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	244	250	240	228	219	219	219	218	218	217	219	218	219	217	218	219	223	225	230	233	238	243	248	252	255	258	258	259	257	257	258
STDV	75	58	49	43	45	52	61	70	75	78	63	69	70	98	96	88	85	89	109	125	166	171	187	173	160	148	143	147	146	138	127
N	50	50	50	50	50	50	50	50	46	40	30	29	23	45	48	50	51	51	51	51	51	51	51	51	51	51	51	51	49	46	36
2	72	--																													
4	53	84																													
6	55	75	82																												
8	46	23	17	47																											
10	30	3	-4	13	85																										
12	32	16	9	19	73	94																									
14	34	21	12	23	68	89	98																								
16	25	8	10	17	56	81	94	98																							
18	26	23	13	10	45	71	85	90	95																						
20	22	14	8	1	38	61	73	78	83	94																					
22	24	14	17	4	30	51	59	59	65	80	93																				
24	22	16	2	-5	9	28	33	28	40	62	84	98																			
26	6	1	0	-1	17	44	55	58	66	76	80	94	96																		
28	0	0	7	1	9	33	45	48	62	74	79	92	94	96																	
30	1	4	14	10	4	19	29	33	53	66	72	86	83	87	95																
32	11	3	11	-1	-22	17	13	-13	7	19	39	59	71	56	73	84															
34	1	10	21	8	23	29	29	29	15	5	18	37	50	32	50	64	88														
36	6	1	16	4	24	35	37	37	31	19	18	30	34	12	29	44	74	85													
38	5	1	8	2	18	28	34	35	30	17	9	15	21	1	13	27	60	78	92												
40	13	6	6	2	14	27	34	37	37	29	3	0	6	19	10	1	37	67	78	92											
42	9	1	4	2	15	32	42	46	50	46	16	16	10	39	29	16	26	51	70	84	93										
44	7	1	1	8	22	39	49	52	59	55	33	35	23	53	46	36	3	34	55	72	82	92									
46	0	6	6	15	30	44	53	56	63	59	36	39	32	52	47	41	3	25	46	60	74	83	93								
48	1	11	10	21	33	47	58	61	65	62	42	41	35	52	46	41	3	24	43	55	67	78	89	96							
50	1	11	17	29	33	44	52	53	57	52	36	36	35	45	42	40	14	11	27	39	52	59	73	81	87						
52	3	13	21	31	29	37	44	44	48	50	32	34	35	43	43	46	22	1	13	22	36	44	59	68	76	92					
54	5	12	20	30	25	31	37	37	43	51	41	46	46	47	49	54	34	13	1	8	22	32	47	56	63	79	93				
56	5	14	22	32	23	23	26	25	33	41	41	45	44	34	38	49	41	24	15	7	4	13	28	36	47	67	81	92			
58	4	14	20	28	16	15	17	15	21	36	43	46	49	28	33	45	45	34	31	27	20	9	10	18	34	59	74	85	95		
60	20	23	24	36	18	14	19	17	24	31	49	51	50	38	38	49	44	40	32	27	22	5	12	22	32	50	61	75	81	87	

**Multiply tabular values by 0.01 to obtain correlation coeffecients.

Table 15-12b. Ft. Churchill—Correlation of July temperatures (K) from surface to 60 km.

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	035	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	284	277	265	252	238	228	224	225	225	225	226	227	229	231	235	238	242	247	252	257	262	268	274	278	279	280	278	276	274	271	269
STDV	53	37	40	47	49	25	49	22	23	22	20	18	16	24	21	25	24	32	30	30	35	37	34	31	39	41	43	45	39	38	43
N	28	28	28	28	28	28	28	28	28	28	28	27	27	28	28	28	28	28	28	28	28	28	28	28	28	28	26	25	25	21	20
2	49	**																													
4	53	89																													
6	49	87	95																												
8	44	82	90	96																											
10	19	35	42	42	55																										
12	24	66	69	76	73	3																									
14	22	69	73	74	74	12	83																								
16	36	82	84	84	82	32	69	78																							
18	22	75	71	69	66	33	46	63	90																						
20	15	66	56	52	50	28	27	45	72	91																					
22	13	71	65	61	59	36	43	54	81	93	97																				
24	13	65	61	53	54	37	37	52	76	80	81	85																			
26	1	37	39	37	39	18	27	42	54	53	48	55	62																		
28	1	32	36	37	33	6	29	44	50	48	38	44	49	92																	
30	19	20	16	15	19	14	19	27	34	25	14	25	37	82	77																
32	26	9	17	11	12	12	15	33	28	27	16	27	31	78	76	82															
34	41	4	10	11	12	6	22	27	26	24	9	21	17	64	67	78	84														
36	22	20	21	23	24	6	32	32	28	17	4	14	27	56	54	71	62	72													
38	19	23	28	33	33	13	17	29	32	25	17	23	23	68	73	66	63	64	66												
40	30	8	16	21	19	8	5	12	34	41	34	35	32	68	70	58	56	63	49	69											
42	39	15	7	1	2	10	6	5	17	17	7	11	6	52	63	57	60	64	38	54	77										
44	23	11	1	6	1	7	18	9	17	11	6	5	0	45	54	46	51	59	47	44	62	80									
46	15	10	8	18	16	7	9	1	16	15	4	6	2	47	50	41	38	51	55	61	51	42	61								
48	29	15	13	12	12	10	0	10	0	1	9	7	0	37	40	38	48	55	51	41	45	38	52	71							
50	40	25	23	26	20	14	4	4	6	10	22	18	0	37	34	49	63	60	57	38	33	38	53	45	80						
52	20	9	3	5	1	9	22	26	15	5	27	19	3	42	43	59	70	58	54	40	25	47	56	40	64	90					
54	2	8	11	9	12	9	27	37	31	12	9	1	10	46	51	55	67	54	52	47	26	50	53	49	65	78	93				
56	12	7	12	7	9	4	24	34	29	11	13	3	8	43	51	57	69	64	57	54	34	54	63	52	68	81	91	95			
58	6	25	16	6	8	6	11	39	44	35	22	24	38	57	59	61	72	67	43	52	43	46	49	37	55	71	80	87	93		
60	9	9	4	6	7	19	2	14	29	19	5	7	26	50	58	54	55	58	36	48	41	47	56	46	60	66	69	77	81	90	

**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 15-12c. Wallops Island—Correlation of January temperatures (K) from surface to 60 km.

KM Kilometers above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	015	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	275	269	260	248	235	222	216	215	211	210	211	214	216	220	223	226	231	236	242	249	255	262	268	270	269	266	263	260	258	256	252
STDV	54	86	79	69	53	33	55	39	44	45	35	35	38	45	54	59	60	63	61	79	95	89	82	80	63	47	66	74	77	90	106
N	44	44	44	44	44	44	44	44	43	43	43	43	43	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	40	34	19
2	74	**																													
4	66	96																													
6	61	87	96																												
8	52	79	88	94																											
10	5	10	14	17	42																										
12	-44	-46	-56	-62	-55	13																									
14	-47	-65	-73	-76	-74	-13	74																								
16	-54	72	-78	-79	-79	-10	63	89																							
18	-49	-79	-82	-83	-81	-20	42	72	82																						
20	-41	-60	-59	-56	-57	-14	28	51	58	72																					
22	-27	-42	-41	-42	-48	17	19	28	37	60	81																				
24	-13	-32	-28	-27	-31	-8	-1	2	12	34	52	64																			
26	7	-19	-19	-14	-18	-7	-2	-6	-9	10	37	48	73																		
28	0	-19	-17	-15	-15	6	0	-3	-12	9	37	42	60	86																	
30	11	-25	-22	-17	-17	0	-5	-4	-12	14	35	38	52	79	83																
32	15	-19	-18	-15	-19	-11	-3	1	-7	11	21	13	31	60	63	82															
34	13	14	-19	-20	-23	-18	11	6	-6	6	3	-9	8	30	29	55	80														
36	3	-15	-23	-29	-33	-30	15	8	7	11	-3	-13	-2	16	15	31	58	81													
38	-4	2	-5	-11	-7	-9	4	-10	-4	-3	-25	-36	-24	-22	-26	-26	-2	30	60												
40	-10	3	-5	-9	-2	0	5	-13	-9	-10	-29	-42	-29	-28	-32	-31	-14	19	56	89											
42	-12	5	5	9	14	8	-3	-28	-20	-27	-43	-44	-23	-29	-29	-30	-17	6	29	68	79										
44	-16	3	5	10	12	8	3	-28	-17	-26	-35	-37	-29	-21	-21	-22	-6	3	26	54	63	85									
46	-23	-1	3	7	9	10	-2	-21	-12	-21	-30	-36	-28	-22	-21	-29	-12	-2	19	52	63	74	87								
48	-17	1	5	18	10	7	-14	-18	-14	-23	-33	-43	-34	-21	-14	-22	2	2	19	42	49	62	69	85							
50	-14	-7	-5	1	1	4	-3	-2	4	-10	-30	-38	-32	-21	-21	-15	16	20	33	45	40	42	49	61	76						
52	15	10	9	13	16	3	-7	-6	-5	-14	-37	-47	-40	-23	-16	-11	10	25	41	47	33	32	26	33	40	66					
54	20	11	10	11	15	6	-7	-4	2	-9	-23	-34	-34	-10	-3	0	15	17	32	25	8	10	16	12	18	36	80				
56	23	8	5	8	11	6	-12	-7	2	-5	-7	-15	-26	-1	0	1	14	11	23	12	0	3	8	10	21	29	58	82			
58	23	23	22	24	28	21	-8	-14	-9	-17	-19	-20	-37	-15	-13	-14	1	-5	7	2	-2	1	9	22	29	33	51	71	89		
60	27	19	22	21	29	51	-3	-23	-19	-29	-18	-9	-35	1	-4	-4	14	-5	10	13	7	3	20	32	35	43	38	59	77	95	

**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 15-12d. Wallops Island—Correlation of July temperatures (K) from surface to 60 km.

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	015	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	297	286	275	264	251	236	220	210	209	212	216	220	223	226	229	233	237	241	246	251	256	262	267	270	271	270	267	264	260	256	251
STDV	34	21	16	14	20	22	22	27	29	23	16	17	16	24	25	30	28	30	27	27	32	37	43	42	38	38	40	47	55	62	88
N	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	34	30	18
2	75	**																													
4	47	68																													
6	36	53	84																												
8	38	48	69	67																											
10	25	37	48	51	81																										
12	12	29	27	27	48	80																									
14	-8	-18	-26	-11	-28	-7	18																								
16	22	-48	-39	-46	-55	-50	-25	36																							
18	-10	-25	-14	-24	-35	-47	-32	31	67																						
20	0	-5	-11	-20	-11	-33	-32	17	48	57																					
22	-13	-11	-8	-5	-4	-15	-9	7	42	50	68																				
24	-22	-12	-3	-4	3	12	25	13	22	11	41	48																			
26	-10	-13	17	22	24	31	25	6	10	1	7	32	48																		
28	1	-1	21	19	33	41	27	-3	13	-7	14	35	46	70																	
30	-8	-8	17	8	18	32	25	-5	8	-18	-2	10	33	49	76																
32	-18	-6	23	10	11	17	20	-11	16	2	10	22	54	44	61	72															
34	-22	-19	15	11	28	37	27	-11	9	-15	7	14	45	64	62	60	63														
36	-32	-23	4	-3	13	25	29	4	31	14	27	50	65	62	67	54	61	68													
38	11	-14	17	25	29	28	6	-1	10	9	-2	19	34	61	62	56	47	49	44												
40	-20	-16	17	17	14	18	8	-32	-11	-20	-26	-6	17	27	23	40	47	50	30	49											
42	-33	-28	1	-1	-5	1	15	-10	22	0	-8	16	16	6	21	48	49	42	32	19	70										
44	-6	11	35	37	30	39	41	6	-2	14	-8	11	29	40	55	52	58	48	38	38	47	53									
46	-18	0	41	41	31	27	20	12	5	-2	-4	13	25	50	54	45	48	51	41	35	30	35	82								
48	-26	-13	18	31	21	25	13	13	-1	-13	-2	16	38	38	53	48	56	42	34	40	34	34	60	66							
50	-25	-15	8	13	23	32	21	-7	-5	-10	2	22	43	51	56	55	55	52	39	42	42	34	48	51	81						
52	-31	-32	-12	-16	9	10	14	-5	9	6	29	40	51	45	49	42	47	52	43	26	27	40	31	30	43	66					
54	-37	-32	-10	-16	9	8	14	-17	6	3	27	44	57	52	45	39	53	57	50	24	32	41	26	23	28	54	88				
56	-23	-16	-7	-18	15	10	15	-23	-1	5	42	39	64	37	30	19	46	47	45	22	26	20	18	3	13	32	66	86			
58	-20	-15	5	6	36	25	17	-17	-9	1	41	44	44	48	41	10	38	56	50	22	29	17	29	19	28	36	61	77	87		
60	-20	-7	8	1	50	48	35	-42	-22	-13	41	56	63	65	72	50	54	79	77	30	37	31	40	25	36	63	86	90	90	95	

**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 15-12e. Kwajalein—Correlation of January temperatures (K) from surface to 60 km.

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	008	2	4	5	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	301	288	279	267	255	241	224	208	195	192	206	212	217	221	225	228	232	237	242	247	253	257	262	267	271	272	272	271	265	265	263
STDV	14	13	13	13	14	14	15	16	15	46	28	26	23	30	27	30	36	36	37	42	42	38	42	48	63	64	51	41	44	49	58
N	42	42	42	42	42	42	42	42	42	42	42	42	41	40	41	42	42	42	42	42	42	42	42	42	42	42	42	42	41	38	34
2	−3	**																													
4	15	39																													
6	11	15	48																												
8	−1	−20	23	28																											
10	−2	8	47	43	70																										
12	5	13	44	52	48	85																									
14	5	−4	22	43	60	71	80																								
16	14	10	−13	6	19	8	15	28																							
18	−6	3	−19	−34	−41	−57	−52	−63	1																						
20	−23	−18	−12	0	4	1	−9	−5	−21	10																					
22	−13	−3	−15	−27	−21	−21	−27	−34	−33	33	37																				
24	−14	3	−12	7	−5	−26	−21	−7	−19	16	3	31																			
26	−4	−15	−1	−18	−2	−7	−18	−15	8	11	−28	−11	6																		
28	0	4	5	−9	1	−12	−18	−21	4	17	−9	−3	15	51																	
30	3	−1	9	−26	1	−16	−29	−19	1	25	−6	−6	−8	50	53																
32	16	−3	24	−13	0	−2	−14	−19	−2	24	−1	6	1	28	32	39															
34	20	19	−11	−25	−22	−23	−26	−29	43	29	−27	−20	−30	34	6	29	9														
36	11	31	6	−17	10	−16	−27	−38	27	38	0	−1	−25	19	19	29	37	47													
38	13	−9	−7	−28	6	3	−10	−19	17	26	25	−8	−39	24	−1	30	26	57	54												
40	−1	−1	−32	−40	8	−12	−19	−28	19	24	−16	2	−2	23	8	18	21	44	29	47											
42	−22	32	−3	−8	−7	−13	−18	−32	26	24	−4	−3	5	8	34	24	30	33	34	24	51										
44	−1	7	−33	−23	−22	−25	−22	−29	19	29	−3	−2	−11	18	14	25	9	51	15	43	54	64									
46	7	−20	−22	−43	−10	−24	−30	−19	−10	16	−6	−2	8	36	1	41	15	35	−4	32	7	8	37								
48	25	−10	−12	−22	−3	−21	−21	−16	26	23	−25	−26	4	32	32	33	38	58	28	49	33	35	42	53							
50	27	8	8	−7	−7	−9	−12	−22	26	11	−26	−43	−33	20	21	32	32	48	43	48	30	36	39	17	63						
52	9	20	13	4	−6	−6	−11	−22	18	18	5	−31	−32	6	27	33	21	49	45	55	24	35	32	6	53	77					
54	17	12	−9	−4	−8	−7	−26	−34	15	13	2	1	−24	3	14	7	11	39	42	38	31	24	30	1	44	54	64				
56	28	22	25	−9	2	−7	−21	−37	4	21	−2	12	−8	7	14	31	20	29	32	30	36	25	26	19	35	29	46	59			
58	33	6	32	−16	−4	−12	−20	−31	1	26	−28	−10	−17	26	25	30	13	23	27	19	16	3	1	16	31	39	48	47	65		
60	21	16	10	−2	−7	−7	−14	−24	13	18	−21	−42	−17	19	14	9	2	32	28	23	11	4	5	2	29	50	64	58	32	71	

**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 15-12f. Kwajalein—Correlation of July temperatures (K) from surface to 60 km.

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	.008	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	[illegible]	2.6	278	266	254	239	222	206	198	203	210	215	219	222	227	230	233	237	241	248	254	259	262	265	266	268	268	265	261	257	253
STDV	12	8	10	12	14	17	16	17	17	24	17	15	20	20	21	34	31	43	42	55	34	35	37	50	48	49	58	69	65	69	61
N	31	31	31	31	31	31	31	31	31	31	31	30	29	25	30	31	32	32	32	32	32	32	32	32	32	32	32	31	31	30	27
2	33	**																													
4	38	72																													
6	23	66	78																												
8	2	49	44	52																											
10	−4	57	50	52	77																										
12	4	64	52	54	72	94																									
14	22	66	48	57	38	52	66																								
16	26	6	11	0	−39	−30	−20	9																							
18	−7	−11	−26	0	−7	−10	−5	7	−22																						
20	−14	−1	12	22	−7	−11	−15	−20	−15	−37																					
22	32	−4	−6	12	−34	−29	−12	20	40	9	15																				
24	28	18	43	42	0	−6	−3	20	18	2	25	14																			
26	2	−3	−1	17	26	4	8	31	7	36	−40	10	9																		
28	7	35	12	25	33	23	24	45	9	6	−38	−20	−19	45																	
30	12	30	19	31	15	8	26	44	7	−6	−15	13	−6	34	47																
32	28	46	32	42	10	24	27	52	5	17	−24	19	22	46	39	52															
34	24	46	43	44	20	27	32	43	5	2	−1	23	5	49	28	42	74														
36	6	32	24	40	23	32	34	39	18	−14	14	29	−10	33	13	38	50	[illegible]													
38	−2	23	29	34	31	20	23	25	7	4	−7	10	5	36	29	28	44	55	53												
40	24	16	18	37	29	19	20	44	11	−6	−10	24	34	37	38	36	46	39	38	52											
42	29	14	17	30	16	2	14	39	26	32	−29	33	49	37	10	19	39	24	25	42	49										
44	1	27	24	31	20	31	40	33	2	17	−23	2	23	25	6	26	42	27	27	33	9	49									
46	0	16	23	27	9	15	23	23	1	−5	−8	7	9	20	13	35	30	27	27	44	18	29	76								
48	5	5	9	21	4	−4	1	18	22	0	−23	39	−18	53	20	39	32	43	54	44	36	36	36	61							
50	−1	−4	−8	17	13	3	5	12	12	24	−29	25	−21	60	29	21	19	23	32	21	15	24	33	46	79						
52	11	22	−1	7	10	7	6	6	−1	25	−29	0	−1	36	0	6	4	14	9	16	12	18	31	48	48	64					
54	21	5	5	7	12	16	21	31	−11	33	−38	1	8	24	1	25	21	25	25	16	26	27	42	43	43	42	76				
56	17	1	2	11	17	22	24	35	−19	33	−51	0	−2	36	25	27	27	18	15	18	41	30	47	42	53	55	59	77			
58	25	20	23	44	9	12	11	37	2	18	−34	20	14	28	50	38	40	28	28	28	46	35	40	53	65	62	47	55	76		
60	34	26	21	43	4	23	21	48	9	18	−22	21	20	24	44	34	36	24	37	22	38	36	43	46	43	50	55	59	72	89	

**Multiply tabular values by 0.01 to obtain correlation coefficients

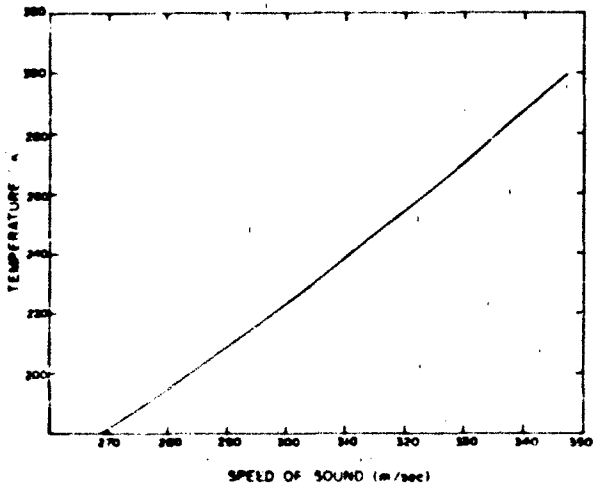


Figure 15-15. Speed of sound vs temperature (K).

where TAR represents the ratio of the thermal admittance of the ground to that of air. For diurnal cycles of net radiation, the ratio $(K k)_{gr}/(K k)_{air}$ is the order of 10^3.

The most extreme surface temperature oscillation occurs over feathery snow where the amplitude may reach approximately four times that over still water or sandy soil, and is at least 100 times as large as that over the turbulent ocean. An amplitude ratio of about 3.5 can be expected for surface temperature over dry vs moist sand surfaces. Theoretically, the penetration of thermal "oscillations" into the soil is inversely proportional to the frequency of the "oscillations" [Lettau, 1954b]. The best insulator is still air or any porous material with air-filled pores, such as feathery snow; materials such as leaf litter have similar insulating properties [Geiger, 1957].

Much information is available on soil-temperature variations in various climatic zones. Table 15-17 gives annual and daily temperature cycles in different soil types. In addition to the type of ground, certain meteorological factors such as rainfall and melting snow have marked effects on the soil temperature. Snow cover is a leading factor in protecting the soil from severe frost. On one extreme occasion with an air temperature of 255 K, the temperature was 272 K under a 13 cm snow cover, whereas on bare soil it was 251 K.

The soil-temperature variations illustrated in Figure 15-16 were obtained at a station cleared of pine trees but in generally wooded country. Topsoil and brown sandy loam (0 to 0.6 m) changed to brown sand and gravel that varied from medium (0.6 to 2 m), to coarse (2 to 4 m), and again to medium (4 to 18 m). The water level was at 15 m. The figure illustrates the amplitude decrease and phase retardation of the annual cycle with depth. Amplitudes of weather disturbances with periods of several days, as illustrated by the temperature curve of the 0.75-m level, decrease with depth more rapidly than the annual amplitudes. Qualitatively, this agrees with the theoretical prediction of an amplitude decrement proportional to the square root of the length of the period of oscillation. The actual half-amplitude depth interval of the annual cycle can be estimated from Figure 15-16 as being nearly 3 m, which is much larger than the depth inferred from experimental values of thermal

Table 15-15. Bolometric records of area-approximately 37 km² surface temperature from an airplane cruising at approximately 370 m along a constant flight path. Apex (544 feet) observed from Albrecht (1952).

				Surface Temperature (K) — Bolometric Data						
Day	Hour	Sun's Elevation (degrees)	Sky Cover	Standard Shelter Temp at Airport K	Baltic Sea	Sand Beach	Down Land	City	Woods	Opening in Woods
9	13 to 13	60.4	10/10	282	275	285	287	281	280	281
11	10 to 23	16	1/10	283	275	280	275	276	276	278
16	19 to 20	1.8	9/10	287	277	281	280	280	277	275
20	05 to 06	0.5	1/10	275	278	271	267	272	270	266
29	14 to 15	42.1	4/10	280	280	319	315	290	289	296

				Wind Speed m/s	Woods	Clear Cutting in Woods	Dry Peat	Swamp	Pond
7	19	—	4/10	0.5	275	271	267	273	274
20	20	—	2/10	1.5	273	272	269	273	273
26	20	—	3/10	2.6	275	270	272	272	274
Albedo values as determined by Albrecht:					5%	8%	8%	7%	5%

Table 15-14. Temperature of the air 10 cm above, and of the soil 0.5 cm below, the earth/air interface measured by thermocouples [Davidson and Lettau, 1957].

Condition	Temperature (K) at Mean Local Time								
	0400	0600	0800	1000	1200	1400	1600	1800	2000
*Air	287.7	289.8	296.0	300.7	303.7	304.7	305.0	301.9	297.5
*Soil	290.6	290.9	296.0	304.1	308.6	309.2	306.8	302.3	298.5
**Air	281.1	282.5	291.5	297.9	301.7	303.0	301.1	296.1	292.3
**Soil	284.8	284.6	291.4	303.1	310.8	310.3	304.3	297.8	293.8

*Mean soil moisture in 0 to 10 cm layer 10% wet weight basis.
**Mean soil moisture in 0 to 10 cm layer about 4% wet weight basis.

diffusivity. The discrepancy may be caused by seepage or downward migration of rain water and the accompanying advection or transfer of heat. This process could increase the apparent or effective thermal diffusivity for annual soil-temperature variations by factors of 4 to 8 times the experimental values obtained in soil of constant moisture. The data in Table 15-17 are more in line with experimental findings than the curves in Figure 15-16. The limitation of Table 15-17 is that the data are for clearly defined and nearly ideal soil types that are seldom matched by actual ground conditions.

Factors that must be investigated and assessed for any one set of soil-temperature observations are (1) type and state of compaction of the soil, (2) moisture content and seepage of the soil during the test, (3) position of the water table during the test, (4) type and color of surface cover, (5) amount and nature of traffic over the site, and (6) local climatic conditions.

Subsoil temperature information is useful in computing thermal stresses and loads. Some examples are the determination of the depth to which a structure should be buried when proximity to natural isothermal conditions is desired to conserve on the air conditioning load, or to dissipate heat generated by power cables. The determination of frost penetration depths is usually the principal concern.

Table 15-15. Comparison of air and soil temperature with surface temperatures of materials exposed on a tropical island with normal trade winds. Air and material surface temperatures at 1.2 m above, soil temperature at 2.5 cm below, the earth/air interface. Exposed surface area about 930 cm² [Draeger and Lee, 1953].

Material	Temperature (K)		
	Highest Recorded	Average Max.	Min
Air (1.2 m)	302	301	299
Soil (2.5 cm)	307	307	299
Wood	314	310	298
Aluminum	313	309	298
Galvanized Iron	318	311	298
Black Iron	324	315	298
Concrete Slab	310	307	298

15.1.7 Degree-Day and Temperature-Wind Combinations

A *degree-day* is a unit adopted to measure the departure of the daily mean temperature from a given standard. In the United States the number of *heating degree-days*, on any one day, is the number of Fahrenheit degrees of the 24-h mean temperature below 65°F (291 K). Cumulated, day by day, over the heating season, the total number of degree days becomes an index of heating fuel requirements. In such cumulation, the days on which the mean temperature exceeds 65°F (291 K) are ignored. When the centigrade scale is used, the base is usually 19°C (292 K). The United States Army Corps of Engineers computes "freezing-degree days" as the departure of the daily mean temperature from 32°F (273 K), a negative departure when above 32°F (273 K). The National Weather Service supplies "normal degree-days," both monthly and annual totals. A few examples of the 30-year annual normals are 9274(F°) for Fargo, N.D., 5634(F°) for Boston, Mass. and 108(F°) for Key West, Florida.

The *wind-chill* concept was introduced in 1939 by the famous antarctic explorer, Paul Siple, to measure the cooling effect of low temperature and strong wind combined. The wind-chill index is the equivalent temperature, in a normal walk (1.9 m/s) in calm air, corresponding to the combination of actual air temperature and windspeed. It can be related to the heat loss H from a nude body in the shade. H is given by

$$H = (10\sqrt{V} + 10.45 - V)(306 - T_a), \quad (15.22)$$

where H is the heat loss in kilogram calories per square meter of body surface per hour, T_a is the air temperature (K), and V is the windspeed (m/s). Neutral skin temperature is roughly 306 K. For windspeeds greater than 1.9 m/s the wind-chill index (T_{wc}) in K is given closely by

Table 15-16. Physical thermal parameters of diverse ground types [Lettau, 1954b].

Ground Type	Thermal Admittance Ratio (TAR), Ground to Air	Half-Amplitude Depth Interval (theoretical)	
		Annual Cycle (n.)	Diurnal Cycle (m)
SOILS			
Quartz sand,			
medium-fine dry	110	1.0	0.05
8% moisture	230	1.6	0.08
22% moisture	360	1.5	0.08
Sandy clay, 15% moisture	280	1.3	0.07
Swamp land, 90% moisture	340	1.0	0.05
ROCKS			
Basalt	350	1.8	0.09
Sandstone	380	2.2	0.12
Granite	440	2.5	0.13
Concrete	440	2.3	0.12
SNOW, ICE, AND WATER			
Feathery snow	10	0.67	0.04
Packed snow	100	1.4	0.07
Still water	280	0.82	0.04
Ice	320	1.4	0.07
Turbulent ocean	10^3 to 10^5	61 to 610	3 to 30

$$T_{wc} = (306 - H/22). \qquad (15.23)$$

This formula gives only an approximation because of individual body variations, incoming radiation, and other factors affecting heat loss from the body. The formula is not not used, or needed, with wind speeds less than 6 km/h (2 m/s).

Extreme temperature-wind combinations are frequently important in thermal equilibrium design problems, requiring estimates of the maximum steady wind speeds likely to be

Table 15-17. Annual and daily temperature cycles. Annual values are averages for the years 1939 through 1946 at Giessen, Germany [Kreutz, 1943]. Daily values are averages of clear weather 10 through 12 August 1893, Finland, after Homen [Geiger, 1957].

	Temperature (K)					
	Annual Means			Daily Means		
	Loam	Sand	Humus	Swamp Land	Sandy Heath	Granite Rock
Surface	282.1	282.3	283.1	289.6	298.0	297.6
1.0 m above	283.8	284.3	284.3	—	—	—
0.6 m above	—	—	—	284.5	287.0	293.4
Surface Amplitude	283.5	283.7	284.4	283.4	290.0	283.2
	Half-Amplitude Depth Interval (m)					
Depth	1.8	1.6	1.4	0.05	0.08	0.15

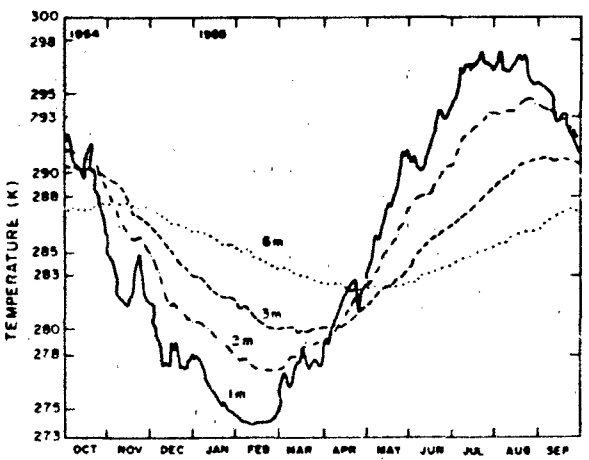


Figure 15-16. Variations of soil temperature at indicated depths, North Station, Brookhaven, Long Island, October 1954 through September 1955, [after Singer and Brown, 1956].

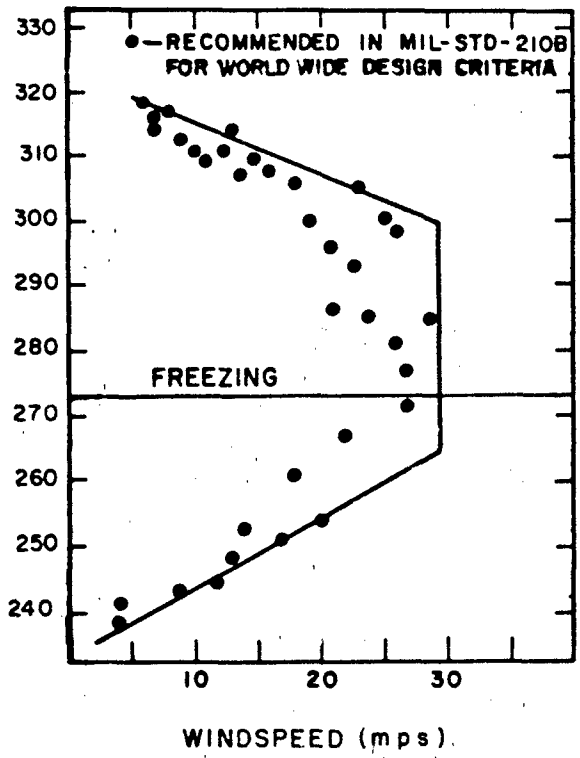


Figure 15-17. Extremes of temperature in combination with windspeed. Windspeeds, in general, were observed 12 to 30 m above the surface. The 35 observations were taken over a 5-year period at some 22 stations widely scattered in the United States. The envelope is for the recommended U.S. design criteria.

encountered at various temperatures. Figure 15-17 was prepared from 4 years of 6-hourly and 1 year of hourly data for 22 stations in the United States [Sissenwine and Court, 1951]. It shows maximum steady (5-min) wind speeds that occurred with temperatures in the range from 236 K to 319 K during this period. The stations used in this study were selected as representative of climatic areas in the United States. Mountainous stations were unrepresentative of generally operational areas and were not among those selected. Also, the high winds of hurricanes and tornadoes were omitted from the figure.

The wind speeds of Figure 15-17 occurred at anemometer heights, usually at 12 to 30 m above ground level during the years of observation. The wind speeds at the 3-m level are approximately 20% less and even 50% less for the extreme low temperature (less than 252 K).

The combinations of values of temperature and windspeed, recommended for extreme U.S. thermal equilibrium design criteria, are shown by the envelope in Figure 15-17. This recommendation is not valid in mountainous areas or in Death Valley. For the latter the criteria are the same as for world-wide criteria, as plotted in Figure 15-17.

15.2 ATMOSPHERIC DENSITY UP TO 90 KM

The density data discussed in this section are from direct and indirect observations obtained from balloon-borne instrumentation for altitudes up to 30 km, and measurements from rockets and instruments released from rockets for altitudes between 30 and 90 km.

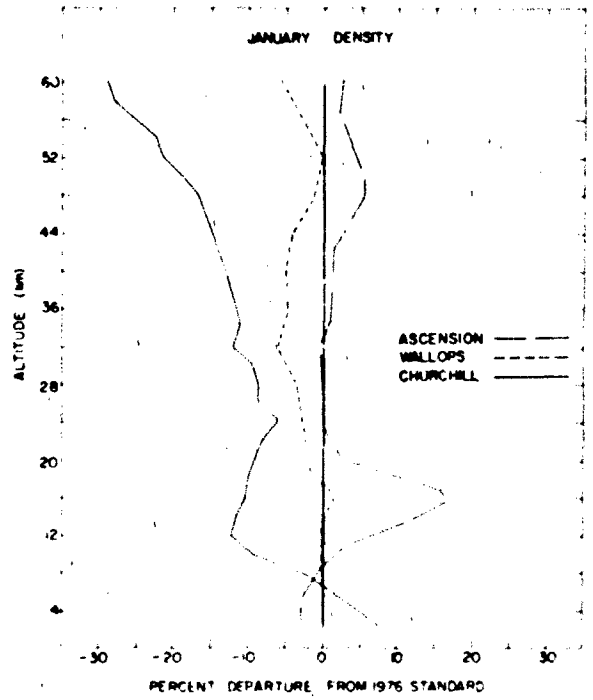
15.2.1 Seasonal and Latitudinal Variations

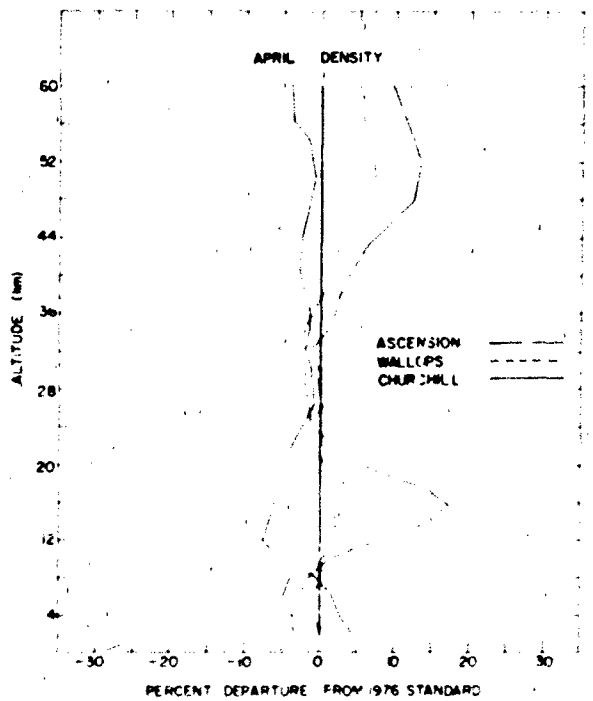
The Reference Atmospheres presented in Chapter 14 provide tables of mean monthly density-height profiles, surface to 90 km, for 15° intervals of latitude between the equator and the North Pole. Densities at altitudes between 10 and 90 km are highest during the months of June and July and lowest in December and January at locations north of 30° latitude. In tropical and subtropical areas seasonal variations are relatively small with highest densities at levels above 30 km occurring in the spring and fall.

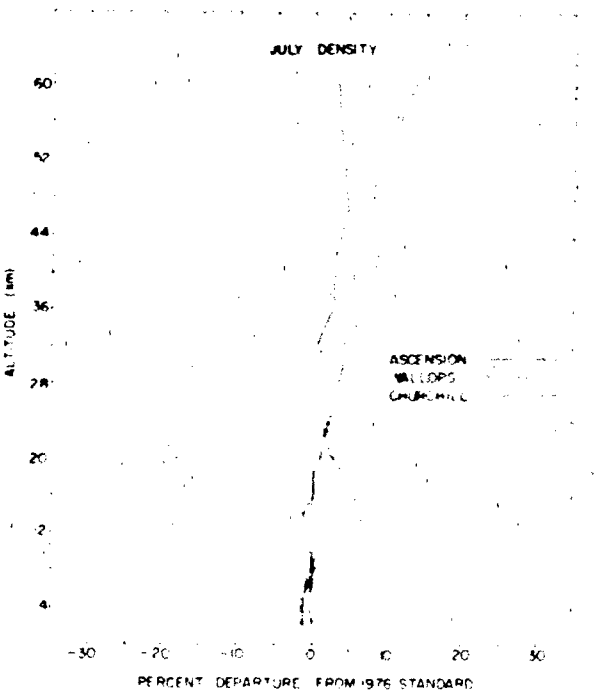
Mean monthly density profiles, surface to 60 km, observed during the midseason months at Ascension Island, 8°S, 14°W, Wallops Island, 38°N, 75°W, and Ft. Churchill, 59°N, 94°W, are plotted in Figure 15-18. Densities are shown as percent departure from the *U.S. Standard Atmosphere, 1976*. The individual mean monthly profiles cross or converge near 8 km and between 22 and 26 km. Both are levels of minimum density variability. The level near 8 km is considered an isopycnic level because mean monthly densities depart from standard by no more than 1% or 2% regardless of the geographical location or season. Between 22 to 26 km, however, there is a marked seasonal variability, even though there is very little longitudinal or latitudinal variability during individual months. Seasonal differences in the density profiles at the same three locations are shown in Figure 15-19. The minimum seasonal variability of the mean monthly values, 1% to 2%, occurs at 8 km, and the maximum seasonal variability occurs above 60 km. The seasonal variations are largest at Ft. Churchill and are smallest at Ascension Island.

15.2.2. Day-to-Day Variations

The density at a specific altitude may differ from the seasonal or monthly mean at that altitude due to day-to-day changes in the weather pattern. The distribution of observed densities in January and July at the most climatically extreme locations for which data are available near 30°, 45°, 60° and 75°N are shown in Table 15-18a to 15-18d for altitudes up to 80 km. Median, and high and low values that are equaled or more severe 1%, 10%, and 20% of the time are given as percent departures from the U.S. Standard Atmosphere at 5-km altitude increments. The 1% values for altitudes







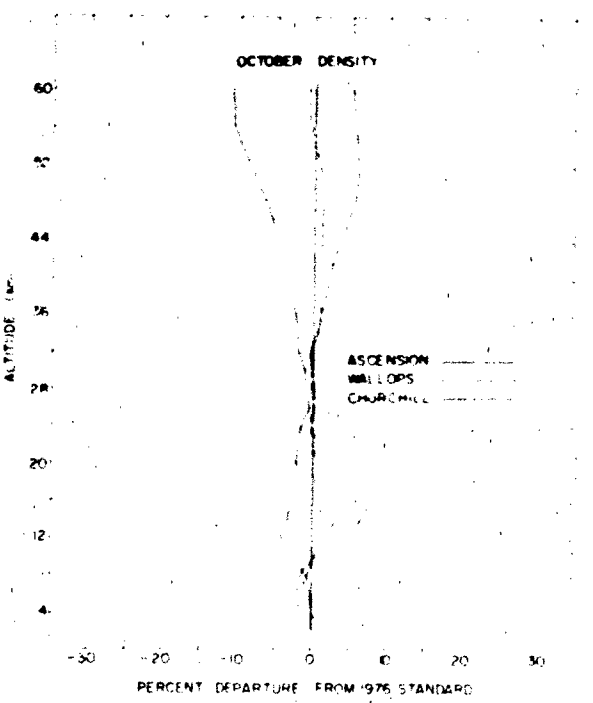
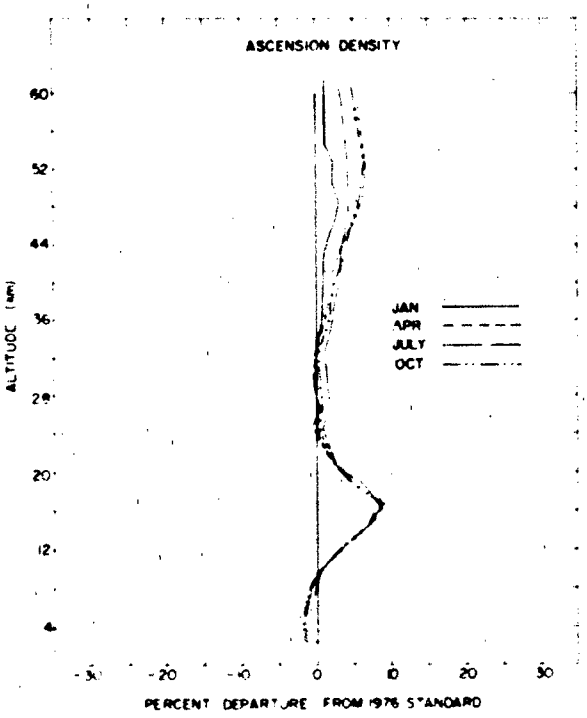
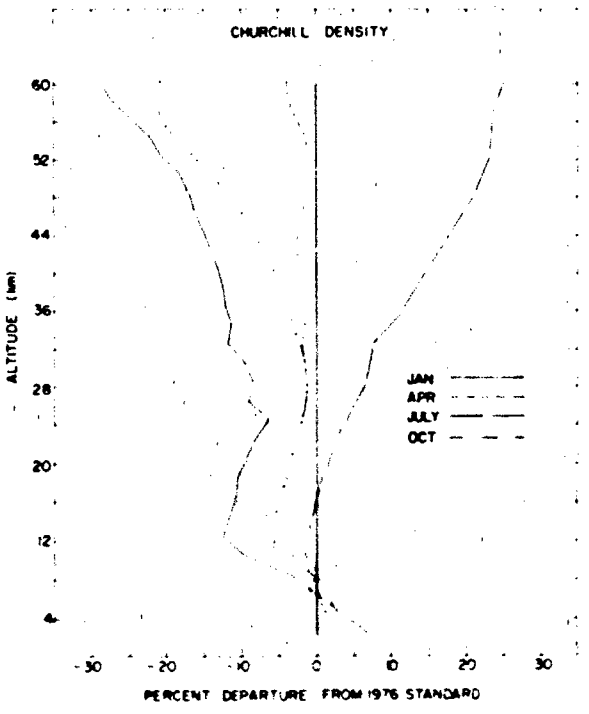


Figure 15-18 Latitudinal differences in the density altitude profiles for the mid-season months at Ascension Island, Wallops Island, and Ft. Churchill





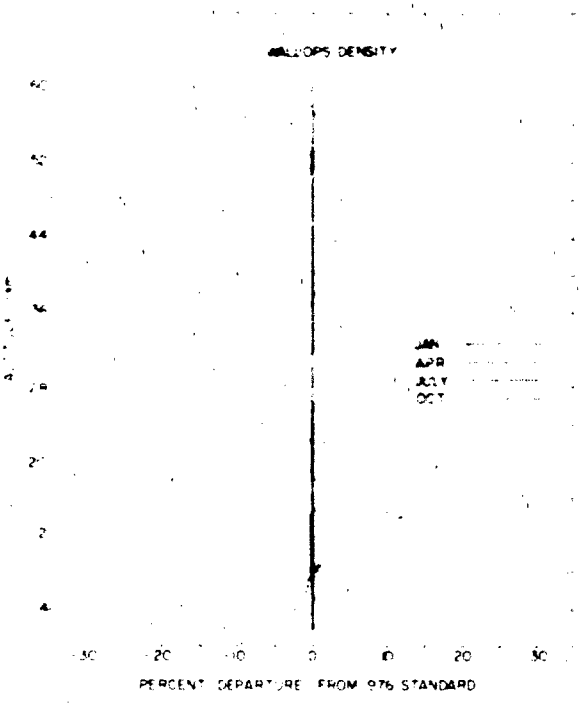


Figure 15-19. Seasonal differences in the density-altitude profiles at Ascension Island, Wallops Island, and Ft. Churchill

above 30 km are considered rough estimates as they are based on the tails of the distributions of observed values plotted on probability paper. Estimates above 60 km are less reliable than those at lower levels because of the paucity of data and larger observational errors at the higher altitudes. In tropical regions the monthly density distributions are nearly normal for altitudes up to 50 km. Consequently, reasonable estimates of the distributions of density in the tropics can be obtained from monthly means and standard deviations. Standard deviations of the observed densities around the mean monthly values at Ascension, given in Table 15-19, are typical of the day-to-day variations found in the tropics [Cole and Kantor, 1980].

15.2.3 Spatial Variation

The rate of decay of the correlation coefficient between densities at two points with increasing horizontal separation is directly related to the scale of the major features of the weather patterns that are experienced at a specific latitude and altitude. Figure 15-20 provides information on the decay of density correlations with distance near 60 N for altitudes up to 60 km. The decay in density correlations below 20 km are based on an interpretation of data from studies of the spatial correlations of pressure, temperature, density and wind at radiosonde levels at locations between 30° and 70° N

Table 15-18a Median, high, and low values of densities given as percentage departure from U.S. Standard Atmosphere 1976 for January and July at 30°N

Altitude (km)	Median (% of Std)	1%		10%		20%		U.S. Std Density (kg/m^3)
		High	Low	High	Low	High	Low	
January								
5	−1	+1	−3	0	−2	0	−2	7.3643 − 1
10	+1	+4	−3	+3	−1	+2	0	4.1351
15	+7	+15	−1	+12	+4	+10	+5	1.9476
20	+3	+7	−2	+5	+1	+4	+2	8.8910 − 2
25	−2	+4	−6	+3	−4	+1	−2	4.0084
30	−4	+2	−10	−2	−8	−3	−6	1.8410
35	−3	+3	−12	0	−8	−1	−6	8.4634 − 3
40	1	+2	−10	+1	−7	0	−5	3.9957
45	0	+8	−10	+3	−7	+2	−5	1.9663
50	+1	+12	−8	+7	−4	+5	−2	1.0269
55	0	+9	−10	+5	−6	+3	−4	5.6810 − 4
60	−2	+12	−15	+5	−9	+2	−6	3.0968
65	−4	+21	−25	+13	−13	+7	−6	1.6321
70	−5	+16	−26	+9	−17	+6	−12	8.2828 − 5
75	−7	+21	−25	+13	−15	+8	−10	3.9921
80	−4	+21	−22	+15	−13	+8	−7	1.8458
July								
5	−3	0	−5	−1	−4	−2	−4	7.3643 − 1
10	+1	+3	−1	+2	0	+2	0	4.1351
15	+16	+20	+11	+17	+13	+17	+14	1.9476
20	+8	+11	+14	+10	+5	+9	+6	8.8910 − 2
25	+4	+9	0	+7	+2	+6	+3	4.0084
30	+3	+7	−1	+5	+1	+4	+2	1.8410
35	+6	+10	+2	+8	+3	+7	+4	8.4634 − 3
40	+9	+15	+2	+11	+5	+10	+7	3.9957
45	+12	+19	+4	+14	+7	+13	+9	1.9663
50	+13	+23	+6	+17	+8	+15	+10	1.0269
55	+11	+20	+2	+15	+5	+13	+7	5.6810 − 4
60	+13	+14	+1	+21	+3	+19	+7	3.0968
65	+15	+43	−6	+38	0	+30	+6	1.6321
70	+15	+32	−9	+23	+1	+20	+8	8.2828 − 5
75	+10	+24	−11	+20	−6	+15	+1	3.9921
80	+6	+22	−15	+17	−6	+14	+1	1.8458

latitude [Bertoni and Lund, 1964]. Information on the spatial correlations at altitudes above 20 km is from a study by Cole [1979]. In that paper, data from constant pressure maps for 5.0, 2.0 and 0.4 mb levels were used together with nearly simultaneous rocket observations at several pairs of stations near 60°N to determine the rates of decay of density correlation at levels between 30 and 55 km. As Figure 15-20 indicates, the rate of decay in density correlation with distance decreases substantially with altitude. At 10 km, for example, zero correlation is attained at about 2000 km at 50 km, zero correlation is attained at more than twice that distance, or 4450 km. This analysis indicates the presence of disturbances with wavelengths of roughly 18,500 km at 50 km, close to planetary wavelength number one at 60°N. Information from Kantor and Cole [1979] on the correlations between densities at points up to 370 km apart in tropical regions is provided in Table 15-20, for levels between 10 and 60 km.

The rms difference between the densities at two points can be estimated by

Table 15-18b. Median, high, and low values of densities given as percentage departure from U.S. Standard Atmosphere 1976 for January and July at 45°N.

Altitude (km)	Median (% of Std)	1%		10%		20%		U.S. Std Density (kg/m^3)
		High	Low	High	Low	High	Low	
January								
5	0	+4	−3	+3	−2	+2	−1	7.3643 − 1
10	−2	+6	−10	+3	−6	+1	−4	4.1351
15	−3	+4	−12	+1	−8	−1	−6	1.9476
20	−2	+2	−8	0	−6	−1	−5	8.8910 − 2
25	−2	+2	−8	0	−6	−1	−5	4.0084
30	−5	+1	−17	−2	−13	−4	−9	1.8410
35	−6	+2	−20	−2	−16	−4	−12	8.4634 − 3
40	−8	+5	−23	0	−17	−4	−13	3.9957
45	−9	+8	−22	+2	−16	−3	−14	1.9663
50	+8	+11	−20	+4	−16	−3	−14	1.0269
55	−9	+9	−25	+2	−18	−4	−16	5.6810 − 4
60	−12	+7	−28	0	−23	−7	−20	3.0968
65	−14	0	−38	−5	−34	−10	−28	1.6321
70	−15	+2	−38	−9	−30	−12	−26	8.2828 − 5
75	−16	−3	−38	−9	−30	−12	−26	3.9921
80	−23	−2	−42	−8	−36	−10	−30	1.8458
July								
5	−2	+1	−5	−1	−4	−1	−3	7.3643 − 1
10	0	+3	−4	+2	−2	+1	−1	4.1351
15	+8	+17	+2	+15	+4	+13	+5	1.9476
20	+6	+11	0	+8	+2	+7	+3	8.8910 − 2
25	+7	+10	+4	+9	+5	+8	+6	4.0084
30	+7	+12	0	+9	+2	+8	+4	1.8410
35	+9	+16	0	+12	+3	+10	+6	8.4634 − 3
40	+13	+21	+4	+16	+3	+14	+10	3.9957
45	+15	+26	+6	+20	+10	+18	+12	1.9663
50	+17	+31	+9	+25	+12	+21	+14	1.0269
55	+17	+32	+8	+25	+11	+22	+14	5.6810 − 4
60	+19	+30	+4	+26	+10	+24	+13	3.0968
65	+20	+40	+4	+35	+10	+30	+13	1.6321
70	+20	+37	0	+32	+9	+27	+12	8.2828 − 5
75	+19	+40	2	+30	+7	+26	+11	3.9921
80	+14	+32	−4	+30	+4	+25	+9	1.8458

$$\sigma_{xy} = \sqrt{\sigma_x^2 + \sigma_y^2 - 2r_{xy}\,\sigma_x\,\sigma_y} \qquad (15\text{-}24)$$

where σ_{xy} is the estimated rms difference between densities at points x and y, σ_x^2 and σ_y^2 are the variances of density around the monthly mean values, and and r_{xy} is the correlation coefficient between the densities at points x and y. For short distances (up to 550 km) σ_x^2 and σ_y^2 can usually be assumed to be equal.

The estimated rms difference between densities that are observed simultaneously at locations 90, 180 and 360 km apart in the tropics are presented in Table 15-21 for altitudes between 10 and 60 km. For a given month, the rms differences provided in Table 15-21 can be considered to represent variability around the mean monthly density gradients, which are given in Table 15-22 [Cole and Kantor, 1975] for the indicated latitudinal differences. Longitudinal differences remain near zero in tropical areas. Information on the spatial variability of density is useful in determining how accurately a density observation taken 75 to 500 km from the point of

Table 15-18c. Median, high, and low values of densities given as percentage departure from U.S. Standard Atmosphere 1976 for January and July at 60°N

Altitude (km)	Median (% of Std)	1%		10%		20%		U.S. Std Density (kg/m^3)
		High	Low	High	Low	High	Low	
January								
5	+1	+6	−3	+4	−1	+2	0	7.3643 − 1
10	−6	+3	−15	+2	−15	−3	−10	4.1351
15	−9	−2	−15	−5	−12	−6	−11	1.9476
20	−8	−1	−15	−5	−11	−6	−10	8.8910 − 2
25	−7	−3	−16	−2	−12	−4	−10	4.0084
30	−10	+7	−32	+2	−18	−2	−15	1.8410
35	−12	+8	−35	−3	−27	−3	−19	8.4634 − 3
40	−15	+10	−36	+5	−30	−4	−20	3.9957
45	−21	+12	−39	+5	−34	−10	−24	1.9663
50	−26	+14	−43	+3	−36	−15	−29	1.0269
55	−32	+9	−48	−10	−39	−20	−35	5.6810 − 4
60	−36	+4	−54	−12	−40	−25	−39	3.0968
65	−36	−5	−50	−16	−46	−27	−42	1.6321
70	−37	−12	−54	−25	−49	−32	−43	8.2828 − 5
75	−35	−10	−53	−24	−47	−30	−42	3.9921
80	−28	−11	−53	−17	−47	−21	−40	1.8458
July								
5	2	+2	−5	+1	−4	0	−3	7.3643 − 1
10	0	+7	−8	+4	−5	+2	−3	4.1351
15	0	+6	−7	+3	−4	+2	−2	1.9476
20	+3	+7	−2	+6	0	+5	+1	8.8910 − 2
25	+5	+8	+1	+7	+2	+6	+3	4.0084
30	+7	+12	−1	+9	+2	+8	+4	1.8410
35	+10	+18	0	+14	+3	+12	+7	8.4634 − 3
40	+15	+23	+5	+19	+10	+17	+12	3.9957
45	+20	+28	+7	+25	+13	+23	+16	1.9663
50	+25	+35	+10	+30	+16	+28	+22	1.0269
55	+27	+35	+11	+30	+16	+29	+22	5.6810 − 4
60	+28	+42	+11	+39	+16	+33	+22	3.0968
65	+35	+50	+11	+44	+18	+39	+28	1.6321
70	+42	+52	+12	+46	+20	+44	+30	8.2828 − 5
75	+44	+58	+12	+52	+20	+48	+35	3.9921
80	+40	+56	+10	+50	+18	+44	+30	1.8458

vehicle reentry represents the conditions encountered in the reentry corridor.

15.2.4 Statistical Applications to Reentry Problems

The relatively large number of available radiosondes and meteorological rocket observations permit a detailed analysis of the characteristics of atmosphere density profiles at altitudes below 60 km. Arrays of means and standard deviations of density at 2-km intervals of altitude from the surface to 60 km, together with interlevel correlation coefficients between levels have been developed for tropical, temperate and arctic regions [Cole and Kantor, 1980]. Tables 15-23a to 15-23f contain statistical arrays of density for the months of January and July at Kwajalein (9°N), Wallops Island (38°N), and Ft. Churchill (59°N).

Variations in the range or deceleration of free falling objects or ballistic missiles that arise from day-to-day changes in atmospheric density can be estimated from Tables 15-23a to 15-23f. The integrated effect, E, of mean monthly density on the trajectory or impact point of a missile can

Table 15-18d. Median, high, and low values of densities given as percentage departure from U.S. Standard Atmosphere 1976 for January and July at 75°N.

Altitude (km)	Median (% of Std)	1%		10%		20%		U.S. Std Density (kg/m^3)
		High	Low	High	Low	High	Low	
January								
5	+2	+6	−1	+5	0	+4	+1	7.3643 − 1
10	−8	+2	−18	−3	−13	−5	−10	4.1351
15	−10	−1	−18	−6	−14	−8	−13	1.9476
20	−12	−1	−22	−6	−17	−8	−15	8.8910 − 2
25	−15	−2	−28	−8	−20	−10	−18	4.0084
30	−21	−4	−36	−9	−26	−16	−24	1.8410
35	−25	0	−43	−10	−32	−16	−30	8.4634 − 3
40	−[illegible]	+4	−48	−9	−38	−16	−38	3.9957
45	−[illegible]	+8	−52	−6	−45	−16	−39	1.9663
50	−[illegible]	+4	−56	−8	−48	−20	−42	1.0269
55	−44	+5	−65	−10	−56	−23	−50	5.6810 − 4
60	−46	0	−70	−16	−60	−32	−55	3.0968
65	−47	+1	−66	−27	−62	−35	−58	1.6321
70	−48	−1	−69	−21	−62	−35	−60	8.2828 − 5
75	−45	−10	−65	+25	−57	−35	−53	3.9921
80	−40	−8	−55	−24	−50	−34	−45	1.8458
July								
5	1	+4	−2	+3	−1	+2	0	7.3643 − 1
10	−4	+5	−12	+3	−10	0	−7	4.1351
15	−4	+2	−9	0	−7	−2	−6	1.9476
20	+1	+6	−4	+4	−2	+3	−1	8.8910 − 2
25	+1	+10	−8	+6	−3	+5	−2	4.0084
30	+7	+13	+2	+10	+5	+8	+6	1.8410
35	+12	+25	+3	+18	+8	+16	+10	8.4634 − 3
40	+19	+27	+6	+23	+13	+21	+16	3.9957
45	+25	+35	+10	+30	+18	+28	+21	1.9663
50	+27	+40	+10	+35	+20	+32	+24	1.0269
55	+32	+42	+10	+39	+20	+35	+25	5.6810 − 4
60	+37							3.0968
65	+48	(Insufficient data above 55 km						1.6321
70	+60	in summer)						8.2828 − 5
75	+67							3.9921
80	+64							1.8458

be determined for a specific location by computer "flights" through mean monthly or seasonal density profiles if the proper influence coefficients, C_i, for the missile at various levels are given. For example, we can write

$$E = \Sigma C_i \bar{\rho}_i \qquad (15.25)$$

where $\bar{\rho}_i$ is the mean monthly density at the ith level. The influence coefficients depend upon aerodynamic characteristics, reentry angle, and the speed of the vehicle. The integrated standard deviation in range or deceleration σ_{im}, due to day-to-day variations from the mean seasonal or the mean monthly density profile can be obtained from

$$\sigma^2_{im} = \Sigma_{ij} C_i \sigma_i r_{ij} C_j \sigma_j \qquad (15.26)$$

where σ^2_{im} is the integrated variance for all layers being considered, C_i and C_j are influence coefficients at the ith and jth levels, σ_i and σ_j are the standard deviations of density at the two levels, and r_{ij} is the correlation coefficient between densities at the two levels. In these computations density is assumed to have a Gaussian distribution at all levels. As a

Table 15-19. Standard deviations (%) of observed day-to-day variations in density around the monthly mean at Ascension Island (8°S).

Altitude (km)	S.D. of Density (% of Monthly Mean)			
	Jan	Apr	July	Oct
5	0.4	0.3	0.3	0.4
10	0.4	0.4	0.4	0.4
15	0.8	0.7	0.8	0.7
20	1.5	1.3	1.8	1.3
25	1.3	1.3	1.2	1.3
30	1.2	1.2	1.4	1.2
35	1.8	1.8	1.4	1.2
40	2.3	2.1	1.8	1.8
45	2.3	2.3	2.6	2.3
50	2.7	2.5	2.6	2.7

result, the error in the CEP (the circle within which 50% of the events are expected to occur) will be generally less than 10%.

15.2.5 Variability with Time

Studies based on radiosonde observations have shown that there are no significant diurnal variations in density at altitudes up to 30 km. The analysis of meteorological rocket observations, however, indicates the presence of a significant diurnal oscillation in density at altitudes between 35 and 60 km. The phases and amplitudes of the diurnal oscillation at these altitudes are best defined in the tropics.

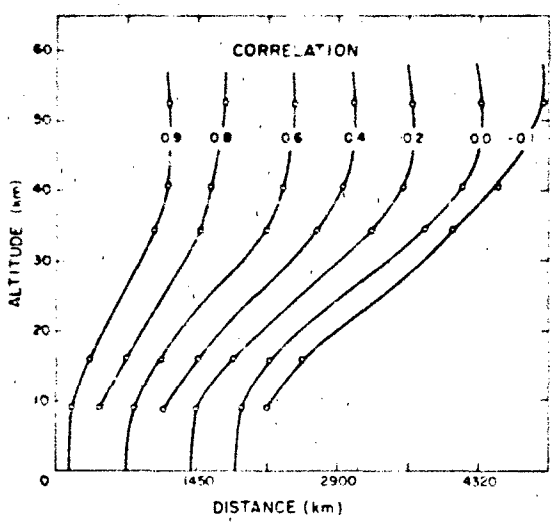


Figure 15-20. Decay of density correlations with distance at various altitudes in midlatitudes.

Table 15-20. Correlation coefficients between densities at points up to 370 km apart in the tropics.

Altitude (km)	Correlation Coefficient		
	90 km	180 km	370 km
10	0.97	0.95	0.90
20	0.98	0.97	0.92
30	0.98	0.97	0.92
40	0.98	0.97	0.92
50	0.98	0.97	0.92
60	0.98	0.97	0.92

The decrease in the number of available observations above 60 km and the larger random observational errors at the higher altitudes make it difficult to obtain reliable estimates of the magnitude of the diurnal variations at altitudes between 60 and 90 km.

The 50-km densities from a series of soundings taken at Ascension during a 48-h period in April 1966 [Cole and Kantor, 1975] are plotted versus local time in Figure 15-21. Densities are given as percent departure from those for the 1976 U.S. Standard Atmosphere. The crosses represent averages of observations taken within two hours of each other. Harmonic analysis of the eight average values produced the solid curve when the first and second harmonics for the 48-h period were added together. An F-test indicates that the second harmonic, which represents the diurnal oscillation in density, has an amplitude of slightly less than 4% (a range of almost 8%) and is significant at the 1% level; it reduces the observed variance by 91%. Maximums occur at 1600 and minimums near 0400 local time. From this analysis it is apparent that the diurnal oscillation is the dominant short-period fluctuation at 50 km.

The rms differences between density observations taken from 1 to 36 hours apart also provide a measure of the rate of change in density with time at a given altitude. Computed rms values from the Ascension series mentioned above are shown as a function of time in Figure 15-22 for altitudes from 35 to 60 km. The number of pairs of observations available for each time interval is also shown. Since at time T = 0 the rms change in space is zero, an estimate of the random observational error can be obtained from the observations themselves by extrapolating curves in Figure 15-22 back to zero hours. This procedure indicates that the random rms errors are approximately 1% at 35 and 40 km, and 1.5% to 2.0% at altitudes between 45 and 60 km.

If there are no well-defined periodic oscillations within a 24-h period, the rms variability would be expected to increase smoothly with time until it reached a value representing the climatic or the day-to-day variability around the monthly mean. However, a well-defined 24-h oscillation can be seen (Figure 15-22) in the rms density variations at

Table 15-21 Estimated rms differences (% of mean) between densities at locations 90, 180, and 360 km apart during the midseason months in the tropics.

Altitude (km)	January 90	January 180 km	January 360	April 90	April 180 km	April 360	July 90	July 180 km	July 360	October 90	October 180 km	October 360
10	0.10	0.13	0.18	0.10	0.13	0.18	0.10	0.13	0.18	0.10	0.13	0.18
15	0.13	0.17	0.25	0.11	0.14	0.21	0.16	0.20	0.30	0.16	0.20	0.30
18	0.50	0.61	1.00	0.34	0.42	0.68	0.30	0.37	0.60	0.34	0.42	0.68
20	0.28	0.34	0.56	0.28	0.34	0.56	0.24	0.29	0.48	0.24	0.29	0.48
25	0.28	0.34	0.56	0.28	0.34	0.56	0.24	0.29	0.48	0.26	0.32	0.52
30	0.30	0.37	0.60	0.30	0.37	0.60	0.28	0.34	0.56	0.30	0.37	0.60
35	0.34	0.42	0.68	0.30	0.37	0.60	0.30	0.37	0.60	0.36	0.44	0.72
40	0.40	0.49	0.80	0.44	0.54	0.88	0.48	0.59	0.96	0.44	0.54	0.88
45	0.46	0.56	0.92	0.40	0.49	0.80	0.60	0.73	1.20	0.52	0.64	1.04
50	0.56	0.69	1.12	0.54	0.66	1.08	0.72	0.88	1.44	0.54	0.66	1.08
55	0.66	0.81	1.32	0.56	0.69	1.12	0.84	1.03	1.68	0.78	0.96	1.56
60	0.84	1.03	1.68	0.66	0.81	1.32	1.00	1.22	2.00	0.82	1.00	1.64

all altitudes between 40 and 60 km with maximums at 12 and 36 hours and a minimum at 24 hours. An analyses of meteorological rocket observations taken at Kwajalein (9°N) and Ft. Sherman (9°N) show similar results [Kantor and Cole, 1981]. The diagram in Figure 15-22 and the results of similar studies show that in the tropics an observation 24 hours old is more representative of actual conditions than one 12 hours old.

The observed rms variations of density with time lags of 1, 2, 4 and 6 hours are shown in Figure 15-23 for levels between 60 and 90 km at Kwajalein. This information, from Cole et al. [1979], is based on a July 1978 series of high-altitude ROBIN falling sphere flights at Kwajalein. The first profile represents the estimated rms observational error.

The rms variations of density with time at the 50-km level are shown in Figure 15-24 for Wallops Island (38°N) and Ft. Churchill (59°N) for the months of January and July. Unlike the tropics, a 24-h oscillation in density is not apparent from this analysis which is based on eight years of data at Ft. Churchill and ten years at Wallops Island. The diurnal oscillation is relatively small and is probably masked by instrumentation errors and changing synoptic patterns. The rms variability at both locations increases with time until the climatic values of day-to-day variations around the monthly means are reached.

15.3 ATMOSPHERIC PRESSURE UP TO 90 KM

Pressure data provided in this section are based on (1) routine radiosonde observations taken by national weather services and extending to approximately 30 km, and (2) measurements from rockets and instruments released from

Table 15-22 Mean monthly latitudinal density gradients (% change per 180 km) in the tropics.

Altitude (km)	January Gradient (%)	April Gradient (%)	July Gradient (%)	October Gradient (%)
10	0.01	0.02	0.03	0.04
15	0.15	0.17	0.08	0.05
20	0.12	0.23	0.08	0.06
25	0.04	0.14	0.10	0.14
30	0.26	0.12	0.14	0.21
35	0.13	0.22	0.16	0.23
40	0.03	0.16	0.16	0.20
45	0.14	0.01	0.17	0.21
50	0.11	0.09	0.12	0.20
55	0.08	0.12	0.04	0.27
60	0.09	0.04	0.10	0.25

Table 15-23a. Kwajalein—Correlation of January density (kg/m) from surface to 60 km.

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	008	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
*MEAN	1167	969	786	639	516	416	332	261	199	141	934	654	464	342	248	182	133	980	725	538	401	303	230	175	234	105	818	640	504	397	310
	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-4	-4	-4	-4	-4	-4	-4	-5	-5	-5	-5	-5	-5	-5	-5	-5	-6	-6	-6	-6	-6
STDV	15	5	4	4	5	4	5	5	10	28	15	13	12	18	15	15	17	18	16	18	21	20	23	28	22	28	28	33	34	37	37
N	42	42	42	42	42	42	42	42	42	42	42	42	41	40	41	42	42	42	42	42	42	42	42	42	42	42	42	42	41	38	34
2	17	**																													
4	17	37																													
6	4	7	22																												
8	-9	-24	2	14																											
10	9	10	17	9	58																										
12	7	-17	1	21	44	74																									
14	9	-9	-25	3	23	27	49																								
16	2	7	-36	-11	-7	10	4	29																							
18	-3	2	-22	-32	-37	-17	-14	9	52																						
20	-1	9	8	19	7	28	17	7	12	15																					
22	1	-2	22	28	25	1	-4	6	21	38	42																				
24	-30	0	-12	16	-6	-16	-5	27	14	7	-5	31																			
26	6	-4	6	-13	-4	4	-2	11	31	15	-5	14	41																		
28	16	14	10	2	-14	-9	1	-9	23	3	9	16	37	66																	
30	25	4	13	-14	-10	-8	-5	17	17	4	6	8	21	61	67																
32	15	6	17	-14	-24	-8	-25	0	-4	-2	7	10	14	30	39	51															
34	0	5	21	-10	33	-8	-11	10	27	-5	-2	-3	10	30	24	41	48														
36	8	21	-1	3	-32	-7	-12	-2	7	-7	10	3	-6	3	18	31	35	52													
38	10	9	3	12	3	33	16	16	-9	-18	-32	0	-1	8	5	32	35	61	55												
40	1	-16	-24	7	-12	16	14	23	-5	-20	0	18	36	-1	8	14	24	45	36	57											
42	-4	14	4	34	-14	7	11	17	-18	-30	6	10	31	-12	24	15	27	33	39	42	70										
44	2	3	-18	23	-20	8	12	29	-14	-23	9	16	26	-2	13	19	21	49	34	59	72	83									
46	-5	15	-5	5	-6	4	3	31	-31	-26	6	16	32	3	4	22	22	31	19	44	47	58	72								
48	9	20	-12	19	-15	8	-3	29	-38	-40	-8	-4	34	-1	19	14	33	39	26	42	45	54	61	70							
50	3	1	12	26	-22	-3	-5	5	-32	-37	-4	-11	14	-1	22	25	34	37	52	44	56	64	63	56	67						
52	-3	7	18	40	17	5	15	9	-43	-42	14	-1	5	-30	3	-1	7	18	40	35	45	64	62	58	58	82					
54	2	8	0	28	-16	0	0	6	-41	-43	9	15	5	-34	-14	-16	2	9	32	25	52	60	63	62	53	67	88				
56	3	5	24	26	-14	-3	2	11	-40	-35	6	14	2	-31	-16	-18	-3	-7	19	14	42	56	55	62	41	52	79	90			
58	-9	6	24	14	-19	5	0	12	-40	-28	7	27	4	-18	-11	-20	0	-6	20	14	35	50	55	67	43	62	81	88	92		
60	-27	5	7	27	-20	6	1	7	-37	-29	4	14	4	-25	-20	-32	-1	6	38	14	29	49	51	58	45	64	83	67	80	91	

*Multiply mean by indicated negative power of 10
**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 15-23b. Kwajalein—Correlation of July density (kg/m) from surface to 60 km.

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	008	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
*MEAN	1170	969	790	642	518	418	334	262	194	134	933	661	473	349	254	187	138	102	754	554	415	313	238	183	142	109	852	669	526	413	321
	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-4	-4	-4	-4	-4	-4	-4	-4	-5	-5	-5	-5	-5	-5	-5	-5	-6	-6	-6	-6	-6
STDV	4	3	4	4	4	4	4	8	12	14	12	9	10	18	11	14	10	16	16	21	20	24	26	27	27	34	36	36	37	37	43
N	31	31	31	31	31	31	31	31	31	31	31	30	29	26	30	30	31	31	31	31	31	31	31	31	31	31	31	30	30	29	26
2	51	**																													
4	65	78																													
6	32	56	71																												
8	0	32	27	40																											
10	14	39	35	30	73																										
12	25	35	24	32	19	21																									
14	4	15	16	26	-20	-14	-4																								
16	-5	-34	-35	-34	-50	-61	-39	-1																							
18	-35	51	-52	-23	-8	-15	-8	4	7																						
20	-24	-29	-9	2	-11	-15	-8	-11	20	6																					
22	-33	-57	-44	-29	-41	-55	-6	15	35	38	56																				
24	-28	-43	-9	-11	-19	-28	-12	-3	29	32	50	30																			
26	-28	-50	-31	-34	7	-7	-39	-37	25	46	15	29	38																		
28	-55	-18	-25	-13	14	4	-31	4	32	25	15	14	2	41																	
30	-29	-20	14	-3	-6	-7	-18	21	12	8	32	31	3	1	52																
32	-22	-21	-17	-19	-24	-1	-12	-17	12	35	31	33	26	36	32	31															
34	-9	-7	8	3	4	-12	-4	-10	14	13	36	37	5	28	26	24	47														
36	-34	-30	-25	18	0	-4	-13	-27	31	11	52	45	12	22	15	17	26	62													
38	-38	-31	-17	-15	18	-6	-13	-31	23	27	37	33	37	44	34	13	20	40	54												
40	-34	-58	-50	-44	-6	-41	-6	-33	37	30	45	42	48	37	16	-4	-1	3	41	49											
42	-35	-55	-46	-44	-9	-33	-5	-18	37	45	29	38	48	27	12	-4	-1	5	47	45	81										
44	-28	-34	-34	-42	-5	-15	2	-29	24	31	31	29	32	20	14	6	8	9	46	35	66	83									
46	-28	-36	-38	-32	-5	-19	-3	-26	24	26	44	38	33	27	24	16	9	16	50	46	69	74	92								
48	-36	-42	-42	-30	-10	-35	-19	-27	42	29	40	54	18	43	22	8	9	24	57	35	76	73	76	81							
50	-38	-40	-41	-34	-2	-19	-8	-29	34	34	32	38	20	39	24	-3	2	16	48	28	75	76	77	77	91						
52	-26	-46	-39	-44	-3	-17	-14	-36	27	32	29	24	26	32	10	-3	-1	13	43	31	78	77	79	79	82	89					
54	-15	-41	-34	-50	-6	-18	-8	-29	20	31	27	19	31	19	3	-7	0	5	36	26	77	79	82	82	86	78	92				
56	-24	-40	-39	-47	-4	-25	-10	-36	20	30	33	24	29	27	19	-6	3	0	33	22	85	81	80	70	76	81	82	87			
58	-29	-38	-32	-29	-20	-47	-16	-27	39	23	44	39	34	15	29	1	13	2	39	29	78	75	69	69	75	77	70	72	87		
60	-51	51	-43	-36	-10	-30	-10	-32	31	29	54	44	37	20	43	27	13	7	59	44	81	82	65	74	73	77	75	77	87	95	

*Multiply mean by indicated negative power of 10
**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 15-23c. Wallops Island—Correlation of January density (kg/m) from surface to 60 km.

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	015	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
*MEAN	1292	1029	821	658	524	411	310	227	168	122	874	638	452	331	241	175	128	936	689	509	379	285	216	187	130	102	601	627	486	378	293
	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-4	-4	-4	-4	-4	-4	-4	-5	-5	-5	-5	-5	-5	-5	-5	-5	-6	-6	-6	-6	-6
STDV	23	26	16	10	15	35	51	43	41	39	29	40	25	22	23	25	29	36	38	48	50	47	42	44	44	46	47	50	52	55	63
N	44	44	44	44	44	44	44	44	43	43	43	43	43	42	43	43	43	43	43	43	43	43	43	43	43	43	43	43	39	33	16
2	74	**																													
4	66	92																													
6	34	28	53																												
8	29	62	50	25																											
10	33	75	74	9	86																										
12	42	72	79	34	81	84																									
14	39	79	83	31	60	84	92																								
16	42	81	84	27	63	82	88	97																							
18	39	82	81	23	63	79	79	92	94																						
20	37	76	72	9	57	70	67	79	83	92																					
22	34	51	47	5	33	46	49	49	52	61	72																				
24	21	53	46	2	36	39	40	48	55	69	76	59																			
26	8	48	37	11	38	40	39	44	48	63	70	57	90																		
28	21	50	41	5	45	45	35	41	43	58	62	51	79	83																	
30	13	45	37	11	46	40	34	41	43	55	53	41	65	73	81																
32	12	28	25	3	21	18	28	33	36	38	29	19	41	46	53	80															
34	17	18	24	21	9	7	25	23	24	24	10	9	23	20	28	60	86														
36	22	17	25	32	3	1	23	22	25	20	4	1	14	4	12	38	69	89													
38	28	8	16	32	5	8	9	6	9	3	12	12	2	16	13	9	42	71	87												
40	29	8	14	23	1	8	6	3	6	0	13	9	2	18	13	9	35	65	81	95											
42	32	6	7	7	6	16	3	5	0	8	18	1	3	17	12	5	26	51	67	80	87										
44	34	7	5	1	6	13	1	1	4	7	16	3	12	22	19	2	19	36	53	64	72	89									
46	30	6	2	3	3	10	1	4	8	3	13	1	12	23	19	5	13	30	45	57	66	79	92								
48	21	6	1	4	1	8	2	7	9	1	13	1	17	22	19	5	8	20	32	42	48	65	79	92							
50	12	7	5	1	1	0	3	12	13	3	12	1	21	23	25	7	1	12	18	24	28	42	60	75	91						
52	3	0	0	3	5	5	2	7	7	3	18	10	29	31	30	13	8	9	15	19	20	34	48	64	82	94					
54	1	0	0	6	6	5	3	7	8	3	16	14	30	32	35	21	19	8	0	3	5	22	42	55	75	88	94				
56	1	6	6	3	12	12	11	1	2	6	16	17	28	29	33	22	23	15	9	8	1	13	33	51	74	86	87	95			
58	2	15	15	4	7	12	10	8	5	17	28	26	39	38	40	18	15	5	1	4	13	27	44	61	78	88	84	91	96		
60	6	26	26	8	47	46	32	29	23	41	52	48	55	59	69	55	39	19	3	10	20	43	59	71	83	92	87	91	93	98	

*Multiply mean by indicated negative power of 10
**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 15-23d. Wallops Island—Correlation of July density (kg/m) from surface to 60 km.

KM Kilometers above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
In Percent of Mean Times 10
N Number of Values at Each Altitude

KM	015	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
*MEAN	1192	980	798	647	522	420	333	254	184	131	934	673	486	360	264	194	143	106	786	588	441	332	253	195	152	119	933	734	574	451	352
	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-4	-4	-4	-4	-4	-4	-4	-4	-5	-5	-5	-5	-5	-5	-5	-5	-6	-6	-6	-6	-6
STDV	14	8	6	5	6	7	9	18	17	14	11	12	12	15	17	18	20	21	22	26	31	32	31	32	36	38	42	42	43	47	33
N	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	34	30	18
2	77	**																													
4	40	72																													
6	26	55	82																												
8	25	39	55	66																											
10	17	32	42	48	73																										
12	0	16	18	20	27	71																									
14	-16	-21	-16	1	-8	11	42																								
16	-26	-31	-22	-16	-22	-10	21	64																							
18	-7	-5	-4	8	-2	-12	-3	47	75																						
20	-1	13	-6	12	7	-2	-3	24	36	61																					
22	-12	11	-9	10	3	3	18	12	21	35	68																				
24	-11	10	0	14	5	18	23	1	-1	13	63	75																			
26	9	17	2	4	-4	-2	9	5	-4	15	38	70	66																		
28	21	34	10	19	8	12	10	-1	1	22	61	77	75	83																	
30	3	7	-16	-9	-18	-4	5	-4	-3	13	53	67	75	70	84																
32	6	18	-2	-1	-13	-8	8	4	10	30	59	73	75	80	84	86															
34	-3	3	-17	-1	-7	-7	-2	0	-1	20	50	63	65	69	71	73	82														
36	5	11	-19	-5	-10	-9	2	13	7	29	51	67	57	71	75	74	83	88													
38	19	10	-19	-3	-12	-19	-12	11	1	30	48	56	50	73	74	73	81	82	90												
40	4	8	-15	-4	-17	-20	-9	2	-2	22	46	58	55	71	70	76	86	87	90	93											
42	2	7	-18	-5	-19	-22	-1	14	10	28	47	59	48	63	67	75	84	82	90	89	95										
44	21	27	-2	6	-7	-9	2	12	-5	16	44	58	48	71	75	68	80	78	87	88	85	89									
46	14	18	-10	-2	-10	-22	-10	13	-4	20	42	54	43	69	68	63	76	79	88	88	84	86	96								
48	7	7	-24	-9	-14	-20	-8	20	4	29	50	58	48	67	69	67	78	79	90	90	86	88	89	92							
50	12	12	-22	-11	-10	-15	-1	14	4	33	52	58	48	70	70	66	78	78	90	90	86	86	85	88	96						
52	11	5	-32	-21	-15	-22	-4	16	8	33	51	55	39	58	61	62	74	77	89	87	82	86	83	87	92	94					
54	8	5	-31	-21	-17	-25	-4	13	8	33	51	58	42	61	61	60	75	78	89	86	83	85	80	84	89	93	98				
56	9	8	-20	-6	-4	-11	-4	12	9	45	69	56	47	56	65	60	75	77	88	85	82	82	81	82	86	93	93	97			
58	12	6	-22	-5	-6	-17	-15	9	6	40	64	53	38	57	65	60	74	81	91	87	85	83	85	88	90	91	93	95	97		
60	-3	-16	-43	-13	-35	-2	35	50	24	31	29	23	32	32	33	45	55	76	85	78	79	78	73	78	83	93	91	91	92	96	

*Multiply mean by indicated negative power of 10
**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 15-23e. Ft. Churchill - Correlation of January density (kg/m) from surface to 60 km.

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	0[15]	2	3	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	1446	1078	848	666	511	375	273	201	148	109	806	591	439	312	227	166	120	881	641	471	347	256	191	142	108	814	625	479	372	286	220
	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	6	6	6	6	6	6
STDV	31	19	15	12	24	26	22	20	20	26	30	38	33	65	79	92	108	118	126	129	138	144	153	150	149	151	154	160	163	138	140
N	50	50	50	50	50	50	50	50	46	40	30	29	23	43	46	46	46	46	46	46	46	46	46	46	46	46	46	46	44	41	33
2	72	**																													
4	36	70																													
6	7	8	60																												
8	6	36	8	55																											
10	25	49	22	33	89																										
12	34	50	24	20	71	94																									
14	46	55	25	12	57	81	92																								
16	49	49	19	0	37	63	76	89																							
18	39	26	6	-3	15	32	42	62	85																						
20	45	22	3	4	2	9	17	35	60	91																					
22	38	18	6	6	11	3	2	18	45	81	96																				
24	26	21	56	41	4	3	1	11	31	72	91	98																			
26	43	28	8	21	2	6	9	26	43	67	83	91	92																		
28	47	33	1	18	0	7	10	29	41	61	76	85	86	98																	
30	43	30	-1	17	3	1	1	20	30	51	62	72	76	94	98																
32	40	30	2	16	4	2	-4	15	24	43	54	65	69	90	96	99															
34	34	25	-4	12	-7	8	11	7	14	35	44	55	57	84	92	97	99														
36	34	29	7	11	5	7	11	7	14	33	33	44	44	81	89	95	97	99													
38	29	29	11	9	3	6	11	5	11	26	22	32	29	76	85	92	95	97	99												
40	25	27	12	8	3	8	14	1	6	19	12	23	20	69	80	87	91	95	97	99											
42	28	29	11	9	2	8	14	-1	3	13	6	19	23	66	77	85	89	92	94	96	98										
44	28	30	13	4	5	10	14	0	3	10	5	18	20	64	75	82	86	89	90	93	94	98									
46	32	32	13	2	5	8	9	5	8	11	3	16	18	60	71	77	80	82	83	85	87	93	97								
48	33	33	14	0	2	3	4	8	9	10	3	16	18	56	67	71	73	74	75	77	79	87	93	98							
50	34	32	16	6	2	2	1	10	10	10	7	17	18	47	56	60	61	62	62	64	66	76	84	92	97						
52	35	32	16	5	1	1	1	10	8	6	4	13	12	38	48	51	52	52	52	54	56	68	76	86	93	98					
54	35	31	15	5	1	0	2	9	7	2	3	12	11	32	41	45	45	45	45	47	50	62	71	82	89	96	99				
56	35	30	12	5	1	1	4	7	5	2	3	12	12	29	35	39	37	37	36	39	41	54	64	76	86	92	97	99			
58	32	30	15	4	4	3	2	7	14	12	3	0	23	29	35	36	37	41	46	51	57	63	73	81	88	93	96	99			
60	49	42	31	6	1	1	1	0	4	10	7	1	10	17	28	37	37	38	44	47	52	57	62	70	76	81	86	90	93	97	

*Multiply mean by indicated negative power of 10
**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 15-23f. Ft. Churchill—Correlation of July density (kg/m) from surface to 60 km.

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	035	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
*MEAN	1236	999	810	654	525	409	306	224	166	122	901	662	487	362	267	197	146	109	813	610	460	349	267	206	161	126	992	765	620	492	389
	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-4	-4	-4	-4	-4	-4	-4	-4	-5	-5	-5	-5	-5	-5	-5	-5	-6	-6	-6	-6	-6
STDV	21	12	11	10	10	19	39	24	22	20	16	14	12	15	13	12	12	13	11	14	15	16	18	21	22	23	23	25	27	29	32
N	28	28	28	28	28	28	28	28	28	28	28	27	27	28	28	28	28	28	28	28	28	28	28	28	28	28	26	25	25	20	19
2	55	**																													
4	59	85																													
6	69	74	89																												
8	39	43	61	78																											
10	8	-32	-19	-17	29																										
12	17	-64	-54	-56	-24	69																									
14	24	-69	58	-56	-26	63	92																								
16	32	-75	-64	-60	-32	56	88	96																							
18	29	-74	60	-54	-30	50	78	92	97																						
20	28	-72	-60	-54	-33	44	71	87	93	98																					
22	33	-72	-63	-57	-39	39	75	86	91	95	98																				
24	-42	-68	62	-56	-44	33	71	81	85	87	89	95																			
26	-8	-42	-40	-40	-23	44	58	71	72	73	71	71	63																		
28	-9	-37	-37	-42	-17	51	51	63	61	60	57	56	49	93																	
30	-1	-33	-27	-34	-25	38	50	56	54	53	48	51	50	84	83																
32	-7	-23	-28	-32	-21	38	37	46	43	41	37	39	37	72	82	79															
34	10	-19	-26	-36	-29	34	37	37	36	32	27	28	25	60	72	77	88														
36	-27	-38	-43	-56	-46	18	36	35	34	30	28	33	39	29	41	48	63	67													
38	-34	-29	-44	-56	-45	4	14	26	27	26	29	33	34	29	46	30	63	53	71												
40	-25	-19	-32	-42	-37	-6	6	16	25	25	27	25	22	22	36	15	48	45	52	78											
42	-21	10	-23	-17	-18	-5	-6	-4	-1	-3	-3	-3	1	-5	17	6	42	37-	44	64	80										
44	-45	-9	-19	-28	-27	-9	-3	-4	-2	-6	-7	-6	3	-25	-3	-12	27	22	52	61	69	85									
46	51	-28	-29	-40	-39	-18	-7	-6	-2	-3	-2	0	8	-19	-3	-11	26	19	57	68	62	64	85								
48	-42	-17	-18	-21	-31	-30	-15	-17	-12	-14	-16	-14	-5	-23	-9	-14	27	23	50	56	57	60	79	92							
50	-35	-7	-10	-10	-24	-30	-11	-12	-12	-14	-17	-14	-2	-23	-12	-7	33	23	51	51	50	60	75	79	89						
52	-31	9	2	-4	-21	-38	-18	-16	-21	-21	-20	-17	-11	-30	-19	-17	22	14	44	50	50	67	79	76	81	94					
54	-51	12	11	-14	-26	-36	-15	-6	-8	-5	-1	1	3	-28	-18	-21	21	9	48	58	56	70	81	81	80	87	95				
56	44	8	-10	-13	-23	-37	-24	-17	-16	-13	-8	-8	-7	-42	-29	-34	7	3	41	58	61	71	83	83	80	84	90	96			
58	49	10	-26	-24	-28	-24	-17	-19	-16	-17	-14	-13	-15	-43	-31	-41	1	3	37	39	55	71	84	78	76	81	87	92	98		
60	66	-11	29	-36	-39	-27	-17	-22	-17	-20	-16	-15	-17	-44	-30	-45	-8	-3	40	41	54	71	85	83	76	74	81	88	95	98	

*Multiply mean by indicated negative power of 10
**Multiply tabular values by 0.01 to obtain correlation coefficients

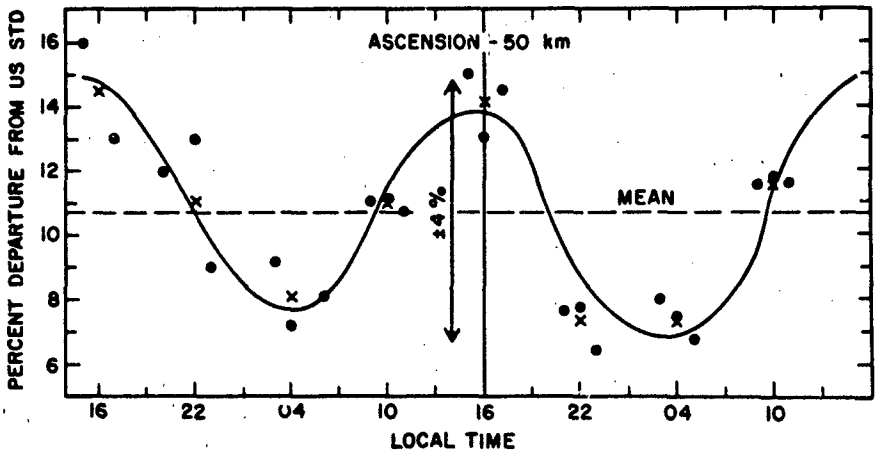


Figure 15-21. Diurnal density (50 km) variation at Ascension Island. (Dots indicate observed values, x's represent 3-h averages, and the solid line depicts the computed diurnal cycle.)

rockets at altitudes between 25 and 90 km. Both data sources are supplemented with pressures derived from measurements made from earth-orbiting satellites. Although atmospheric pressure above radiosonde altitudes is occasionally measured directly, it normally is calculated hydrostatically (as discussed in Chapters 14) from observed temperatures or densities for altitudes above 30 km. These data are intended for use in design problems involving variations in the heights of constant pressure surfaces and/or changes in pressure at specific altitudes.

15.3.1 Sea-Level Pressure

The variations of sea-level pressure normally have little effect on the operation of surface equipment. However, in the design of sealed containers that could possibly explode or collapse with pressure changes, the range of surface pressures likely to be encountered should be considered. Surface pressures vary with the height of the station above sea level as well as with changing weather patterns. Standard atmospheric pressure at sea level is 1013.25 mb, but there

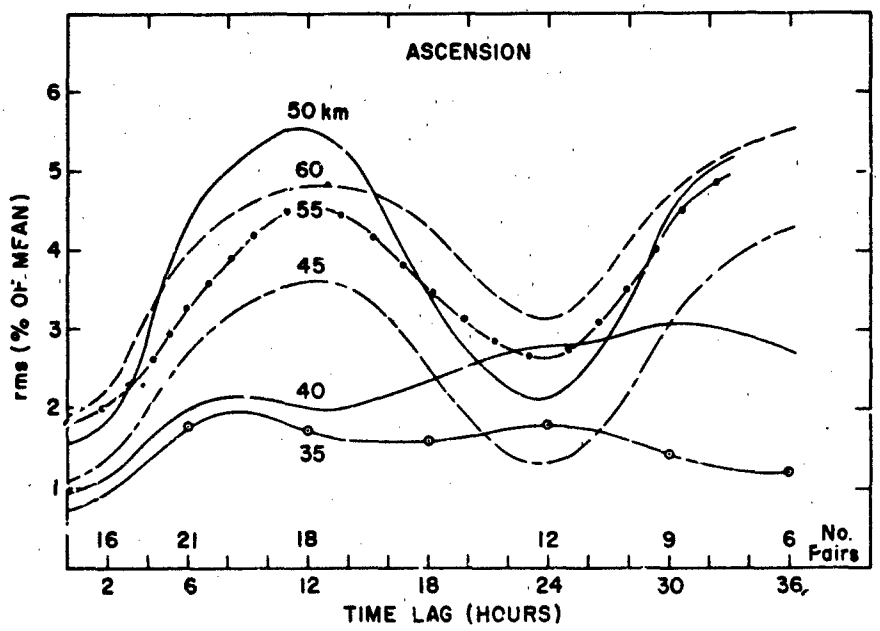


Figure 15-22. Root mean square (rms) lag variability of density with time at Ascension Island.

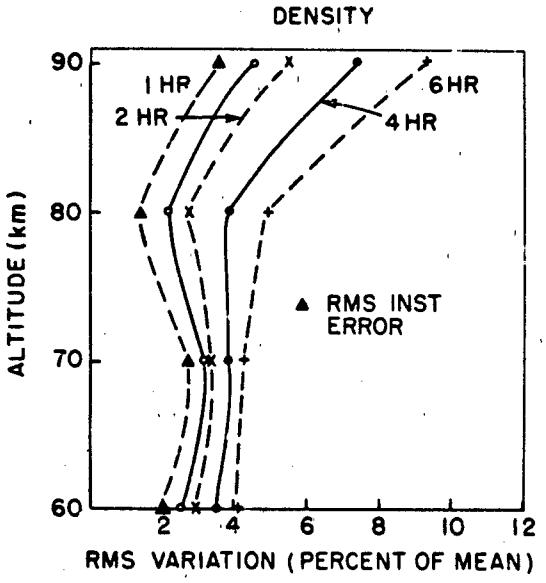


Figure 15-23. The rms variation in density for time lags of 1 to 6 h at Kwajalein.

are sizable variations from this value with both time and location.

Table 15-24 indicates extreme sea-level pressures that may be encountered in the Northern Hemisphere. During the month of January, pressures exceeded 99% of the time are given for areas under the influence of semipermanent cyclones, and pressures exceeded 1% of the time are given

Table 15-24. Sea-level pressures exceeded 99% and 1% of the time in January.

Location	Pressure (mb)
	Exceeded 99% of time
Aleutian low	965
Icelandic low	953
	Exceeded 1% of time
Siberian high	1057
Pacific high	1038
Canadian high	1052

in areas under the influence of anticyclones. In the Northern Hemisphere extreme values, excluding tropical cyclones and tornadoes, are most likely to occur in these regions during January. Table 15-25 lists for comparison actual worldwide pressure extremes, including those resulting from storms of tropical origin. Examples of mean sea-level pressures and typical fluctuations are given in Table 15-26 which contains mean sea-level pressures for the four midseason months and the standard deviations of daily values around these means for 16 specific locations in the Northern Hemisphere.

15.3.2 Seasonal and Latitudinal Variations

The Reference Atmospheres presented in Chapter 14 provide tables of mean monthly pressure-height profiles,

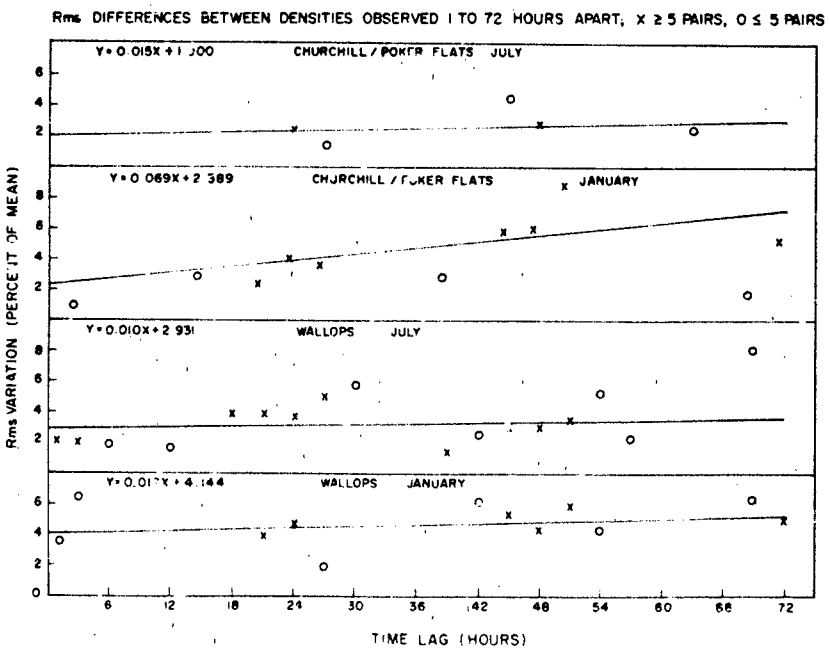


Figure 15-24. Rms differences between densities observed 1 to 72 h apart at 50 km (x ≥ 5 pairs, o ≤ 5 pairs).

Table 15-25. Worldwide pressure extremes.

	Pressure (mb)	Location	Date
LOW			
World	870*	17°N, 138°E, Typhoon Tip [Wagner, 1980]	12 Oct 1979
	876*	13°N, 141°E, Typhoon June [Holliday, 1976]	19 Nov 1975
	877*	19°N, 135°E. Typhoon Ida [Riordan, 1974]	24 Sep 1958
	877*	15°N, 128°E, Typhoon Nora [Holliday, 1975]	6 Oct 1973
No. America	892.3	Matecumbe Key, Florida, hurricane [Riordan, 1974]	2 Sep 1935
HIGH			
World	1083.8	Agata, Siberia [Riordan, 1974]	31 Dec 1968
	1075.2	Irkutsk, Siberia, [Valley, 1965]	14 Jan 1893
No. America	1067.3	Medicine Hat, Alberta [Riordan, 1974]	24 Jan 1897

*Dropsonde measurements

surface to 90 km, for 15° intervals of latitude from the equator to the north pole. These atmospheric models describe both seasonal and latitudinal variation of pressure. Figure 15-25 contain the vertical pressure profiles for the midseason months at each of four latitudes: 15°, 30°, 45° and 60°N. The profiles are displayed as percent departures from the *U.S. Standard Atmosphere, 1976*. Pressures at altitudes between 2 km and 70 or 80 km are highest in summer and lowest in winter over regions poleward of 30°N. In tropical latitudes, seasonal differences are small and do not display a systematic seasonal pattern. Departures from standard generally increase with latitude. Thus, largest seasonal differences are shown at 60°N where mean monthly pressures at 60 to 70 km are nearly 40% greater than standard in July and 30% to 36% less than standard in January. Consequently, July values are roughly twice those in January between 60 and 70 km. Pressures at these levels at 75°N (not shown) are roughly 10% lower than these values in winter and 15% higher in summer.

Table 15-26. Mean monthly sea-level pressures and standard deviations of daily values.

Location		January		April		July		October	
Latitude	Longitude	Mean (mb)	S.D. (mb)	Mean (mb)	S.D. (mb)	Mean (mb)	S.D. (mb)	Mean (mb)	S.D (mb)
10°N	70°W	1013	2	1012	1	1012	1	1011	1
20°N	70°W	1018	2	1017	2	1018	2	1013	2
30°N	70°W	1022	6	1019	5	1021	3	1018	4
40°N	70°W	1018	10	1017	9	1016	5	1018	8
50°N	70°W	1016	12	1014	10	1011	6	1013	11
60°N	70°W	1008	11	1014	10	1008	7	1008	10
70°N	70°W	1004	11	1014	10	1009	6	1006	10
80°N	70°W	1011	11	1020	9	1011	6	1013	8
10°N	20°E	1012	4	1008	3	1009	2	1009	2
20°N	20°E	1017	4	1011	3	1008	2	1012	2
30°N	20°E	1019	5	1014	4	1012	3	1015	3
40°N	20°E	1018	9	1013	6	1013	3	1016	5
50°N	20°E	1020	12	1013	7	1013	5	1017	8
60°N	20°E	1015	16	1012	10	1011	7	1011	11
70°N	20°E	1004	15	1010	10	1012	7	1005	11
80°N	20°E	1008	15	1016	10	1014	7	1010	10

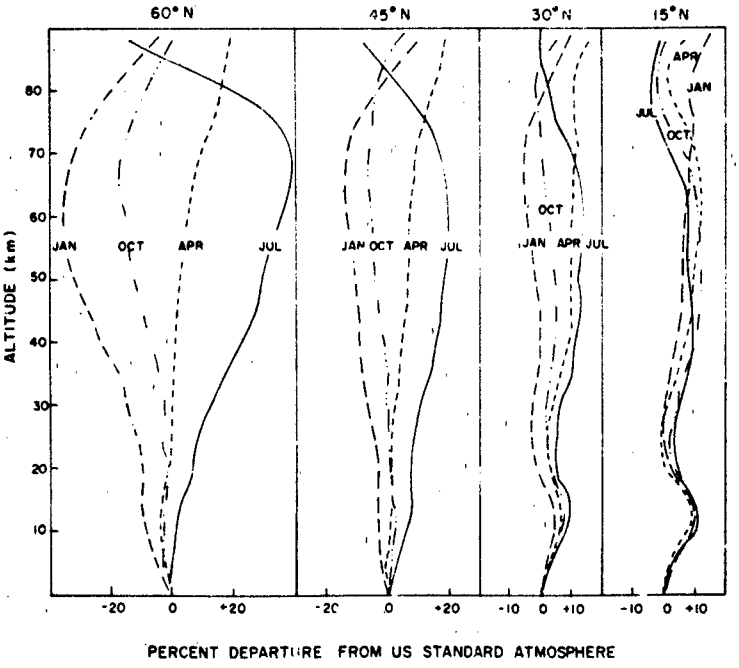


Figure 15-25. Pressure-altitude profiles for midseason months.

Table 15-27. Average height and standard deviation at standard pressure levels over North America, 90 to 100°W

Average Height and Standard Deviation														
	20°N		30°N		40°N		50°N		60°N		70°N		80°N	
Pressure (mb)	Mean (km)	S.D. (m)	Mean (km)	S.D. (m)	Mean (km)	S.D. (m)	Mean (km)	S.D. (m)	Mean (km)	S.D. (m)	Mean (km)	S.D. (m)	Mean (km)	S.D. (m)
January														
700	3.165	30	3.115	55	3.015	85	2.865	100	2.770	100	2.710	90	2.690	75
500	5.845	40	5.745	85	5.565	125	5.340	145	5.180	150	5.075	140	5.055	125
300	9.595	55	9.425	125	9.150	175	8.825	195	8.585	230	8.425	210	8.380	180
200	12.280	70	12.090	130	11.765	165	11.430	175	11.180	190	10.995	200	10.920	195
100	16.455	55	16.325	95	16.110	110	15.890	145	15.655	175	15.400	195	15.195	190
50	20.540	130	20.500	200	20.415	215	20.280	215	20.075	200	19.775	180	19.440	180
25	24.900	210	24.865	245	24.790	335	24.555	275	24.380	245	23.905	245	23.425	230
15	28.100	245	28.050	335	28.000	365	27.750	365	27.650	350	27.000	335	26.350	305
10	30.600	250	30.550	380	30.500	400	30.250	380	30.150	380	29.500	320	28.750	260
July														
700	3.185	15	3.190	20	3.170	35	3.080	55	3.005	60	2.975	65	2.930	65
500	5.890	20	5.910	25	5.875	50	5.720	85	5.600	90	5.540	90	5.465	95
300	9.675	30	9.705	35	9.630	80	9.405	130	9.215	125	9.125	125	9.010	130
200	12.395	40	12.430	50	12.345	100	12.080	135	11.870	135	11.790	135	11.710	135
100	16.570	45	16.625	50	16.605	65	16.515	80	16.455	85	16.420	80	16.390	75
50	20.765	75	20.865	90	20.940	105	20.975	105	21.005	105	21.045	90	21.050	75
25	25.180	150	25.330	150	25.440	175	25.530	175	25.625	165	25.715	150	25.765	150
15	28.300	170	28.450	170	28.650	175	28.800	175	28.900	175	29.100	170	29.200	165
10	30.890	190	30.950	190	31.200	190	31.400	190	31.550	190	31.750	185	31.950	190

Table 15-28. Departures from mean monthly pressures (%) exceeded less than 5% of the time in January and July. Values below 30 km are based on radiosonde observations. Those values above 30 km are based on rocketsonde observations

Height	January					July				
(km)	75°N	60°N	45°N	30°N	15°N	15°N	30°N	45°N	60°N	75°N
0	±2.5	±3	±2.5	±1	±0.4	±0.4	±0.5	±1	±1	±1.5
10	7	4	3	2	0.8	0.7	0.8	2	3	4
20	10	10	10	7	4	2	2	3	3	3
30	20	16	14	12	7	4	4	4	5	5
40		25	20	15	8	7	8	8	10	
50		30	25	18	10	10	12	13	14	
60		35	30	20	12	12	14	16	18	
70		30	25	18	10	10	12	15	16	
80		20	16	12	8	8	9	10	10	

15.3.3 Day-to-Day Variations

Changing synoptic situations, which include movements of high and low pressure centers and their associated ridges and troughs, and variations in the energy absorbed directly by the atmosphere, cause day-to-day changes in the height of constant pressure surfaces. Information on the magnitude of day-to-day variations in the heights of such surfaces are provided in this section. Detailed information for specific levels and locations should be requested if conditions appear critical.

Table 15-27 lists monthly mean heights of pressure surfaces in January and July and their standard deviations for middle North America. These data indicate the variation in the mean heights of constant pressure surfaces between 700 and 10 mb with latitude and season, and the estimated distributions of day-to-day variability around the monthly means. Estimated departures from mean monthly pressures, which are equaled or exceeded less than 5% of the time between 15° and 75°N, are shown in Table 15-28 as percentages of the mean January and July values, surface to 80 km.

As can be seen in Table 15-28, day-to-day variability generally increases with latitude and altitude in both January and July, although to a much smaller extent in July. The estimated 5% extremes are largest at 60 to 65 km at all latitudes, reaching ±35% during 60°N winter. Variability appears to decrease above 65 km to a probable minimum value near 85 km. The estimated departures shown in Table 15-28 include some diurnal and semidiurnal fluctuations due to solar influences, particularly since the basic pressure data were not observed at the same time every day. Envelopes of these estimated 95% values should not be used as profiles since such pressures would not necessarily be found at all altitudes at any one given time and/or location. Decreases in atmospheric pressure in one layer, for example, normally are associated with increases in another layer.

Table 15-29. Amplitudes of systematic pressure variations and time of maximum at Terceira, Azores [Harris et al, 1962].

		Diurnal		Semidiurnal	
Pressure Level (mb)	(m)*	Ampl. (mb)	Time (h)	Ampl. (mb)	Time (h)
Sfc	0	0.10	2100	0.50	0948
1000	122	0.10	1904	0.53	0950
950	570	0.12	1824	0.46	0956
900	1033	0.16	1612	0.49	1002
850	1454	0.18	1604	0.47	1002
800	2027	0.20	1612	0.44	1002
750	2569	0.20	1616	0.38	1010
700	3127	0.25	1548	0.37	1002
650	3731	0.18	1608	0.40	1030
600	4365	0.25	1608	0.33	1020
550	5051	0.27	1508	0.33	1034
500	5782	0.28	156	0.29	1032
450	6587	0.27	1424	0.24	1036
400	7449	0.31	1504	0.24	1046
350	8409	0.31	1504	0.20	1046
300	9482	0.32	1444	0.18	1108
250	10708	0.33	1420	0.16	1102
200	12149	0.32	1408	0.14	1110
175	12991	0.32	1352	0.13	1120
150	13948	0.30	1348	0.11	1100
125	15066	0.28	1328	0.11	1124
100	16423	0.26	1304	0.09	1128
80	17776	0.24	1300	0.09	1120
60	19547	0.23	1256	0.07	1124
50	20668	0.21	1256	0.07	1114
40	22077	0.20	1244	0.06	1116
30	24012	0.18	1256	0.05	1110
20	26673	0.16	1256	0.04	1128
15	28005	0.15	1252	0.03	1136
10	30507	0.12	1304	0.01	1204

*Estimated mean annual height

15.3.4 Diurnal and Semidiurnal Variations

Mean hourly sea-level pressures follow a systematic diurnal and semidiurnal periodicity somewhat variable in amplitude and phase according to location and season. The sea-level pressure cycle is generally characterized by minima near 0400 and 1600 hours and maxima near 1000 and 2200 hours local time. The amplitude approaches 1 mb, which is small relative to synoptic changes in middle latitudes. In the tropics only minor synoptic changes occur from day to day, so that interdiurnal pressure changes are small compared to the systematic daily variations in these latitudes.

Upper-air pressures appear to follow a systematic diurnal/semidiurnal cycle similar to that at sea level; however, extremes occur at somewhat different hours. Table 15-29 lists amplitudes and times of occurrence of diurnal and semidiurnal maxima to 10 mb (roughly 30 km) over Terceira, Azores, which provides an estimate of mean annual systematic pressure variations at a maritime location near 40°N. The *semidiurnal* variations at climatically and geographically different locations such as Washington, D.C. and Terceira, Azores, appear to be similar [Valley, 1965]. The *diurnal* maxima and minima, however, that result from solar insolation and terrestrial radiation, may differ considerably in time of occurrence and amplitude at various locations, particularly at or near surface levels.

CHAPTER 15

REFERENCES — CHAPTER 15

Albrecht, F., "Micrometeorologische Temperaturmessungen von Flugzeug aus," *Ber. Deutsch Wetterd, Bad Kissingen*, **38:** 332, 1952.

Bertoni, E.A. and I.A. Lund, "Winter Space Correlations of Pressure, Temperature and Density of 16 km," AFCRL-64-1020, ADA611002, 1964.

Cole, A.E. "Review of Data and Models of the Middle Atmosphere," *Space Research*, **20** 1979.

Cole, A.E. and A.J. Kantor, "Tropical Atmospheres, 0 to 90 km," AFCRL TR-75-0527, ADA019940, 1975.

Cole, A.E. and A.J. Kantor, "Interlevel Correlations of Temperature and Density, Surface to 60 km," AFGL TR-80-0163, ADA090515, 1980.

Cole, A.E., A.J. Kantor, and C.R. Philbrick, "Kwajalein Reference Atmospheres, 1979," AFGL TR-79-0241, ADA081780 1979.

Davidson, B. and H. Lettau (eds.), *Great Plains Turbulence Field Program*, Pergamon Press, New York, 1957.

Draeger, R.H. and R.H. Lee, "Meteorological Data Eniwetock Atoll," *Naval Medical Res. Inst. Memo Rept. 53-8*, Bethesda, Md. 1953.

Geiger, R., *The Climate Near the Ground*, Harvard University Press, Translated by M.N. Stewart et al., 1957.

Gringorten, I.I., "Probability Models of Weather Conditions Occupying a Line or an Area," *J. Appl. Meteorol.*, **18:** 957, 1979.

Harris, M.F., F.G. Finger, and S. Teweles, "Diurnal Variation of Wind, Pressure and Temperature in the Troposphere and Stratosphere over the Azores," *J. Atmos Sci.*, **19:** 136, 1962.

Holliday, C.R., "An Extreme Sea-level Pressure Reported in a Tropical Cyclone," *Monthly Weather Rev.*, **103:** 163, 1975.

Holliday, C.R., "Typhoon June — Most Intense on Record," *Monthly Weather Rev.*, **104:** 1188, 1976.

Kantor, A.J. and A.E. Cole, "Time and Space Variations of Density in the Tropics," AFGL TR-79-0109, ADA074472, 1979.

Kreutz, W., "Der Jahresgang der Temperatur in Verschiedenen Boeden unter Gleichen Witerungs Verhaeltnisser," *Zeitschr. f. angewandte Meteorologie*, **60:** 65, 1943.

Lettau, H., "A Study of the Mass, Momentum and Energy Budget of the Atmosphere," *Archiv fur Meteorologie, Geophysik and Bioklimatologie* Serie A, **7:** 133, 1954a.

Lettau, H., "Improved Models of Thermal Diffusion in the Soil," *Trans. AGU*, **35:** 121, 1954b.

National Weather Service, "Surface Observations," *Federal Metecrological Handbook 90-1*, Government Printing Office, Washington, D.C., 1979.

Riordan, P., Weather Extremes Around the World, Report ETL-TR-74-5, U.S. Army Engineering Topographic Laboratories, Ft. Belvoir, Va., 1974.

Singer, I.A. and R.M. Brown, "Annual Variation of Sub-Soil Temperatures About a 600-Foot Diameter Circle," *Trans. AGU*, **37:** 743, 1956.

Sissenwine N. and A. Court, "Climatic Extremes for Military Equipment," Env. Protection Br., Rpt. No. 146, Dept. of Army, Washington, D.C., 1951.

Tattelman, P., "Duration of Cold Temperature Over North America," AFCRL-68-0232, AD669567, 1968.

Tattelman, P. and A.J. Kantor, "Atlas of Probabilities of Surface Temperature Extremes: Part I, Northern Hemisphere," AFGL TR-76-0084, ADA027640, 1976a

Tattelman, P. and A.J. Kantor, "Atlas of Probabilities of Surface Temperature Extremes: Part II, Northern Hemisphere," AFGL TR-77-0001, ADA038237, 1976b.

Tattelman, P. and A.J. Kantor, "A Method for Determining Probabilities of Surface Temperature Extremes," *J. Appl. Meteorol.*, **16:** 175, 1977.

Tattelman, P., N. Sissenwine, and R.W. Lenhard, Jr., "World Frequency of High Temperature," AFCRL-69-0348, AD696094, 1969.

Valley, S. (ed.), *Handbook of Geophysics and Space Environments*, AFCRL, 1965.

Wagner, A.J., "Weather and Circulation of October 1979," *Monthly Weather Rev.*, **108:** 119, 1980.

WMO, "Proceedings of the World Climate Conference, Geneva, 12–23 Feb 1979," WMO Report No. 537, 1979.

Chapter 16

WATER VAPOR, PRECIPITATION, CLOUDS, AND FOG

Section 16.1 D.D. Grantham, I.I. Gringorten, and A.J. Kantor
Section 16.2 R.M. Dyer, I.I. Gringorten, and P. Tattelman
Section 16.3 A.A. Barnes, E.A. Bertoni, I.D. Cohen, Capt. M. Glass, D.D. Grantham, I.I. Gringorten, K.R. Hardy, Y. Izumi, A.J. Kantor, and J.I. Metcalf
Section 16.4 H.A. Brown and B.A. Kunkel

One of the most important constituents of the atmosphere is that all-pervasive substance, water, which appears in the solid state as snow, hail, frost, and cloud particles; in the liquid state as rain, clouds, fog, and dew; and in the gaseous state as water vapor. In addition to water, the atmosphere contains in suspension many types and sizes of aerosols of varying composition, including ions, salt, dust, smoke particles and other contaminants. This chapter deals with a variety of conditions or events that are related to atmospheric water content, including surface rainfall rates; the vertical distribution of precipitation intensity; drop-size distributions in rain, clouds and fog; types of cloud data; and water and water-vapor content in and out of clouds.

16.1 ATMOSPHERIC WATER VAPOR

There are various ways to express the water vapor content of the atmosphere. The *dew point* is the temperature to which a parcel of air, at constant pressure and water vapor content, must be cooled for saturation to occur. Further cooling produces condensation to liquid water. Condensation could be delayed as the parcel is cooled below the dew point, in which case the air becomes supersaturated. When the air is supersaturated, the condensation of water vapor can take place quite rapidly and the released latent heat of vaporization will raise the air temperature. If the water vapor content is low enough, then the parcel of air might be cooled below 273 K until saturation with respect to ice occurs, at which point there usually is sublimation directly from water vapor into ice crystals. This temperature is called the *frost point*. Sometimes the air can be cooled below 273 K, to as low as 233 K without sublimation, but there could be condensation onto existing supercooled water drops. Because the saturation vapor pressure over water is greater than that over ice, the dew point for a given water vapor content is always less than the frost point. The frost point-dewpoint difference, although essentially zero at 273 K, becomes increasingly large as the temperature lowers.

Several other measures of water vapor content are related in a one-to-one correspondence to the dew point. Table 16-1 shows this correspondence between dew point and frost point, vapor pressure, and mixing ratio. The table also gives the absolute humidity (vapor density) corresponding to the other measures when the air is saturated.

In a unit volume of air, the water vapor exerts a *vapor pressure*, which in itself is another measure of the water vapor content of the air. When the air is saturated, the vapor pressure is a direct function of air temperature.

The *absolute humidity* is the mass of water vapor in a unit volume of air; it is actually the water vapor density within the air. The Gas Law shows that at a given air temperature, there is a *saturation water vapor content* per unit volume that the air can hold, which roughly doubles for each 10 K increase of air temperature. As a consequence of the Gas Law, an isobaric increase of the temperature will decrease the absolute humidity (vapor density) while the other measures in Table 16-1 remain the same.

If the mass of water vapor is given in proportion to the mass of the ambient dry air, typically in grams per kilogram, the resulting measure is called the *mixing ratio*. To relate mixing ratio to the other measures of water vapor content, the atmospheric pressure must be given. When the mixing ratio remains constant, as it does in adiabatic lifting of unsaturated air, the dew point will decrease with increasing altitude.

Relative humidity is the ratio of the actual vapor pressure of the air to the saturation vapor pressure at the ambient air temperature; it is given in percent. Relative humidity can be derived with sufficient accuracy as the corresponding ratio of the mixing ratio to saturation mixing ratio. At a relative humidity of 100% the dew point is equal to the ambient temperature. This can occur by the addition of moisture into the air and/or by the decrease of temperature to the dew point. Thus, relative humidity usually varies considerably during the course of a day, increasing during the cool of the night and decreasing in the heat of the day.

Precipitable water, W, is the amount of water contained in a vertical air column of unit cross-section extending between two specified levels. While the physical unit is the

CHAPTER 16

Table 16-1. The correspondence between the several measures of water vapor content (based on Smithsonian Meteorological Tables [List, 1968]).

Dew Point (K)	Frost Point (K)	Vapor Pressure (mb)	Absolute‡ Humidity (g/m³)	Mixing Ratio (g/kg)								
				1000 mb	850 mb	700 mb	500 mb	400 mb	100 mb	50 mb	10 mb	1 mb
313		7.378 + 1	5.119 + 1	4.98 + 1	5.941 + 1	7.361 + 1	1.080 + 2	1.411 + 2	*	*	*	*
308		5.624 + 1	3.963 + 1	3.725 + 1	4.427 + 1	5.456 + 1	7.910 + 1	1.020 + 2	8.008 + 2	*	*	*
303		4.243 + 1	3.038 + 1	2.769 + 1	3.282 + 1	4.029 + 1	5.786 + 1	7.399 + 1	4.590 + 2	*	*	*
298		3.167 + 1	2.305 + 1	2.044 + 1	2.417 + 1	2.959 + 1	4.219 + 1	5.363 + 1	2.886 + 2	*	*	*
293		2.337 + 1	1.730 + 1	1.495 + 1	1.766 + 1	2.156 + 1	3.059 + 1	3.870 + 1	1.899 + 2	5.462 + 2	*	*
288		1.704 + 1	1.283 + 1	1.083 + 1	1.278 + 1	1.557 + 1	2.201 + 1	2.775 + 1	1.279 + 2	3.217 + 2	*	*
283		1.227 + 1	9.399	7.762	9.146	1.113 + 1	1.569 + 1	1.973 + 1	8.707 + 1	2.024 + 2	*	*
278		8.719	6.797	5.495	6.471	7.870	1.107 + 1	1.389 + 1	5.946 + 1	1.314 + 2	*	*
273	273.0	6.108	4.847	3.839	4.519	5.492	7.710	9.664	4.049 + 1	8.659 + 1	9.764 + 2	*
268	268.6	4.215	3.407	2.644	3.112	3.780	5.300	6.637	2.739 + 1	5.728 + 1	4.533 + 2	*
263	264.1	2.863	2.358	1.794	2.110	2.562	3.590	4.492	1.834 + 1	3.779 + 1	2.495 + 2	*
258	259.6	1.912	1.605	1.197	1.408	1.709	2.393	2.993	1.213 + 1	2.474 + 1	1.470 + 2	*
253	255.1	1.2540	1.074	7.847 − 1	9.227 − 1	1.120	1.568	1.960	7.903	1.601 + 1	8.919 + 1	*
248	250.5	8.070 − 1	7.047 − 1	5.048 − 1	5.936 − 1	7.204 − 1	1.008	1.260	5.603	1.021 + 1	5.461 + 1	*
243	245.8	5.088 − 1	4.534 − 1	3.182 − 1	3.742 − 1	4.540 − 1	6.352 − 1	7.538 − 1	3.183	6.397	3.335 + 1	6.443 + 2
238	241.2	3.139 − 1	2.856 − 1	1.963 − 1	2.308 − 1	2.801 − 1	3.918 − 1	4.896 − 1	1.960	3.931	2.016 + 1	2.846 + 2
233	236.5	1.891 − 1	1.757 − 1	1.183 − 1	1.390 − 1	1.687 − 1	2.360 − 1	2.948 − 1	1.179	2.362	1.199 + 1	1.450 + 2
273.0	273	6.107	4.847	3.839	4.518	5.492	7.709	9.668	4.048 + 1	8.658 + 1	9.759 + 2	*
267.3	268	4.015	3.246	2.518	2.963	3.599	5.047	6.322	2.604 + 1	5.433 + 1	4.722 + 2	*
261.8	263	2.597	2.139	1.627	1.913	2.324	3.255	4.075	1.660 + 1	3.409 + 1	2.182 + 2	*
256.2	258	1.652	1.387	1.034	1.216	1.476	2.067	2.592	1.045 + 1	2.126 + 1	1.231 + 2	*
250.8	253	1.032	8.835 − 1	6.456 − 1	7.592 − 1	9.214 − 1	1.289	1.613	6.490	1.311 + 1	7.158 + 1	*
245.3	248	6.323 − 1	5.521 − 1	3.955 − 1	4.650 − 1	5.643 − 1	7.895 − 1	9.872 − 1	3.961	7.969	4.199 + 1	*
239.9	243	3.798 − 1	3.385 − 1	2.375 − 1	2.792 − 1	3.388 − 1	4.740 − 1	5.926 − 1	2.373	4.763	2.456 + 1	3.809 + 2
234.6	238	2.233 − 1	2.032 − 1	1.396 − 1	1.642 − 1	1.993 − 1	2.787 − 1	3.483 − 1	1.393	2.791	1.42 + 1	1.788 + 2
229.3	233	1.283 − 1	1.192 − 1	8.026 − 2	9.434 − 2	1.144 − 1	1.600 − 1	2.001 − 1	7.996 − 1	1.601	8.084	9.154 + 1
224.1	228	7.198 − 2	6.836 − 2	4.503 − 2	5.293 − 2	6.422 − 2	8.981 − 2	1.122 − 1	4.483 − 1	8.970 − 1	4.510	4.824 + 1
	223	3.935 − 2	3.821 − 2	2.463 − 2	2.895 − 2	3.512 − 2	4.910 − 2	6.135 − 2	2.450 − 1	4.901 − 1	2.457	2.548 + 1
	218	2.092 − 2	2.078 − 2	1.309 − 2	1.539 − 2	1.867 − 2	2.611 − 2	3.261 − 2	1.302 − 1	2.604 − 1	1.304	1.329 + 1
	213	1.080 − 2	1.098 − 2	6.761 − 3	7.947 − 3	9.640 − 3	1.347 − 2	1.684 − 2	6.723 − 2	1.344 − 1	6.725 − 1	6.791
	208	5.606 − 3	5.627 − 3	3.386 − 3	3.979 − 3	4.826 − 3	6.749 − 3	8.427 − 3	3.365 − 2	6.728 − 2	3.362 − 1	3.381
	203	2.615 − 3	2.784 − 3	1.639 − 3	1.926 − 3	2.336 − 3	3.265 − 3	4.076 − 3	1.628 − 2	3.254 − 2	1.627 − 1	1.631
	198	1.220 − 3	1.334 − 3	7.646 − 4	8.986 − 4	1.090 − 3	1.524 − 3	1.902 − 3	7.593 − 3	1.518 − 2	7.590 − 2	7.597 − 1
	193	5.472 − 4	6.138 − 4	3.423 − 4	4.023 − 4	4.882 − 4	6.828 − 4	8.530 − 4	3.406 − 3	6.810 − 3	3.404 − 2	3.405 − 1
	188	2.353 − 4	2.710 − 4	1.472 − 4	1.736 − 4	2.099 − 4	2.936 − 4	3.668 − 4	1.465 − 3	2.928 − 3	1.464 − 2	1.464 − 1
	183	9.672 − 5	1.144 − 4	6.051 − 5	7.111 − 5	8.679 − 5	1.207 − 4	1.508 − 4	6.020 − 4	1.204 − 3	6.016 − 3	6.016 − 2

‡ At saturation only.
* Atmospheric saturation is not possible at this ambient temperature and pressure.

mass per unit area, it is commonly reported as the height at which the liquid would stand if it were all condensed in a vessel of the same unit cross-section as the air column. For example, a height of 1 cm corresponds to 1 gm/cm² of precipitable water.

A programmable computer can use a set of equations or algorithms to generate the numbers given in Table 16-1. Where T_D represents dew point (K), T_F frost point (K), e vapor pressure (mb), p atmospheric pressure (mb), ρ, absolute humidity (g/m³) and r mixing ratio (g/kg), then in terms of the frost point, the vapor pressure is approximated by

$$\log_{10} e = -2485.0/T_F + 3.5665 \log_{10} T_F - 0.0032098\, T_F + 2.0702 \quad (16.1)$$

In terms of the dew point, the vapor pressure is approximated by

$$\log_{10} e = -2949.1/T_D - 5.0281 \log_{10} T_D + 23.832 \quad (16.2)$$

In terms of the vapor pressure and the air temperature T(K), the absolute humidity (g/m^3) is approximated by

$$\rho_v = 216.68\ e/T\ . \quad (16.3)$$

This will be the same as in Table 16-1 for saturated air only.

In terms of the vapor pressure (mb) and the atmospheric pressure (mb), the mixing ratio (g/kg) is approximately by

$$r = 621.97\ e/(p-e)\ . \quad (16.4)$$

16.1.1 Surface

Water vapor at the earth's surface has a range of more than three orders of magnitude. The maximum amount of water vapor that the atmosphere can contain is regulated by its temperature. However, the amount actually present in the atmosphere is a complex function of various atmospheric and geophysical features. The highest value of atmospheric water vapor, observed in the Persian Gulf, is a dew point of 307 K, translating to 38 g/m^3 of water vapor or a mixing ratio of about 35 g water vapor per kg dry air. For extremely cold air (213 K) the moisture content, at best, can be no more than 0.02 g/m^3, or, at sea level, 0.01 g water vapor per kg dry air. Even this amount, however, can be important in the transmission of infrared radiation over long path lengths.

Information on the distribution of water vapor in the northern hemisphere was collected in the *Atmospheric Humidity Atlas* [Gringorten et al., 1966], which maps surface mixing ratios and dew points from the 5th to the 95th percentile for the midseason months. Two atlases, for the northern and southern hemispheres [Crutcher and Meserve, 1970; Taljaard et al., 1969] have been published that include mean monthly dew points and standard deviations of the monthly averages.

Figure 16-1, taken from the *Atmospheric Humidity Atlas*, presents 95th percentile surface dew points in July to illustrate high values of dew points over the northern hemisphere. The peak values are found in the Persian Gulf above exceptionally warm waters. Over deserts the water vapor content is lower, and coupled with high temperatures, the relative humidity drops to as low as 5% or less. Figure 16-2, also from the *Atlas*, presents 5th percentile surface dew points in January to illustrate low moisture content during the coldest time of the year. As expected, the dew point is lowest in polar and subpolar regions, but it is also low on high mountains. On the other hand, relative humidity can be high since it would not require much moisture to saturate the air at low temperatures.

World-wide extremes of humidity and the diurnal cycles of humidity in several distinctive climates are presented in a Department of Defense Military Standard, *Climatic Extremes for Military Equipment* [U.S. DOD, 1973] and are further discussed in the background studies document [Sissenwine and Cormier, 1974]. Table 16-2 gives details of the diurnal cycle of humidity, temperature, and solar insolation associated with the upper 1% dew point (304 K) in a hot coastal climate. In this kind of weather the dew point remains nearly constant throughout 24 hours, while the relative humidity changes from nearly 90% to 60%, respectively, as the temperature climbs from 304 K in the morning to 314 K in the afternoon. Table 16-3 presents a repeated diurnal cycle for an extreme moisture-laden case of 30 days duration. The conditions are maritime, with temperatures 300 K to 303 K, dew points 299 K to 301 K and relative humidity at 91% to 94%. Table 16-4 presents a diurnal cycle that has high relative humidity associated with high temperature. This case differs from that in Table 16-2 by having less absolute humidity, less insolation and lower temperature, but greater relative humidity throughout the 24 hours. Relative humidity varies from 75% to a more persistent 100% as the temperature varies from an afternoon maximum of 308 K to a nighttime minimum of 299 K. Table 16-5 depicts the persistent high relative humidity of a steaming jungle or rain forest. While the temperature

Table 16-2. Diurnal cycle of dew point, temperature, relative humidity, and solar insolation associated with the 1% high operational extreme of moisture content (dew point 304 K) in a coastal desert location.

Time (LST)	Dew Point (K)	Temperature (K)	R.H. (%)	Insolation L/h
0000	302	304	88	0
0300	302	304	88	0
0600	302	304	88	4
0900	304	310	73	68
1200	304	313	63	93
1500	304	314	60	61
1800	304	310	69	1
2100	303	306	85	0

Table 16-3. A repeated diurnal cycle during a month of sustained extreme sustained moisture content in August at Belize City, Belize.

Time (LST)	Dew Point (K)	Temperature (K)	R.H. (%)	Insolation L/h
0000	300	301	91	0
0300	299	301	92	0
0600	299	300	94	4
0900	300	302	92	54
1200	301	303	91	83
1500	301	302	91	68
1800	300	302	91	20
2100	300	302	91	0

Figure 16-1. Dew point (K-273) at the surface in July, 95th percentile [Gringorten et al., 1966].

consistently remains moderate at 297 K, the relative humidity persists above 95%. Such conditions may prevail several days a month throughout the year.

16.1.2 Troposphere

Climatic humidity data are relatively plentiful for altitudes up to 7 or 8 km (approximately 400 mb). Information on the distribution of water vapor at these levels in the northern hemisphere is provided in the *Atmospheric Humidity Atlas* [Gringorten et al., 1966]. Dew points aloft are mapped at 850, 700, 500 and 400 mb in percentiles from 5th to 95th for mid-season months. Also, two northern and southern hemispheric atlases [Crutcher and Meserve, 1970; Taljaard et al., 1969] have been published showing mean monthly dew points and standard deviations of the monthly averages for altitudes up to 500 mb.

Aloft, relative humidities of 100% can occur at any altitude below the tropopause over the entire globe. How-

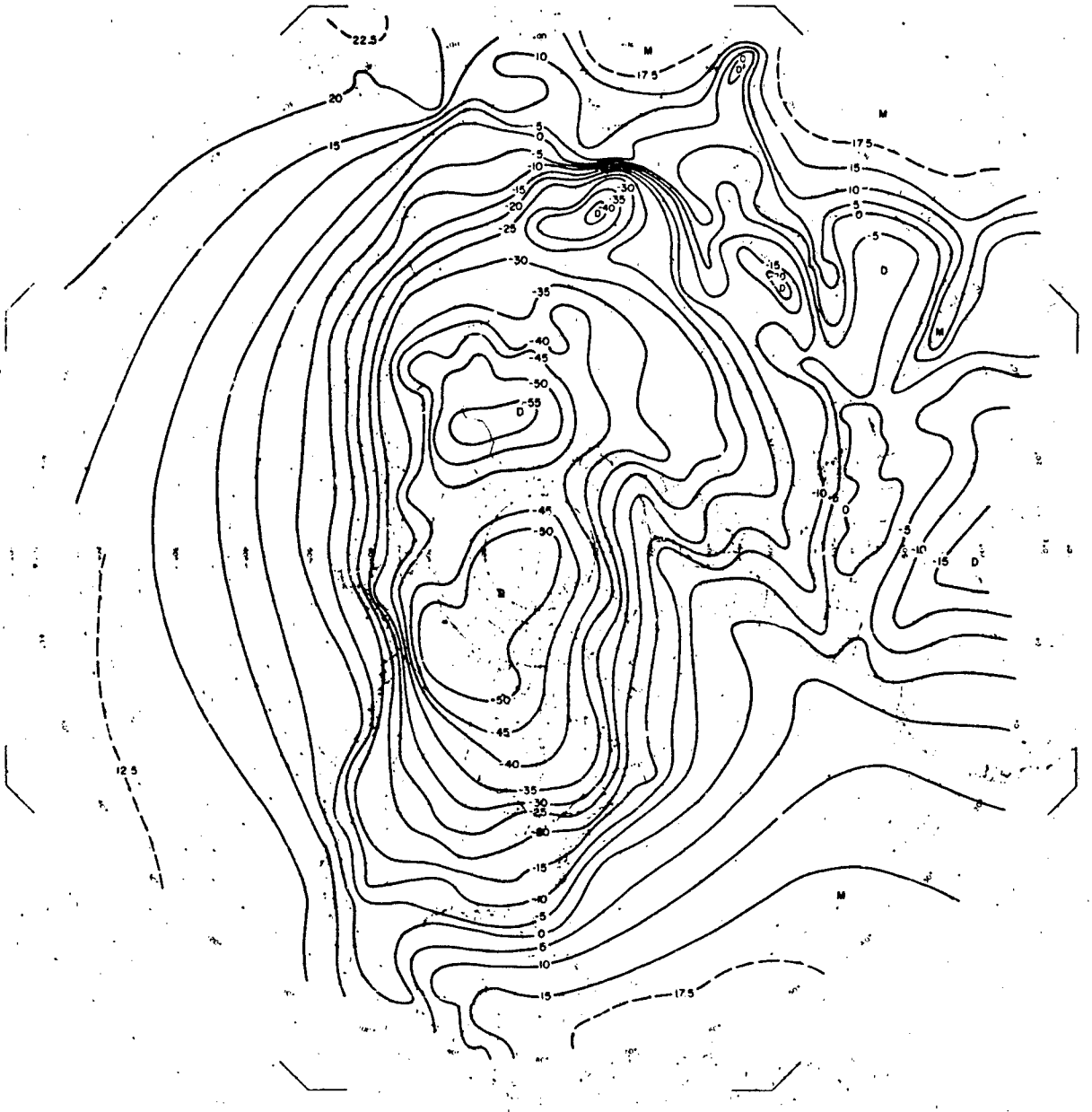


Figure 16-2. Dew point (K-273) at the surface in January, 5th percentile [Gringorten et al., 1966].

ever, the greatest moisture content is imbedded in the monsoonal flow over northern India. Table 16-6 presents high dewpoints with 1%, 5%, 10%, and 20% estimated probabilities in July from 1 to 8 km over northwestern India. Table 16-7 presents the low frost points with 1%, 5%, 10%, and 20% estimated probabilities in January from 1 to 8 km in the Canadian Northwest Territories.

Water vapor measurements above radiosonde humidity altitudes (roughly 8 km) are not made on a regular basis because of a lack of radiosonde instruments capable of sensing low concentrations. Consequently, it is not possible at this time to describe spatial and temporal variations of water vapor in the upper troposphere in a manner comparable to that given for altitudes below 8 km.

For altitudes from 8 km up to the tropopause, atmospheric moisture can be estimated using information obtained from several different experimental sensors. In the upper troposphere the median mixing ratio appears to decrease approximately logarithmically with increasing altitude up to the tropopause. As a result, median dewpoints, given at 400

Table 16-4. A diurnal cycle associated with temperature and high relative humidity in the rainy tropics.

Time (LST)	Dew Point (K)	Temperature (K)	R.H. (%)	Insolation L/h
0000	300	300	100	0
0300	299	299	100	0
0600	299	299	100	4
0900	300	304	82	54
1200	302	307	75	83
1500	302	308	74	68
1800	302	305	82	20
2100	300	301	95	0

Table 16-5. A typical diurnal cycle in a tropical jungle with high relative humidity. (Such conditions may occur several days a month throughout the year.)

Time (LST)	Dew Point (K)	Temperature (K)	R.H. (%)
0000	297	297	100
0300	297	297	100
0600	297	297	100
0900	296	297	95
1200	296	297	95
1500	296	297	95
1800	296	297	95
2100	297	297	100

Table 16-6. High absolute humidity (dew point) extremes, 1 to 8 km aloft (typical of northern India in July).

Altitude (km)	Dew Point (K)			
	1% Probable	5%	10%	20%
1	302	300	299	298
2	297	295	294	293
4	289	286	283	282
6	276	273	272	271
8	265	262	261	260

Table 16-7. Low absolute humidity (dew point) extremes, 1 to 8 km aloft (typical of Canadian Northwest Territories in January).

Altitude (km)	Frost Point (K)			
	1% Probable	5%	10%	20%
1	223	231	233	236
2	222	228	231	234
4	217	221	222	224
6	208	212	213	214
8	202	202	204	205

mb (7 to 8 km) in the *Atmospheric Humidity Atlas* [Gringorten et al., 1966], can be converted to median mixing ratios and extended upward logarithmically to medians of 2 to 4 ppmm (parts per million by mass) at the appropriate tropopause height. The altitude of the tropopause ranges from 16 or 17 km in the tropics to 9 or 10 km near the pole [Cole and Kantor, 1978]. Estimated median mixing ratios at various latitudes are shown in Table 16-8.

World maps of the mean precipitable water above the surface and above 850, 700, and 500 mb have been published for the midseason months [Bannon and Steele, 1960]. More recently, reports on "Precipitable Water Over the United States" were compiled by the National Weather Service [Lott, 1976; Ho and Riedel, 1979]. Based on radiosonde observations, the range of the maximum precipitable water content over the contiguous United States is about 1 cm over the Rocky Mountains to nearly 7 cm along the Gulf of Mexico in the surface-to-400 mb layer. Table 16-9 presents mean monthly maximum precipitable water content for Key West, Florida and Lander, Wyoming.

16.1.3 Stratosphere and Mesosphere

Water vapor measurements in the stratosphere and mesosphere are not made at regular intervals because standard instrumentation is not available for use at these altitudes. Measurements at very low concentrations typical of the stratosphere can be made only with research quality instruments. Such sensors include frost point hygrometers, optical devices (such as spectrometers, radiometers), cryogenic air samplers and absorption devices. Roughly two dozen investigators have each provided from a few to over 100 observations during the last 35 years, with the bulk of stratospheric humidity measurements coming since 1965.

The earliest measurements, made by British aircraft [Brewer, 1949], provided evidence that the tropopause acts as a cap to water vapor, preventing it from penetrating into the stratosphere. General atmospheric circulation theory indicates that tropospheric air passes into the stratosphere over equatorial regions from where it moves northward. It sinks back into the troposphere at higher latitudes, closing the cycle. Tropical tropospheric air following this route into the stratosphere would have to pass through the very cold tropopause of the tropics (193 K) where the water vapor density would be forced to extremely low concentrations due to condensation and sublimation. As this air continues upward from this level, the mixing ratio in the stratosphere must remain constant if there is neither source nor condensation/sublimation of water vapor. The mixing ratio that most closely depicted these British findings is 2 mg per kg which may be interpreted as 2 ppmm. Some investigators report their measurements in terms of volume mixing ratio, denoted as ppmv. The relationship between ppmv and ppmm is ppmv = 1.607 ppmm. Other possible sources that have been suggested for water vapor to be introduced into the

Table 16-8. Estimated distribution of median mixing ratios in the upper troposphere.

Altitude (km)	15°N		30°N		45°N		60°N		75°N	
	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
	(ppmm)		(ppmm)		(ppmm)		(ppmm)		(ppmm)	
400 mb*	433*	1050*	295*	722*	155*	540*	90.1*	401*	63.4*	295*
7	—	—	—	—	148	—	57.5	--	34.2	—
8	346	1010	213	510	38.5	368	11.3	48.7	8.3	24.0
9	205	513	130	239	9.8	136	2.0	13.2	2.0	2.0
10	122	250	80.0	113	2.5	50		2.0		
11	71.2	119	49.0	52.5		18.3				
12	42.0	57.2	29.8	24.6		6.7				
13	24.6	27.5	18.3	11.5		2.5				
14	14.5	13.2	11.2	5.4						
15	8.6	6.3	6.8	2.5						
16	5.3	3.0	4.1							
17	3.0		2.5							

*400-mb mixing ratios calculated from Gringorten et al. [1966].

stratosphere are the transport of tropospheric air upward through the overlapping tropical and polar tropopauses, cumulonimbus clouds penetrating the tropopause, and chemical reactions such as the oxidation of methane. The extent to which such mechanisms influence the stratospheric humidity regime continues to be debated. However, the combined effects of the Hadley cell circulation and tropical thunderstorms, both of whose inputs of moisture into the stratosphere must penetrate the tropical "cold trap," are responsible for the bulk of stratospheric mixture. In fact, most investigators agree that all other sources are smaller by a factor of 5 to 10.

In the early 1960s there was a significant controversy concerning the amount and distribution of water vapor in the stratosphere [Gutnick, 1961]. Much uncertainty continues, but the extent of uncertainty has narrowed. The bulk of credible observations of stratospheric mixing ratio now range from about 0.5 to 5 ppmm. The stratospheric water vapor measurements made over the previous decade or so have been summarized in two recent articles [Harris, 1976 and Penndorf, 1978]. The conclusions are that the mean midlatitude stratospheric mixing ratio is 2.5 ± 0.7 ppmm and is nearly constant with altitude up to at least 30 km. There is some evidence that from 30 km the mean may gradually increase to 3 ppmm at 45 km. This increase could be the result of an increase in water vapor due to the oxidation of methane in the middle atmosphere. There is also limited evidence for the existence of weak, shallow layers of water vapor with mixing ratios up to about 4 ppmm occurring in 1- to 2-km layers at altitudes up to about 30 km.

Based on three independent series of field measurements, an annual cycle for stratospheric water vapor has been suggested, in which the amplitude of the seasonal variation is largest in the lowest part of the stratosphere, about 0.4 ppmm at 15 km, and decreases with altitude to about 0.1 ppmm above 20 km. The phase of the annual cycle also changes with altitude, with the maximum occurring in November at 15 km and in April at 20 km. This annual cycle may be related to the annual variations in the temperature and altitude of the tropical tropopause.

In the lower stratosphere (below 20 km) there seems to be a small but discernible latitudinal variation in mixing ratio, with higher values, about 3 ppmm, occurring near the equator and decreasing with latitude to somewhat less than 2 ppmm near 45° latitude in both hemispheres. The lack of data in polar regions does not allow estimates of variation at higher latitudes.

There have been even fewer observations of water vapor in the mesosphere than in the stratosphere; consequently, these distributions and variations continue to be controversial. However, there does seem to be reasonably consistent hypotheses of the general envelope of the vertical profile of mixing ratio from the stratopause to the mesopause. Deguchi and Muhleman [1982] compared their ground-based radio telescope measurements to rocket measurements by Scholz

Table 16-9. Mean monthly maximum precipitable water content—two extremes.

Station	Month	Layer, Surface to 850 mb	700 mb	500 mb	400 mb
Key West, Fla.	Sep	2.7 cm	4.2 cm	5.2 cm	5.4 cm
Lander, Wyo.	Jan	—	0.4	0.7	0.8

et al. [1970], Ehhalt et al. [1975] and Rogers et al. [1977]; airborne measurements by Farmer et al. [1980] and Waters et al. [1977], and current photochemical models by Crutzen [1974] and Allen et al. [1981]. A concensus of these observations and models provides a midlatitude mesospheric water vapor profile of 4 to 5 ppmm at 50 km decreasing logarithmically with altitude to roughly 1 ppmm at 90 km. Such a profile is consistent with microphysical observations of noctilucent clouds. An experiment in which noctilucent cloud particles, captured by a rocket at the 80 km mesopause level included ice [Soberman, 1963, and Michaels, 1965], implies the existence of water vapor through a layer at this high altitude. Such clouds are not found at tropical latitudes and rarely found at middle latitudes where temperatures are seldom colder than 173 K [Cole and Kantor, 1978] at this level. They are normally seen at high latitudes in the summer where extremely cold mesopause temperatures prevail. During one experiment [Anonymous, 1963] in which such clouds were physically sampled, the temperature observed was 130 K. During another sampling, when the clouds had dissipated, a temperature of 153 K was measured. Assuming the temperature and pressure structure of the *Air Force Reference Atmospheres* [Cole and Kantor, 1978], frost points of 130 K and 153 K at 80 km convert to saturation mixing ratios of about 0.01 ppmv and 6 ppmv, respectively.

It is interesting to note that there is a region of the middle atmosphere, roughly 40 to 65 km, in which the combined temperature and pressure structure does not support the existence of clouds [Schilling, 1964]. The relatively warm temperature structure at these altitudes requires a saturation vapor pressure greater than the total atmospheric pressure—an obvious impossibility. This condition is denoted by the asterisk entries in Table 16-1.

16.2 PRECIPITATION

Knowledge of the frequency of occurrence of given rates of precipitation and associated vertical distributions of precipitation particles is important to the design and operation of many types of equipment. It is a critical input for evaluating and deploying systems which employ surveillance radar, and for the selection of microwave frequencies used in communicatins systems. The frequency and intensity of rainfall must also be considered in the design and operation of helicopter rotor blades, and leading edges of aircraft and missiles which are subject to erosion by precipitation particles. Water ingested by jet engines can cause flame-outs and penetrate protective coverings of exposed electronic and mechanical material. Precipitation in the form of hail is a serious hazard for aircraft, whereas snow loads and ice accretion on surface structures can destroy towers, radomes, and buildings. This section contains examples of the type of precipitation information available.

16.2.1 Surface Rates of Rainfall

Frequency distributions of various rainfall rates from actual data are available for only a small number of meteorological stations. Precipitation data for most areas of the world are limited mainly to average monthly and annual totals, and to the number of days on which precipitation fell. Clock hourly precipitation data (totals on the hour every hour) are available for numerous stations in the United States and Europe, but for only a few stations elsewhere. Monthly tabulations of the amounts of precipitation which fall in 5 to 180 min periods are also available for National Weather Service first-order stations in the United States and its island possessions. Special 1-, 2-, and 4-min rainfall rates have been taken at a number of locations in Southwest Asia, Europe and the United States.

These data have been used by scientists to develop mathematical models for estimating the frequency of occurrence of various duration-rainfall rates at locations where only routine rainfall observations are available. Some selected methods for estimating clock hour and instantaneous (1 min) rainfall rates are presented in the following sections.

16.2.1.1 Clock Hourly Rates. Several researchers have used available clock-hour rainfall data to derive methods for estimating the number of hours that specific rates occur at other locations. In 1965, Cole et al., [1965] presented three linear equations for determining the frequency of 3 clock-hour precipitation rates (1.5, 3.0, and 4.6 mm/h) vs an index of rainfall intensity. The index is the average per day of measurable precipitation obtained by dividing the total annual precipitation by the number of days with measurable precipitation of ≥0.25 mm. Winner [1968] elaborated on this and other earlier studies to develop techniques for estimating clock-hours of rainfall at rates of trace, 0.25, 0.5, 2.5, 6.4, 12.7, 25.4 and 50.8 mm/h, based on data from 123 National Weather Service stations with 5 to 10 year periods of record. The technique incorporates a moisture index, Y, where

$$Y = 100\left[\frac{P/25.4}{PE} - 1\right] \qquad (16.5)$$

P is the mean annual rainfall in mm, and PE is the mean annual potential evapotranspiration determined by the method developed by Thornthwaite et al. [1958].

Figures 16-3, 16-4, and 16-5 are nomograms for estimating the mean annual clock-hours per year that a rainfall rate of 6.4, 12.7, and 25.4, mm/h is equaled or exceeded in temperate latitudes. Potential evapotranspiration data for Equation (16.5) can be calculated using Thornwaite et al., [1958]. It is desirable to use the same period of record for all variables in the nomograms. If these publications are not available, clock-hour estimates of more limited value can be made by assuming Y to be zero.

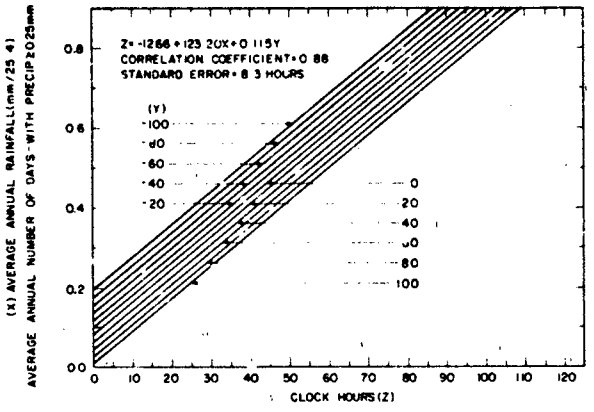


Figure 16-3. Nomogram for determining average number of clock hours per year a rainfall rate of 6.4 mm/h is equaled or exceeded.

The ratio of the average annual rainfall to the number of days with precipitation ≥0.25 mm for a location is then obtained from climatic data. Knowing this value and the value for Y, the number of clock-hours can be determined from the nomogram.

The technique described above provided poor results when applied on independent data from tropical locations. Therefore, Wiriner [1968] developed an alternate method for use in the tropics using data from 32 stations in the Panama Canal Zone. The best results were obtained from a linear correlation of the annual rainfall vs the number of clock-hours at specified rates. Figure 16-6 presents the clock-hours per year that rates of 6.4, 12.7, 19, and 25.4 mm/h are equaled or exceeded in the tropics.

16.2.1.2 Instantaneous Rate. For most applications, the distribution of rainfall rates for short periods, usually for 1 min, are of the greatest importance since these can greatly

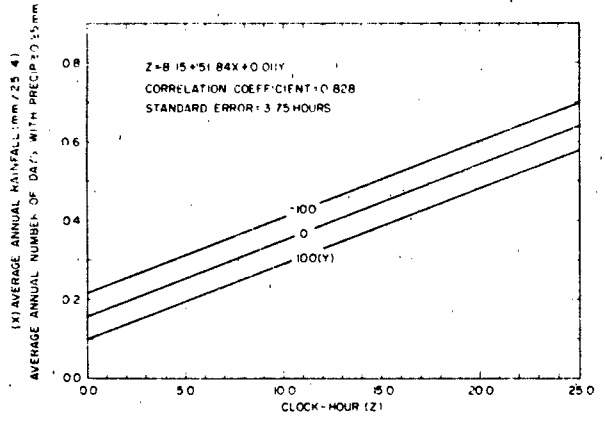


Figure 16-4. Nomogram for determining average number of clock hours per year a rainfall rate of 12.7 mm/h is equaled or exceeded.

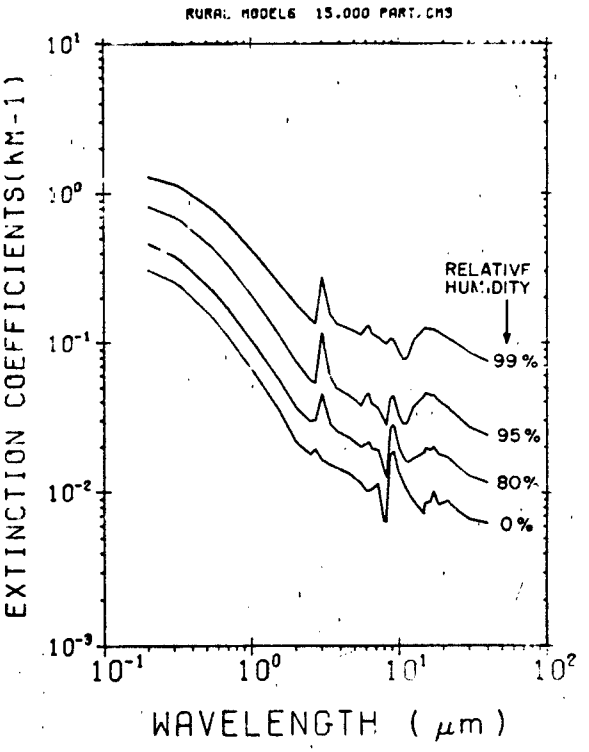


Figure 16-5. Nomogram for determining average number of clock hours per year a rainfall rate of 25.4 mm/h is equaled or exceeded.

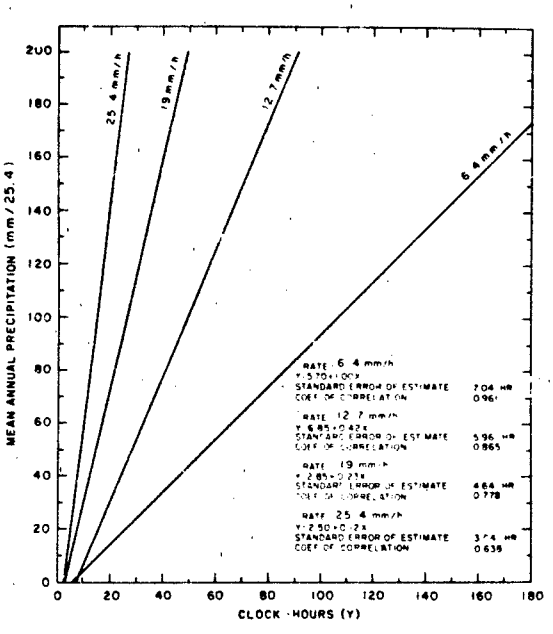


Figure 16-6. Nomogram for use in tropical regions to determine mean number of clock hours occurrence of a specified precipitation rate when the mean annual precipitation amount is known

Table 16-10. Percent contribution of instantaneous precipitation rate to clock-hour precipitation rates for Urbana, Illinois (annual).

Clock-hour Rates mm/h*	Instantaneous Rates (mm/h)* 0.00 – 0.76	1.0 – 2.3	2.5 – 6.1	6.4 – 12.4	12.7 – 25.1	25.4 – 50.5	50.8 – 126.7	127 – 254	>254	Tot No Clock-h
Trace	96.02	3.43	0.51	0.04	0.0	0.0	0.0	0.0	0.0	576
0.25	87.56	10.38	1.96	0.06	0.04	0.0	0.0	0.0	0.0	114
0.5 – 2.3	59.17	29.16	10.17	1.16	0.25	0.08	0.01	0.0	0.0	581
2.5 – 6.1	27.34	19.24	37.58	11.82	2.82	1.05	0.15	0.0	0.0	179
6.4 – 12.4	23.50	11.28	23.01	23.14	12.16	5.33	1.58	0.0	0.0	61
12.7 – 25.1	20.45	8.53	14.74	16.28	16.99	13.53	8.05	0.83	0.0	26
25.4 – 50.5	6.67	5.56	8.89	13.33	19.44	20.00	25.56	0.56	0.0	3
50.8 – 76	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
76.2 – 101.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
101.6 – 126.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
≥127	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0

*This work was originally presented in English units. Direct conversion to metric is the reason for small gaps between each range of rates.

exceed the hourly rate. A 1-min rainfall rate is generally referred to as being "instantaneous." Bussey [1950] found that about 20% of the time during an hour the instantaneous rate was zero, or a very low amount called a "trace." About 35% of the time the hourly rate was exceeded, and it was common for the hourly rate to be exceeded by 5 or 6 times for a few minutes.

Davis and McMorrow [1976] developed tables of clock-hour rate versus instantaneous rate for 6 locations, and clock-hour rate versus 4-min rate for 7 locations from original data. They indicate that these can be used to predict the 1- or 4-min rate for any region having a similar precipitation regime as determined from available clock-hour data. A variety of climates, including the wet tropics, is represented. Tables 16-10 and 16-11 from this study give the percent contribution of the instantaneous rate to clock-hour rate for Urbana, Illinois and Miami, Florida.

A more direct method for determining instantaneous rainfall rates was developed by Jones and Sims [1978]. They used instantaneous rainfall rate data for 9 observation sites, plus 4-min rainfall rate data at another 10 sites to evaluate the frequency of rainfall rates based on a world rain climate classification used by Herbstritt [1973]. Figure 16-7 from the Jones and Sims report shows the climatic classifications and the data stations. The data for stations within each of four of the five climatic zones were summarized and are presented in Figure 16-8. Although the curves in this figure are intended to represent instantaneous rainfall rates, mixing of the 1- and 4-min data has the net effect of giving a slightly low estimate for less than 10^{-1} exceedance frequency.

Table 16-11. Percent contribution of instantaneous precipitation rate to clock-hour precipitation rates for Miami, Florida (annual).

Clock-hour Rates mm/h*	Instantaneous Rates (mm/h)* 0.00 – 0.76	1.0 – 2.3	2.5 – 6.1	6.4 – 12.4	12.7 – 25.1	25.4 – 50.5	50.8 – 126.7	127 – 254	>254	Tot No Clock-h
Trace	93.55	4.60	1.49	0.33	0.04	0.0	0.0	0.0	0.0	92
0.25	85.31	9.44	4.01	1.17	0.06	0.0	0.0	0.0	0.0	27
0.5 – 2.3	65.67	19.97	11.13	2.30	0.66	0.26	0.01	0.0	0.0	170
2.5 – 6.1	39.21	16.40	29.82	9.84	2.21	2.14	0.38	0.0	0.0	74
6.4 – 12.4	29.81	11.52	18.64	22.42	9.47	5.87	2.23	0.04	0.0	44
12.7 – 25.1	34.04	6.32	16.67	10.61	10.00	11.14	10.18	1.05	0.0	19
25.4 – 50.5	15.50	12.17	7.50	12.33	13.17	13.67	20.83	4.83	0.0	10
50.8 – 76	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
76.2 – 101.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
101.6 – 126.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
≥127	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0

*This work was originally presented in English units. Direct conversion to metric is the reason for small gaps between each range of rates.

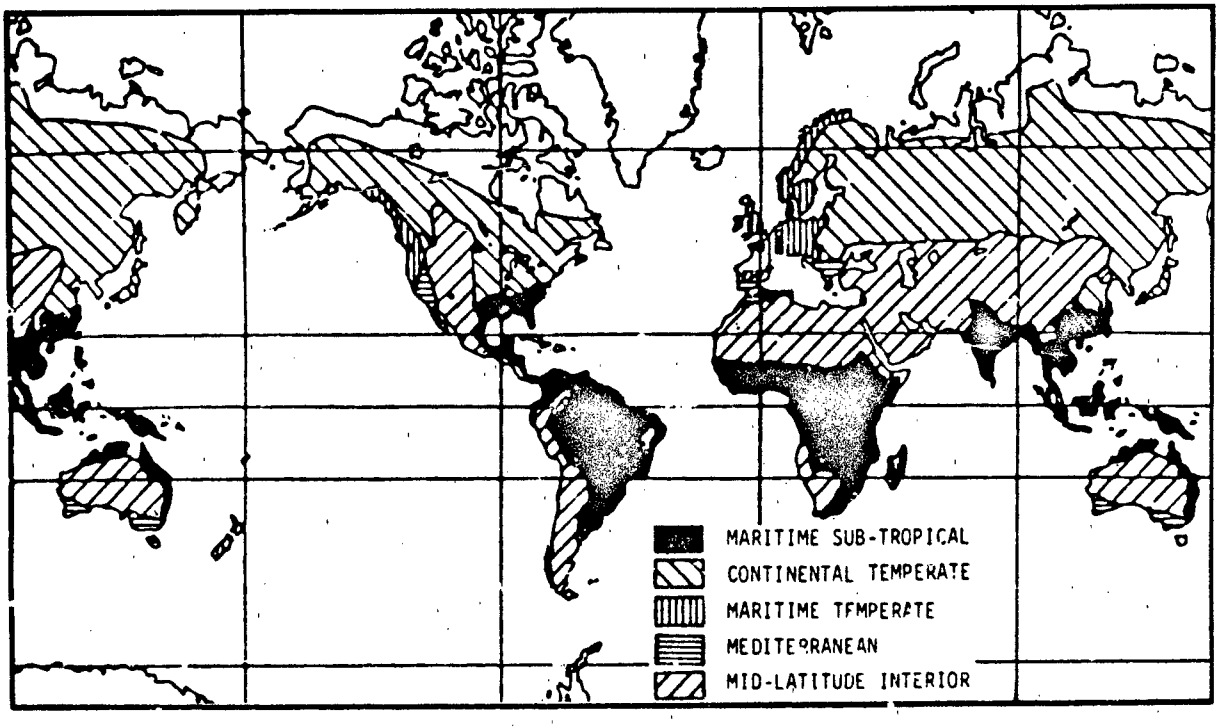


Figure 16-7. Rain climates of the world [after Herbstritt, 1973].

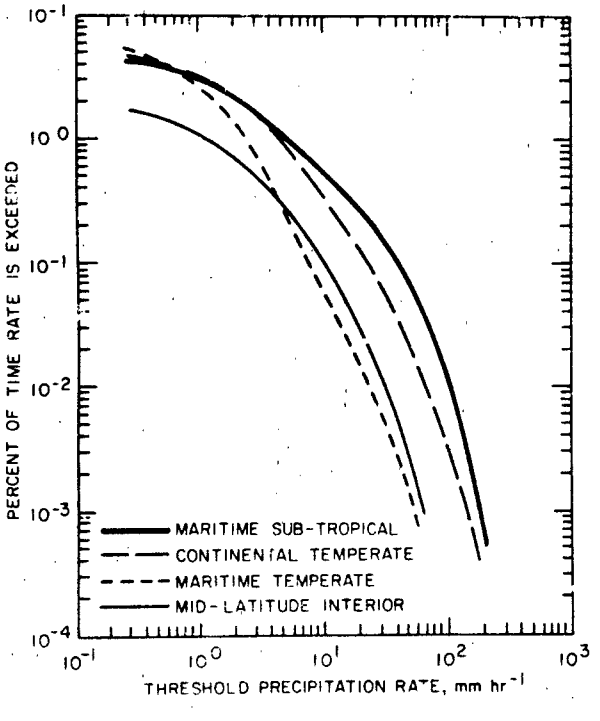


Figure 16-8. Average rainfall rate—frequency relationships for four rain climates.

A more comprehensive model for determining instantaneous precipitation rates exceeded for a specified percent of a year was developed by Crane [1980]. The world was divided into 8 regions, as shown in Figure 16-9, on the basis of total rain accumulation, and maps showing the number of thunderstorm-days published by Landsberg [1974]. Additional guidance was taken from the Köppen world climate classification. Boundaries were adjusted to accommodate variations in terrain, predominant storm type and motion, general atmospheric circulation, and latitude. Satellite and precipitation frequency data were used to extend the rain climate regions over the ocean; however Feldman [1979] presents an elaborate expansion of the model for ocean areas. The U.S. is spanned by 5 regions; but one of them is divided into 3 sub-regions as shown in Figure 16-10.

The measured instantaneous rain rate distributions available for each of the 7 regions and 3 sub-regions were pooled to construct median rain rate distributions. These are presented in Table 16-12. No data were available for Region A, and the author does not explain how its distribution was derived.

Tattelman and Grantham [1982] discuss sources of 1-min data and present a survey and comparison of available models for estimating 1-min rates. They found that the best estimates of annual statistics on 1-min rates are made using the Davis and McMorrow model for locations where clock-

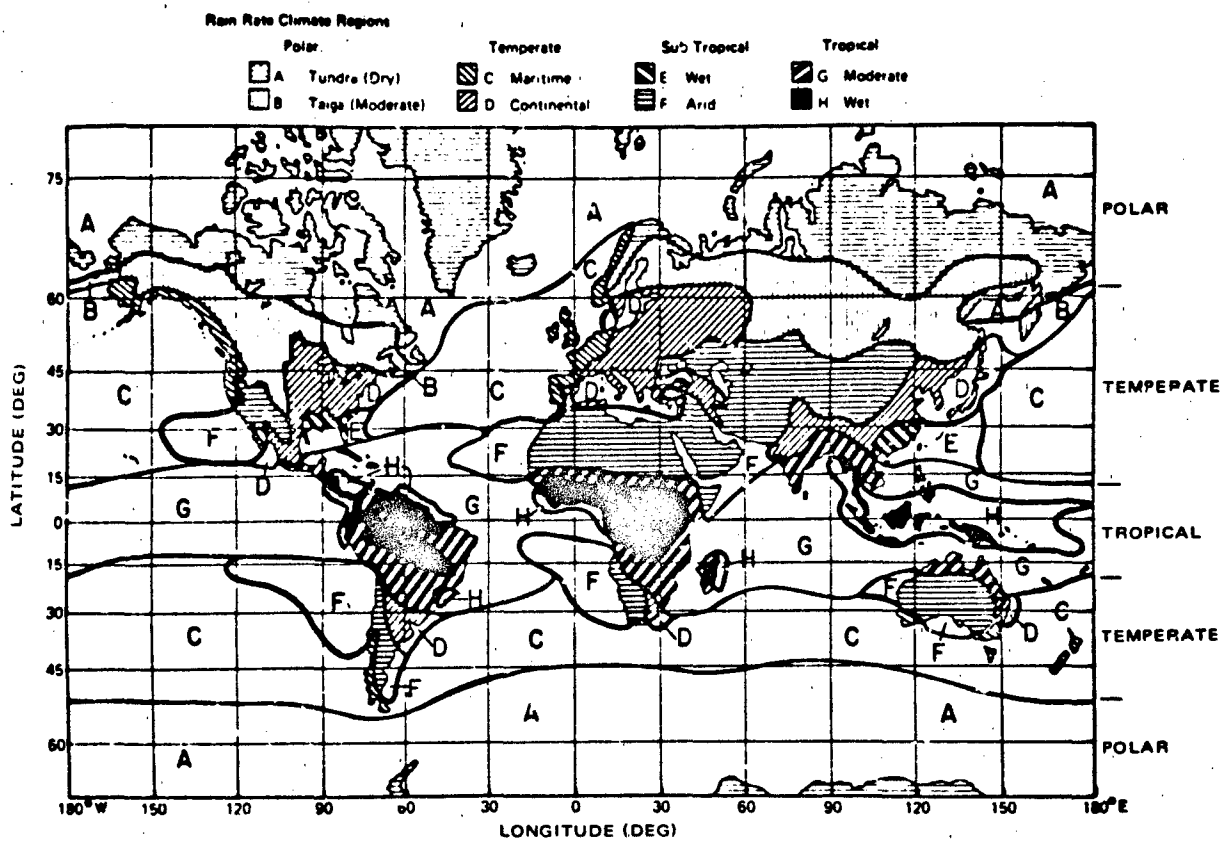


Figure 16-9. Rain rate climate regions [Crane, 1980]. (Reprinted with permission from IEEE © 1980)

hour data are available. Tattelman and Grantham also found that good estimates from clock-hour data could be made using a model developed by Huschke et al. [1982]. Huschke et al. consolidated the tables for the individual locations from Davis and McMorrow into 3 tables for all locations. Two tables represent "convective" and "stratiform" precipitation. The third table, which was found to be most accurate, simply includes all the data. Where clock-hour data are not available, Tattelman and Grantham found the Crane model produced estimates of annual 1-min rain rate statistics at least as good as any of the other available models.

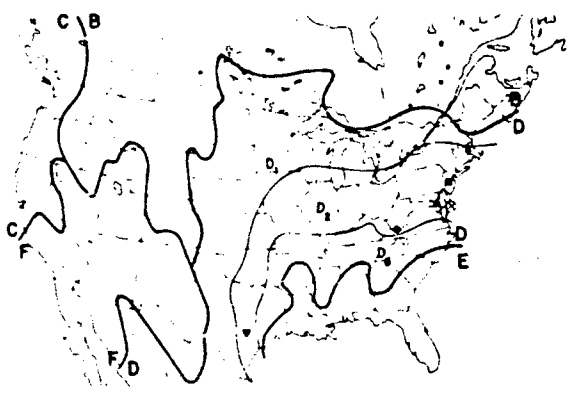


Figure 16-10. Rain rate climate regions within the United States [Crane, 1980]. (Reprinted with permission from IEEE © 1980)

A model for estimating instantaneous rainfall rates equaled or exceeded 2.0%, 1.0%, 0.5%, 0.1%, 0.05% and 0.01% of the time during a month from readily available climatological data was developed by Lenhard [1974] and improved by Tattelman and Scharr [1983].

The Tattelman-Scharr model was developed using stepwise multiple regression analysis for each of six exceedance levels (p = 2.0%, 1.0%, 0.5%, 0.1%, 0.05% and 0.01% of time during the month). The multiple regression equation is expressed by

$$R_p = A_p + B_pT + C_pI + D_p\ f(L,T) \quad (16.7)$$

where R_p is the estimated precipitation rate (mm/min) for exceedance level p, T is the monthly mean temperature (°F or 1.8 K − 459), I is a precipitation index (monthly mean precipitation divided by the monthly mean number of days

Table 16-12. Point rain rate (R_p) distribution values (mm/hr) versus percent of year rain rate is exceeded.

Percent of Year	Rain Climate Region									
	A	B	C	D_1	D_2	D_3	E	F	G	H
0.001	28	54	80	90	102	127	164	66	129	251
0.002	24	40	62	72	86	107	144	51	109	220
0.005	19	26	41	50	64	81	117	34	85	178
0.01	15	19	28	37	49	63	98	23	67	147
0.02	12	14	18	27	35	48	77	14	51	115
0.05	8.0	9.5	11	16	22	31	52	8.0	33	77
0.1	5.5	6.8	7.2	11	15	22	35	5.5	22	51
0.2	4.0	4.8	4.8	7.5	9.5	14	21	3.2	14	13
0.5	2.5	2.7	2.8	4.0	5.2	7.0	8.5	1.2	7.0	13
1.0	1.7	1.8	1.9	2.2	3.0	4.0	4.0	0.8	3.7	6.4
2.0	1.1	1.2	1.2	1.3	1.8	2.5	2.0	0.4	1.6	2.8
Number of Station Years of Data	0	25	44	15	99	18	12	20	2	11

with precipitation), and f(L,T) is a latitude-temperature term defined as

$$f(L,T) = \begin{cases} 0 & L \leq 23.5° \\ (L - 23.5)T & 23.5° < L \leq 40° \\ (40 - 23.5)T & L > 40° \end{cases} \quad (16.8)$$

A_p is the constant for exceedance level p, and B_p, C_p, and D_p are multiple regression coefficients for T, I and f(L,T), respectively, for exceedance level p.

The minimum threshold amount of precipitation to define a rainy day varies with country. Three of the most common threshold values used worldwide to define a rainy day are 0.25, 1, and 2.54 mm. The number of days per month with another frequently used threshold called a "trace" differs only slightly with the number of days equal to or greater than 0.25 mm. Results of the stepwise multiple regression analysis for each of the six exceedance levels, including regression coefficients and constants to be used in Equation (16.7), are provided in Table 16-13, 16-14 and 16-15 for indices based on rainy day threshold values of 2.54 mm, 1 mm, and 0.25 mm, respectively.

The Tattelman-Scharr model was subjectively evaluated by estimating rates at independent locations representing a wide variety of the earth's climates. Results indicted several circumstances when the model is either invalid or should be used with discretion. This occurred for very dry or cold months for which there were little or no data among the dependent stations. The model was found to be generally invalid when any of the following conditions exist for a specific month at a location:

1. $T \leq 273$ K
2. $I < 2$mm/day
3. Number of rainy days < 1.

Inconsistencies such as negative rates, or increasing rates with increasing exceedance level occasionally occur when

Table 16-13. Results of stepwise multiple regression analysis for exceedance levels p = 0.01%, 0.05%, 0.10%, 0.50%, 1.0%, and 2.0% based on a threshold value of 2.54 mm for I. (The regression coefficients are given for each independent variable.)

P	Constant A_p	T B_p	$I_{(2.54)}$ C_p	f(L,T) D_p	R	SEE (mm/min)
0.01	−0.91	2.8×10^{-2}	2.3×10^{-2}	-3.4×10^{-4}	0.83	0.43
0.05	−0.50	1.6×10^{-2}	1.9×10^{-2}	-3.1×10^{-4}	0.86	0.24
0.10	−0.31	1.1×10^{-2}	1.4×10^{-2}	-3.0×10^{-4}	0.85	0.19
0.50	−0.01	2.5×10^{-3}	5.4×10^{-3}	-1.5×10^{-4}	0.76	0.09
1.0	0.03	7.4×10^{-4}	2.9×10^{-3}	-7.6×10^{-5}	0.67	0.06
2.0	0.04	-2.0×10^{-4}	1.5×10^{-3}	-3.2×10^{-5}	0.64	0.02

Table 16-14. Results of stepwise multiple regression analysis for exceedance levels p = 0.01%, 0.05%, 0.10%, 0.50%, 1.0%, and 2.0% based on a threshold value of 1.00 mm for I. (The regression coefficients are given for each independent variable.)

P	Constant A_p	T B_p	$I_{(1.00)}$ C_p	f(L,T) D_p	R	SEE (mm/min)
0.01	−1.00	2.8×10^{-2}	3.6×10^{-2}	-2.2×10^{-4}	0.84	0.41
0.05	−0.56	1.6×10^{-2}	2.5×10^{-2}	-2.4×10^{-4}	0.88	0.23
0.10	−0.36	1.1×10^{-2}	2.0×10^{-2}	-2.4×10^{-4}	0.87	0.18
0.50	−0.03	2.4×10^{-3}	7.8×10^{-3}	-1.2×10^{-4}	0.79	0.09
1.0	0.02	6.9×10^{-4}	4.2×10^{-3}	-6.2×10^{-5}	0.71	0.05
2.0	0.04	-1.8×10^{-4}	2.0×10^{-3}	-2.6×10^{-5}	0.67	0.02

T is between 273 K and 277 K. When T is in this range, estimated rates are acceptable if they are positive and decrease with increasing exceedance level. If the estimated rates are consistent, but a negative rate is estimated for p = 2%, this may indicate that it rains less than 2% of the time during that month.

When there are between 1 and 3 rainy days in the month, the model may estimate rates for each of the exceedance levels which, when integrated, indicate a total rainfall two or more times greater than the monthly mean precipitation. Under these circumstances heavy, but infrequent, convective precipitation accounts for just about all of the precipitation in the month. Therefore, the most accurate estimates are for the two or three lowest exceedance levels, and significant rainfall might not occur more often than 0.5% of the month. These possibilities should be considered before accepting rate estimates in arid locations.

16.2.1.3 Extreme Rates. As noted in the previous sections, there are many stations in the world where precipitation is measured, but there are few outside of the U.S. and Europe where amounts are tabulated and published for time periods of less than 24 hours. In tropical regions, where the most intense precipitation occurs, data on extremes of short-duration rainfall are even rarer than for the middle latitudes. United States Weather Bureau Technical Paper No. 25 [1955] presents rainfall intensity-duration curves for 200 stations in the contiguous United States plus 1 station each in Alaska, Hawaii, and Puerto Rico. Lenhard and Sissenwine [1973] found Pensacola, Florida to have the most intense rainfall of the 203 stations; the curves for Pensacola are reproduced in Figure 16-11 for 2- to 100-yr return periods. Smoothed analyses of the data from individual stations were subsequently published by the National Weather Service because the variable nature of rainfall can result in substantial differences in rainfall intensity statistics between stations separated by a relatively short distance [for example, Frederick et al., 1977]. A return period represents the time interval that an event would be expected to occur at least once. Mathematically, a return period of n years is the reciprocal of the probability of the event occurring in one year. If we denote the probability of the event not occurring in one year by P, the P = (2 − 1/n). The probability that it will not happen in n years is P^n. That is, the probability of the event occurring at least once in its return period approaches 63% asymptotically as n increases. Comparision of the these curves with estimates of world-wide extremes using the model presented in the Lenhard and Sissenwine study indicate that, except for tropical mountainous areas with orographically enhanced rainfall, rainfall intensities at Pensacola are comparable to those at some of the more extreme locations in the world. This is not surprising, since thunderstorms occurring in tropical maritime air over the United States are considered as severe as most places in the world.

The most extreme rainfall intensities that have ever been recorded in the world for time periods of 1 minute to 1 year are shown in Figure 16-12, with an envelope of their extremes. Data used were obtained from Riordan [1974].

Table 16-15. Results of stepwise multiple regression analysis for exceedance levels p = 0.01%, 0.05%, 0.10%, 0.50%, 1.0%, and 2.0% based on a threshold value of 0.25 mm for I. (The regression coefficients are given for each independent variable.)

P	Constant A_p	T B_p	$I_{(0.25)}$ C_p	f(L,T) D_p	R	SEE (mm/min)
0.01	−1.00	2.8×10^{-2}	4.2×10^{-2}	-2.2×10^{-4}	0.85	0.41
0.05	−0.56	1.6×10^{-2}	3.0×10^{-2}	-2.3×10^{-4}	0.88	0.22
0.10	−0.36	1.1×10^{-2}	2.4×10^{-2}	-2.3×10^{-4}	0.88	0.17
0.50	−0.03	2.3×10^{-3}	1.0×10^{-2}	-1.2×10^{-4}	0.82	0.08
1.0	0.01	5.6×10^{-4}	6.0×10^{-3}	-5.6×10^{-5}	0.77	0.05
2.0	0.03	-2.3×10^{-4}	2.8×10^{-3}	-2.4×10^{-5}	0.74	0.02

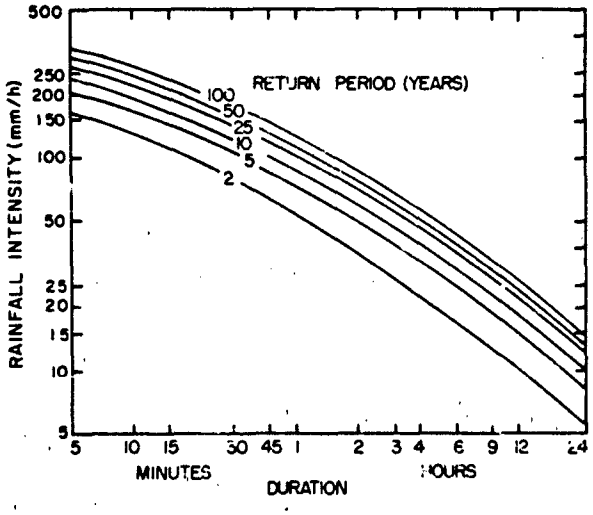


Figure 16-11. Rainfall intensity-duration-frequency curves for Pensacola, Florida.

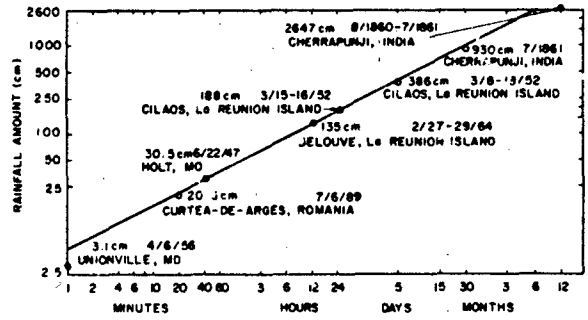


Figure 16-12. World record rainfall values.

16.2.1.4 Duration of Rainfall Rates. The question of how long a given rate is likely to persist was investigated by Sims and Jones [1973]. They derived duration tables for Miami, Fla.; Majuro Island; Island Beach, N.J.; Franklin, N.C.; and Urbana, Ill. based on an analysis of at least one year of 1-min rainfall rates at each location. The tables present the conditional frequency distribution of occurrences of various durations from 1 to 90 minutes for rainfall rates from 0.2 to 3.0 mm/min. Table 16-16 presents the results for Miami, Fla. based on 1 year of data. The table indicates, for example, that when it was raining the rate exceeded 0.8 mm/min, it continued at that rate or greater for at least 4 min 52.1% of the time. It equaled or exceeded that rate continuously for 25 min only 1.1% of the time. Note that the table does not indicate how often various rates occur, but how long various rates continued at that rate or higher without dropping to a lower rate.

16.2.1.5 Rainfall Rates Along Lines. Jones and Sims [1971] present data for 3 rain gage networks, one near Miami and two in England. Data for a larger network in central Illinois are presented by Sims and Jones [1973]. Average intensities along paths of various lengths, based on 2-min rates, were derived for all the networks. The cumulative frequencies of average rates for 2-gage (5-km), 8-gage (33-km), and 14-gage (62-km) north-south lines in the central Illinois network are shown in Figure 16-13. This figure illustrates the effect of line length on the frequency of the line-averaged rates. For rates greater than 0.2 mm/min, the highest rates occur most often on the shortest line. This is because larger paths are likely to extend beyond the high rates at the core of a rain cell into an area of light or no rain. This situation reverses during light rain (<0.2 mm/min in the figure) since a short path is more likely to be associated with the fringe of a rain cell.

Table 16-16. Conditional frequencies of durations of various precipitation rates for about one year of data from Miami, Florida.

Duration Equaled or Exceeded (min)	Rate Equaled or Exceeded (mm/min)										
	0.2	0.4	0.6	0.8	1.0	1.25	1.50	1.75	2.0	2.5	3.0
1	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2	85.7	67.1	72.3	77.7	81.9	69.0	62.8	74.1	59.1	45.5	50.0
4	53.3	47.9	43.1	52.1	43.1	29.3	20.9	22.2	22.7	0.0	0.0
6	37.3	27.7	26.2	28.7	19.4	8.6	11.6	11.1	4.5		
8	25.4	20.7	20.0	16.0	8.3	6.9	4.7	3.7	4.5		
10	19.7	14.9	11.5	11.7	6.9	5.2	4.7	0.0	0.0		
15	11.9	8.0	6.9	6.4	5.6	1.7	0.0				
20	7.8	4.8	4.6	4.3	0.0	0.0					
25	5.7	1.6	0.8	1.1							
30	4.1	0.5	0.8	0.0							
45	1.2	0.0	0.0								
60	0.4										
90	0.0										

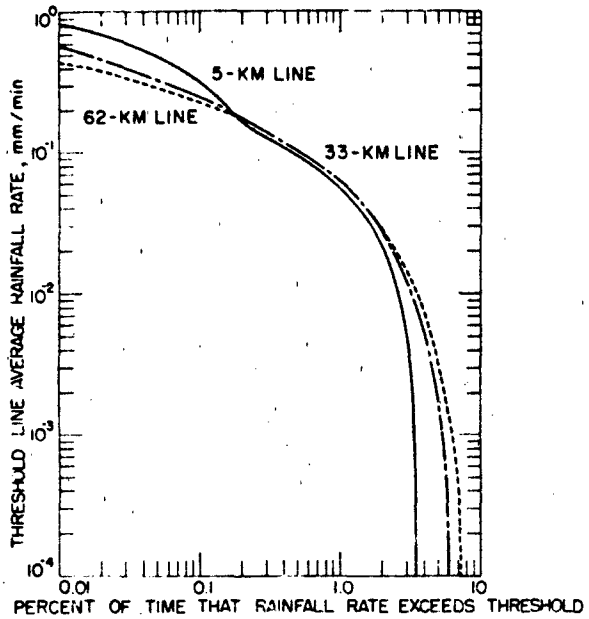


Figure 16-13. Cumulative frequency distributions of average rainfall rates for 5-, 33-, and 62-km north-south lines in central Illinois.

16.2.2 Hail

Hail consists of ice particles with a diameter of 5 mm or more. Smaller particles are classified as ice pellets. Generally transparent globular ice pellets are called sleet, while generally translucent snow pellets encased in a thin layer of ice are called graupel or soft hail. If the diameter of the ice particles is 2 cm or more, it is called large hail.

The potential for damage of hailstones depends on the size, density, hardness, number density, impact velocity, and duration of the hailstorm. For a vehicle in flight the horizontal extent of the hailstorm also becomes important. However, it is difficult or impossible to take all factors into account. Consequently, for this review the criterion of hailstorm intensity is arbitrarily restricted to maximum stone size.

Hailstone diameters of 1.9 to 2.5 cm are the minimum for damage to aircraft, especially in flight when the impact is greater due to airspeed. The metal surface of a transport aircraft at rest on the ground resists damage from hailstones up to 7.5 cm in diameter. Damage to crops reportedly occurs with hailstones exceeding 3.5 cm in diameter. However, a great number density of smaller 6 mm stones will damage crops of wheat, corn and soybeans [Gringorten, 1971]. Terminal velocity, an important factor in the destructive power, is related directly to the size of the hailstones.

16.2.2.1 Hailstone Characteristics. Although hailstone diameters are usually less than 1.3 cm, there is a distribution frequency from 0.25 cm to several centimeters. In September 1970 a hailstone diameter of 14 cm was reported to have fallen in Coffeyville, Kansas. In hailstorms, the conditional probability of occurrence of maximum-size hailstones of 7.5 cm or greater is approximately 0.5%, of 5 cm or greater 2%, of 2.5 cm or greater 7% or 8%. Diameters of less than 3 mm occur roughly 5% of the time [Gringorten, 1971].

The density (weight per unit volume) of hailstones is a variable with little documentation. For some large natural hailstones in several storms in the midwestern U.S., estimates ranged from 0.828 to 0.867 g/cm^3. Hence, a rounded value of 0.9 g/cm^3 is deemed acceptable in calculations of impact energy. Such large hailstones have an onion-like structure, attesting to the up-and-down motion to which they have been subjected during their growth in severe storms.

Terminal velocity (w), according to most authors, can be related directly to the hailstone diameter (d) by

$$w = K\sqrt{d} \qquad (16.10)$$

For w in cm/s and d in cm, the value of K at the surface has been found to range from 1150 to 1990, with a "best" estimate at 1630.

16.2.2.2 Hailstorm Characteristics. The ratio of hailstorms to thunderstorms varies greatly with geography, so that thunderstorm frequency is no criterion of hail frequency. The month of greatest hailstorm frequency also varies. In the United States the hail season extends from April to September, with June and July the most severe months. Most hail occurs between 1200 and 1900 LST with a peak for damaging hail at 1600 LST.

The width of a swath of hailstone incidence has been variously estimated from a few yards to 120 km, but rarely more than 8 km. The length of the path is considerably greater than the width, but even more difficult to define. The term "hailstreaks" has been coined for hailstorms in the midwest, which are normally 2 to 25 km long in Illinois. The median width is 1.5 to 3 km, the length is 10 km in Illinois and 25 km in South Dakota. The areal extent of these hailstreaks varies from 2 to 2000 km^2.

Large hail (greater than 2 cm), is found only in well-developed thunderstorms whose cloud tops may reach 15 km. Small hail and soft hail are thought by some meteorologists to be an essential feature of all thunderstorms. However, ice particles less than 1 cm in diameter, which may be present aloft in the cores of the thunderstorms, are likely to melt completely before reaching the ground. Although thunderstorms are most frequent in the tropics and subtropics, occurring up to 180 days per year in several places, hail is rarely found on the ground in the tropics. Hail on the ground is most likely found in midlatitude mountainous and adjacent areas.

16.2.2.3 Frequency and Duration at the Ground. The most severe hail activity in the world occurs in the United States, and the most significant center within the United States is in western Nebraska and southeastern Wyoming. Although the greatest number of hailstorm days is at Cheyenne, Wyoming, averaging 9.4 days per season, it is so localized that an average of 7 days per year may be considered an acceptable extreme. By comparison, in some areas thunderstorms average 40 to 50 days per year. In the most severe month and in the most severe location, the average number of hailstorms is 2.9 per month. However, a more acceptable number for our use, especially for application to flight paths, is 2 hailstorms per month.

Using the average of 7 hailstorms per year in the Poisson distribution for the frequencies of zero, one, and two hailstorms per year, together with the conditional frequency distribution of the maximum hailstone size in a storm (Section 16.2.2.1), we obtain the cumulative probability of the annual largest hailstone diameter in the most severe location shown in Figure 16-14. Application of the Gumbel distribution further leads to estimates of the greatest diameters in periods of 1 to 25 years (Table 16-17).

The duration of hailstorms has been reported variously from a few minutes to a few hours. The best estimate for the average time, at the point of most severe occurrence and in the most severe month, is 10 min [Changnon, 1970]. When this is coupled with the average of 2 hailstorms per month, the single-point probability of encountering a hailstorm becomes 0.000448. The probability of encounter in terms of hailstone size is shown in Table 16-18.

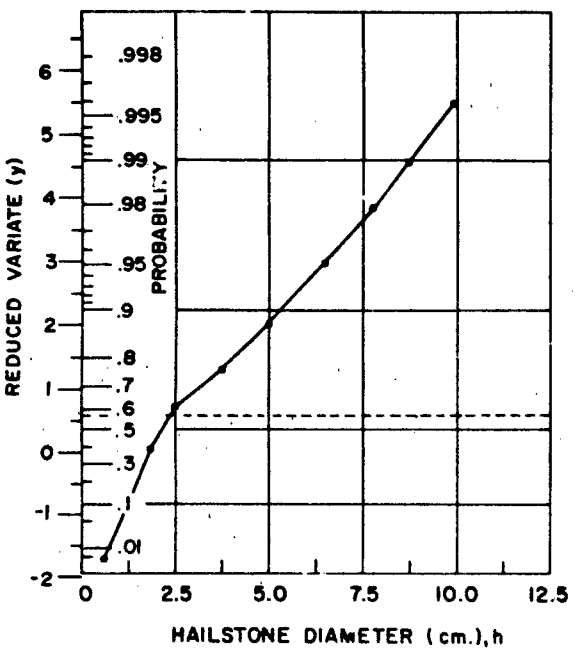


Figure 16-14. Cumulative probability of the annual largest hailstone diameter in the most severe location, plotted on extreme probability paper.

Table 16-17. Diameter of the maximum-size hailstone in period of 1 to 25 years in the most severe location.

Period of Years	Hailstone Diameter (mm)	
	Average	10% Extreme
1	25	55
2	35	66
5	53	76
10	63	89
25	76	102

16.2.2.4 Hail Aloft. Nearly every thunderstorm has some hail, snow, or ice pellets at one stage or another of its development, particularly at levels between 6 and 8 km. At any given time, however, hail is considered to occur in a shallow layer. Above 9 km hail has been encountered infrequently, but cannot be discounted.

In the absence of sufficient and objective data a certain amount of inference must be used in considering the frequency of hail as a function of altitude. Hailstones must form and grow in size above the freezing level in the atmosphere. Over northeastern Colorado, western Nebraska, or southeastern Wyoming, the average height of the 273 K isotherm is 4.1 km during the hail season (summer). Once formed, the hailstones will fall, be buffeted vertically up and down, or become suspended at a "balance" level. This should become a level of high concentration and, therefore, a level with a relatively high probability of hail occurrence. Balance levels may be either at or below the level of updraft maximum, but above the freezing level.

With our present knowledge, it is expedient to assume

Table 16-18. Estimates of the probability of encountering hailstones of given diameter at a single-point location from ground to 1500 km in the most severe location in the worst month.

Hail Diameter (cm)	Single-station Probability
Any size	0.448×10^{-3}
≥ 0.6	0.354×10^{-3}
≥ 1.3	0.161×10^{-3}
≥ 1.9	0.605×10^{-4}
≥ 2.5	0.314×10^{-4}
≥ 5.1	0.851×10^{-5}
≥ 7.6	0.170×10^{-5}
≥ 10.2	0.25×10^{-6}

that the probability of encountering hail is uniformly the same at any level from 3 to 6 km, and that any level in this interval can become a level of hail concentration. Concomitantly, it is assumed that the probability of hail encounter decreases from 3 downward to 1.5 km and also from 6 upward to 14 km. Since hailstones do not form or grow at the 1.5 km level, and since hailstones do not melt appreciably when falling from that level to the surface, the probability of hail encounter at levels at and below 1.5 km is approximately the same as that found at the surface (0.000448).

For the significantly large hailstones (>2 cm diameter), it has been found that their diameters remain virtually unchanged when falling from aloft to the ground. As a result, the conditional probability distribution of hailstone diameters aloft can be adequately estimated by the conditional probability distribution at the surface.

Estimates of the ratio of probabilities of hail aloft to hail at a lower level are available for thunderstorms in Ohio and Florida. In these storms hail was encountered about 7 times more often between 3 and 6 km than at 1.5 km. These data also indicated that at 7.6 km the probability of encountering hail is about 3 times as great as at 1.5 km. Above 7.6 km the probability of encountering hailstones diminishes steadily to nearly zero at 13.7 km (Table 16-19).

The probability of hail occurring at the ground increases with increasing height of the radar-echo tops of the associated thunderstorm. For example, hail has been reported on the ground for about half of the New England thunderstorms when radar-echo tops have exceeded 15 km. In addition, hailstorms in the southern and midwestern U.S. sometimes have echo tops up to 18 and 21 km [Kantor and Grantham, 1968]. These extremely high echo tops extend well above the cirrus anvil and penetrate several kilometers into the stratosphere. Because these giant storms must be accompanied by exceedingly high vertical velocities, hail above the tropopause is possible.

Table 16-19. Estimates of the probability of encountering hail of any size at a single-point location by altitude.

Altitude (km)	Probability
Ground level	0.000448
1.5	0.000448
3.0	0.00314
4.6	0.00314
6.1	0.00314
7.6	0.00134
9.1	0.00100
10.7	0.00067
12.2	0.00034
13.7	0.000

Table 16-20. Estimates of the probability of occurrence of a hailstorm in an area (A), in the worst area during the worst period, when 2 hailstorms per month are expected at any single point.

Area A (km^2)	Probability (in %)
0	0.045
26	0.06
260	0.17
2600	0.58
26 000	3.1
260 000	29

16.2.2.5 Area and Line Coverage. Ten years of observations at a network of some 50 stations in a 400 km^2 area around Denver [Beckwith, 1960] gave the average ratio of occurrence of hailstorm days in that area to the occurrence at a single station as 4.4 to 1. Thus, if the probability of occurrence at a single station is 0.000448, then the probability of one or more occurrences in a 400 km^2 area would be 0.00197. A subsequent investigation in Illinois and South Dakota [Changnon, 1971] has led to the conclusion that the area-to-point ratio of hail-day frequencies is related to the area size, but is reasonably independent of geography.

With the above information a recent model of areal coverage [Gringorten, 1979] has yielded an estimate of the correlation coefficient of 0.99 for hail events at points 1.3 km apart. With this estimate the probability of a hailstorm occurrence somewhere in an area (A) as a function of the size of that area becomes as shown in Table 16-20. The model also has yielded the probability of encountering hail along a flight path (at an altitude where the single-point probability is 0.00314), as 0.025 in a path length of 160 km, and 0.042 in a path of 320 km. The probability of encountering hailstones of at least 2.5 cm in diameter is smaller by 7/100, becoming 0.0018 for 160 km and 0.0029 for 320 km.

16.2.2.6 Hail in Design and Operation of Equipment. Since the probability of hail of even the smallest size is very slight between the ground and 1500 m and since the effect of hail on most surface equipment will not result in endangering human life, there is no need to specify a hail-size extreme as an operational design criterion. If hail interferes with a surface operation, postponement until the storm ends is advisable.

Aloft (between 3 and 6 km), probabilities of hail are as high as 0.00314, which is still sufficiently low so that there is no need to specify a hail-size extreme for equipment whose failure would not endanger human life. If such danger exists, however, the estimates provided in this section should be considered.

16.2.3 Snowfall

Snow is important to the design of equipment and structures that must bear an increased structural load imposed by snow accumulations. It can also impede movement over uncleared ground. In areas where the mean winter temperatures are near or below freezing, the maximum depth of snow during the season is usually the accumulation from many individual snows. In warmer climates extremely heavy snowstorms can occur, but these are usually followed by melting and amounts on the ground at any one time are not as great as in colder areas farther north. The most favored locations for heavy snowstorms and excessive snow depths are mountain locations, especially those in proximity to large bodies of water.

A survey on North American snowfall extremes through the 1969-70 season by Ludlum [1970] indicates that each state in the U.S., except Florida and Hawaii, has recorded 24-h snowfalls of at least 45 cm. North America's heaviest 24-h accumulation of 193 cm occurred at Silver Lake, Colorado, a mountain location (elevation 3115 m). A study of intense snowfalls by Riordan [1973] indicates that at least 75 cm of snow has been recorded in a 24-h period in 23 states in the U.S. including locations as far south as Arizona, New Mexico, North Carolina, and Virginia. The greatest Canadian 24-hour snowfall was 112 cm at Livingston, Alberta. North America's greatest single storm accumulation of 480 cm occurred at Mt. Shasta Ski Bowl in the California mountains [Riordan, 1974]. The greatest measured snow depth in the U.S. was 1145 cm at Tamarack, California.

Snow loads on structures are not routinely measured and must be estimated. This requires information not only on the depth of snow, but also its density. Even then, the shape and type of structure, and its wind exposure must be taken into account. The density of freshly fallen snow is most commonly between 0.05 and 0.15 g/cm^3. Consequently an accumulation of 76 cm with a "normal" density of 0.1 g/cm^3 would result in a snow load on the ground of about 76 kg/m^2. Depths in much of North America can greatly exceed 76 cm, and rain on top of the snow significantly increases its density. Boyd [1961] analyzed snowfall statistics for more than 200 Canadian stations. His analysis of the maximum depth of snow is shown in Figure 16-15 for a 30-year return period, the time interval in which the indicated

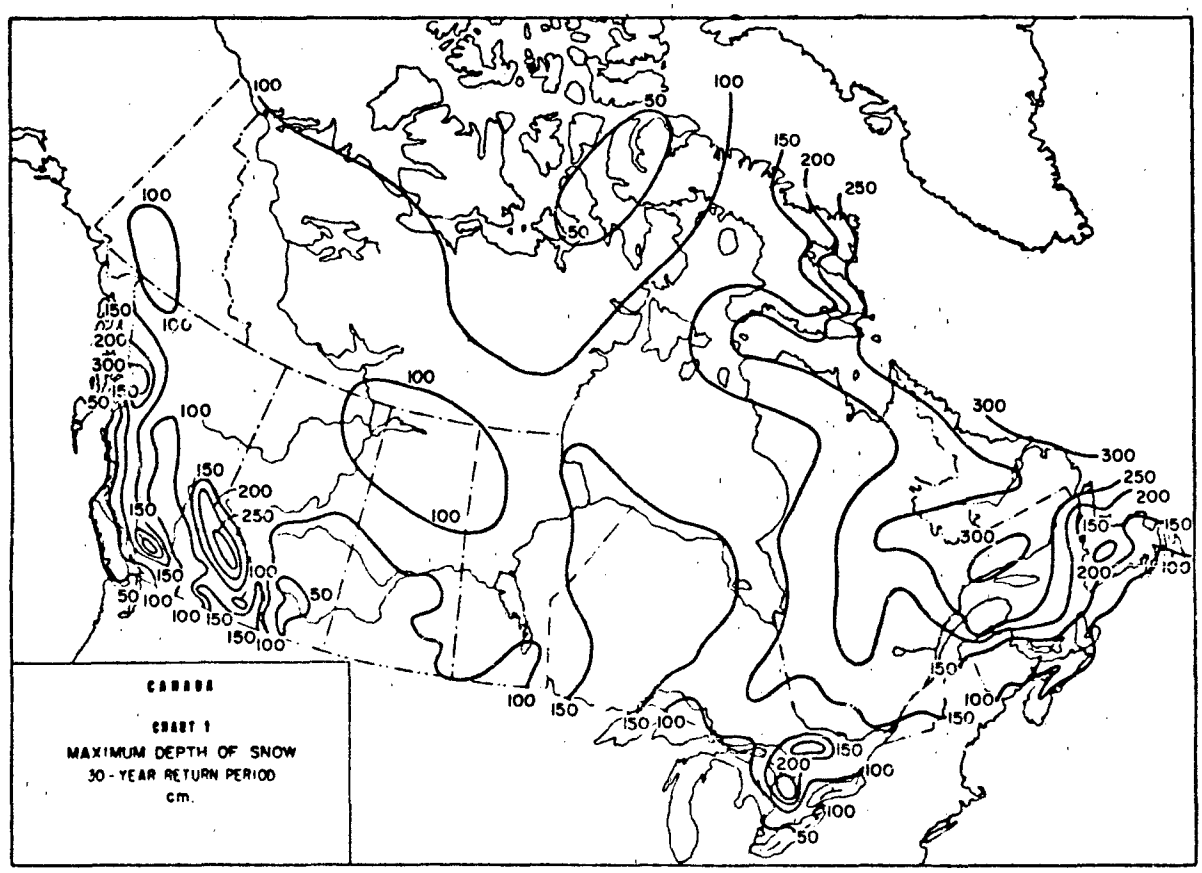


Figure 16-15. Maximum depth of snow for a 30-year return period for Canada expressed in metric units [Boyd, 1961]

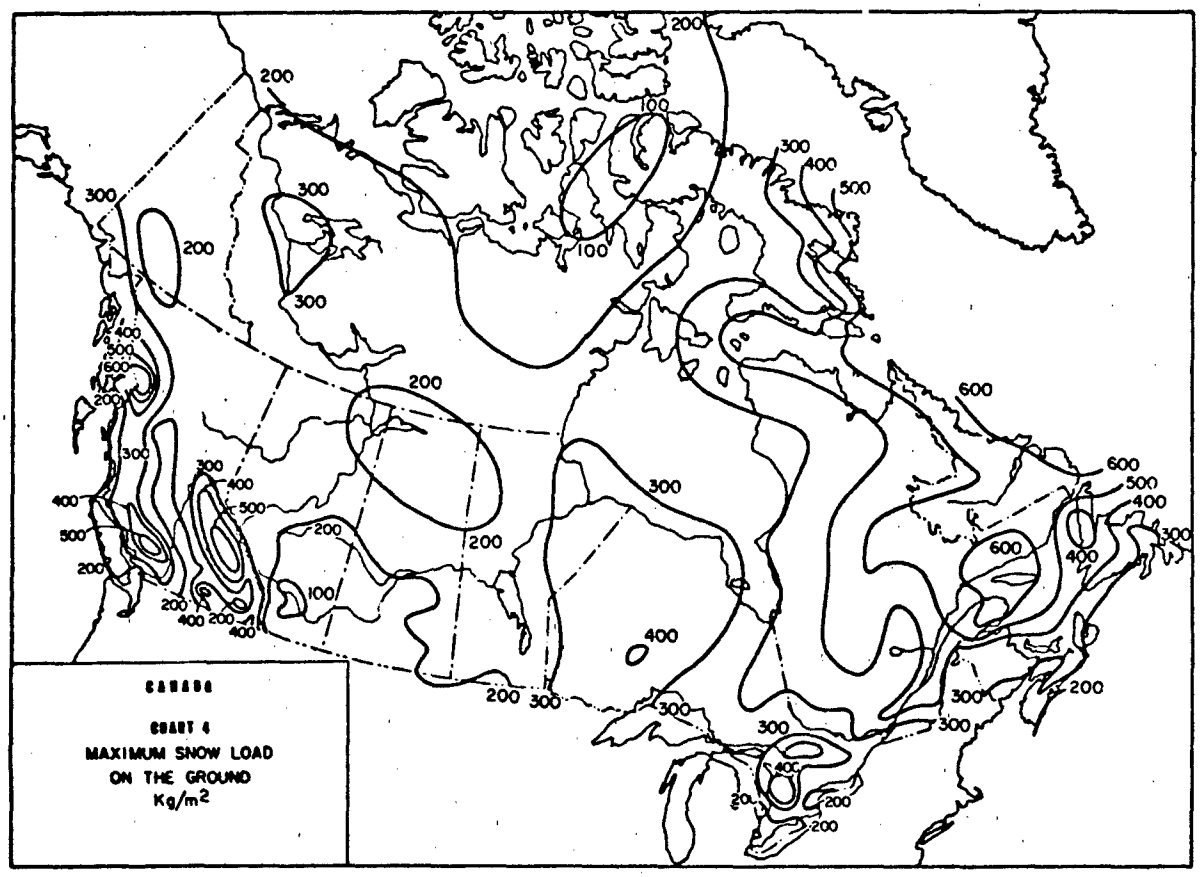


Figure 16-16. Estimated maximum snowload in Canada in a 30-year return period expressed in metric units [Boyd, 1961].

maximum depth would be expected to occur (a 63% chance) at least once. He also estimated the amount of rain that could occur and be retained by the snow. These findings are incorporated into Figure 16-16, which indicates that a maximum snow load on the ground of 600 kg/m^2 can occur in non-mountainous areas in eastern Canada. He tentatively indicates that snow loads on flat or low-slope roofs would be about 80% of the ground load based on limited observations. In a 1951 U.S. Weather Bureau study, the maximum non-mountainous ground snow load in the U.S. is estimated at 300 kg/m^2, and the 1 year in 10 maximum at 200 kg/m^2. Snowfall extremes in North America are considered to be as severe as anywhere in the world, so that these values are considered representative of non-mountainous worldwide extremes.

16.2.4 Ice Accretion on Surface Structures

Ice accretion, or icing, refers to ice accumulating on stationary objects located near the earth's surface. Such ice can disrupt transportation, cause power and communication outages, and result in severe damage to structures bearing the burden of accumulated ice. Nevertheless, it remains one of the few meteorological conditions not quantitatively observed on a routine basis by any national weather service.

There are three basic types of ice formed by accretion in the atmosphere: glaze, hard rime, and soft rime. Glaze occurs most commonly when rain or drizzle freezes on objects. It is clear and nearly as dense as pure ice (which has a density of 0.917 g/cm^3). Hard rime is less transparent than glaze, and at times is opaque, depending on the quantity of air trapped in the ice. Its density varies from about 0.6–0.9 g/cm^3. Soft rime, which is white and opaque, is feathery or granular in appearance, with a density less than 0.6 g/cm^3. Rime ice, both hard and soft, is most often formed by the freezing of supercooled cloud or fog droplets. For this reason it is a localized phenomenon, occurring most frequently in mountainous areas or hilltops exposed to low, stratus type clouds, or on land areas adjacent to bodies of water.

Most of the information available on surface icing is for glaze; it is usually recorded because of extensive human distress or a large amount of damage. Even then, measurements have been subjective, non-standardized, and difficult to interpret. The most comprehensive study on the geographical distribution and frequency of glaze, worldwide,

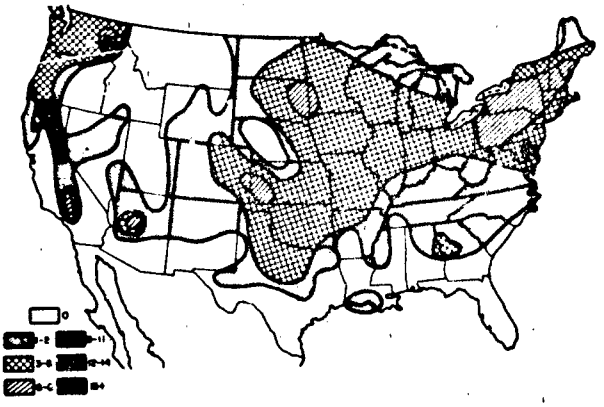


Figure 16-17. Number of times ice at least 1.25 cm thick was observed during the 9-year period of the Association of American Railroads Study [Bennett, 1959].

was done by Bennett [1959]. The information presented indicates that glaze is as frequent and as severe an occurrence in parts of North America as it is anywhere in the world. Figure 16-17 from that report shows the number of times that ice at least 1.25 cm in thickness was observed in the U.S. during a 9-yr period in a study by the Association of American Railroads. Although the length of record is short for climatological purposes, the relatively high frequency of heavy ice storms in the central and northeastern U.S. is supported by other data.

Tattelman and Gringorten [1973] reviewed many sources for information on ice storms in the contiguous U.S. for the 50 winter seasons from 1919-20 through 1968-69. The data were analyzed to determine the probability of occurrence of ice thicknesses in each of eight regions in the U.S. The regions shown in Figure 16-18 are areas with similar glaze characteristics subjectively determined on the basis of latitude, geography, climatology and the distribution of ice storms. Figure 16-19 shows ice thickness versus probability

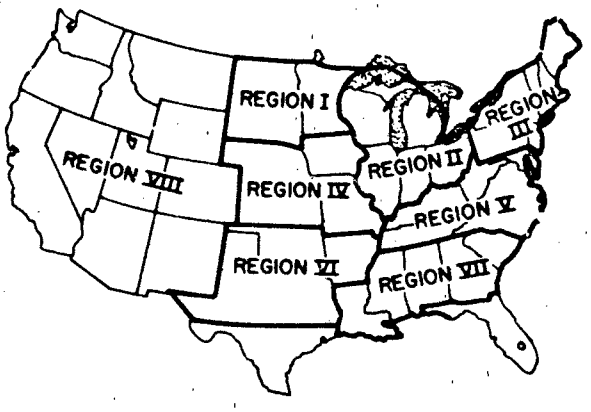


Figure 16-18. Regions of similar glaze characteristics [Tattelman and Gringorten, 1973]

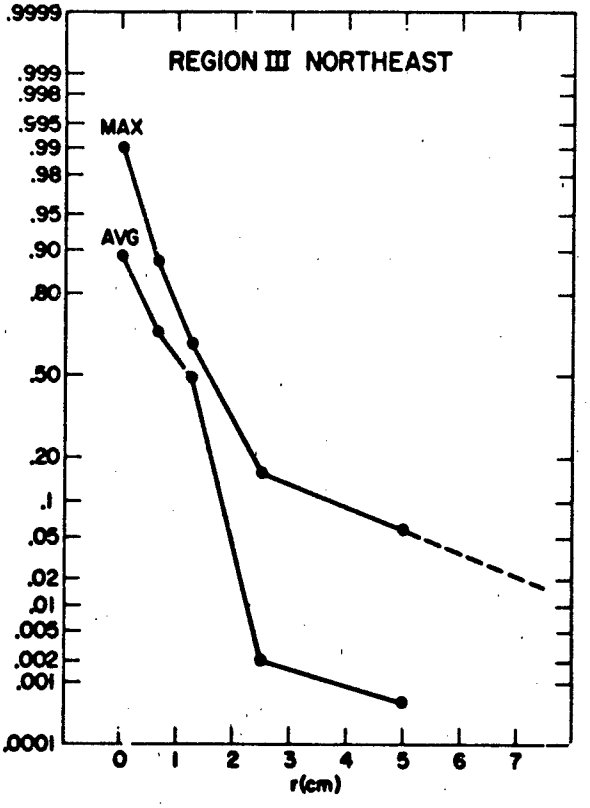


Figure 16-19. Estimated probability of ice thickness, r, occurring at least once in any year at a representative point in region III (AVG), and in the most severe part of region III (MAX) [Tattelman and Gringorten, 1973].

of occurrence in a year for the northeast (region III in Figure 16-18), which had the greatest frequency of ice storms. The curve labelled AVG represents an average for all states in the region, whereas the curve labelled MAX represents the estimate for the state in the region with the highest frequency of ice storms. The authors stress the subjectivity involved in arriving at their estimates.

Eastern Canada, from Nova Scotia and Newfoundland northward to southern Baffin Island, has the highest frequency of freezing rain and drizzle in North America [McKay and Thompson, 1969]. Extreme ice storms in this area have produced accumulations of ice 20 to 40 cm in thickness. This is not surprising because the area is frequently subject to the passage of slow moving "mature" cyclones, and associated winds can exceed hurricane force (33 m/s). Strong winds increase the number of precipitation drops impinging on a vertical surface, so that the accumulation of ice can exceed the depth of precipitation by a factor of 3 or more.

Ice accretion is difficult to measure because the shape and size of the collecting surface, and its orientation to the wind can result in large differences. Furthermore, the configuration and density of ice on a surface are dependent upon a number of atmospheric variables. Tattelman [1982]

Table 16-21. Mean and standard deviation of the ratios of the mass of ice on the 3-, 13-, and 50-mm diameter cyclinders to that on the 25 mm diameter cylinder.

Category	Cylinder Diameter (mm)	Mean	Standard Deviation
In-cloud icing	3.2	0.21	0.059
In-cloud icing	12.7	0.59	0.062
In-cloud icing	50.8	1.47	0.116
Freezing rain	3.2	0.18	0.041
Freezing rain	12.7	0.54	0.056
Freezing rain	50.8	1.65	0.091
Both combined	3.2	0.20	0.054
Both combined	12.7	0.57	0.063
Both combined	50.8	1.56	0.139

gives an overview on the dynamics of ice accretion and research in the field. He also presents an analysis of icing measurements on cylinders, for differing synoptic conditions, made during climatic chamber tests of an ice measurement system. Results include a comparision of the mass of ice that accumulates on cylinders which are 3, 13, 25, and 50 mm in diameter (Table 16-21).

16.2.5 Distributions of Precipitation Elements

Precipitation parameters vary appreciably with type of storm, geographic location, and even from storm to storm of the same type and in the same geographic region; for this reason, no model storms are presented in this section. However, individual profiles or averages that are derived from observations at several locations are given, and, wherever possible, the applicability and representativeness of the data are indicated. Great care must be taken in extrapolating the results to geographical regions that are characterized by a climatology which differs from that of the region from which the data were obtained.

16.2.5.1 Raindrop Size Distributions. Numerous equations have been proposed to express the size distribution of raindrops measured at the ground as a function of rainfall rate. They can all be expressed in the form

$$N_D\,\Delta D = N_0 D^{\alpha}\exp(-\Lambda D^{\beta}), \qquad (16.11)$$

where $N_D\,\Delta D$ is the number of drops per unit volume with diameters between D and $D + \Delta D$, N_0 is the value of N_D where the curve crosses the $D = 0$ axis, Λ is a parameter that depends on the type and intensity of the precipitation, and α and β are parameters which determine the shape of the distribution.

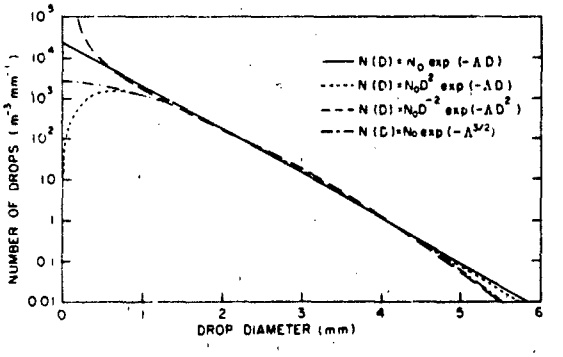


Figure 16-20. Four typical shapes of raindrop size distributions measured at the ground.

Examples of some drop size distributions are shown in Figure 16-20. When $\alpha = 0$ and $\beta = 1$, Equation (16.11) reduces to the simple exponential distribution first proposed by Marshall and Palmer [1948]. This is shown in the solid line of Figure 16-20. This distribution was derived from data obtained in stratiform rain, and tends to overestimate the number of small drops. When $\alpha = 2$ and $\beta = 1$, the resulting expression is an approximation to the log-normal distribution proposed by Levin [1954]. In this case, there is a realistic decline in the number of drops at the lower end of the spectrum. Setting $\alpha = -2$ and $\beta = 2$ results in a gamma distribution first derived by Best [1950]. This distribution holds if, as some measurements indicate, the liquid water content per drop size interval follows an exponential distribution. The fourth distribution is one proposed by Litvinov [1956] for drizzle-type rains, setting $\alpha = 0$ and $\beta = 3/2$. Although all the distributions in Figure 16-20 have been extended to 5-mm diameter drops (the usual maximum drop size observed), the distribution for drizzle-type rains should be truncated at approximately 2-mm diameter.

The total liquid water content M, the Rayleigh radar reflectivity factor Z, and the median volume diameter D_0 can be expressed as weighted integrals of the size distribution. The weighting factor in deriving Z is D^6, while that for deriving M is $\pi/6\,\rho D^3$, where ρ is the density of water in g/cm³. Assuming an exponential distribution ($\alpha = 0$, $\beta = 1$), and integrating between $D = 0$ and $D = \infty$ (an approximation sufficiently accurate for most purposes), we obtain

$$D_0 = 3.67/\Lambda \quad \text{(mm)} \qquad (16.12)$$

$$M = 10^{-3}\,\pi(\rho N_0/\Lambda^4)\ \text{(g/m}^3\text{)} \qquad (16.13)$$

$$Z = 720\,N_0/\Lambda^7 \quad \text{(mm}^6\text{/m}^3\text{)}. \qquad (16.14)$$

Other parameters (for example, the median power diameter) can also be found using the appropriate weighting factor. Such integrations, however, become cumbersome for distributions in which $\alpha \neq 0$ and $\beta \neq 1$. For this reason,

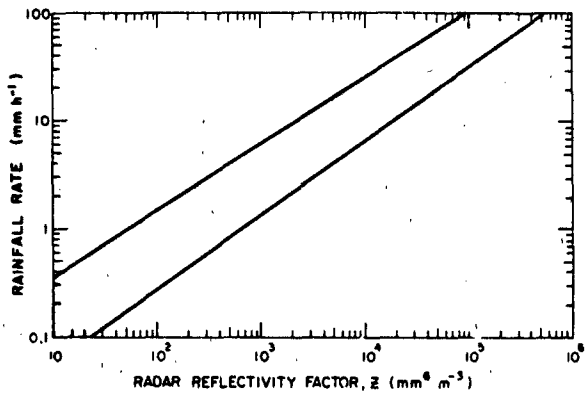


Figure 16-21. Radar reflectivity factor vs rainfall rate (shaded area is the range of the measured relations).

the simple exponential distribution is often used, even when it is known to deviate from measured size distributions.

Numerous empirical relations have been proposed expressing radar reflectivity as a function of rainfall rate or liquid water content, or rainfall rate as a function of liquid water content. The relations differ significantly, according to the location and type of precipitation. Figures 16-21, 16-22, and 16-23 summarize the results of measurements from several observers. The shaded area of Figure 16-21 encompasses the range of measured relations between radar reflectivity and rainfall rates, and illustrates the variability that can be encountered. Figure 16-22 expresses the same data in terms of radar reflectivity versus liquid water content. If rainfall rate is the measured quantity, Figure 16-23 can be used to find the expected range of liquid water content.

16.2.5.2 Snowflakes. Gunn and Marshall [1958] found that an exponential law similar to that shown by the solid line of Figure 16-20 was applicable to the size distribution of aggregate snowflakes. The spectral parameters in snow are related to the precipitation rate by

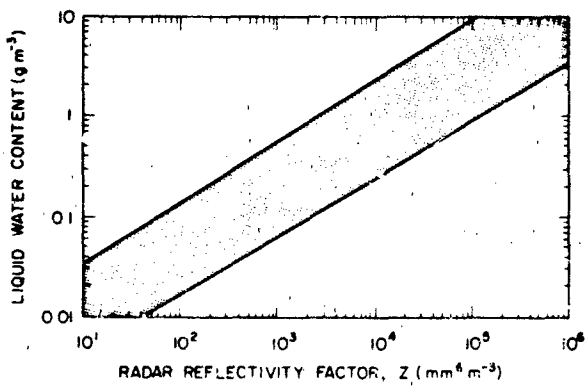


Figure 16-22. Radar reflectivity factor vs liquid water content of rain (shaded area is the range of observed relations).

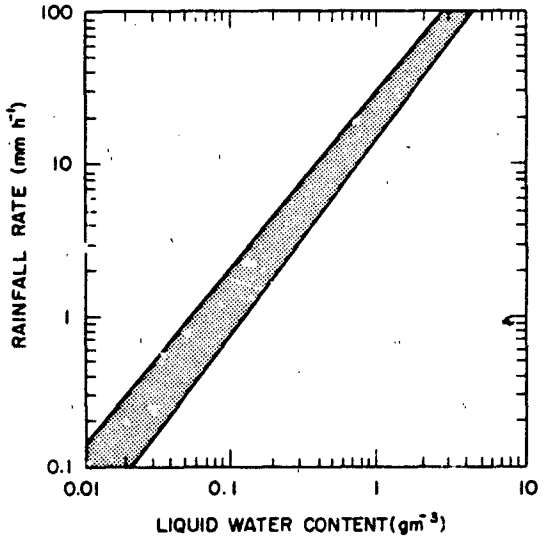


Figure 16-23. Liquid water content of precipitation vs rainfall rate.

$$N_D \, \Delta D = N_0 \exp(-\Lambda D) \qquad (16.15)$$

where $N_0 = 3.8 \times 10^3 \, R^{-0.87} (m^{-3} mm^{-1})$ (16.16)

$$\Lambda = 2.55 R^{-0.48} \quad (mm^{-1}), \qquad (16.17)$$

where R is precipitation rate in millimeters of water per hour. For snowflakes, the diameter D refers to the melted spherical diameter of the snowflake.

Recent measurements [Lo and Passarelli, 1982] of the evolution of snow size spectra with decreasing altitude demonstrate the existence of three growth regimes. At the top of the snow region, growth is by deposition only. The size distribution is typified by curve 1 in Figure 16-24. Progressing down through the atmosphere, N_0 increases while Λ remains constant. Then aggregation begins, and there is a rapid decrease in both N_0 and Λ. Curve 2 of Figure

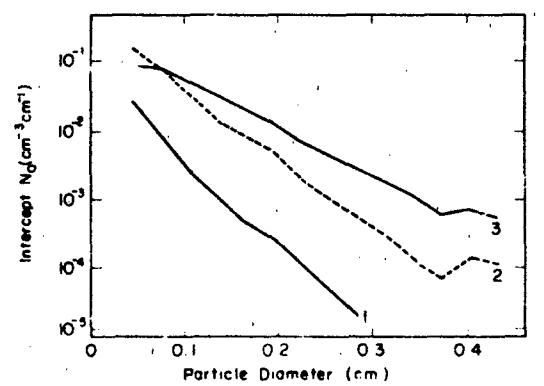


Figure 16-24. Snow size spectra as a function of growth regime and altitude.

16-24 illustrates a typical spectrum during this phase. Finally, breakup counteracts the effects of both deposition and aggregation, and the distribution remains constant, with a slope and intercept typical of curve 3 in Figure 16-24. The minimum value of Λ observed in the atmosphere is approximately 10 cm^{-1}.

16.2.5.3 Distribution of Precipitation Elements with Height. The vertical distribution of precipitation parameters can be inferred from the vertical distribution of radar reflectivity, using the appropriate conversions. When the precipitation particles are ice crystals or snowflakes, it is necessary to determine the type of particle and general size regime before choosing the equation to convert the particle size or mass to the equivalent diameter or mass of a liquid drop. Table 16-22 shows factors for converting from a physical length, L, of the solid particle to an equivalent spherical diameter, D. A power relation of the form $D = a L^b$ has been assumed, and a and b are listed for general particle types.

The maximum size of hydrometeors is related to the slope of the distribution by the empirical relation

$$D_m \Lambda \approx C , \qquad (16.18)$$

where D_m is the maximum equivalent diameter, Λ is the slope of the exponential distribution, and C is a constant dependent on the type of hydrometeor. Using data collected over several years, AFGL scientists have found that average values of C range from 12 for ice crystals to 10.5 and 9 for small and large snow, respectively, and 7.5 in rain.

If an exponential distribution is assumed, the particle size distribution of hydrometeors can be completely described from the slope of the distribution (Λ), the intercept (N_0), and the total number of drops (N_t), which depends upon the minimum and maximum particle size assumed. The size is dependent upon the slope of the distribution as shown in Equation (16.18).

Table 16-23 gives the characteristics of average distributions for various hydrometeor types. In each case, N_t, N_0, and Λ are expressed as power functions of M, the total water content in g/m^3. The units of Λ are $1mm^{-1}$; of N_0, $m^{-1}mm^{-1}$; and of N_t, m^{-3}. Expressions connecting other sets of parameters may be obtained from this table and manipulation of Equations (16.12) through (16.14), or integration of Equation (16.11) with the appropriate weighting factor.

Table 16-22. Factors for converting from physical length, L, to equivalent diameter, D.

Particle type	Size range	a	b
Ice crystals	<0.2mm	0.438	1.000
	≥0.2mm	0.256	0.667
Small snow	<0.5mm	0.370	0.670
Large snow	≥0.5mm	0.400	0.782

Height variations of gross parameters such as water content or radar reflectivity factor are functions of type of storm and geographic location. However, when there is a sufficient body of data for a particular region, it is possible to define a climatological model. Two such models are shown in Figures 16-25 and 16 26. Figure 16-25 shows the results for stratiform storms at Wallops Island, Virginia. The dotted line may be considered an average vertical profile of liquid water content for stratiform storms at this location, and the solid lines indicate the variability of the data. Figure 16-26 shows the same analysis for convective storms at Wallops Island. The major difference between these two figures is the high water content between 2 and 4 km in convective storms. This storage zone aloft is characteristic of most convective storm situations, regardless of location. Stratiform storms, on the other hand, are characterized by relatively smooth vertical profiles of liquid water content, with only a slight bulge at the melting zone.

16.2.5.4 Extreme Values of Liquid Water Content. Measurements of the maximum concentration of liquid water in severe convective storms have not been made on a systematic basis. Isolated reports of concentrations as high as 30 g/m^3 [Sulakvelidze et al., 1967] and even 44 g/m^3 [Roys and Kessler, 1966] are found in the literature. There have been occasions when investigators have suspected the occurrence of abnormally high concentrations but lacked the

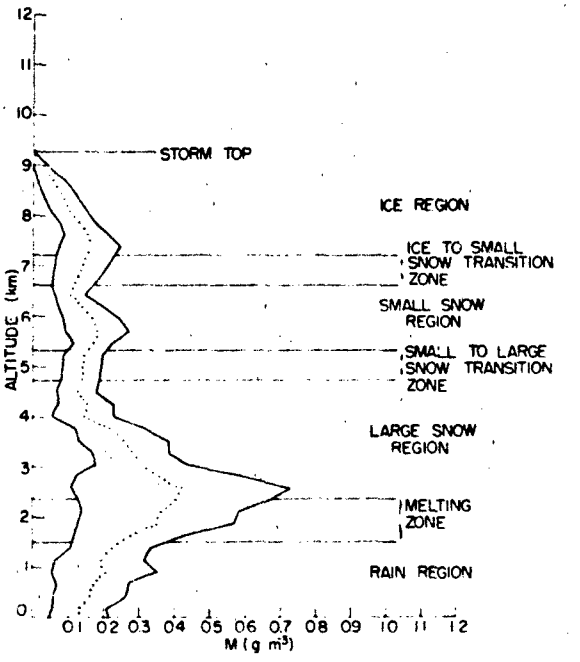


Figure 16-25. Liquid water content in grams per cubic meter vs altitudes in kilometers for stratified-storm situations (dotted line is the mean and the solid lines represent plus and minus 1 standard deviation).

Table 16-23. Distribution parameters as a function of hydrometeor type and of total mass, M (N_t is the total number of particles per cubic meter, N_0 and Λ are the zero intercept and slope for an exponential size distribution).

Hydrometeor Temp. Range Min. Diameter Sub Types	$N_t = e\,M^f$		$N_O = i\,M^j$		$\Lambda = k\,M^q$	
	e	f	i	j	k	q
Ice Crystal <248 K d = 0.055 mm						
Tropical	61522	0.073	514541	−0.583	6.343	−0.394
Temperate	98719	0.303	1104536	−0.201	7.675	−0.299
Small Snow 258 to 243 K d = 0.065 mm						
Tropical	50723	0.032	403471	−0.739	5.968	−0.431
Temperate	19727	0.266	109002	−0.148	4.295	−0.287
Large Snow 273 to 253 K d = 0.065 mm						
Tropical	4122	−0.373	10839	−1.089	2.403	−0.522
Temperate						
Agg. of Plates	6833	0.384	23966	0.096	2.930	−0.226
Agg. of Stellars & Dendrites	1649	0.155	3251	−0.202	1.778	−0.301
Agg. of Columns & Bullets	14005	0.455	67147	0.164	3.791	−0.209
Rain > 273 K d = 0.1 mm						
Tropical	4033	0.443	12619	0.161	2.471	−0.210
Temperate						
Widespread	1934	0.133	4382	0.018	1.897	−0.245
Drizzle	5850	0.366	21509	0.017	2.824	−0.246
Thunderstorm	649.9	0.310	945.6	0.017	1.293	−0.245

equipment for measuring them. In considering many aspects of the storm, Roys and Kessler could not find evidence which indicated that the measured concentration 44 g/m^3 should be either rejected or accepted. Omitting the value of 44 g/m^3, their 26 other measurements of the maximum liquid water content in Oklahoma thunderstorms fit very well the distribution illustrated in Figure 16-27. According to this distribution, the probability that the maximum water content, M, in an Oklahoma thunderstorm will exceed a given value, x, is

$$P(M > x) = \exp(-x^2/64), \qquad (16.19),$$

where M and x are in g/cm^3.

It is dangerous to extrapolate from such a small sample. Nevertheless, the actual occurrence of extreme values of liquid water concentrations probably follows the general shape of Figure 16-27 with possible modification of the constants in the distribution function. Figure 16-27 indicates that 75% of Oklahoma thunderstorms have maximum liquid water contents exceeding 4.3 g/m^3, and that liquid water contents in excess of 9.4 g/m^3 occur in only 25% of the storms. Substitution into Equation (16.19) shows the values of 30 g/m^3 or higher are literally "one in a million" occurrences in Oklahoma thunderstorms, and the occurrence of a value of 44 g/m^3 has a probability of 10^{-12}.

It is of interest to compare these extreme values of water content to the water content corresponding to a record rain-

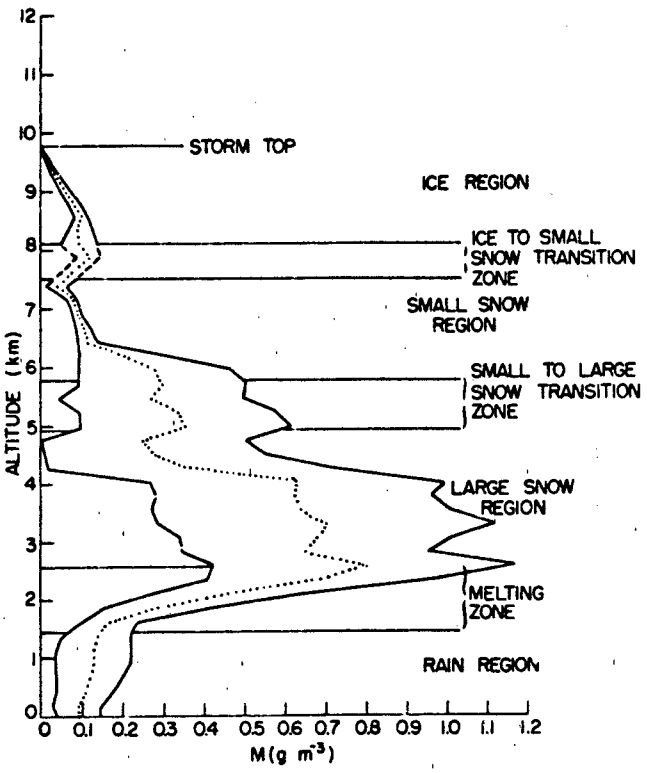


Figure 16-26. Liquid water content in grams per cubic meter vs altitudes in kilometers for convective-storm situations (dotted line is the mean and the solid lines represent plus and minus 1 standard deviations).

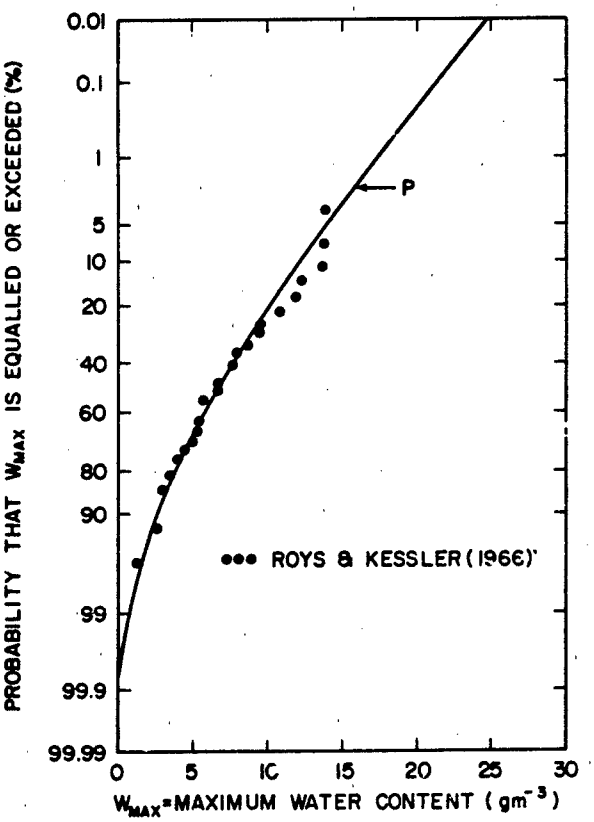


Figure 16-27. Probability of maximum liquid water content within Oklahoma thunderstorms.

fall rate. Over a 1-min interval, the world record rainfall amount is 3.1 cm; this fell at Unionville, Maryland, in 1956 [Riordan 1974]. Assuming a Marshall-Palmer distribution, this rainfall rate corresponds to a liquid water content of approximately 55 g/m^3. Consequently, extremely large values of water content (for example, greater than 30 g/m^3) may occur on rare occasions either at the surface or aloft in severe thunderstorms.

16.3 CLOUDS

16.3.1 Surface Observations

Cloud observations taken regularly by ground observers at weather stations throughout the world primarily contain the following information:

(1) Visual estimate of total amount of sky covered by clouds,

(2) Cloud ceiling, which is the estimated or measured height above ground of the lowest layer of clouds that cover more than half of the sky.

Secondarily, visual estimates of amount and height of bases of individual cloud types and layers are reported. But because low clouds obscure higher clouds, it is often impossible to obtain accurate information on the distribution of clouds at the higher levels. Studies of the accuracy of visual cloud observations show that the greatest errors occur in estimating cloud amounts when the sky is 3/10 to 7/10 covered. These errors however, tend to average out over long periods. The average error in visual estimations of cloud heights ranges from 300 m for clouds at 750 m, to 1500 m for clouds with bases near 7 km. At many points, however, particularly at airport stations, ceiling heights can be measured with a higher degree of accuracy with ceilometers, clinometers, or balloons.

Observed values for a particular time give amount of sky covered but not the cloud distribution. For example, if 5/10 cloud cover is reported, one large cloud formation may cover half of the sky or small individual cloud cells may be equally distributed over the sky. There is no direct way to obtain the geometry of clouds in the sky from the reported standard observations.

There can be considerable variation in the frequency of occurrence of given cloud amounts and ceiling heights between stations located within several kilometers of each

other. Because local topographical effects are the primary cause of such variations, caution must be used in applying climatological data on clouds of one or more stations to an entire region, particularly in mountainous and coastal regions.

16.3.1.1 Summaries of Surface Observations. Tabulations of surface cloud observations are available for stations in most areas of the world. A *Guide to Standard Weather Summaries and Climatic Services* [NOC, 1980] has been prepared by the Naval Oceanography Command in Asheville, North Carolina, which lists some 23 types of cloud summaries. The station list includes locations in various countries, continents, islands, and oceans. There are close to 10 000 weather stations around the world, although not all of them provide routine observations of clouds. Among the better summaries of cloud information are those prepared by the United States Air Force Environmental and Technical Applications Center (USAF/ETAC) as *Revised Uniform Summaries of Surface Weather Observations* (RUSSWO). Similar summaries are issued by the Naval Weather Service as *Summary of Meteorological Observations, Surface* (SMOS) and *Summary of Synoptic Meteorological Observations* (SSMO).

Table 16-24 is a sample of the frequencies (in tenths) of *sky cover*, from clear (0/10) to overcast (10/10) at a single station (Bedford, Mass.) in a single month (January) for eight 3-h periods of the day. It also shows the average sky cover, but does not subdivide the cloud cover by height of clouds. Table 16-24 also shows the frequencies for the mid-season months for all hours combined.

Table 16-24. Sample percent frequencies of sky cover, by tenths. (example is for Bedford, Mass).

Hours	January Sky Cover (tenths)											Avg. Sky Cover (tenths)
	0	1	2	3	4	5	6	7	8	9	10	
00–02	34	2	3	3	2	1	2	3	2	2	46	5.5
03–05	32	2	3	3	2	2	3	3	3	2	45	5.7
06–08	21	5	4	4	3	2	2	3	5	4	47	6.3
09–11	17	5	5	4	4	2	2	3	6	4	48	6.5
12–14	16	4	5	4	4	3	3	4	6	5	46	6.6
15–17	16	6	5	5	3	3	3	4	5	5	45	6.4
18–20	23	3	5	4	3	2	3	3	4	2	43	5.7
21–23	32	3	4	4	2	1	3	3	4	2	42	5.5
Mid-season Month	Seasonal Sky Cover											
Jan	24	4	4	4	3	2	3	4	4	3	45	6.0
Apr	18	4	4	5	4	3	3	4	7	4	44	6.4
Jul	15	5	6	7	6	5	5	6	9	5	31	5.9
Oct	30	5	5	5	4	3	3	4	5	4	32	5.0

In the RUSSWO, as in many summaries, the frequencies are given for ceiling heights jointly with the visibility in matrix form. Table 16-25 is a sample of ceiling height information as gleaned from eight such matrices for Bedford, Mass., one for each 3-hour period. Table 16-25 also shows the cumulative percent frequencies of ceilings for the mid-season months for all hours combined.

Table 16-25. Sample percent frequencies of ceiling heights equal to or higher than, including ceiling unlimited (example is for Bedford, Mass).

	January Ceiling (m)											
Hour	6100	4900	3700	3000	2400	1800	1200	900	600	300	150	30
00–02	55	55	58	60	62	66	72	76	80	86	92	98.4
03–05	51	52	55	57	60	63	70	74	80	86	93	98.3
06–08	48	48	53	56	60	63	70	74	79	86	94	98.9
09–11	50	51	56	59	63	66	71	75	80	86	94	99.8
12–14	50	51	56	60	64	66	70	75	80	86	94	99.4
15–17	52	53	58	61	64	67	72	77	81	86	93	99.2
18–20	56	57	60	62	65	67	71	76	81	87	94	99.2
21–23	57	57	60	62	64	67	72	77	81	87	94	99.1
Mid-season month	Seasonal Ceiling (m)											
	6100	4900	3700	3000	2400	1800	1200	900	600	300	150	30
Jan	52	53	57	60	63	66	69	76	78	86	94	99.0
Apr	49	50	53	56	60	64	73	77	81	88	95	99.5
Jul	61	62	68	72	75	77	82	85	88	92	97	99.5
Oct	62	62	65	68	70	72	77	81	84	89	95	98.6

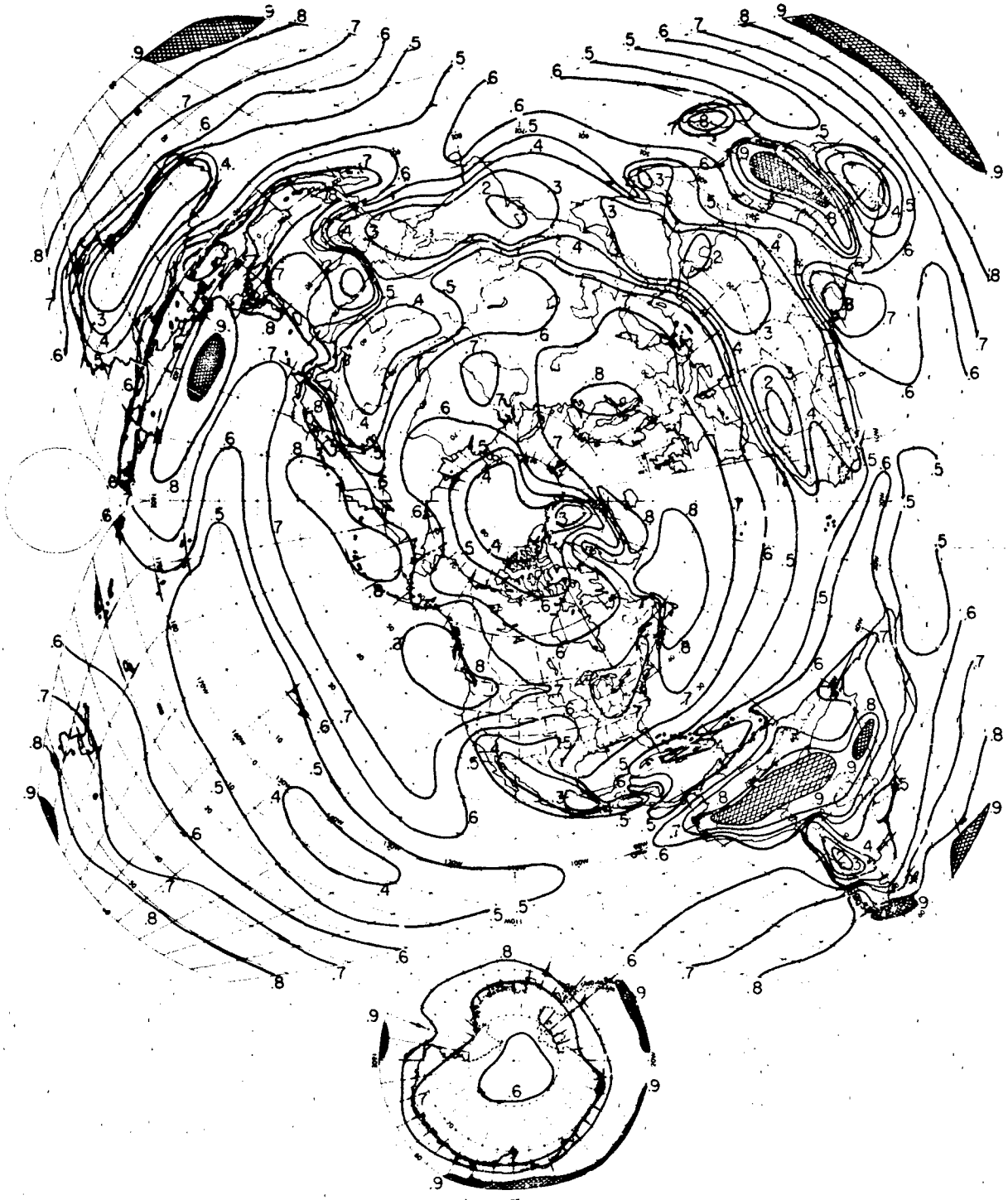


Figure 16-28. Mean sky cover, January, 1200–1400 LST.

Figure 16-29. Mean sky cover, July, 1200-1400 LST.

16 3.1.2 Global Sky Cover. While there are numerous tabulations of sky cover compiled from surface observations, there are relatively few maps. One of the sources for these maps is the chapter "Climatology" by Landsberg [1945] in the *Handbook of Meteorology*. The chapter has six world maps of sky cover, one for every other month beginning with January. Figures 16-28 and 16-29 show average cloudiness for 1200–1400 LST in January and July, respectively, based on Landsberg's charts; subsequently published tables of the National Intelligence Survey, RUSSWO, SMOS and SSMO; and data collected at the Woods Hole Oceanographic Institute for the Atlantic and Indian oceans.

Mean sky cover is significantly correlated with the frequency of clear and of overcast conditions. Table 16-26 presents a sample of RUSSWO information for some U.S. stations; it contains the percent frequencies of sky cover, from clear to overcast in January at noontime. Among these stations the mean sky cover ranges from 4.7 to 8.4 tenths. For low mean sky cover the frequency of clear skies is high and the frequency of overcast low. The opposite is true for a high mean sky cover. In extra-tropical regions the incidence of 1/10 to 7/10 sky cover remains generally low. However, the U-shaped distribution is less marked in the summer months (Table 16-24). In tropical and sub-tropical regions partial sky cover also is more frequent at the expense of clear and overcast. This is evident from the frequencies for Key West, Florida, and is particularly well illustrated by the data from Diego Garcia in the Indian Ocean at latitude 7° south (Table 16-26).

16.3.1.3 Modeling of Cloud Distributions. In the following presentations a distinction is made between the terms *cloud cover* and *sky cover*. Sky cover is an observer's view of cover of the sky dome, whereas cloud cover can be used to describe areas that are smaller than the floor space of the sky dome, or larger.

Cloud Cover Models. There have been several alternative mathematical formulas proposed for the probability distribution of sky cover. Each uses the variable (x) ranging from zero, for clear, to 1.0 for overcast. Each model is claimed to have versatile statistical characteristics to simulate U-shaped curves of sky cover. That is, large frequencies of clear and/or overcast skies, with small likelihoods of partial cover, can be fitted; yet the models also will fit the bell-shaped distributions in which the sky cover is mostly scattered to broken.

Alternative 1: The Beta distribution is an early cloud model [Falls, 1973] whose density function is given by

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\, x^{a-1}(1-x)^{b-1}; \qquad 0 \le x \le 1,\ a,b > 0\,. \tag{16.20}$$

Falls gives pairs of values of the two parameters in some 29 regional types that cover the world, for the four mid-season months, for two times of the day.

Alternative 2: More recently, Somerville and Bean [1979] have used a model called the S-distribution in which the cumulative probability F(x) of sky cover (x) is estimated by

Table 16-26. Percent frequencies of sky cover, by tenths, at a representative sample of stations. (For uniformity, the cases are all for January 1200 LST.)

Station	Sky Cover (tenths) 0	1	2	3	4	5	6	7	8	9	10	Mean Sky Cover (tenths)
Extra-tropical												
Phoenix, Ariz.	33	6	4	1	3	3	4	5	7	5	26	4.7
Washington, D.C. (Dulles Airport)	22	8	6	3	3	3	2	3	6	4	40	5.7
Bridgeport, Conn.	21	3	4	5	4	2	3	5	6	6	41	6.2
Austin, Texas	18	3	4	3	3	3	3	5	6	6	44	6.4
Chicago, Ill.	14	4	4	3	3	3	2	5	6	6	50	6.9
Great Falls, Mont.	11	4	3	3	3	2	3	5	8	8	50	7.3
Boise, Idaho	11	4	3	3	2	2	2	4	5	6	58	7.5
Niagara Falls, N.Y.	3	3	3	3	3	3	2	4	9	8	59	8.2
Seattle, Wash.	5	2	2	2	1	2	2	4	9	7	64	8.4
Sub-tropical and tropical												
Key West, Florida	14	9	10	12	8	8	4	7	9	7	12	4.7
Diego Garcia 7°18'S, 72°24'E	0	0	0	4	5	7	6	16	18	14	30	7.8

$$\hat{F}(x) = 1 - (1 - x^{\alpha})^{\beta}; \quad 0 \leq x \leq 1,\ \alpha,\beta > 0. \tag{16.21}$$

Pairs of values of the two parameters (α, β) have been determined to make the distribution $F(x)$ fit the data in the sky-cover summaries; they have been published for some 23 stations around the world, for each of eight periods of the day in each month of the year. The procedure has been to use 10 climatic cumulative frequencies $F(x_i)$ of sky cover for $x_i = 0.05(0.1)0.95$. A first guess is made for α, and β is then estimated by method of least squares:

$$\beta = \frac{\Sigma\{\ell n\,[1 - F(x_i)] \cdot \ell n[1 - x_i^{\alpha}]\}}{\Sigma\{\ell n[1 - x_i^{\alpha}]\}^2}. \tag{16.22}$$

After the pair of values (α, β) are entered into the basic equation to obtain the estimates $\hat{F}(x_i)$, the sum of squares of the differences between $\hat{F}(x_i)$ and $F(x_i)$ is obtained. The best values for the parameters (α, β) are those that minimize the sum of squares of errors. These were determined by trial and error, such that the best pair of values for the sky cover at Bedford, Mass., in January at noontime, were found to be $\alpha = 0.1468$, $\beta = 0.1721$. These estimates are different from those in Table 16-24 by rmse = 0.01.

In the above-mentioned models, strictly speaking, the probabilities of absolutely clear and of complete overcast are each computed to be zero; hence they are not truly realistic. The difficulty is overcome, in practice, by stipulating that the integration for clear be taken from zero cover to 0.5 tenths, for overcast from 9.5 tenths to 10 tenths. There is another difficulty, however; in that each solution is given for the sky dome as seen by a ground observer, considered to have a radius of 28 km. The above alternatives have not been developed to give cloud-cover distribution for varying areal coverage. However, simplicity is in their favor, especially with the S-distribution, which is readily programmed on a desk-top computer.

Alternative 3: A *"Model B"* for linear and areal coverage of a weather element has been described recently [Gringorten, 1979] which, like the above models, requires two parameters for a description of the probability distribution of cloud cover. The parameters in Model B have physical meaning. One parameter (P_o) is the mean cloud cover as given in RUSSWO and other climatic summaries; it is taken to be the single-point probability of a cloud intercept when looking up from the ground. The second parameter (r), known as the *scale distance*, is the distance between two stations whose correlation coefficient of cloud covers is 0.99. At Bedford, Mass., it varies from 0.5 km in summer noontime to as much as 10 km in winter midnight. Regrettably, as developed so far, estimations of the probability of cloud cover by this model must be made by use of 11 charts as published [Gringorten, 1979]. It does not yield estimates succinctly by formula. On the other hand, Model B avoids the difficulties of the previous two alternatives by providing finite estimates of all-clear and full overcast for varying areal extent. As an example, Table 16-27 shows the percent probability of cloud cover at Bedford, Mass., January noontime; the parameter values are $P_o = 0.66$, and $r = 2.6$ km. In Table 16-27 the cloud cover is given, not only as a ground observer sees it, but also for a small area of 100 km^2 and for a large area of 100 000 km^2. A small area is likely to be either all-clear (0/10) or overcast (10/10), but a large area is nearly always partially covered. For the ground observer's sky cover, the estimates differ from those in Table 16-24 by rmse = 0.02.

Ceiling Model. The *Burr curve* provides one of the best models for ceiling height cumulative distributions:

$$\hat{F}(h) = 1 - \{1 + (h/c)^a\}^{-b};\ a,b,c > 0, \tag{16.23}$$

where h is the ceiling height and a,b,c are parameters.

At the University of Florida [Bean et al., 1979] sets of values for a, b, and c have been determined (for each of eight periods of the day in each month of the year at some 23 stations around the world) to make the estimated distributions, $\hat{F}(h)$, fit the data for some 30 ceiling heights as

Table 16-27. An example of model estimates of percent probabilities of cloud cover, by tenths, depending upon size of the area (example is for Bedford, Mass., January 1200 LST).

Areal Size (km^2)	Cloud Cover (tenths)											Mean Cloud Cover (tenths)
	0	1	2	3	4	5	6	7	8	9	10	
100	29	1	1	1	2	1	1	2	2	1	59	6.6
2424 (Observer's sky dome)	16	4	5		3	3	5	4	6	9	41	6.6
100 000	0.5	3	4	7	9	9	12	13	16	19	8	6.6

given in the climatic summaries. In this procedure, a rounded figure is chosen for c, such as 300, 600, or 3 000 m. Trial estimates for a and b are then made to minimize the sum of squares of the differences between the model estimates $\hat{F}(h)$ and the cumulative frequencies F(h) of the climatic summaries. An initial trial estimate of a can be made by a linearization of the Burr curve through Taylor expansion, resulting in

$$a = \frac{\left[N \sum x_i y_i - \sum x_i \sum y_i\right]}{\left[N \sum x_i^2 - (\sum x_i)^2\right]}. \quad (16.24)$$

where $x_i = \ln(h_i/c)$, $y_i = \ln F(h_i)$, and N is the number of ceiling heights for which the climatic tables give the frequencies $F(h_i)$. Corresponding to each trial estimate of a,

$$b = \frac{\left|-\sum \ln\{1 - F(h_i)\} \ln\{1 + (h_i/c)^a\}\right|}{\sum\{\ln(1 + (h_i/c)^a\}^2}. \quad (16.25)$$

For Bedford, Mass., January noontime, values found were a = 1.1678, b = 0.1927 when c = 0.305 km, giving the estimates of probability that differed from those in Table 16-25 for 1200–1400 hours by rmse = 0.01.

16.3.2 Cloud Cover in Layers Aloft

Nephanalysis, concerning the type, amount, and scatter of clouds both in the horizontal and vertical directions, is important, particularly in connection with line-of-sight requirements. For example, knowledge of the probability of cloud-free lines-of-sight (CFLOS) from one level of the atmosphere to another at selected angles of elevation, can be vital to electro-optical communication systems. Although there is now a sizable body of literature on this subject, the statistics are none too satisfactory. Clouds are not only volatile but are often hidden by other clouds, so that an analysis of cloud presence, or a climatology of cloud frequencies as a function of level or layer above the ground is crude and speculative, especially if information is limited to the observations from the ground. An observer sees the "sky dome" with an uncertain radius, assumed to be 27 km in the following discussion and analysis.

A valuable study on the frequency of occurrence of clouds in the vertical was based on observations from aircraft over Germany [deBary and Moller, 1963]. It resulted in estimates of the average cloud cover as a function of altitude and the thickness of the atmospheric layer. Recently a model has been developed that yields similar estimates of cloud cover, without using aircraft observations. The model uses only RUSSWO statistics on the total sky cover and ceiling heights [Gringorten, 1982]. Figure 16-30 illustrates the output of the model in December, noontime for Grissom AFB, Indiana. The abscissa is for the cloud cover from 0 (clear) to 1.0 (overcast). The ordinate is for elevation above the ground, from 0 (surface) to 10 km aloft. The broken curve is drawn for the mean cloud cover at levels, and the solid curves are drawn for the mean cloud cover in layers. The broken curve shows, for example, mean cloud cover of 0.1 at 1 km, 0.14 at 2 km, 0.07 at 5 km. From each level the solid curve, when followed to a higher level, gives the average cloud cover in the layer between the two levels. For example, in the layer from 2 km to 4 km, the cloud cover averages 0.32, from 2 km to 8 km it averages 0.38, and so on.

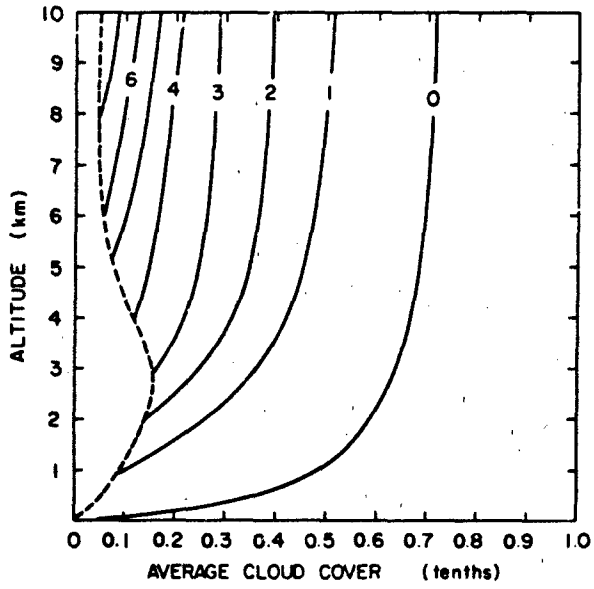


Figure 16-30. Average cloud cover by layers, December, 1200–1400 LST, Grissom AFB, Indiana

In similar manner, the distribution of cloud cover in layers aloft can be reconstructed for any station that has RUSSWO or similar records. Grissom AFB, Indiana, was chosen for illustration because of its relatively high average sky cover. With a lower average, a station should have lower frequencies in most or all of its layers aloft.

16.3.3 Satellite and Radar Observations

16.3.3.1 Satellites. Weather satellite images are providing detailed information on the structure and distribution of clouds over large areas of the earth. Figure 16-31 illustrates a variety of cloud and fog formations over Europe and Northern Africa; the image was obtained from one of the satellites of the Defense Meteorological Satellite Program (DMSP). In addition to images in the visible portion of the spectrum as shown in Figure 16-31, corresponding images

Figure 16-31. DMSP I-F low-enhancement IR ISCAT 13 October 0755 Local (1255 GMT)

in the thermal infrared are available. These simultaneous images or brightness fields can be used to deduce the height of the cloud tops. From the accumulation of such data, a regional and, eventually, a global cloud climatology will be developed.

16.3.3.2 Radar. Because of the electromagnetic scattering characteristics of water drops and ice particles, detection of non-precipitating (less than 200 μm diameter), hydrometeors in clouds is effectively limited to the shorter radar wavelengths (less than a few cm). For example, radars operating at 0.86-cm wavelength have been used for operational cloud detection and continue to be used in research for deducing physical and dynamical characteristics of clouds prior to the development of precipitation. Precipitation-sized particles are readily detected by longer wavelength (3–10 cm) radars, as discussed in Section 16.2.5. Optical and infrared wavelength (0.5 to 15 μm) radars, or lidars, have also been used to observe clouds; theoretical comparisons of optical and microwave cloud detection capabilities have been made by Derr [1978].

Observations by the vertically pointing TPQ-11 radar were recorded by the Air Weather Service between 1964 and about 1975 at 43 locations around the world. The multiple-layer ceiling observations are archived at the National Climatic Center, Asheville, N.C. Two of these radars, acquired by the NOAA Wave Propagation Laboratory and the University of Washington, respectively, have been modified for improved reliability and sensitivity and are used in multisensor field programs. Early results were presented by Weiss et al. [1979], and by Pasqualucci and Miller [1981]. The Wave Propagation Laboratory subsequently built a transportable fully-scanning coherent 0.86-cm radar with a dual-polarization receiver which was first used in the spring of 1981. The dual polarization capability permits derivation of hydrometeor microphysical characteristics, while the coherent reception permits measurement of Doppler winds in clouds. There is increasing interest in the detection of clouds and aerosols by 0.3 cm radar [Lhermitte, 1981].

16.3.4 Clear and Cloud-Free Lines-of-Sight

To determine the utility of various communications, surveillance, and weapons systems, the probability of clear/cloud-free lines-of-sight (CLOS/CFLOS) must be known for the geographical areas and climatic regimes in which these systems must operate. (CLOS is defined as LOS not obscured by either haze or clouds). Since CLOS and CFLOS are neither directly observed nor measured on a routine basis, various methods have been developed to estimate their frequency of occurrence. Some of these methods are described in this section and examples of data, analyses, and results are provided on the frequency of occurrence of CLOS and CFLOS.

16.3.4.1 Aircraft Observations. In a test program, Bertoni [1967] determined that realistic estimates of the probability of clear and cloud-free lines-of-sight could be derived from a large sampling of aircraft in-flight observations. This resulted in a major effort in which observations were made from aircraft, at various angles to the sky ($+30°$ and $+60°$), the horizon (0°), and the earth's surface ($-30°$ and $-60°$). Observations were acquired over most of the northern hemisphere during all months, but were limited to the daylight hours. (Nighttime observations tended to overestimate clear conditions).

Figure 16-32 is an example of the analyses shown in the resulting report [Bertoni, 1977a,b]. The large numbers in the figure are percent probabilities of CFLOS; the smaller numbers, in parentheses, indicate the total number of observations taken within each of the 10° latitude-longitude sectors. Figure 16-33 is similar to Figure 16-32 in that it shows probabilities of CLOS which are generally smaller than probabilities of CFLOS. This is illustrated more dramatically in Figure 16-34 in which frequencies of CLOS and CFLOS are compared for an area around Columbia, Mo, about 1000 km^2. Values for CLOS are reduced by some 20% to 30%, due mostly to haze. Figure 16-34 also confirms that the probabilities of CLOS and CFLOS depend on the angle of the line-of-sight as well as the percent of sky cover. CLOS and CFLOS probabilities increase as the lines-of-sight approach the vertical.

16.3.4.2 Surface Observations. Whole-sky photographs at Columbia, Missouri and conventional National Weather Service observations of sky cover were used by Lund and Shanklin [1973] to develop a universal model for estimating probabilities of CFLOS through the entire atmosphere. Earlier CFLOS studies based on these same observations are described by Pochop and Shanklin [1966], Bundy [1969], Shanklin and Landwehr [1971], and Lund and Shanklin [1972].

Probabilities of CFLOS based on the Lund and Shanklin model can be estimated through the use of the following formula:

$$_{\alpha}\hat{P}_1 = {}_{\alpha}C_s K_1, \qquad (16.26)$$

where $_{\alpha}\hat{P}_1$ is a column vector of α rows for each angle considered; $_{\alpha}C_s$ is a matrix for α rows and s columns, one row for each elevation angle, one column for each sky cover category; and $_sK_1$ is a column vector of s rows. The P̂ values are estimates of probabilities of CFLOS through the atmosphere, the C values are probabilities of CFLOS at angles α given k tenths of sky cover and the K values are probabilities of each k tenths of sky cover. The $_{\alpha}C_s$ matrix is contained in Table 16-28.

Using the historical records for Columbia, Missouri as an example, the following probabilities of 0/10, 1/10, . . . , 9/10, 10/10 sky cover at Columbia are 0.187, 0.047, 0.047, 0.039, 0.037, 0.031, 0.045, 0.045, 0.055, 0.065, 0.392.

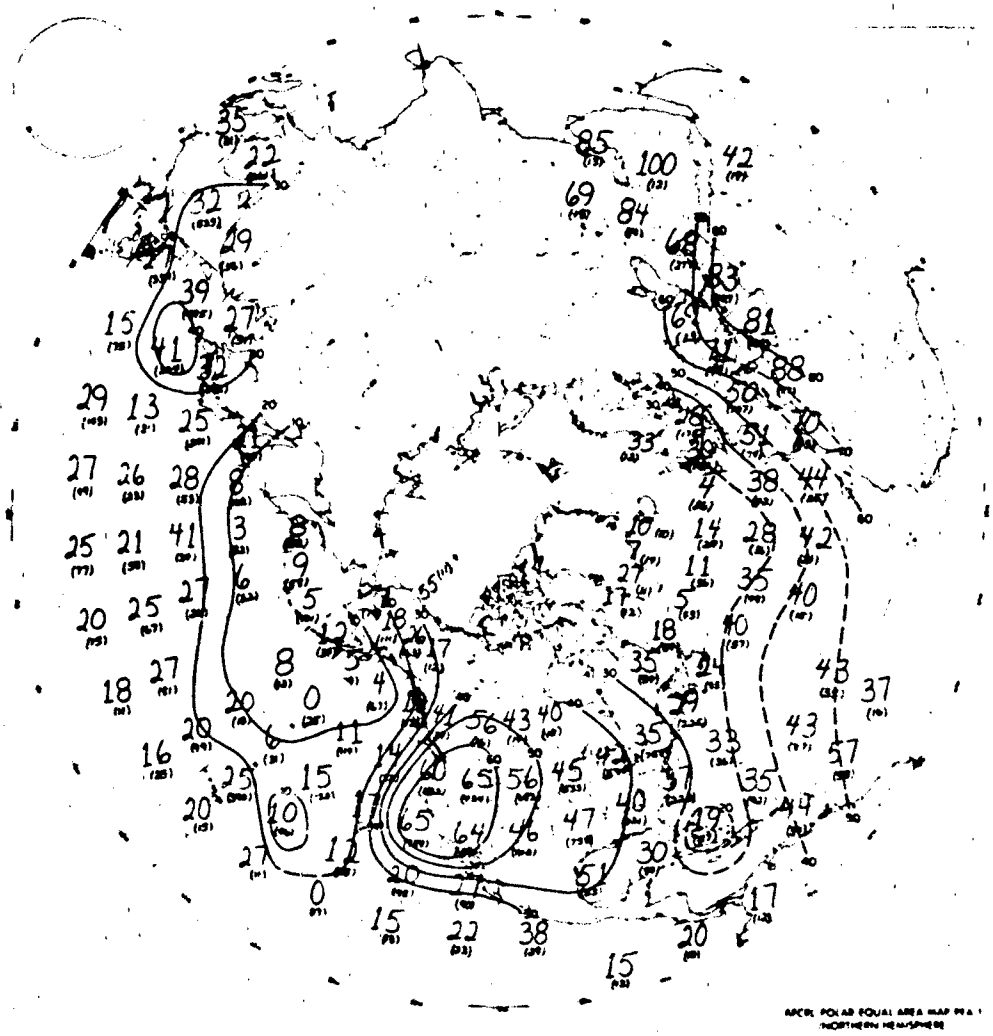


Figure 16-32. Estimates of the probability of a cloud-free line-of-sight to the ground at 30° below the horizon (−30°) from altitudes between 7.6 and 10.6 km in summer

Then the matrix multiplication indicated in Equation (16.26) yields

Elevation Angle (deg)	Probability
90	0.509
80	0.506
70	0.504
60	0.500
50	0.496
40	0.483
30	0.463
20	0.437
10	0.392

Consequently, there is an estimated 50.9% probability of a CFLOS at Columbia looking toward the zenith (90°), and a 50% and 46.3% probability of a CFLOS at 60° and 30° elevation angles, respectively. (See Section 16.3.4.3 for discussion of a correction to a suspected bias toward higher CFLOS probabilities produced by the Lund-Shanklin model.)

Using this model, atlases of CFLOS probabilities have been compiled for many parts of Europe, Asia and Africa, as well as for the U.S. and the USSR. (Lund et al., 1975, 1976, 1977, 1978, and 1979). A sample page from the atlas for the United States of America (Lund et al., 1977) is shown in Figure 16-35. This shows the probability of having a CFLOS in July at 1200 to 1400 LST at a 30° elevation angle. This atlas is based on at least a 15-yr period of record for the majority of U.S. stations shown by the dots in Figure 16-35.

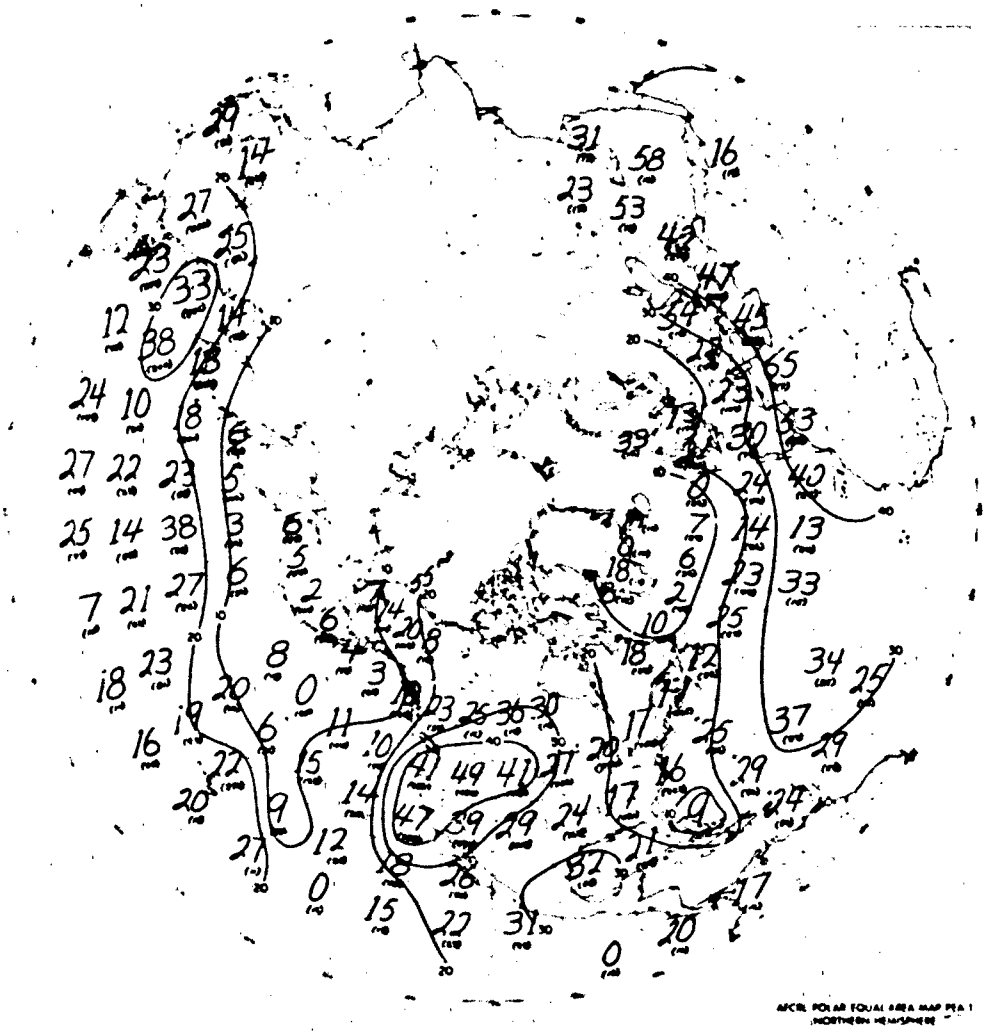


Figure 16-33. Estimates of the probability of a clear line-of-sight to the ground at 30° below the horizon (−30°) from altitudes between 7.6 and 10.6 km in summer.

Table 16-28. Probabilities of CFLOS as a function of elevation angle and observed Total Sky Cover [Lund and Shanklin, 1973]

Elevation Angle (deg)	Sky Cover (Tenths)										
	0	1	2	3	4	5	6	7	8	9	10
90	1.00	0.97	0.92	0.87	0.81	0.77	0.70	0.62	0.48	0.31	0.08
80	0.99	0.97	0.92	0.87	0.81	0.77	0.69	0.61	0.47	0.31	0.08
70	0.99	0.97	0.91	0.86	0.80	0.76	0.68	0.61	0.47	0.30	0.08
60	0.99	0.96	0.90	0.85	0.80	0.75	0.66	0.60	0.46	0.29	0.08
50	0.99	0.96	0.90	0.85	0.78	0.73	0.64	0.58	0.45	0.29	0.08
40	0.99	0.95	0.88	0.83	0.76	0.71	0.62	0.55	0.42	0.27	0.07
30	0.98	0.93	0.86	0.80	0.73	0.66	0.57	0.50	0.38	0.24	0.06
20	0.98	0.90	0.83	0.75	0.67	0.59	0.50	0.42	0.33	0.21	0.05
10	0.97	0.86	0.76	0.65	0.55	0.47	0.39	0.32	0.24	0.16	0.03

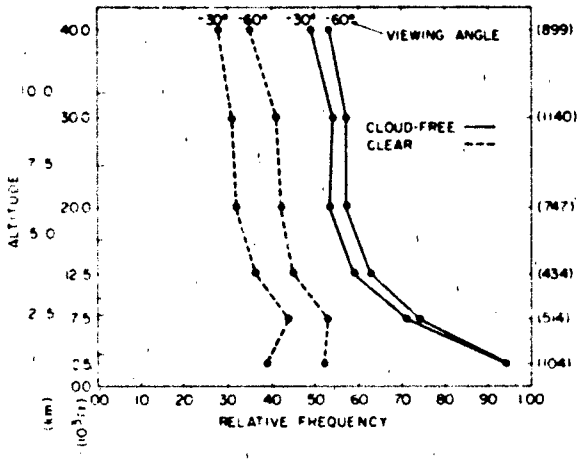


Figure 16-34. Relative frequencies of clear or cloud-free lines-of-sight observed in the area between 34° to 43°N and 88° to 97°W as a function of altitude. (The number of observations is shown in parentheses on the right hand margin.)

Rapp, Schutz, and Rodriques [1973] extended the Lund-Shanklin CFLOS model to include finding the probability of a CFLOS between the earth and any given altitude in the atmosphere. They accomplished this by using 3-h synoptic reports of clouds in addition to total sky cover.

16.3.4.3 Comparison of Estimates Between Aircraft and the CFLOS Model Based on Surface Observations. CFLOS probabilities computed from sky-cover observations taken at 6 stations (Columbia, Mo., Atlanta, Ga., Fort Worth, Tx., North Platte, Neb., Alburquerque, N.M., and Ely, Nev.) have been compared with CFLOS probabilities based on observations taken from aircraft by Lund and Bertoni [1980].

CFLOS probabilities were computed at selected altitudes for each of the 6 stations, each midseason month, and each depression angle. The probabilities for the four midseason months were averaged and plotted, and were usually higher than those based on inflight observations. Figures 16-36 and 16-37 show the model and inflight CFLOS probabilities for

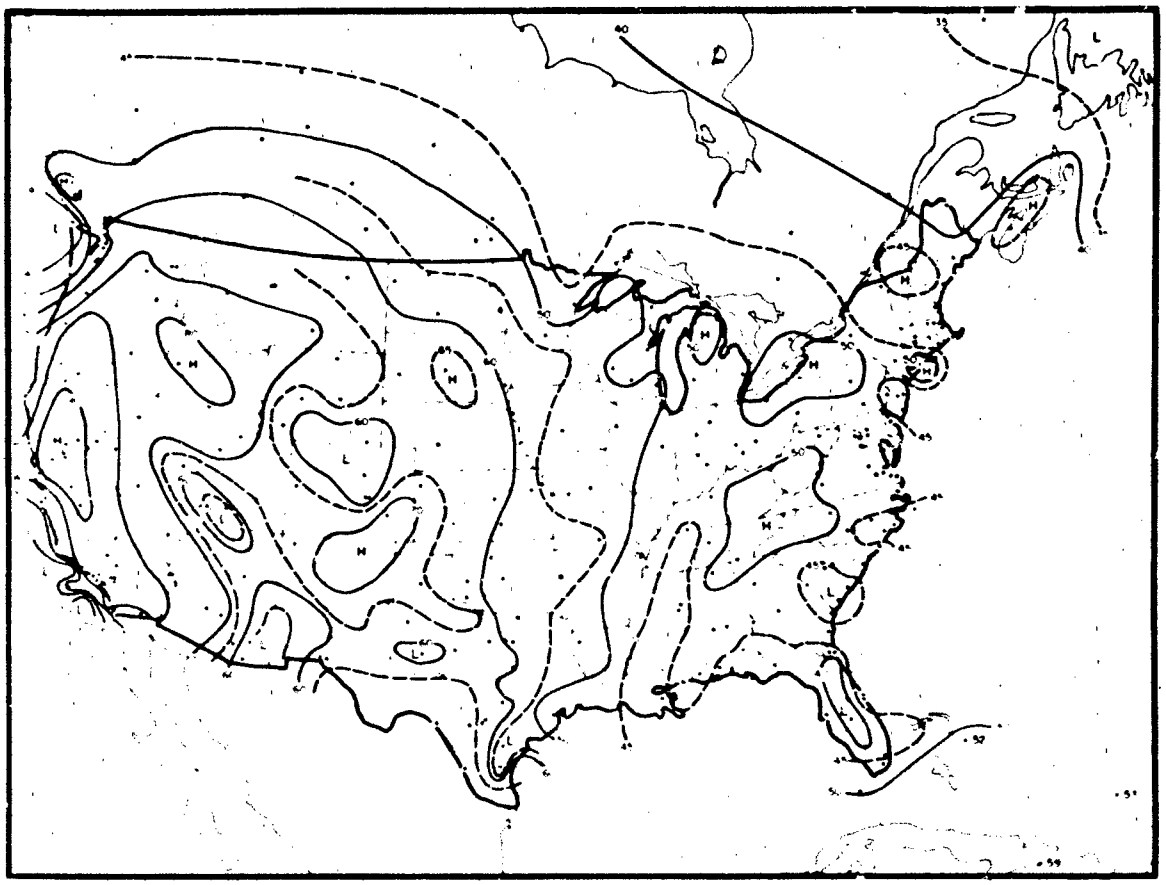


Figure 16-35. CFLOS probabilities for July, 1200-1400 LST, 30° elevation.

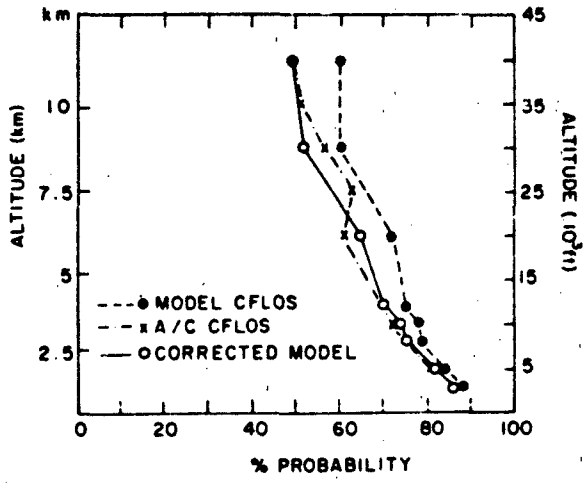


Figure 16-36. Six-station average probabilities of a cloud-free line-of-sight determined from aircraft observations (X's), from the CFLOS model (●'s), and from the model corrected for bias (○'s) at a depression angle of 30°.

depression angles 30° and 60°, respectively. The differences suggest that the Lund-Shanklin model CFLOS probabilities are biased toward higher values. This apparent bias can be essentially eliminated by a linear reduction of the model estimates as follows:

$$\hat{P}' = (0.99 - 0.0045a)\hat{P} \; ; \; 0 \leq a \leq 45, \quad (16.27)$$

where "a" is the altitude in kilometers and $\hat{P}$ is the model estimate. Figures 16-36 and 16-37 show the corrected CFLOS

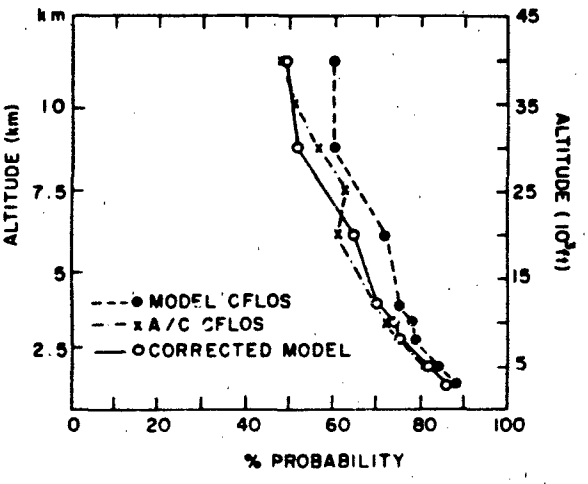


Figure 16-37. Six-station average probabilities of a cloud-free line-of-sight determined from aircraft observations (X's), from the CFLOS model (●'s), and from the model corrected for bias (○'s) at a depression angle of 60°.

estimates as well as the model and inflight CFLOS probabilities.

16.3.4.4 Cloud-Free Fields-of-View. An instantaneous CFLOS is often not sufficient for acquiring and tracking a target, or for estimating the probability of detecting a target when there is an opportunity to search for favorable cloud conditions. Therefore, probability distributions of cloud-free areas of fields-of-view (FOV) are needed. Grantham et al. [1979] and Lund et al. [1980] completed studies for estimating the probability of cloud-free fields-of-view (CFFOV) between earth and an airborne or space platform. Using whole-sky photographs taken at Columbia, Mo., the average cloud-free fraction of each of 185 annular sectors [see Figure 16-38] was computed, plotted, and analyzed to determine how it varied with position on the photograph. Figure 16-38 depicts the average cloud-free fraction of each sector when 2/10 total cloudiness was observed. Clearly, the average cloud-free fraction decreases from the center of the picture, which corresponds to directly overhead, to low elevation angles of the edges of the picture. The maxima (highs) are not exactly centered at the zenith (the center of the picture) as one would expect. The failure of the isolines to be symmetric about the center of this figure is due in part to sampling variability. However, other studies of the photographs have suggested that sun angle or local effects, perhaps from the city of Columbia, may also contribute to the lack of symmetry.

Circular fields-of-view centered at a ground observers zenith were investigated. Figure 16-39 shows circles with

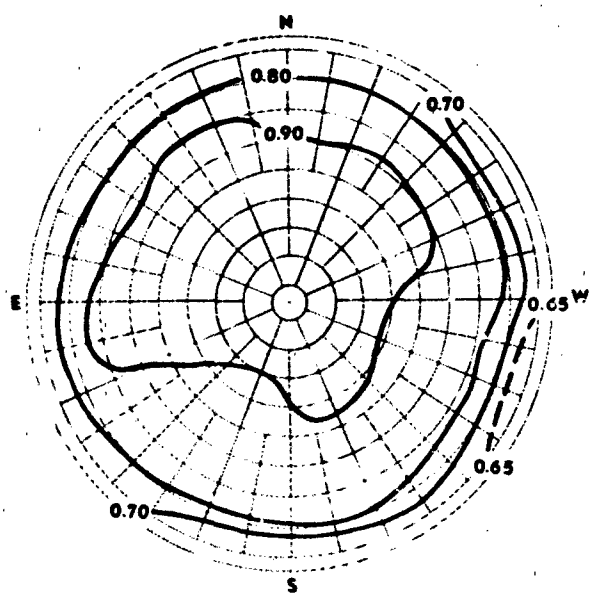


Figure 16-38. Average fraction of each annular sector that was cloud-free on the 131 photographs when 2/10 sky cloud cover was observed.

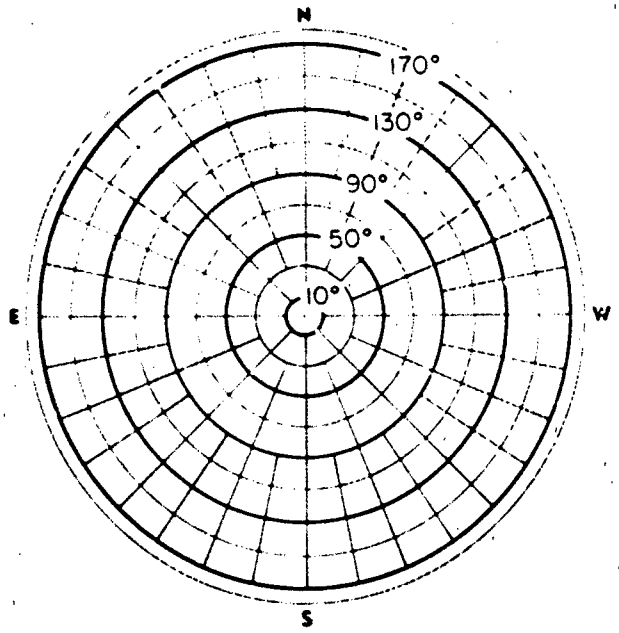


Figure 16-39. Circular fields-of-view with angular diameters of 10°, 50°, 90°, 130°, and 170° centered at the ground observer's zenith.

angular diameters, centered at the observer's zenith, of 10°, 50°, 90°, 130°, and 170°. Because the probability distributions are functions of the sky cover, they were stratified on that basis.

Figure 16-40 depicts cumulative frequency distributions of the fraction of the area that is cloud free in the 10° circle as a function of the observed sky cover. The points on the figure were obtained from the data and the curves were subjectively drawn through the points. This figure shows, for example, that if one wants to be 80% sure that at least 0.5 of the 10° FOV is cloud free, operation can be allowed only when 0/10, 1/10, 2/10, 3/10, or 4/10 sky cover is observed. If he relaxes his requirements to a 50% probability that at least 0.2 of the 10° FOV is cloud free, he can additionally operate when 5/10 or 6/10 sky cover is observed.

Figure 16-41 shows how the cloud-free fraction varies for five fields-of-view (10°, 50°, 90°, 130° and 170°) when the weather observer reports 5/10 sky cover. Let us assume that at least 0.2 of a FOV must be cloud free. The curve labeled 10° shows there is a 73% chance that at least 0.2 of the FOV is cloud free. The curve labeled 90° shows a 94% chance, and the curve labeled 170° shows a 100% chance that at least 0.2 of the 170° FOV will be clear when the observer reports 5/10 sky cover. If, however, the requirement is that at least 0.8 of the area must be cloud free, the probability will be near zero for the large 170° FOV and will increase to 44% for the small 10° FOV.

Lund et al. [1980] modified the Lund-Shanklin CFLOS model described in Section 16.3.4.2 to estimate climatic

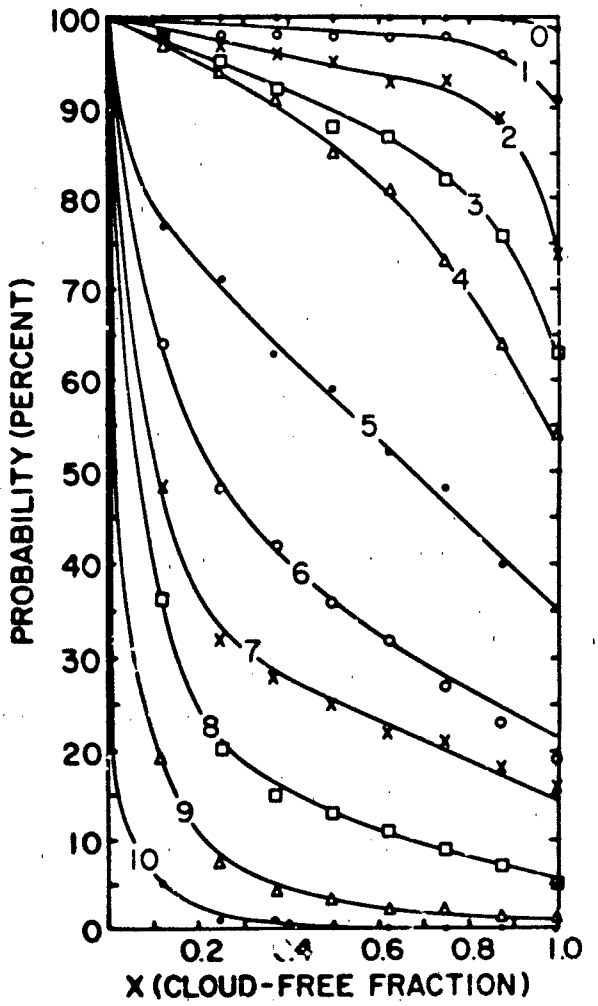


Figure 16-40. Probability that a 10° FOV at zenith will have a cloud-free fraction ≥x (curves for each tenth of sky cover were subjectively drawn to the data points).

probabilities of CFFOV from climatic probabilities of the tenths of cloudiness. Figure 16-42 shows a comparison of the probability of a CFLOS as a function of elevation angle when two-tenths sky cover is reported, with corresponding CFFOV probabilities when the FOV is 10° and 20°. The figure shows good agreement between CFLOS probabilities and FOV probabilities for cloud-free fractions ≥0.5 and ≥0.75. However, there is only a 61% chance that a 10° FOV centered at a 30° elevation angle will be totally cloud free, and only a 48% chance for a 20° FOV to be totally cloud free. These differences between the CFLOS and CFFOV model probabilities are larger than expected. Therefore, users of the CFFOV model and the CFLOS model should understand that the models yield probability estimates which are believed to be good first approximations. Better approximations await higher quality whole-sky photographs or more precise observing techniques.

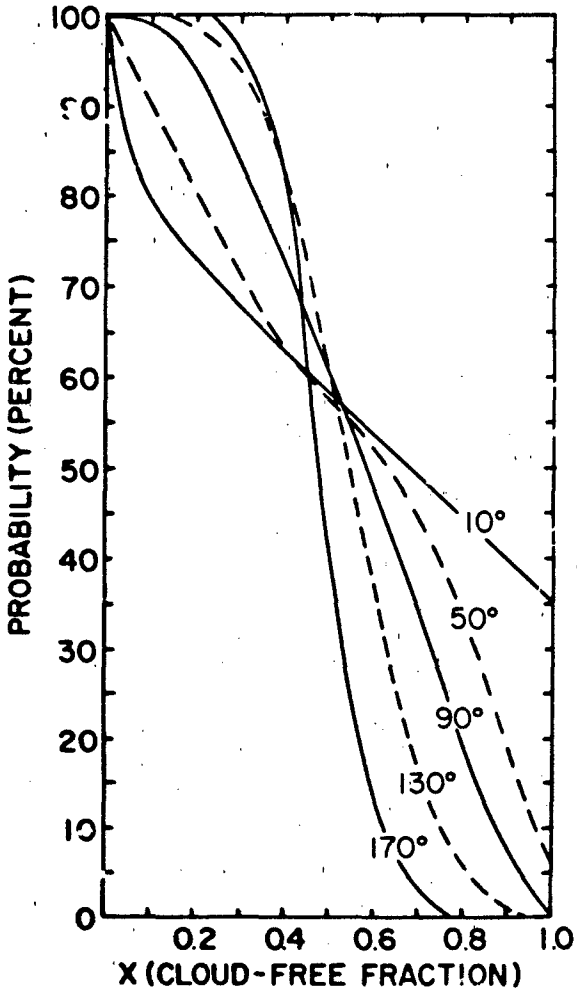


Figure 16-41. The probability that a 10°, 50°, 90°, 130°, and 170° FOV at zenith will have a cloud-free fraction ≥x when the weather observer reports 5/10 sky cover.

16.3.4.5 Joint Probabilities of CFLOS. Lund [1973] extended the Lund-Shanklin [1973] CFLOS model described in Section 16.3.4.2 to address the problem of estimating joint probabilities of cloud-free lines-of-sight from more than one observing sight. This model requires a climatic record of sky cover observations taken simultaneously from all sites.

Table 16-29 gives estimated CFLOS probabilities for Minot, N.D., $P(C_M)$: three-way joint probabilities of CFLOS for Grand Forks, Minot, and Fargo, $P(C_G\ C_M\ C_F)$; and, three-way joint probability for none of the three sites, $P(\overline{C}_G\ \overline{C}_M\ \overline{C}_F)$. The probabilities apply to August, the month when the highest CFLOS probabilities occur, and the elevation angle is 40°. Table 16-30 is the same as Table 16-29 except that probabilities apply to November, the month when the lowest CFLOS probabilities occur.

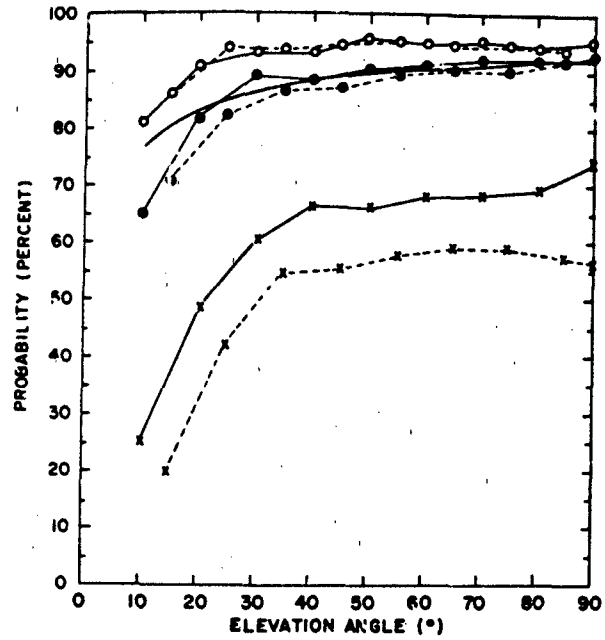


Figure 16-42. The probability of a CFLOS as a function of elevation angle when 2/10 sky cover is reported (heavy curve), and corresponding CFFOV probabilities when the FOV is 10° (solid lines) and 20° (dashed lines.).

16.3.5 Water Content of Clouds

Moisture in clouds can exist as a vapor, liquid, or solid. Some water vapor is always present in the atmosphere. The amount of water vapor a parcel of air can hold is a function of the temperature [Haurwitz, 1941]. Under conditions where the atmosphere becomes saturated the water vapor condenses and forms clouds.

Knowledge of the amounts of liquid water and ice contained within clouds is important to aviation due to its direct and indirect effects. The quantity "liquid water content" is the concentration, expressed in mass per unit volume, precipitation. Although many methods of determining liquid water content have been used, none are completely satisfactory.

Table 16-29. Estimated 3-way joint probabilities of CFLOS at 40° elevation angle for August at Minot $P(C_M)$; Grand Forks, Minot, and Fargo, $P(C_G\ C_M\ C_F)$; none of the three sites, $P(\overline{C}_G\ \overline{C}_M\ \overline{C}_F)$.

	Hour							
	0	3	6	9	12	15	18	21
$P(C_M)$	0.73	0.74	0.65	0.62	0.59	0.58	0.60	0.67
$P(C_G\ C_M\ C_F)$	0.48	0.48	0.37	0.35	0.34	0.31	0.35	0.41
$P(\overline{C}_G\ \overline{C}_M\ \overline{C}_F)$	0.07	0.08	0.12	0.13	0.15	0.14	0.13	0.10

Table 16-30. Same as Table 16-29 except the probabilities apply to November, the month when the lowest CFLOS probabilities occur.

	Hour							
	0	3	6	9	12	15	18	21
$P(C_M)$	0.50	0.49	0.50	0.36	0.38	0.39	0.40	0.49
$P(C_G\ C_M\ C_F)$	0.23	0.22	0.20	0.14	0.13	0 14	0.17	0.23
$P(\bar{C}_M\ \bar{C}_M\ \bar{C}_F)$	0.27	0.28	0.28	0.40	0.38	0.37	0.34	0.29

16.3.5.1 Liquid Water Content in Clouds. The actual amount of liquid water from droplets or ice crystals in a cloud will vary, depending on the type, altitude, and temperature of the cloud. Clouds formed by strong vertical motion will generally contain large amounts of liquid water. Lamb et al. [1976] measured as many as 1.5 g/m^3 in synoptic-scale winter clouds that contained convective elements. Generally, clouds of this type range in amounts that are illustrated in Figure 16-25. Aufm Kampe and Weickmann [1957] suggest 10.0 g/m^3 as the maximum value of liquid water content in cumulonimbus clouds. Higher values have been reported. The range of values and probability of occurrence are discussed in detail in Section 16.2.5.4.

In large-scale storm systems, liquid water content in most of the clouds is in the range 0.1 g/m^3. The highest values of liquid water content (see Figure 16-25) occur near the freezing level (or melting layer).

The values quoted in the preceeding paragraphs have been computed from particle-size distributions and assumptions relating crystal size and equivalent melted diameter. They should be considered as best estimates. Increased accuracy will be possible only when a reliable method of directly measuring water content is developed.

16.3.5.2 Measuring Liquid Water Content. Three methods of measuring liquid water content have been tried: direct measurement, computation from particle distributions, and attenuation of radar or lidar signals.

Although several possible instruments for direct measurement have been devised, only one, the Johnson-Williams (JW) indicator has been used on an operational basis. This instrument operates on the principle that the change in resistance of a heated wire as it evaporates impacting particles is proportional to the liquid water content. The probe is sensitive only to liquid droplets having diameters of ≤ 30 μm.

The most frequently used method of determining liquid water content is by integration of the particle-size spectra obtained by use of an optical array spectrometer. This procedure is straightforward when particles are liquid. However, when particles are ice crystals, assumptions concerning the relation of crystal size and their equivalent melted diameter are necessary. The AFGL Cirrus study [Cohen, 1981a] and Large Scale Cloud Systems study [Cohen, 1981b] used this method. Knollenberg [1972] notes that this method compared favorably to results given by direct measurement.

During recent years, attempts to relate liquid water content to the intensity of radar returns have been made. Stickel and Seliga [1981] describe one such attempt. At this time, however, radar is not able to give the definition that either direct measurements or computation from particle distributions can provide.

16.3.5.3 Supercooled Clouds and Aircraft Icing. The principal meteorological parameters associated with aircraft icing are the ambient temperature, the cloud liquid water content, and the drop size distribution. The extent to which ice forms on an aircraft (that is, icing intensity) and whether or not rime or glaze ice is formed depend additionally on the aerodynamic shape of the airplane structures, its true air speed, and the length of time the aircraft is in the icing environment.

The source of most meteorological information collected on icing parameters, including the data presented here, are from stratiform clouds (stratus, stratocumulus, and altostratus). Ninety percent of these clouds are 1 km thick [Figure 16-43]. Stratiform clouds are of principal concern with respect to icing since convective clouds (cumulus, cumulocongestus, and cumulonimbus) are generally avoided by aircraft for other reasons. Cumulus clouds are also relatively limited in horizontal dimension and duration. However, conditions in cumuliform clouds can result in severe icing

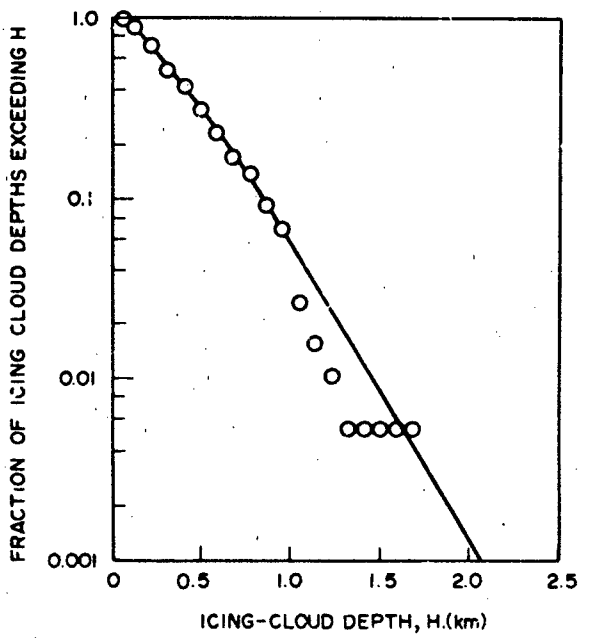


Figure 16-43 Cumulative frequency distribution of depth of icing-cloud layer [after Perkins, 1978].

of aircraft since supercooled water in excess of 2.0 g/m³ [Mason, 1971] have been observed as well as drop sizes occurring over the complete range of the size spectrum. Additional information concerning specific cloud types, associated synoptic conditions, and their relation to icing is given by Lewis [1951] and the Air Weather Service [AWS, 1980].

Clouds containing supercooled droplets (that is, water in the liquid state at temperatures below freezing) can exist in the atmosphere to temperatures as low as 233 K. However, the probability that supercooled droplets will remain unfrozen below the temperature range of 258 K to 253 K decreases rapidly since the number concentration of active freezing nuclei becomes significant at these temperatures. Perkins [1978] observed that approximately 90% of the clouds sampled that contained supercooled water were at temperatures warmer than 253 K. The likelihood of encountering clouds containing supercooled droplets as a function of temperature is shown in Figure 16-44.

The liquid water content (LWC) in wintertime stratiform clouds is a highly variable parameter. This is due principally to the horizontal variations in cloud structure. LWC averaged over long path lengths (>25km) usually do not exceed 0.3 g/m³. Altostratus and altocumulus clouds forming at higher altitudes have proportionately lower LWC amounts due to the temperature dependency of saturation vapor pressure of water. In individual stratocumulus cloud elements, with horizontal dimensions of 0.5 km to 1.0 km, peak LWC values of 1.3 g/m³ have been observed, although values not exceeding 0.8 g/m³ are more typical. The distribution (in percent) of supercooled water content amounts is shown in Figure 16-45. These data are based on sample measurements [Glass and Grantham, 1981] from a day considered representative of clouds during winter icing conditions. The distribution of median volume diameters for these data is shown in Figure 16-46.

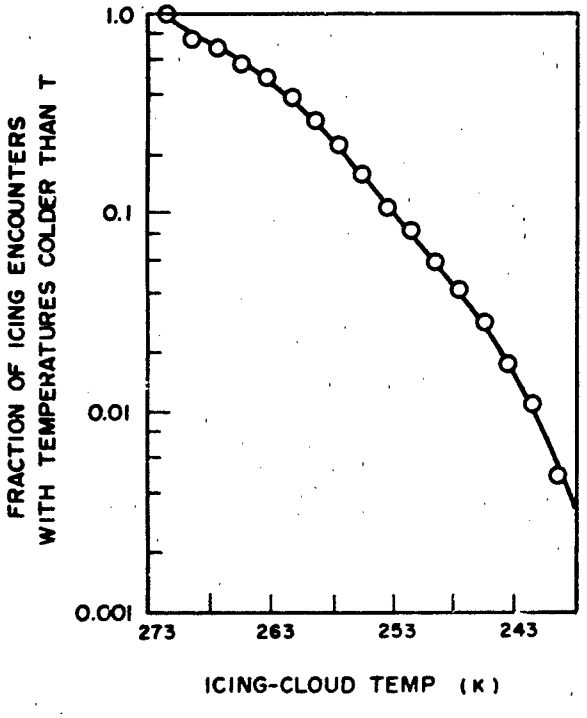


Figure 16-44. Cumulative frequency distribution of temperature of icing-clouds [after Perkins, 1978].

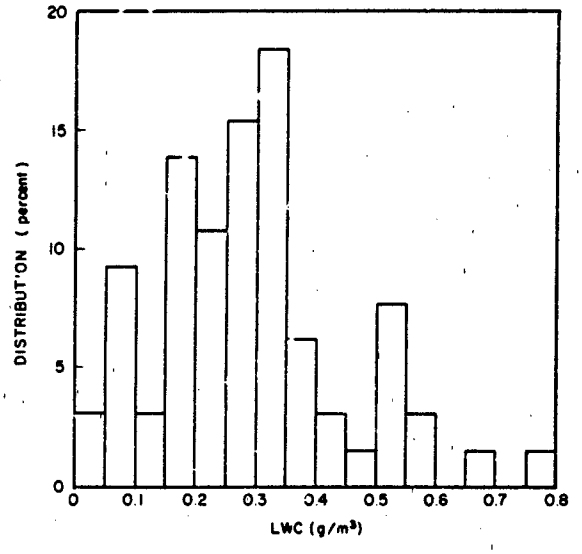


Figure 16-45. Distribution of LWC from 65 sample cases [Glass and Grantham, 1981].

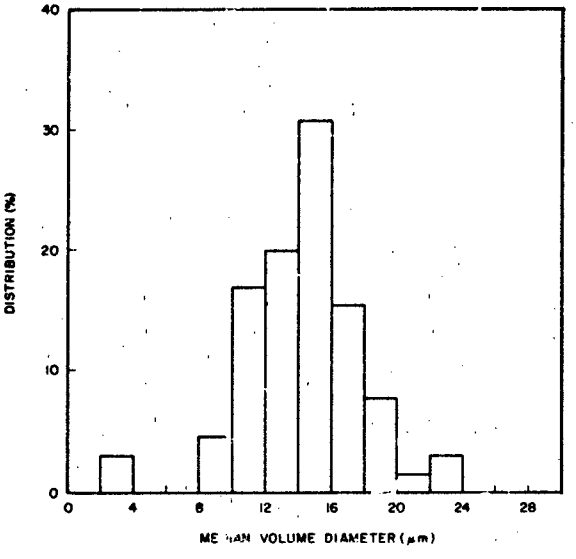


Figure 16-46. Distribution of average median volume diameter during each of 65 sample cases [Glass and Grantham, 1981].

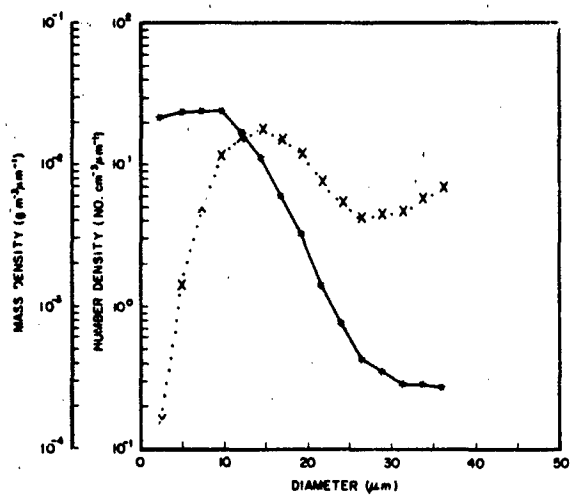


Figure 16-47. Number (*) and Mass (..x..) distribution with size for data sets with ≥0.2 g/m³ liquid water.

Distributions of the number and mass density of droplets with respect to size obtained from the same data set are shown in Figures 16-47 and 16-48. Both figures are based on long-duration data sets. Figure 16-47 is drawn from data sets where the average LWC was equal to or greater than 0.2 g/m³, whereas Figure 16-48 is based on data sets from which LWC values did not exceed 0.1 g/m³. Some parameters obtained from these distributions are shown in Table 16-31. Past observations [Perkins, 1978] of droplet-size distribution in supercooled stratiform clouds indicated that values can range up to 50 μm. More recent observations have suggested that large supercooled drops (approaching 300 μm in diameter) can be found in these clouds. Their number density, however, is much less than 1.0/cm³.

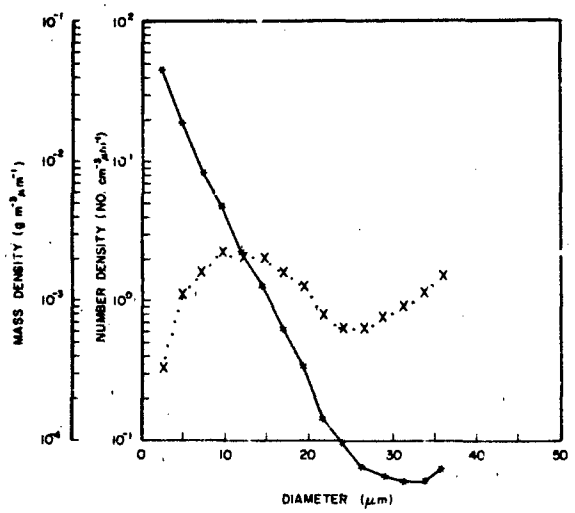


Figure 16-48. Number (*) and mass (..x..) distribution with size for data sets with <0.1 g/m³ liquid water.

Table 16-31. Characteristics of size and mass spectra for the indicated figures.

	Figure 16-47	Figure 16-48
LWC(g/m³)	0.27	0.04
Number density (cm⁻³)	321	192
Mean Vol Diameter (μm)	12	8
Median Vol Diameter (μm)	16	14

Techniques developed by the Air Weather Service [AWS,1980] for forecasting icing conditions are based on radiosonde observations of temperature and the dew-point spread (difference between temperature and dew point). These parameters are used since formation and maintenance of clouds containing supercooled droplet spectra require a saturated or near-saturated environment. This report quotes studies indicating that when only dew-point spread is considered, there is an 80% probability of icing when the spread is ≤3 K, and 84% probability of no icing when the spread exceeds 3 K. A northern hemisphere climatology of icing probabilities, utilizing the Air Weather Service analysis method [AWS,1980] and applied to 380 northern hemisphere radiosonde stations, has been published by the Air Weather Service [Heath and Cantrell, 1972].

16.3.6 Clouds above 6 Km

Almost all clouds found above 6 km are cirriform clouds composed of ice crystals. Cirriform clouds have been observed at altitudes in excess of 15 km.

16.3.6.1 Observations of Cirrus Clouds. In 1896–97, various observatories made double-theodolite measurements of cirriform clouds. Table 16-32a lists the average heights of the cloud bases in different seasons and at different latitudes. Table 16-32b shows the maximum base heights of cirrus and cirrostratus clouds at these locations [Suring, 1941; Appleman, 1961]. This table indicates that the bases of cirriform clouds are higher at lower latitudes and during the summer, and that the bases of cirrus clouds tend to be higher than those of cirrostratus layers. These observations were confirmed by aircraft flights in the 1950s. The flights made at that time also provided data on the heights of tops, thickness, and frequency of occurrence of cirriform clouds.

Cirriform clouds in the temperate zone are more common in cold weather and at high latitudes. Both bases and tops vary in height, being higher in summer and lower in winter. Tops of cirriform clouds are closely associated with the tropopause. Generally, the tops occur 1.2 to 1.5 km below the tropopause. The thickness of high clouds averages 2.1 to 2.2 km and shows little or no variation with season.

Table 16-32a. Mean heights of cirrus bases for international cloud year 1896–1897.

Latitude	Place	Northern summer (km)	Northern winter (km)
78°N	Cape Thordsen	7.3	—
70°N	Bossekop (Sweden)	7.5	—
60°N	Pavlovsk plus Uppsala	7.9	7.1
51°N	Potsdam plus Trappes (average)	8.7	7.7
40°N	Blue Hill plus Washington (average)	10.2	9.1
35°N	Mera (Japan)	11.0	9.1
14°N	Manila	12.0	11.1
7°S	Batavia	11.0	—
35°S	Melbourne	8.5	9.6

Table 16-32b. Maximum heights of cirrus bases for international cloud year 1896–1897.

Latitude	Place	Cirrus (km)	Cirrostratus (km)
78°N	Cape Thordsen	8.6	—
70°N	Bossekop (Sweden)	11.8	10.4
60°N	Pavlovsk plus Uppsala	11.7	10.1
51°N	Potsdam plus Trappes (average)	12.7	11.9
40°N	Blue Hill plus Washington (average)	15.0	13.6
35°N	Mera (Japan)	16.8	15.5
14°N	Manila	20.4	17.1
7°S	Batavia	18.6	14.2
35°S	Melbourne	—	—

In the 1970s, both particle density and liquid water content of high-altitude clouds were investigated in airborne studies of the microphysics of clouds. Aircraft studies of cirriform clouds have been reported by Heymsfield [1975], Ryan et al. [1972], and McTaggert-Cowan et al. [1970]. In more recent years, AFGL has published a series of reports on this subject [Cohen, 1981a,b].

AFGL research flights have shown that cirrus cloud particles are almost always smaller than 1200 μm in diameter. Heymsfield did find occasional particles as large as 1900 μm, but these were very rare, and generally their number density was less than 1/cm^3. Most cirrus particles are smaller than 600 μm. Varley et al. [1980] found some larger particles in the vicinity of a surface storm, but this was an exceptional case.

Although comparison of the work of different investigators is risky due to differences in instrumentation and presentation, an analysis of the results of three studies are

Table 16-33. Water content (melted ice) and particle density observations in cirriform clouds.

Observer	Date	Cloud Type	Altitude (km)	Water Content (Melted Ice) (gm/m^3)	Maximum Particle Density (Number/cm^3)
Heymsfield	9 Jul 70	Cirrus	7	0.30	Not reported
Heymsfield	4 Apr 72	Cirrostratus	8	0.03	0.20
Ryan	12 Jun 70	Cirrus	9	Not Reported	3.60
McTaggart-Cowan	18 Jun 69	Cirrus	12	Not Reported	0.53
AFGL	28 Jan 79	Heavy Cirrostratus	8	0.23	0.13
AFGL	2 Feb 79	Cirrostratus	9	0.03	0.11

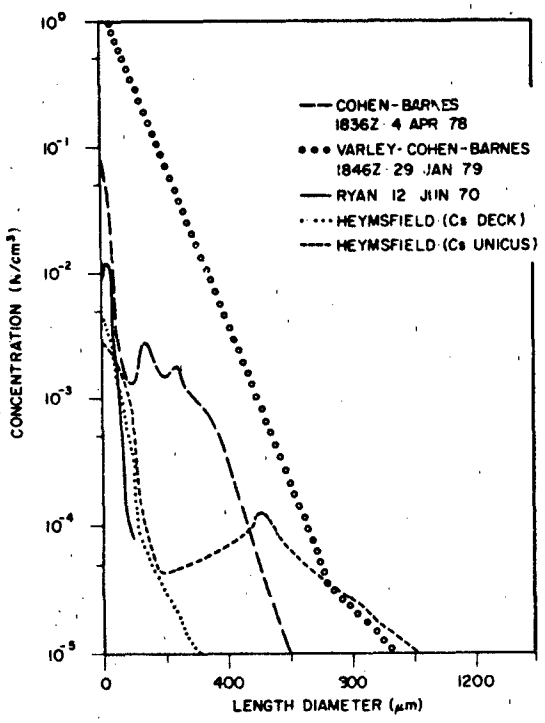


Figure 16-49. Cirrus particle concentrations seen by different observers.

shown in Figure 16-49. This figure gives an idea of different size distributions and is not intended to compare the results of the different studies.

Typical water content (melted ice) and particle densities reported by different observers are shown in Table 16-33.

16.3.6.2 Subvisible Cirrus. Cirrus-like ice particles occurring in apparently cloudless air are called subvisible cirrus. Thin cirrus clouds that are only visible from certain directions or at certain times may also be called subvisible cirrus. Most subvisible cirrus is found in the 3 or 4 km immediately below the tropopause. The maximum occurrence of cirrus, both visible and subvisible, is usually 1 to 2 km below the tropopause. Recent observations suggest that subvisible cirrus is present more than 75% of the time; one rarely finds cloudless areas in the cirrus region in which there are fewer than one particle greater than 2 μm in diameter per cm³.

There are two kinds of subvisible cirrus. The most common consists of small particles with diameters of less than 50 μm, with most particles being less than 20 μm in diameter (see Figure 16-50). Barnes [1980] describes the second type of subvisible cirrus as consisting of isolated ice crystals with diameters of 100 μm to 3000 μm. These particles may occur with the first type (see Figure 16-51), or they may be found in layers devoid of the smaller particles, having fallen from higher-level clouds (often cirrus unicus), where they were generated. In Arctic regions, these crystals

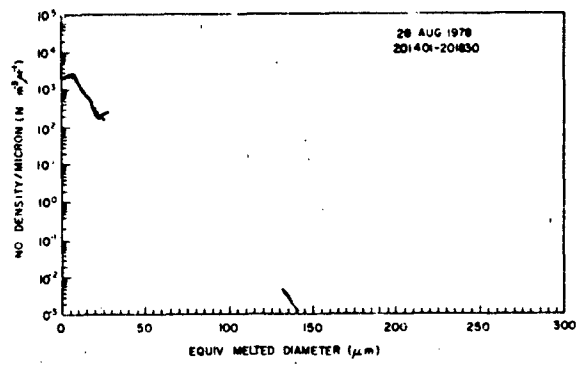


Figure 16-50. Particle size distribution of small-particle subvisible cirrus.

can fall to the ground [Kikuchi and Hogan, 1979] as "diamond dust" snowfalls. Cirrus particles can survive falls in excess of 2 km in subsaturated air [Hall and Prupacher, 1976].

Hall [1964] implies that subvisible cirrus can reduce the lift on laminar flow wings. Barnes [1978] notes that cirrus particles cause erosion of hypersonic vehicles, such as reentry vehicles, and that the presence of unreported thin and subvisible cirrus reduces the efficiency of solar collectors and attenuates energy in laser beams [Barnes, 1982].

16.3.6.3 Clouds Above the Tropopause. While almost all of the water and ice clouds in the atmosphere are in the troposhere, there are some occasions in which clouds penetrate into the stratosphere. This is generally the result of strong convective activity in which severe thunderstorms, caused by unstable conditions and high moisture content, have enough energy to break through the tropopause (Figure 16-52). In this vein, the frequency and percent occurrence of radar precipitation echoes have above 15 km have been estimated [Kantor and Grantham, 1968]. This climatology of radar precipitation echoes also provides information on the month of occurrence and estimated vertical extent of

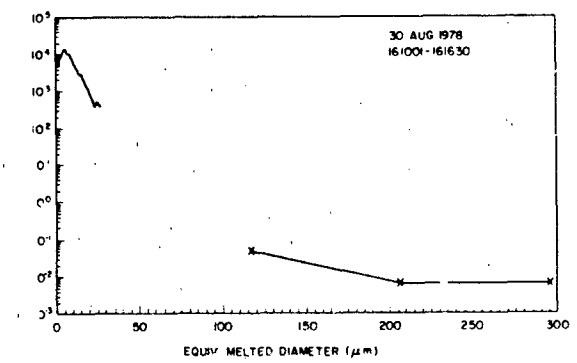


Figure 16-51. Particle size distribution showing both types of subvisible cirrus.

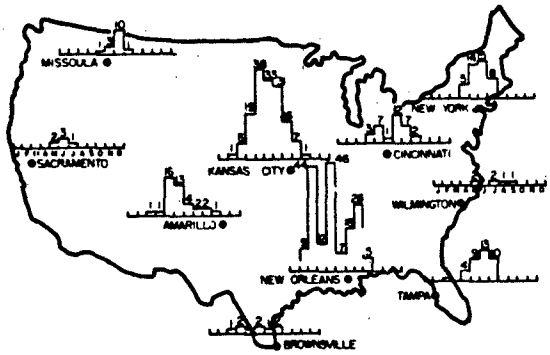


Figure 16-52. Average monthly tropopause penetrations by thunderstorms, 1961–1964 [Long et al., 1965].

cumulonimbus clouds over the United States. As a further example, Table 16-34 indicates that cirrus tops can build well beyond the tropopause. The cirrus produced by thunderstorms can often extend about 185 km downwind of the cell that produced it. Thus, stratospheric clouds, while not common, do exist and can be expected in and near regions of strong convective activity.

The size and shape of high altitude convective clouds have been estimated [Long et al., 1965; Roach, 1967]. For thunderstorms that have penetrated the tropopause in the midwestern United States, the shape of the cloud is a cone with a diameter-to-height ratio of roughly five to one for the portion of the cloud above the tropopause. However, much of the cirrus blowoff (shield) that is observed from the ground remains below the tropopause. Mean diameter of high altitude convective cells, including the associated anvil, has been estimated at 32 to 50 km.

16.3.6.4 Tropopause Height. Because there are seasonal, latitudinal, and daily variations in the height of the tropopause, both the latitude and season of the year must be considered in studying the height of the tropopause. Table 16-35 gives mean seasonal heights of the tropopause in the Northern Hemisphere; however, mean seasonal heights may be several kilometers higher than indicated in Table 16-35 on the east side of semistationary lows such as the Icelandic and Aleutian lows during the winter and spring months.

In middle latitudes the day-to-day variations in the height of the tropopause exceeds the seasonal variations; large variations in height occur as a result of horizontal oscillations in the tropopause. The arctic tropopause from the north is brought southward in the rear, and the tropical tropopause from the south is projected northward in advance of well-developed migratory cyclones. Although data are not sufficient to establish exact limits, available data indicate that in middle latitudes (40° to 60°N) the tropopause can fluctuate between about 6 and 15 km. North of 60° and south of 30° latitude, the range is much smaller, the order of 900 to 1800 m around the seasonal mean heights.

Table 16-34. Distribution of cirrus tops relative to the tropopause.

Distance From Tropopause (m)	Frequency of Occurrence (%)
British Isles	
>1200 above	10
1200 to 0 above	10
0 to 1200 below	50
1200 to 2400 below	16
>2400 below	14
Canada	
>1500 above	1
1500 to 0 above	12
0 to 1500 below	52
1500 to 3000 below	23
<3000 below	12

16.4. FOG

Fog, as well as clouds, affect the propagation of electromagnetic radiation through the atmosphere, especially in the frequency spectrum >30 GHz (<1 cm wavelength). The absorption and scattering of electromagnetic radiation by water spheres and fogs have been extensively studied. Deirmendjian [1964], Herman [1962], and Platt [1970] have discussed extinction and scattering by water droplets in the visible, infrared and submillimeter regions of the spectrum, respectively. Gunn and East [1954] have discussed absorption and scattering of microwaves by clouds and rain. The amounts of absorption and scattering are functions of the microphysical structure of the fog.

16.4.1 Microphysical Structure

16.4.1.1 Condensation Nuclei. The atmosphere contains a large concentration of dry aerosol particles ranging in size from about 10^{-3}μm to 10μm radius. Mason [1971] divides the spectrum of particles into three parts: (1) Aitken

Table 16-35. Mean seasonal heights of the tropopause.

Latitude	Altitude			
	Dec–Feb (m)	Mar–May (m)	Jun–Aug (m)	Sep–Nov (m)
60° to 80°N	8 400	8 800	9 400	9 200
50° to 60°N	9 200	10 000	11 300	10 000
40° to 50°N	10 000	11 000	12 700	11 800
30° to 40°N	14 000	13 800	14 000	14 200
10° to 30°N	17 000	16 800	15 700	16 300

nuclei ($r < 0.1\mu m$), (2) large nuclei ($0.1 < r < 1.0\mu m$), and (3) giant nuclei ($r > 1.0\mu m$). All of these particles can act as condensation nuclei provided the supersaturation is large enough (300%). However, in the real atmosphere there is a sufficient number of hygroscopic particles so that condensation occurs at supersaturations of less than 0.1%. These hygroscopic particles are either large or giant nuclei. In other words, the Aitken nuclei are not an important factor in the formation of fog even though concentrations of Aitken nuclei up to $10^6/cm^3$ are possible in urban industrial areas. Concentrations of large and giant nuclei are considerably less, averaging about $10^2/cm^3$, with a maximum concentration of about $10^3/cm^3$.

Several workers have derived analytical expressions describing the nuclei size spectrum. The most popular expression is that derived by Junge [1963] for particles greater than $r = 0.1\ \mu m$

$$\frac{dN}{d(\log r)} = Cr^{-B}, \qquad (16.30)$$

where N is the number of nuclei per cm^3 in the radius interval d (log r), and B and C are functions of the drop-size distribution. Measurements from several different parts of the world show, on average, an r^{-3} law, especially in the size range 0.05 μm through 1.0 μm. Below a radius of 0.05 μm, the number of Aitken nuclei begins to decrease with decreasing size.

There are 3 primary sources of origin of nuclei: (1) soil and vegetation, (2) sea spray, and (3) combustion. Depending on their origin, the nuclei normally contain various quantities of SO_x, NH_x, NO_x, Cl, and Na. The SO_x and NO_x are primarily from industrial or combustion sources, Cl is from industrial sources as well as the sea, NH_x is from decaying organic matter, and Na is from the sea and exists primarily as sodium chloride.

Depending on the size and hygroscopicity of the particles, condensation may begin when the relative humidity is as low as 70%. As the humidity rises due either to additional moisture or a lowering of the temperature, the particles continue to grow and form a haze. When a slight supersaturation occurs, the particles quickly grow into fog droplets. The type and concentration of hygroscopic particles and the amount of superasaturation are the primary factors that affect the fog drop size distribution and resulting visibility.

16.4.1.2 Drop Size Distributions. There are two primary techniques for measuring fog drop-size distributions: impaction and light scattering. In the impaction technique, droplets are impacted on a glass slide which is usually coated with gelatin or some oily substance, the droplets leaving impressions on the slide. The impressions are then later counted and sized to determine the drop size distribution. The 3 primary disadvantages of this technique are (1) the tedious and time consuming process of counting and sizing the droplets, (2) the low collection efficiencies of the smaller droplets resulting in unreliable counts below about 4 μm radius, and (3) fog-laden air must be drawn through a small opening, thus creating problems with obtaining representation samples. In recent years the light-scattering technique has become much more reliable and popular. The advantages of this technique are the near real time readout of the drop size distribution and the ability to measure the smaller droplets. The main disadvantage of most light scattering instrumentation is that, like the impaction technique, droplets must be drawn into the sample volume through a narrow opening.

Most impaction techniques show droplet radii ranging from 2 to 40 μm with a peak or mode in the number concentration occurring around 5 to 10 μm. Recent light-scattering techniques, however, reveal that the droplet count frequently continues to increase at radii below 5 to 10 μm. An example of this is described in Figure 16-53 which shows four comparisons of fog droplet spectra at Otis AFB, Mass., each measured simultaneously with a Calspan droplet sampler (which uses gelatin coated slides), and a Particle Measuring System forward scatter spectrometer probe (FSSP-100) [Kunkel, 1981].

Using the FSSP-100, Kunkel [1982] shows three basic types of distributions in advection fogs at Otis AFB (Figure 16-54). The common feature of the three distributions is that the maximum concentration per micron diameter interval occurs below 2 μm, apparently the result of inactive or haze nuclei. The middle spectrum (Type B) was the most typical of the distributions observed at Otis AFB, showing a primary mode between 0.5 and 2 μm, a secondary mode or plateau between 5 and 10 μm and a plateau between 15 and 30 μm. Type A has a mode in the 15 to 30 μm plateau. This type of distribution was quite common in radiation fogs observed at Hanscom AFB, Mass. Type C spectra show no plateau or peak but a steady decrease in concentration which can be represented by a power law curve [Junge, 1963].

16.4.1.3 Liquid Water Content. The liquid water content (LWC) of fog can be measured through a variety of techniques. One technique is to collect the liquid water on some type of absorbent filter paper over a specific time period. The paper is weighed before and after collection, and the increase in weight can then be converted to LWC. There are various types of hot-wire devices in which the droplets impact on electrically heated wire and subsequently evaporate and cool the surface of the wire. The magnitude of the cooling is a measure of the LWC. The LWC also can be determined by integrating the drop-size distribution curves, assuming the volume of air sampled is known. A small wind tunnel or blower is necessary with all of these techniques, in order to draw the sample into the sampling surface or through the sampling area.

Measurements of LWC in fog have been very sparse. Therefore, it is difficult to obtain any type of statistics on

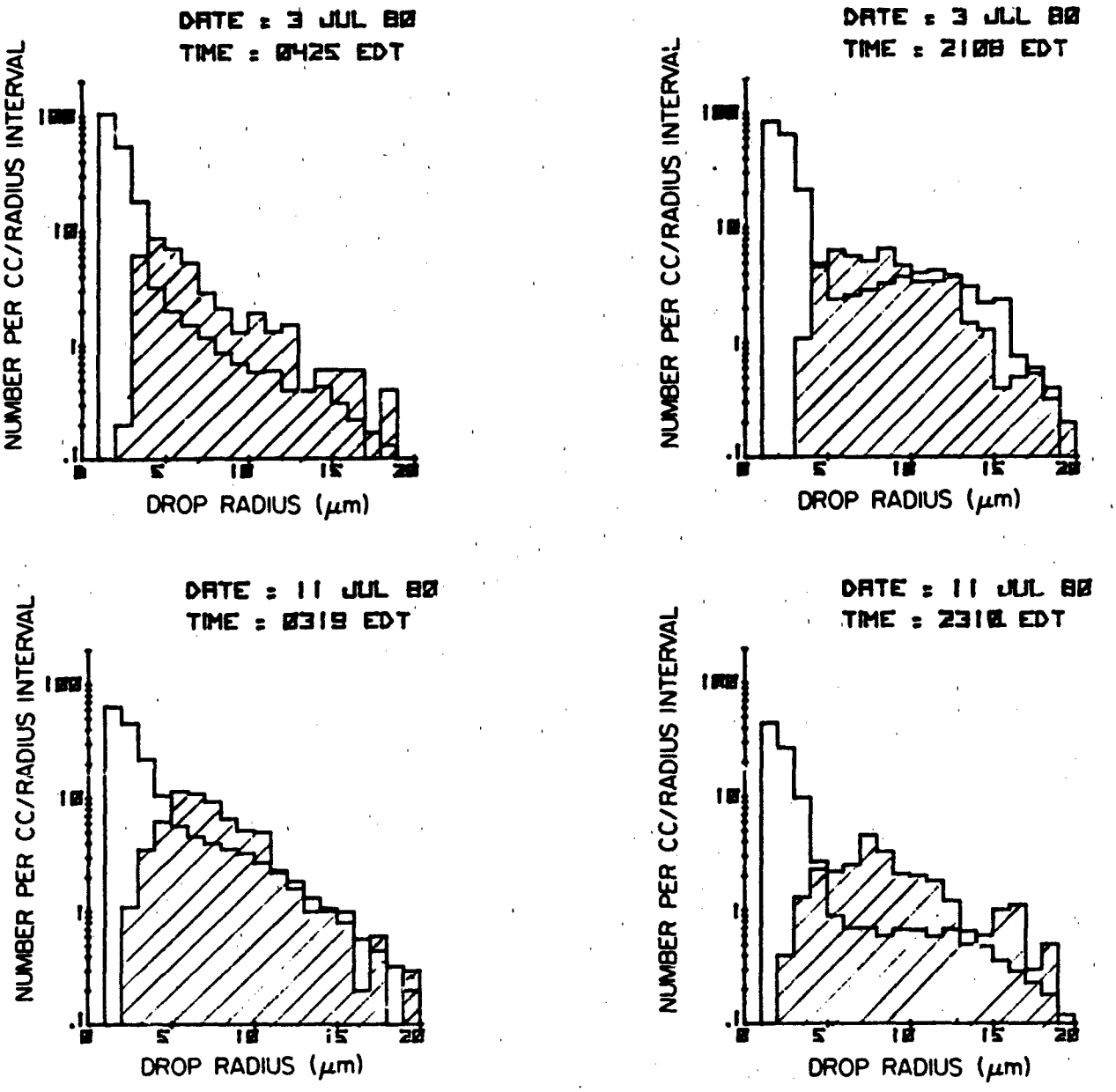


Figure 16-53. Four examples of droplet concentration spectra from the Calspan sampler (hatched area) and the FSSP-100.

the frequency of LWC amounts. Recently however, Kunkel [1982] determined the LWC from 400 samples during 7 advection fogs at Otis AFB by integrating the drop-size distributions which were averaged over 5-min periods. Figure 16-55 shows the percentage of time that the LWC at 5 m and 30 m above the ground is below a given value. The data include all samples with a liquid water content of at least 0.01 g/m^3 at the 5 m level. The average LWCs were 0.03 and 0.17 g/m^3 and the maximum LWCs were 0.33 and 0.48 g/m^3 at the 5- and 30-m levels, respectively.

16.4.2 Visibility

16.4.2.1 Definitions. There are several terms commonly used to describe the opacity of the atmosphere. Some of these terms and their definitions are

Visibility—A subjective evaluation of the greatest distance in a given direction at which it is just possible to see and identify with the unaided eye in the daytime a prominent dark object against the sky at the horizon, and at night a known, preferably unfocused, moderately intense light source.

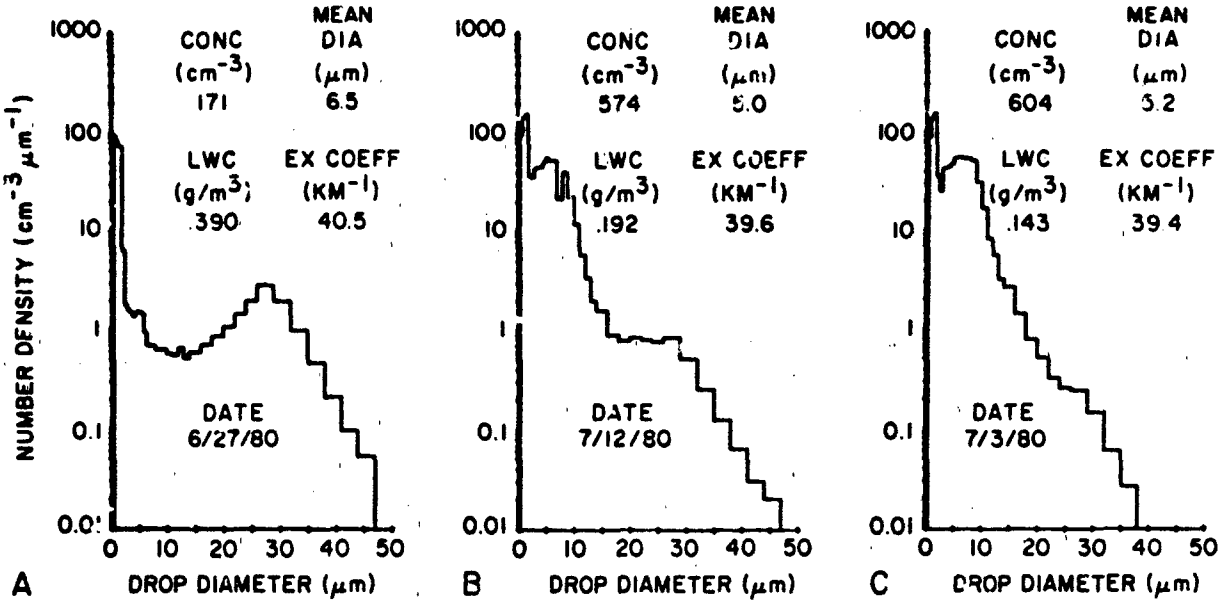


Figure 16-54. Example of three types of drop size distributions in advection fogs at Otis AFB, Mass. [Kunkel, 1981].

Prevailing Visibility—The greatest horizontal visibility that is equaled or surpassed throughout half of the horizon circle. This value is entered in surface weather observations.

Daytime Visual Range—The distance at which the apparent contrast between a specified type of target and its background becomes just equal to the threshold contrast of an observer. Daytime visual range V is a function of the extinction coefficient of the atmosphere k, the threshold contrast of the observer's eyes ε, and the viewing angle. The relationship can be described by Koschmieder's Law [1924]

$$V = \frac{1}{k} \ln \frac{1}{\varepsilon}. \qquad (16.31)$$

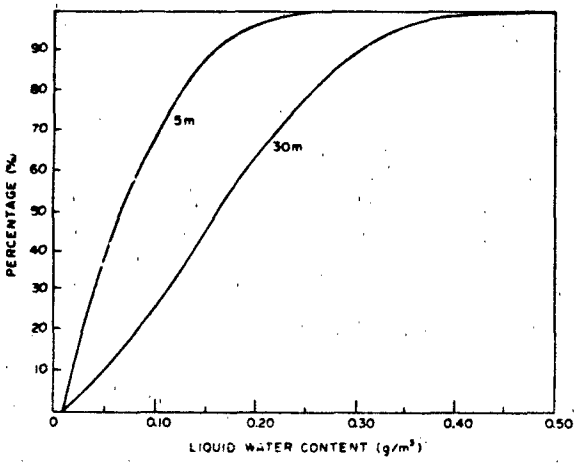


Figure 16-55. The percentage of time that the liquid water content at 30 and 5 m above the ground is below a given value [Kunkel, 1981].

Nighttime Visual Range—The greatest distance at which a point source of light of a given candle-power can be perceived by an observer under given atmospheric conditions. Nighttime visual range V is a function of extinction coefficient of the atmosphere k, the candle-power of the point source I and the threshold illuminance for the observer's eyes E. The relationship can be described by Allard's Law [1876]

$$V^2 = \frac{I}{E} \exp(-kV). \qquad (16.32)$$

Threshold Contrast (ε)—The smallest contrast of luminance that is perceptible to the human eye. The threshold contrast value can vary from 0.005 to 5.0 depending on the observer and the brightness and visual angle of the target. An average threshold contrast of 0.055 is normally used.

Threshold Illuminance (E)—The lowest value of illuminance which the eye is capable of detecting under specified conditions of background luminence and degree of dark adaption of the eye. A value of 10^{-1} lumens per km^2 is normally used.

Extinction Coefficient (k)—A measure of the space rate of extinction of any transmitted light. The extinction coefficient is a function of the concentration N and radius r of particles in the atmosphere; it can be calculated by the following.

$$k = \pi \sum_{i=1}^{n} K_i N_i r_i^2, \qquad (16.33)$$

where K is the scattering area ratio for a given particle size i. K is normally given a value of 2.

16.4.2.2 Variability. The variability of visibility in fog has been the subject of numerous recent studies, such as Chisholm and Kruse [1974a], Chisholm [1974], Chisholm and Kruse [1974b], Brown [1979], and Jiusto and Lala [1980]. The first four papers present analytical results of data collected in AFGL research projects during the last ten years. For example, an automated mesoscale weather network was in operation in eastern Massachusetts from 1972 to 1976 and was followed by the establishment of a weather facility at Otis AFB, Mass. on Cape Cod. All visibility measurements were obtained using forward-scatter-meters (FSM) whose characteristics have been previously described [Muench et al., 1974]. The visibility measurements in this section will be expressed either in terms of atmospheric extinction coefficient ($\times 10^{-4}$/m) as directly measured by the FSM, or a daytime sensor equivalent visibility (SEV) as defined by George and Lefkowitz [1972]. The daytime SEV is equivalent to the daytime visual range defined in the previous section.

The presentation of the variability of visibility in fog will be divided into three parts: (1) *temporal variability;* (2) *horizontal variability* and (3) *vertical variability*. Measures of variability are the simple correlation coefficient r, and the standardized error of estimation (SE) or persistence error computed from

$$SE = [(1 - r^2)\,\sigma^2]^{1/2}, \qquad (16.34)$$

where σ^2 is the percent variance of the dependent variable.

Temporal Variability—Chisholm and Kruse [1974a] examined the question of obtaining visibility at an airport with less than three small-volume visibility instruments along a runway and produced an insight into the temporal variability of fog. Figure 16-56 shows the resulting plots of persistence error for a single instrument (W) and a three-instrument mean based on autocorrelations with time lags from 1 to 90 min. They restricted the data to night-time observations to avoid sunrise-sunset effects. The instruments were situated on both sides of the runway. Distances between the central site and the two extremes were 1240 and 1280 m; the distance between the two extreme sites was 2400 m.

The single-station and three-station mean curves are nearly identical for advection fog. They show a rapid increase in persistence errors during the first 10 min and consistently lower persistence errors than the radiation fog cases. The similarity of the single- and three-instrument means validates existence of a higher degree of homogeneity in advection fog. On the other hand, in the radiation fog cases, the single- and three-station means differ widely and have consistently higher persistence errors than advection fog,

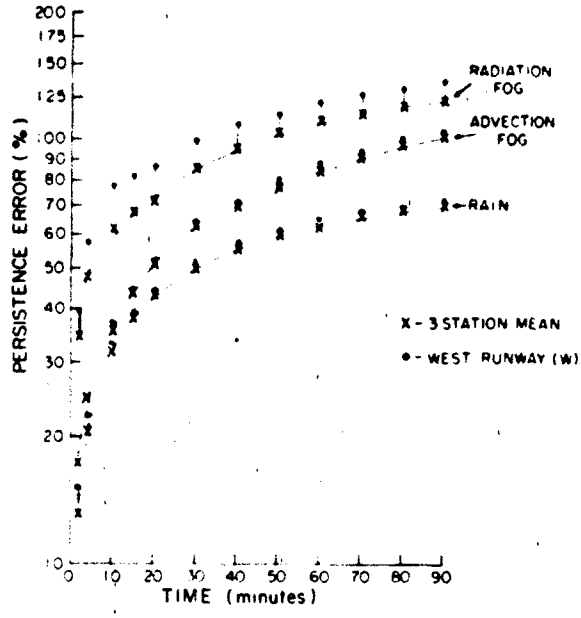


Figure 16-56. Temporal variability of SEV persistence error as a function of restriction cause [Chisholm and Kruse, 1974a].

indicating more rapid changes in visibility over short distances and time intervals.

Figure 16-57 is a time section of visibility (SEV) observed at the three instruments at Hanscom AFB during a radiation fog. The rapid and erratic fluctuations of visibility in time and space are clearly shown.

A 45-m tower located about 1000 m north of the west runway instrument (W) was also instrumented with forward-

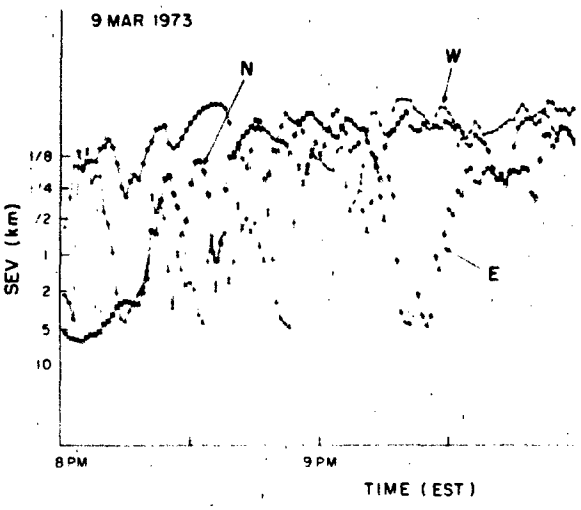


Figure 16-57. Short time plot of SEV at three locations along Hanscom AFB runway in a radiation fog episode [Chisholm and Kruse, 1974a].

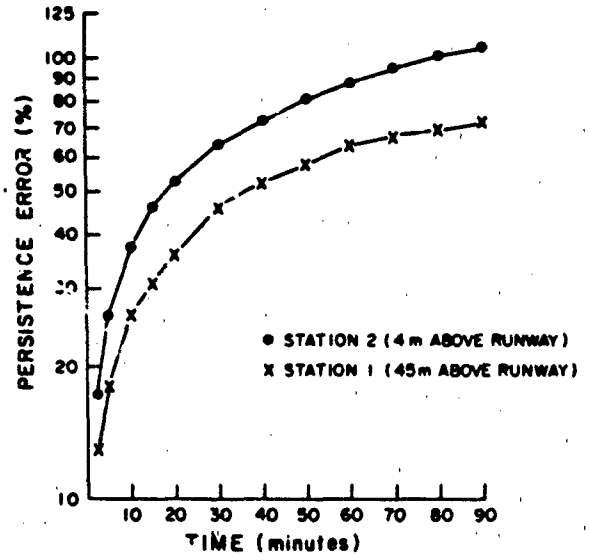


Figure 16-58. Time dependent persistence error of SEV as a function of sensor height above ground [Chisholm and Kruse, 1974b].

scatter-meters. Figure 16-58 shows for advective fog cases that the persistence error is consistently less at higher levels indicating a greater degree of homogeneity in advection fog with height.

Horizontal Variability—Chisholm and Kruse [1974a] compared simultaneous readings of two forward-scatter-meters separated horizontally by 3 m to obtain a measure of the accuracy and representativeness of the individual sensor observations. They found that the advective and radiation fog variability at this scale is on the order of 8% and 13%, respectively. Figure 16-59 shows a 1-h sample of the simultaneous observations during a period of radiation fog when visibility varied from 10 km to less than 100 m.

Figure 16-60 is a plot of forward-scatter-meter extinction coefficient measurements made at the Otis Weather Test Facility at Otis AFB, Mass., during an advection fog on 5 Nov 1978. For comparative purposes, an extinction coefficient of 30×10^{-4}/m denotes a daytime SEV of approximately 1 km while a reading of 210×10^{-4}/m represents an SEV of about 140 m, assuming a contrast threshold of 0.055. These measurements were made by instruments located at the 30-m level of four towers (A, B, P and Q). Towers P, A, and Q were oriented southeast-northwest at 500-m intervals. Tower B was located about 85 m southwest of Tower A, the center tower of the line. The graph shows the degree of homogeneity that existed over these distances in one particular case.

Chisholm and Kruse [1974a] also evaluated the degree of homogeneity in radiation and advection fogs along a runway (locations W, N, and E; see *temporal variability*) by examining observations at two points on the runway given the conditions at a third. Table 16-36 presents a summary of the two distributions, labeled midpoint and rollout,

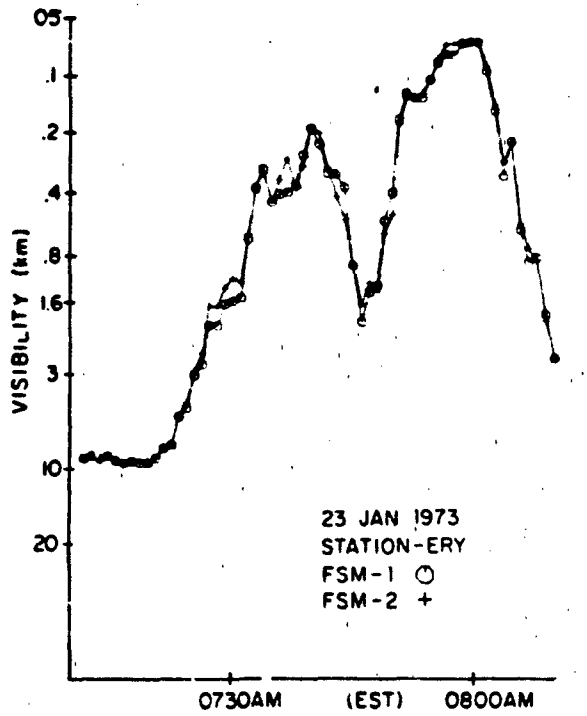


Figure 16-59. Time plot of SEV from two sensors collocated at east end of runway in a radiation fog episode [Chisholm and Kruse, 1974a].

for 5 classes of visibility restriction at the touchdown point. Data from the radiation and advection fogs have been further stratified so that each end of the runway (W and E) becomes a touchdown point. The principal diagonal in each table represents the homogeneous condition, that is conditions at mid- and endpoints of runway are the same as at the touch-

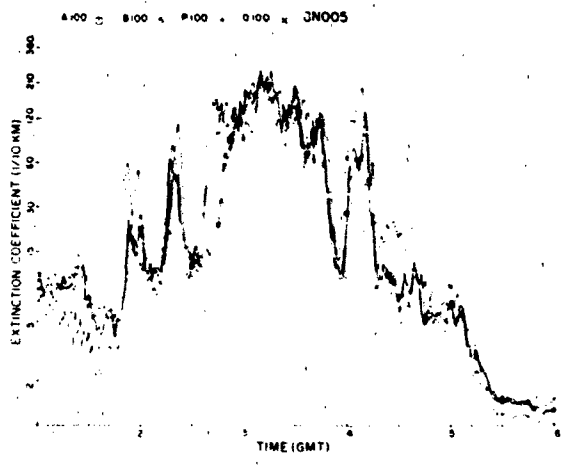


Figure 16-60. Time series of atmospheric extinction coefficient (1-min averages) measured at the 29 m levels at towers A, B, P, and Q on 5 November 1978 during fog.

Table 16-36 Conditional frequency distribution along Hanscom AFB Runway of equivalent daytime SEV

Restriction Cause	Touchdown Location	SEV Range (m)	Midpoint Distribution					Rollout Distribution				
			> 1600	800–1600	400–800	200–400	0–200	> 1600	800–1600	400–800	200–400	0–200
Radiation Fog	W	>1600	80	9	3	5	4	77	8	2	5	7
		800–1600	49	14	12	8	17	54	8	8	8	22
		400–800	23	5	18	21	32	42	12	6	21	20
		200–400	10	*	7	32	50	16	11	23	35	15
		0–200	8	0	7	31	55	9	5	7	42	37
	E	>1600	84	5	3	4	4	84	6	3	3	3
		800–1600	49	10	13	16	11	62	6	7	14	12
		400–800	26	6	8	34	27	24	8	5	40	23
		200–400	23	5	4	24	44	19	3	6	23	49
		0–200	11	15	12	17	45	28	9	6	10	47
Advection Fog	W	>1600	96	4	*	*	0	93	7	*	0	0
		800–1600	16	78	5	1	0	12	78	10	1	0
		400–800	1	18	78	3	0	2	22	63	13	0
		200–400	0	0	10	85	5	0	0	4	79	17
		0–200	0	0	0	24	76	0	0	0	2	98
	E	>1600	98	2	*	*	0	96	4	*	0	0
		800–1600	19	76	5	0	0	20	71	9	0	0
		400–800	0	21	75	3	0	*	25	71	4	0
		200–400	0	0	22	76	2	0	2	16	81	*
		0–200	0	0	0	47	53	0	0	0	44	56
All Restrictions	W	>1600	94	4	1	1	*	90	7	1	1	1
		800–1600	18	74	7	1	*	11	79	8	1	1
		400–800	3	15	76	4	2	5	25	57	12	1
		200–400	2	*	11	73	14	4	3	10	71	12
		0–200	4	0	4	24	68	7	4	5	32	52
	E	>1600	96	3	*	1	*	95	3	1	*	*
		800–1600	23	67	9	1	*	22	66	10	1	1
		400–800	5	17	69	6	3	6	19	64	8	3
		200–400	5	1	19	62	12	6	2	14	62	16
		0–200	6	7	6	28	53	16	5	4	22	53

*less than 1%

down point. The significantly higher variability of the radiation fog is very apparent, for example, with a touchdown visibility between 400 and 800 m the corresponding visibility at the midpoint and rollout are in the same range only 18% and 6% of the time for the West (W) touchdown and 8% and 5% for the East (E) touchdown. For advective fog the midpoint frequencies are in the same range 78% and 76% and rollout frequencies 63% and 57%. Disparity between these results and those reported by Schlatter and Lefkowitz [1968], which showed more variability at airports located at Atlantic City, N.J. and Los Angeles, Calif., can be attributed to different climatic regimes and to different observing systems. The transmissometer used in the Schlatter and Lefkowitz [1968] report was compared with the FSM by Muench et al. [1974] and was found to have an inherent overall error of ± 20% due to calibration drift and instrument noise as compared to the ± 4% of the FSM.

Figure 16-61 shows a comparison of spatial variability at Hanscom AFB with variabilities reported by Hage and Entrekin [1965] for Kennedy Airport, N.Y., and by Hage and Wilson [1967] for Atlantic City, N.J. Chisholm and Kruse [1974a] imply that a significant part of the variability at Kennedy and Atlantic City is due to use of the transmissometer as a measuring system rather than to natural causes. Figure 16-61 illustrates horizontal variability over distances ranging from 0.4 to 3 km in radiation and advection fog. Figure 16-62 represents a compilation of space correlations in the range 5 to 40 km for station pairs carefully

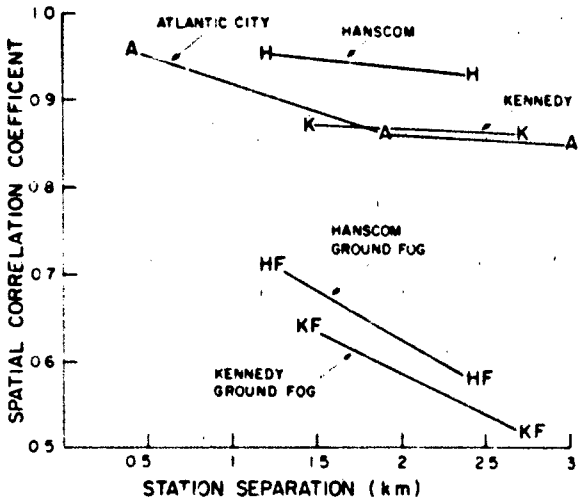


Figure 16-61. Comparison of horizontal variability of SEV at three airport locations [Chisholm and Kruse, 1974a].

selected from the AFGL Mesonetwork to eliminate coastal and hilltop effects and to insure similar elevations. The decay in correlation as a function of station separation is clearly evident in the advection fog and quite dramatic in the case of radiation fog. Because radiation fog rarely is found in the eastern part of the network, the station pairs for this class restriction were limited to 15 km.

Vertical Variability—The AFGL Mesonetwork, in addition to numerous surface weather stations, contained one tower at Hanscom AFB which was instrumented at the 30 and 45 m levels. Figure 16-63 from Chisholm and Kruse [1974a] shows the cumulative frequency distribution of visibility in advective fogs (and rain cases) at three levels: ground, 30 m, and 45 m. The figure shows a dramatic increase of reduced visibility with height. For example, during advective fog, visibilities (SEV) greater than 400 m occur about 90% of the time at the surface, while at the 45 m level, they occur only 65% of the time.

Figure 16-64 shows a typical time section of the vertical variation of visibility (extinction coefficient) at Otis AFB, Mass., on 5 Nov 1978 during an advection fog. Five levels, from 3 m(X) to 57 m(A200), of visibility were plotted. The striking decrease in visibility (increase in extinction coefficient) with height is readily apparent. Figure 16-65, on the other hand, shows a time section of visibility when a typical ground fog occurred. The same levels of the tower are shown. The striking feature in this figure is the complete reversal of visibility gradient, with the surface instrument (X) recording the lowest visibility.

16.4.2.3 Relationship Between Extinction Coefficient and Liquid Water Content. A direct relationship exists between the extinction coefficient and the amount of liquid water in a cloud. However, the extinction coefficient in the visible part of the spectrum is also inversely related to the radii of the droplets. In other words, a large-droplet fog with a given liquid water content will yield a lower extinction (higher visibility) than a small-droplet fog with the same liquid water content. Eldridge [1966] suggested that the extinction coefficient (σ) and the liquid water content (W) can be related by the empirical formula

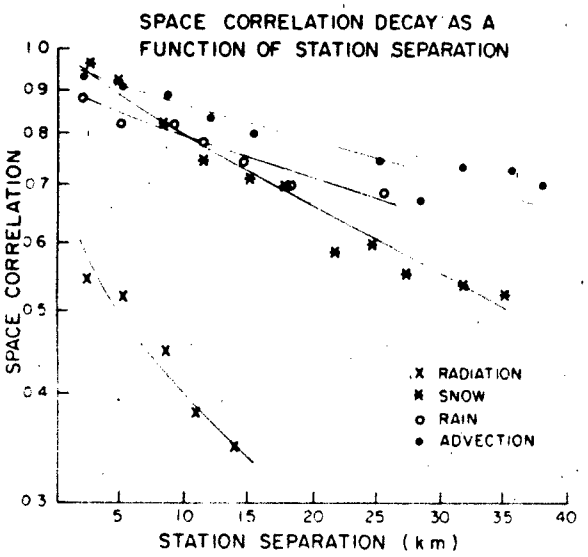


Figure 16-62. Average spatial correlation of SEV in the Hanscom mesonetwork during various types of restrictions [Chisholm and Kruse, 1974b].

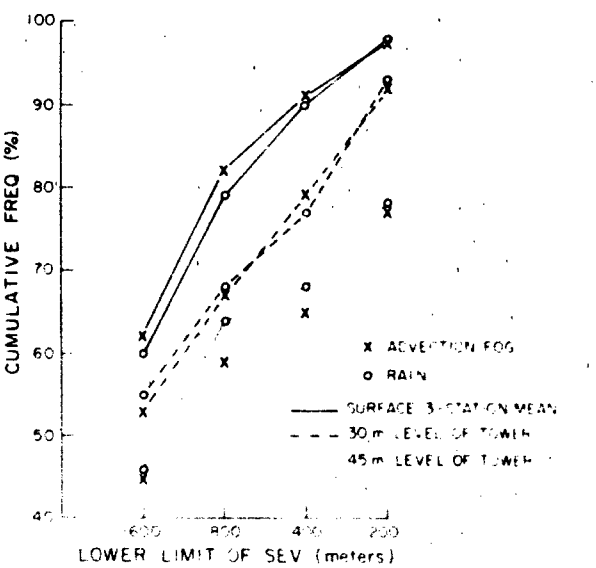


Figure 16-63. Cumulative frequency distribution of SEV at three levels above the Hanscom AFB runway [Chisholm, 1974].

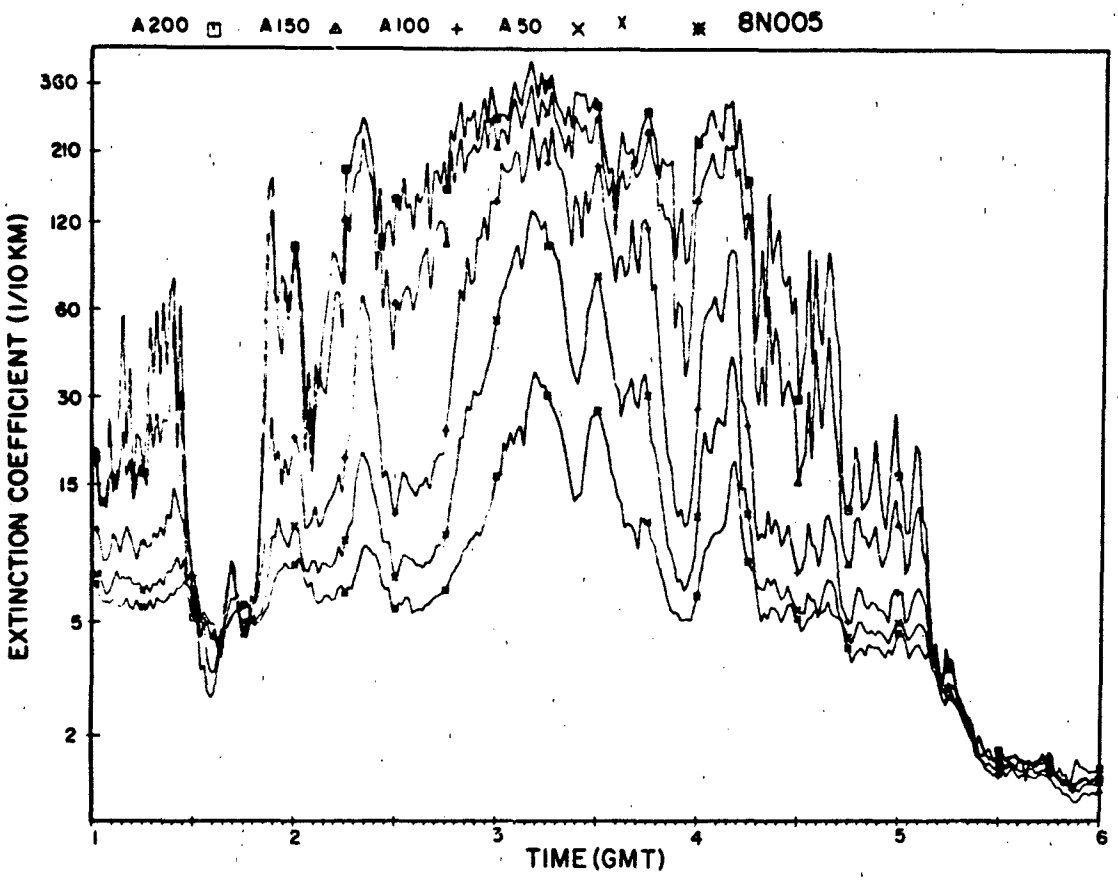


Figure 16-64. Time series of atmospheric extinction coefficient (1-min averages) measured at the 57-m(A-200), 48-m(A-150), 29-m(A-100) and 16-m (A-50) levels of tower A and at the 3-m (x) level of site x on 5 November 1978 during fog.

$$\sigma = 163W^{0.65}, \qquad (16.35)$$

where σ is in 1/km and W in g/m^3. The Eldridge relationship was derived from somewhat questionable drop-size data and would probably hold true only in a highly polluted, foggy environment such as might occur near an industrial area.

Kunkel [1982], using 1525 drop-size distributions from eight advection fog cases, arrived at the following formula:

$$\sigma = 155\ W^{0.94}. \qquad (16.36)$$

He found, however, that the coefficient varied from 110 in a large-droplet fog with drizzle to 253 in a small-droplet fog and the exponent varied from 0.96 to 1.14. For all practical purposes a linear relationship can be assumed, but the slope will depend on the drop size distribution.

Chylek [1978] found that at a wavelength of 11 μm there is an unambiguous relation between the extinction coefficient and the liquid water content, independent of the size distribution, in fogs with droplets up to 28 μm diameter. That relationship is as follows:

$$\sigma = 128W. \qquad (16.37)$$

Since the relationship is unique to all fogs, one can determine the extinction coefficient at a wavelength of 11 μm from the LWC measurements and vice versa.

16.4.2.4 Modeling of Visibility Distributions. The Weibull Distribution has been used for visibility

$$\hat{F}(x) = 1 - \exp(-\alpha x^{\beta}), \qquad (16.38)$$

where x stands for visibility, and α and β are parameters. Estimates for α and β are made to minimize the sums of squares of the differences between the model estimates $\hat{F}(x)$ and the cumulative frequencies $F(x)$ of the climatic summaries.

At the University of Central Florida [Somerville et al., 1979] sets of values for α and β have been determined for each of the eight periods of the day in each month at some 22 stations around the world, to make the estimated cumulative distributions $\hat{F}(x)$ fit the data for some 15 visibil-

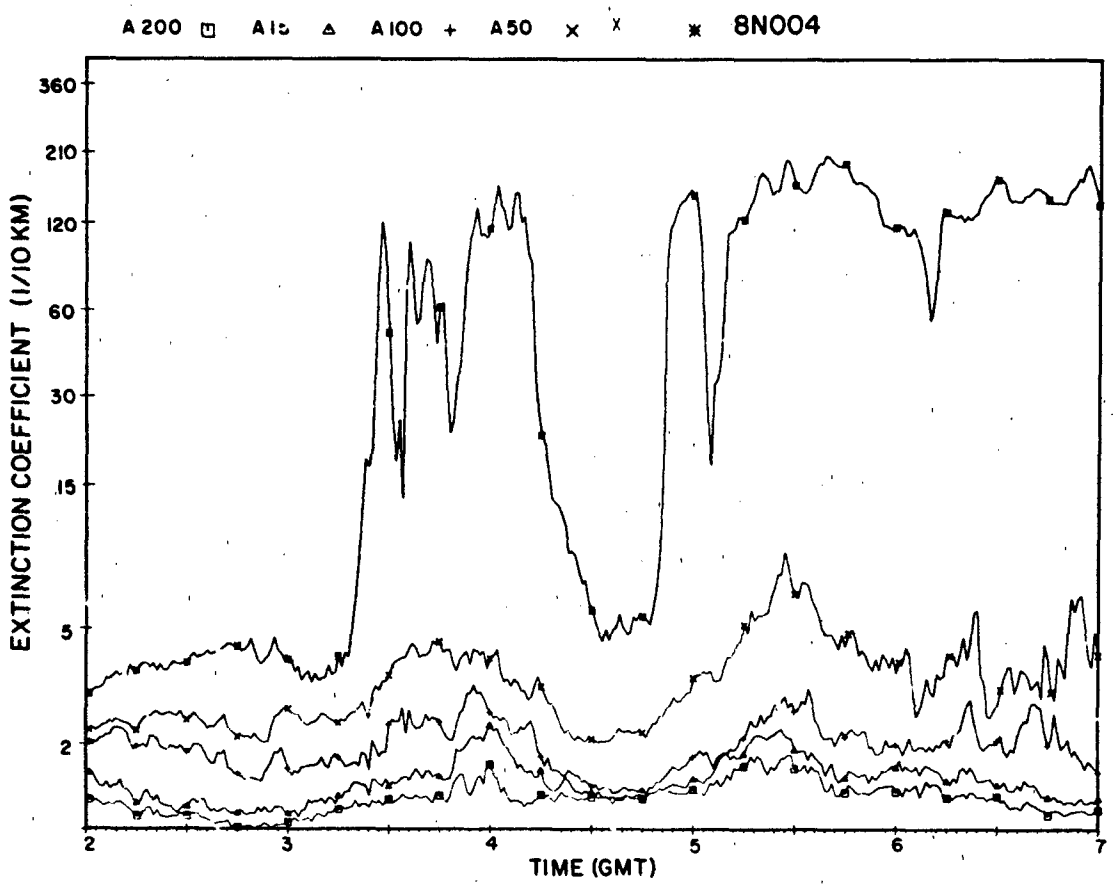


Figure 16-65. Time series of atmospheric extinction coefficient at the same levels as shown in Figure 16-64 (observations taken on 4 November 1978 during ground fog).

Table 16-37. A sample of cumulative frequency distribution of visibility in kilometers. (The station is Bedford, Mass., January noontime. Probability estimates were made by the Weibull distribution, with parameter values $\alpha = 0.04679$, $\beta = 0.81856$.)

Visibility (km)	RUSSWO Frequency	Model Probability
0.4	0.018	0.022
0.5	0.028	0.026
0.8	0.034	0.038
1.0	0.053	0.046
1.2	0.056	0.053
1.6	0.072	0.067
2.0	0.094	0.080
2.4	0.099	0.092
3.2	0.116	0.115
4.0	0.135	0.136
4.8	0.146	0.156
6.4	0.184	0.193
8.0	0.221	0.227
9.7	0.254	0.259
16	0.377	0.365

ities (in kilometers) corresponding to those in the climatic summaries.

A simpler method for solving for α and β utilizes the linear relationship between $\ell n(-\ell n[1 - F(x)])$ and ℓnx. This permits a solution for β and ℓnd by linear regression. As an example, at Bedford, Mass., in January at noontime, α and β were found to be 0.04679 and 0.81856, respectively; the resulting visibility probabilities are compared with RUSSWO data in Table 16-37.

CHAPTER 16

REFERENCES

Air Weather Services, "Forecasters' Guide on Aircraft Icing," AWS-TR-80-0001 Scott AFB, Ill., 1980.

Allard, E., *Memoire sur L'intensite et la Portee des Phares*, Paris, Dunod, 1876, quoted in Middleton, W.E., *Vision Through the Atmosphere*, University of Toronto Press, 1963.

Allen, M., Y.L. Yung, and J.W. Waters, " Vertical Transport and Photochemistry in the Terrestrial Mesosphere and Lower Thermosphere (50–120 km)," *J. Geophys. Res.*, **86:** 3617–3627, 1981.

Anonymous, "Lowest Measured Temperature in the Earth's Atmosphere Recorded in U.S. - Sweden Experiments," *Bull. Am. Meteorol. Soc.*, **44:** 806, 1963.

Appleman, H.S., "Occurrence and Forecasting of Cirrostratus Clouds," Tech. Note 140, World Meteorological Organization (WMO), 1961.

aufm Kampe, H.J., and H.K. Weickmann, "Physics of Clouds," *Meteorol. Monographs*, **3:** 182, 1957.

Bannon, B.A. and J.P. Steele, "Average Water Vapor Content of the Air," Geophys. Memoirs No. 102, Meteorological Office, Air Ministry, London, 1960.

Barnes, A.A., "New Cloud Physics Requirements," *Preprints 4th Symp. on Meteorological Observations and Instrumentation*, American Meteorological Society, Boston, Mass., 1978.

Barnes, A.A., "Observations of Ice Particles in Clear Air," *Journal de Recherches Atmospheriques* **14:** 311–315, 1980.

Barnes, A.A., "Ice Particles in Clear Air," *Preprints 8th International Conf. on Cloud Physics*, Clermont-Fermad, France, 15–19 July 1980, pp 189–190, AFGL TR-81-0009, AD094444, 1981.

Barnes, A.A., "The Cirrus and Subvisible Cirrus Background," *Preprints 2nd Symp. on Composition of Non-Urban Troposphere*, American Geophysical Union, Williamsburg, Va., May 1982.

Bean, S.J., P.N. Somerville, M. Heuser, "Some Models for Ceiling," AFGL TR-79-0221, ADA078033, 1979.

Beckwith, W.B., "Analysis of Hailstorms in the Denver Network, 1949–1958," *Physics of Precipitation*, edited by H. Weidkman, American Geophysical Union, Washington, D.C., 348–353, 1960.

Bennett, I., "Glaze, Its Meteorology and Climatology, Geographic Distribution and Economic Effects," Tech. Rept. EP-105, Quartermaster Research and Engineering Center, Envir. Prot. Res. Div., Natick, Mass., 1959.

Bertoni, E.A., "Clear Lines-of-Sight from Aircraft," AFCRL 67-0435, AD657801, 1967.

Bertoni, E.A., "Clear and Cloud-Free Lines-of-Sight from Aircraft," AFGL TR-77-0141, ADA046352, 1977a.

Bertoni, E.A., "Clear and Cloud-Free Lines-of-Sight from Aircraft," AFGL TR-77-0141, ADC 015057 1977b.

Best, A.C., "The Size Distribution of Raindrops," *Quart. J. Roy. Meteorol. Soc.*, **76:** 16, 1950.

Boyd, D.W., "Maximum Snow Depths and Snow Loads on Roofs in Canada," Research Paper 142, Div. of Building Research, National Research Council, Ottawa, Canada, 1961.

Brewer, A.W., "Evidence for a World Circulation Provided by Measurement of Helium and Water Vapor Distribution in the Stratosphere," *Quart. J. Roy. Meteorol. Soc.*, **75:** 351–363, 1949.

Brown, H.A., "Preliminary Assessment of an Automated System for Detecting Present Weather," AFGL TR-79-0137, ADA07831, 1979.

Bundy, D.S., "A Two-Year Study of the Probability of Cloud-Free Lines-of-Sight at Columbia, Mo.," Unpublished Masters Dissertation, University of Missouri, Columbia, 1969.

Bussey, H.E., "Microwave Attenuation Statistics Estimated from Rainfall and Water Vapor Statistics," *Proc. IRE*, **38:** 781, 1950.

Changnon, S.A., "Hailstreaks," *J. Atmos. Sci.*, **27:** 109–125, 1970.

Changnon, S.A., "Means of Estimating Areal Hail-Day Frequencies," *J. Wea. Modifica.*, **3:** 154–158, 1971.

Chisholm, D.A., "The Mesoscale Variability of Visibility in Time and Space," Proceedings 5th Conference on Weather Forecasting and Analysis, St. Louis, Mo., American Meteorological Society, Boston, Mass., 1974.

Chisholm, D.A. and H. Kruse, "The Variability of Airfield Visibility: A Preliminary Assessment," AFCRL TR-74-0027, AD777219, 1974a.

Chisholm, D.A. and H. Kruse, "The Variability of Airfield Visibility: A Preliminary Assessment," AFCRL TR-74-0265, AD784791, 1974b.

Chylek, P., "Extinction and Liquid Water Content of Fogs and Clouds," *J. Atmos. Sci.*, **35:** 290–300, 1978.

Cohen, I.D., "Development of a Large Scale Cloud System, 23–27 Mar 78," AFGL TR-81-0127, ADA106417, 1981a.

Cohen, I.D., "Cirrus Particle Distribution Study, Part 8," AFGL TR-81-0316, ADA118715, 1981b.

Cole, A.E., R.J. Donaldson, D. Atlas, A.J. Kantor, R.K. Soberman, and J.E. Manson, "Precipitation, Clouds, and Aerosols," Chapter 5, *Handbook of Geophysics and Space Environments*, AFCRL, 1965.

Cole, A.E. and A.J. Kantor, "Air Force Reference Atmospheres," AFGL TR-78-0051, AD058505, 1978.

Crane, A.K., "Prediction of Attenuation by Rain," *IEEE Trans. Comm.* **COM-28:** 1717–1733, 1980.

Crutcher, H.L. and J.M. Meserve, "Selective Level Heights, Temperatures and Dew Points for the Northern Hemisphere," Naval Weather Service Command, NAVAIR 50-IC-52, U.S. Government Printing Office, Washington, D.C., 1970.

Crutzen, P.A., "A Review of Upper Atmospheric Photochemistry," *Can. J. Chem.*, **52:** 1569–1581, 1974.

Davis, A.R. and D.J. McMorrow, "Stochastic Models for Deriving Instantaneous Precipitation Rate Distributions," USAF Air Weather Service, AWS-TR-76-263, 1976.

deBary, E. and F. Moller, "The Vertical Distribution of Clouds," *J. Appl. Meteorol.*, **2:** 806–808, 1963.

Deirmendjian, D., "Scattering and Polarization Properties of Water Clouds and Hazes in the Visible and Infrared," *Appl. Opt.*, **3:** 187–196, 1964.

Derr, V.E., "A Basic Comparison of Lidar and Radar for Remote Sensing of Clouds," Tech. Rept. ERL 397-WPL 52, Wave Propagation Laboratory, NOAA, Boulder, Colo., 1978.

Deguchi, S. and D.O. Muhleman, "Mesospheric Water Vapor," *J. Geophys. Res.*, **87:** 1343–1346, 1982.

Ehhalt, D.H., L.E. Heidt, R.H. Lueb, and E.A. Martell, "Concentrations of CH_4, CO, CO_2, H_2, H_2O, and NO_2 in the Upper Stratosphere," *J. Atmos. Sci.*, **32:** 163–169, 1975.

Eldridge, R.G., "Haze and Fog Aerosol Distributions," *J. Atmos. Sci.*, **23:** 605–613, 1966.

Falls, L.W., "The Beta Distribution: A Statistical Model for World Cloud Cover," NASA TM X-64714, Marshall Space Flight Center, Ala., 1973.

Farmer, C.B., O.F. Raper, B.D. Robbins, R.A. Toth, and C. Muller, "Simultaneous Spectroscopic Measurements of Stratospheric Species O_2, CH_4, CO, CO_2, N_2O, H_2O, HCL, and HF at Northern and Southern Mid-Latitudes," *J. Geophys. Res.*, **85:** 1621–1632, 1980.

Feldman, N.E., "Rain Attenuation Over Earth-Satellite Paths," Science Applications, Inc., Final Report on Contract N00039-79-C-0136 for Naval Electronic Systems Command, ADA075390, 1979.

Fett, R.W., J.J. Bates, W.A. Bohan, R.J. Englebretson, J. Rosenthal, and G.A. Stevenson, *Navy Tactical Applications Guide, Vol. 3, North Atlantic and Mediterranean Weather Analysis and Forecast Applications*, NEPRF Tech. Report 80-07, Monterey, Calif., 1981.

Frederick, R.H., V.A. Meyers, and E.P. Auciello, "Five- to 60-Minute Precipitation Frequency for the Eastern and Central United States," NOAA Tech. Memo. NWS HYDRO-35, 1977.

George, D.M., and M. Lefkowitz, "A New Concept: Sensor Equivalent Visibility," Proceedings International Conference on Aerospace and Aeronautical Meteorology, Washington, D.C., American Meteorological Society, 1972.

Glass, M. and D.D. Grantham, "Response of Cloud Microphysical Instruments to Aircraft Icing Conditions," AFGL TR-81-0192, ADA112317, 1981.

Grantham, D.D., I.A. Lund, and R.E. Davis, "Estimating the Probability of Cloud-Free Fields-of-View Between Earth and Airborne or Space Platforms," Air Weather Service Tech. Rept. 79-001, 1979.

Gringorten, I.I., "Hailstone Extremes for Design," AFCRL 72-0081, AD743831, 1971.

Gringorten, I.I., "Probability Models of Weather Conditions Occupying a Line or an Area," *J. Appl. Meteorol.*, **18:** 957-977, 1979.

Gringorten, I.I., "Climatic Probabilities of the Vertical Distribution of Cloud Cover," AFGL-TR-82-0078, ADA118753, 1982.

Gringorten, I.I., H.A. Salmela, I. Solomon, and J. Sharp, "Atmosphere Humidity Atlas—Northern Hemisphere," AFCRL 66–621, AD642429, 1966.

Gunn, K.L.S. and T.W.R. East, "The Microwave Properties of Precipitation Particles," *Quart. J. Roy. Meteorol. Soc.*, **80:** 522–545, 1954.

Gunn, K.L.S. and J.S. Marshall, "The Distribution with Size of Aggregate Snowflakes," *J. Meteorol.*, **15:** 542, 1958.

Gutnick, M., "How Dry is the Sky?", *J. Geophys. Res.*, **66:** 2867–2871, 1961.

Hage, N. and H. Entrekin, "Further Studies of Space and Time Variations in Atmospheric Transmission Along Airport Runways," TR7417-143, The Travelers Research Center, Inc., Hartford, Conn., 1965.

Hage, N. and J. Wilson, "Analyses of the Variability of Clouds, Fog, Wind and Rain in the Atlantic City Mesonetwork," TN7481-262, The Travelers Research Center, Inc., Hartford, Conn., 1967.

Hall, G.R., "On the Mechanics of Transition Produced by Particles Passing Through on Initially Laminar Boundary Layer and the Estimated Effect on the LFC (Laminar Flow Control) Performance of the X-21 Aircraft," Northrop Aircraft Report, October 1964.

Hall, W.D. and H.R. Pruppacher, "The Survival of Ice Particles Falling from Cirrus Clouds in Saturated Air," *J. Atmos. Sci.*, **33:** 1995–2006, 1976.

Harris, J.E., "The Distribution of Water Vapor in the Stratosphere," *Rev. Geophys. Space Phys.*, **14:** 565–575, 1976.

Haurwitz, B., *Dynamic Meteorology*, McGraw-Hill, New York, 1941.

Heath, E.D., and L.M. Cantrell, "Aircraft Icing Climatology for the Northern Hemisphere," Tech. Rept. 220, Air Weather Service (MAC) USAF, 1972.

Herbstritt, R.L., "Rain Climate Considerations for Domestic Satellites—Continental U.S.," Report R-7301, FCC, Washington, D.C., 1973.

Herman, B.M., "Infrared Absorption, Scattering and Total Attenuation Cross-Sections for Water Spheres," *Quart. J. Roy. Meteorol. Soc.*, **86:** 153–250, 1962.

Heymsfield, A.H., "Cirrus Unicus Generating Cells and the Evolution of Cirriform Clouds, Part I. Aircraft Observations of the Growth of the Ice Phase," *J. Atmos. Sci.*, **32:** 799–808, 1975.

Ho, F.P. and J.T. Riedel, "Precipitable Water over the United States, Vol. II, Semimonthly Mariver," U.S. Dept. of Commerce, NOAA/NWS, Silver Springs, Md., 1979.

Huschke, R.E., H.T. Bailey, R.J. Nelson, and J.P. Bibbo, "An Improved Weather-Attenuation Model for Earth-Satellite Communication Links," Science Applications Inc., Final Report on Contract DAAK30-80-C-0049, U.S. Army Satellite Communications Agency, Ft. Monmouth, N.J., 1982.

Jiusto, J.E. and G.G. Lala, "Thermodynamics of Radiation Fog Formation and Dissipation—A Case Study," *Proceedings 8th International Cloud Physics Conference*, Clermont-Ferrand, France, 15–19 June 1980.

Jones, M.A. and A.L. Sims, "Climatology of Instantaneous Precipitation Rates," AFCRL TR-72-0429, AD748283, 1971.

Jones, M.A. and A.L. Sims, "Climatology of Instantaneous Rainfall Rates," *J. Appl Meteorol.*, **17:** 1135–1140, 1978.

CHAPTER 16

Junge, C.E., *Air Chemistry and Radioactivity*, Academic Press, New York, 1963.

Kantor, A.J. and D.D. Grantham, "A Climatology of Very High Altitude Radar Precipitation Echoes," AFCRL 68-0630, AD681107, 1968.

Kikuchi, K. and A.W. Hogan, "Properties of Diamond Dust Type Ice Crystals Observed in the Summer Season at Amundsen-Scott South Pole Station, Antarctica," *J. Meteorol. Soc. of Japan*, **57:** 180–190, 1979.

Knollenberg, R.G., "Comparative Liquid Water Content Measurements of Conventional Instruments with an Optical Array Spectrometers," *J. Appl. Meteorol.*, **11:** 501–508, 1972.

Koschmieder, H., "Theorie Der Horizontalen Sichtewite," *Beitr. Phys. Atmos.*, **12:** 33–53, 1924. (Cited in Middleton W.E.K., *Vision Through the Atmosphere*, University of Toronto Press, 1963.)

Kunkel, B., "Comparison of Fog Drop-Size Spectra Measured by Light Scattering and Impaction Techniques," AFGL TR-81-0049, ADA100252, 1981.

Kunkel, B., "Microphysical Properties of Fog at Otis AFB, MA," AFGL TR-82-0026, ADA119928, 1982.

Lamb, K., K.W. Nielson, H.E. Klieforth, and J. Hallett, "Measurements of Liquid Water Content in Winter Conditions Over the Sierra Nevadas," *J. Appl. Meteorol.*, **15:** 763–775, 1976.

Landsberg, H., "Climatology," Chapter XII, in *Handbook of Meteorology*, 1945.

Landsberg, H.E., *World Survey of Climatology*, Vols. I–XV, Elsevier, Amsterdam, The Netherlands, 1974.

Lenhard, R.W., "Precipitation Intensity and Extent," *Journal de Recherches Atmospheriques*, **8:** 375–384, 1974.

Lenhard, R.W. and N. Sissenwine, "Extremes of 1, 12 and 24 Hour Rain for MIL-STD-210B," AFCRL TP-73-0329, AD766210, 1973.

Levin, L.M., "Size Distribution Function for Cloud Droplets and Raindrops," TR-263R, DSNS, Canada, 1954.

Lewis, W., "Meteorological Aspects of Aircraft Icing," *Compendium of Meteorology*, American Meteorol. Society, Boston, Mass., 1951.

Lhermitte, R.M., "Millimeter Wave Doppler Radar," *Preprints 20th Conf. Radar Meteorol.*, American Meteorol. Society, Boston, Mass., 1981.

List, R.J. (ed.), *Smithsonian Meteorological Tables*, Smithsonian Institution Press, Washington, D.C., 1968.

Litvinov, I.V., "The Function of the Distribution of Particles of Liquid Precipitates," *IZV. Acad. Sci.—Geophys.*, **12:** 1474–1483, 1956.

Lo, K.K. and R.E. Passarelli, "The Growth of Snow in Winter Storms: An Airborne Observational Study," *J. Atmos. Sci.*, **39:** 697–706, 1982.

Long, M.J., H.H. Hanks, and R.G. Beebe, "Tropopause Penetrations by Cumulonimbus," AFCRL 65-561, AD621573, 1965.

Lott, G.A., *Precipitable Water Over the United States, Vol. I, Monthly Means*, U.S. Dept. of Commerce, NOAA/NWS, Silver Springs, Md., 1976.

Ludlum, D.M. (ed.), "Extremes of Snowfall; United States and Canada," *Weatherwise*, **23:** 286–294, 1970.

Lund, I.A., "Haze Free and Cloud-Free Lines-of-Sight Through the Atmosphere," AFCRL 72-0540, AD751264, 1972.

Lund, I.A., "Joint Probabilities of Cloud-Free Lines-of-Sight Through the Atmosphere at Grand Forks, Fargo, and Minot, N. Dak.," AFCRL TR-73-0178, AD763083, 1973.

Lund, I.A. and M.D. Shanklin, "Photogrammetrically Determined Cloud-Free Lines-of-Sight Through the Atmosphere," *J. Appl. Meteorol.*, **11:** 773–782, 1972.

Lund, I.A. and M.D. Shanklin, "Universal Methods for Estimating Probabilities of Cloud-Free Lines-of-Sight Through the Atmosphere," *J. Appl. Meteorol.*, **12:** 28–35, 1973.

Lund, I.A. and E.A. Bertoni, "Some Comparisons Between Probabilities of Cloud-Free Line-of-Sight Estimated from Aircraft and from Sky Cover Observations," AFGL TR-80-0046, ADC021847, 1980.

Lund, I.A., D.D. Grantham, and R.E. Davis, "Estimating Probabilities of Cloud-Free Fields-of-View from the Earth Through the Atmosphere," *J. App. Meteorol.*, **19:** 452–463, 1980.

Lund, I.A., D.D. Grantham, and C.B. Elam Jr., "Atlas of Cloud-Free Line-of-Sight Probabilities", in several parts: *Part 1, Germany*, AFCRL TR-75-0261, ADA013784, 1975. *Part 2, U.S.S.R.*, AFGL TR-77-0005, ADA040705, 1976. *Part 3, U.S.A.*, AFGL TR-77-0188, ADA051112, 1977. *Part 4, Europe*, AFGL TR-78-0276, ADA065167, 1978. *Part 5, North Africa and the Middle East*, AFGL TR-79-0275, ADA088156, 1979.

Marshall, J.S. and W. M. Palmer, "The Distribution of Raindrops with Size," *J. Meteorol.*, **5:** 165, 1948.

Mason, B.J., *The Physics of Clouds*, 2nd Ed., pp 113–121, Clarendon Press, Oxford, England, 1971.

McKay, G.A. and H.A. Thompson, "Estimating the Hazard of Ice Accretion in Canada from Climatological Data," *J. Appl. Meteorol.*, **8:** 927–935, 1969.

McTaggert-Cowan, J.D., G.G. Lala, and B. Vonnegut, "The Design, Construction and Use of an Ice Crystal Counter for Ice Crystal Studies by Aircraft," *J. Appl. Meteorol.*, **9:** 294–299, 1970.

Michaels, D.W., "Noctilucent Cloud Research," *Foreign Sci. Bull.*, **6:** 51-55, Library of Congress, 1965.

Muench, H.S., E.Y. Moroz, and L.P. Jacobs, "Development and Calibration of the Forward-Scatter-Meter," AFCRL TR-74-0145, AD4783270, 1974.

NOC (Naval Oceanography Command), *Guide to Standard Weather Summaries and Climatic Services*, NAVAIR 50-1C-534, Asheville, N.C., 1980.

Pasqualucci, F., B.W. Bertram, R.A. Kropfli, and W.R. Moninger, "A Millimeter Wavelength Dual-Polarization Doppler Radar for Cloud and Precipitation Studies," *J. Clim. Appl. Meteorol.*, **22:** 758–765, 1983.

Pasqualucci, F. and L.J. Miller, "Dual Wavelength Radar Observation in Clouds and Precipitation," *Preprints 20th Conference Radar Meteorology*, American Meteorol. Society, Boston, 1981.

Penndorf, R., "Analysis of Ozone and Water Vapor Field Measurement Data," FAA Tech. Rept., FAA-AEE-78-29, ADA072721, Nov 1978.

Perkins, P.J., *Proceedings, Second Annual Workshop on Meteorological and Environmental Inputs to Aviation Systems*, NASA Conf. Publication 2057: 85–99, 1978.

Plank, V.G., "Hydrometeor Data and Analytical-Theoretical Investigations Pertaining to the SAMS Missile Flights of the 1972–73 Seasons at Wallops Island, Virginia," AFGL TR-77-0149, ADA051192, 1977.

Platt, C.M.R., "Transmission of Submillimeter Waves Through Water Clouds and Fog," *J. Atmos. Sci.*, **27:** 421–425, 1970.

Pochop, L.O. and M.D. Shanklin, "Sky Cover Photograms, A New Technique," *Weatherwise*, **19:** 198–200, 1966.

Rapp, R.R., C. Schutz, and E. Ronriques, "Cloud-Free Line-of-Sight Calculations," *J. Appl. Meteorol.*, **12:** 484–493, 1973.

Riordan, P., "Weather Extremes Around the World," U.S. Army Engineer Topographic Laboratories, Ft. Belvoir, Va., Report ETL-TR-74-5, 1974.

Riordan, P., "Extreme 24-Hour Snowfalls in the United States: Accumulation, Distribution, and Frequency," U.S. Army Engineer Topographic Laboratories, Ft. Belvoir, Va., Special Report ETL-SR-78-4, 1973.

Roach, W.T., "On the Nature of the Summit Areas of Severe Storms in Oklahoma," *Quart. J. Roy. Meteorol. Soc.*, **93:** 318–336, 1967.

Rogers, J.W., A.T. Stair, T.C. Degges, C.L. Wyatt, and K.D. Baker, "Rocketborne Measurement of Mesospheric H_2O in the Auroral Zone," *Geophys. Res. Lett.*, **4:** 366–368, 1977.

Roys, G.P. and E. Kessler, "Measurements by Aircraft of Condensed Water in Great Plains Thunderstorms," NSSL Tech. Note 49, ESSA, U.S. Dept. of Commerce, 1966.

Ryan, R.T., H.H. Blau, P.C. Von Thuna, M.L. Cohen, and G.D. Roberts, "Cloud Microstructure as Determined by an Optical Cloud Particle Spectrometer," *J. Appl. Meteorol.*, **11:** 149–156, 1972.

Schilling, G.F., "Theoretically Permissible Altitudes and Seasons for the Occurrence of Clouds Near the Mesopause," Rand Corp. Paper P-2899, AD435637, 1964.

Schlatter, E. and M. Lefkowitz, "Evaluation of Multi-Transmisometer Systems," RD-68-49, ESSA-Weather Bureau, Atlantic City, N.J., 1968.

Shanklin, M.D. and J.B. Landwehr, "Photogrammetrically Determined Cloud-Free Lines-of-Sight at Columbia, Mo.," Final Report under Contract F19628-68-C-0140, University of Missouri, AFCRL TR-71-0273 AD725758, 1971.

Scholz, T.G., D.H. Ehhalt, L.E. Heidt, and E.A. Martell, "Water Vapor, Molecular Hydrogen, Methane, and Tritium Concentrations Near the Stratopause," *J. Geophys. Res.*, **75:** 3049–3054, 1970.

Sims, A.L. and D.M.A. Jones, "Climatology of Instantaneous Precipitation Rates," AFCRL TR-73-0171, AD760785, March 1973.

Sissenwine, N. and R. Cormier, "Synopsis of Background Material for MIL-STD-210B, Climatic Extremes for Military Equipment," AFCRL-TR-74-0056, AD780508, 1974.

Soberman, R.K., "Noctilucent Clouds," *Sci. Am.*, pp 51–59, June 1963.

Somerville, P.N., and S.J. Bean, "A New Model for Sky Cover," AFGL-TR-79-0219, ADA078368, 1979.

Somerville, P.N., S.J. Bean, and S. Falls, "Some Models for Visibility," AFGL-TR-79-0144, ADA075490, 1979.

Stickel, P.G., and T.A. Seliga, "Cloud Liquid Water Content Comparisons in Rain Using Radar Differential Reflectivity Mesurements and Aircraft Measurements," *Preprints 10th Conference Radar Meteorology*, American Meteorol. Society, Boston, 30 Nov-3 Dec 1981.

Sulakvelidze, G.K., N.Sh. Bibilashvili, and V.F. Lapcheva, "Formation of Precipitation and Modification of Hail Process," Gidrometeorologicheskoe Izdatel'-stvo, Leningrad, Translation by Israel Program for Scientific Translations, Tel-Aviv (No. 1806) 1967.

Suring, R., *Die Wolken*, 2nd Ed., Leipzig, 1941.

Taljaard, J.J., H. vanLoon, H.L. Crutcher, and R.L. Jenne, *Climate of the Upper Air—Southern Hemisphere, Vol. 1, Temperatures, Dew Points and Heights Selected Levels*, U.S. Dept. of Commerce, ESSA, Environmental Data Service, NAVAIR 50-IC-55, 1969.

Tattelman, P., "An Objective Method for Measuring Surface Ice Accretion," *J. Appl. Meteorol.*, **21:** 153–166, 1982.

Tattelman, P. and D.D. Grantham, "A Survey of Techniques for Estimating Short-Duration Precipitation Rate Statistics," AFGL-TR-82-0357, ADA125705, 1982.

Tattelman, P. and I.I. Gringorten, "Estimated Glaze Ice and Wind Loads at the Earth's Surface for the Contiguous United States," AFCRL-TR-73-0646, AD775068, 1973.

Tattelman, P. and K. Scharr, "A Model for Estimating 1-Minute Rainfall Rates," *J. Climate Appl. Meteorol.*, **22:** 1575–1580, 1983.

Thornwaite, C.W., J.R. Mather, D.B. Carter, and C.E. Molineux, "Estimating Soil Moisture and Tractionability Conditions for Strategic Planning," AFCRC TN-58-201 (1 and 2), AD146789, 1958.

U.S. Dept. of Defense, "Military Standard Climatic Extremes for Military Equipment," *MIL-STD-210B*, U.S. Government Printing Office, 1973.

U.S. Weather Bureau, "Rainfall Intensity-Duration-Frequency Curves," Tech. Paper 25, 1955.

U.S. Weather Bureau, *Determination of Snow Loads for Building Construction*, Div. of Climatological and Hydrological Services, Washington, D.C., 1951.

Varley, D.J., I.D. Cohen, and A.A. Barnes, "Cirrus Particle Distribution Study, Part 7," AFGL-TR-80-0324, ADA100269, 1980.

Waters, J.W., J.J. Gustincic, R.K. Kakar, A.R. Kerr, H.K. Roscoe, and P.N. Swanson, "The Microwave Limb Sounder Experiment: Observations of Stratospheric and Mesospheric H_2O in Interhemispheric Survey of Minor Upper Atmospheric Constituents during October-November 1976," NASA Tech. Memo TMX 73630, 1977.

Weiss, R.R., J.D. Locatelli, and P.V. Hobbs, "Simultaneous Observations of Cloud and Precipitation Particles with Vertically Pointing X-Band and K_a-Band Radars," *IEEE Trans. Geosci. Electr., V.* **GE-17:** 151, 1979.

Winner, D.C., "Climatological Estimates of Clock-Hour Rainfall Rates," Air Weather Service Tech. Rept. 202, 1968.

CHAPTER 17

WINDS

Section 17.1	**D.D. Grantham and I.I. Gringorten**
Section 17.2 and 17.3	**A.E. Cole and A.J. Kantor**
Section 17.4	**I.I. Gringorten and P.I. Tattelman**
Section 17.5	**F.A. Marcos**

The atmosphere's motions defy rigorous classification or modeling. The application of a particular feature of wind structure to a given engineering problem should be dictated by its physical dimensions. Extrapolation of data beyond its indicated limitations is risky. Local conditions may not always be well represented by the data in this chapter, and may produce extreme wind variations in excess of those presented, even when the local wind structure is free from perturbations such as fronts, thunderstorms, or squalls. The practicing engineer should avail himself of applicable local meteorological records whenever possible. Information other than that given below is available through the various national weather services and the World Meteorological Organization [NOCD, 1980]. Special studies prepared by Environmental Technical Applications Center (ETAC) or Air Force Geophysics Laboratory (AFGL) might provide the best answers for specific design problems.

The measure of wind, in speed and direction, presents an immediate problem in the time interval for the observation. A conventional observation of wind speed is the wind travel in 1 min, 5 min, or 1 h, that is, the 1-min, 5-min, or 1-h average. The current standard averaging time period in the United States is 1 min. In England and Canada a 10-min period is customary when wind speed records are available, otherwise the averaging period is something over 15 s. However, in published climatic data, hourly (60 min) averaged winds are often given.

As might be expected, the variance of the 1-min wind is greater than the 5-min, which in turn is greater than that of the 1-h wind, but not significantly so. On the other hand, wind speeds of shorter duration than 1 min are subject to significantly greater variability. When the wind speeds peak between lulls, in 20 s or less, they are conventionally termed *gusts*.

17.1 WIND AS A FUNCTION OF HEIGHT

A major problem with pooling the various surface wind observations is in the determination of the best method or model for adjusting wind speeds to a common height above the surface. Once established, however, the model formula can be used to describe wind speed and gusts along a vertical profile in the lowest 50 or 100 m of the atmosphere.

Wind flows in response to pressure gradients in the atmosphere. Such pressure gradients change slowly with altitude, negligibly within the first 100 m. Yet the changes in wind speed with height are pronounced. Air motion near the surface does not obey the simple pressure gradient law. Anemometers near the ground may be hardly turning, whereas those on tall buildings or towers may show moderately strong, gusty winds. Kites may be difficult to launch, but once several hundred feet high, they may fly without difficulty.

Friction caused by terrain is one of the main factors affecting the horizontal wind speed up to an altitude known as the gradient level. At this height, 300 or 600 m, the pressure gradient is said to be dynamically balanced against two other influences: the earth's rotation and the curvature of the wind path. A theoretical wind speed that closely approximates the observed wind at gradient height can be computed from the isobaric spacing and curvature on surface weather maps.

The height of the gradient level and the velocity profile up to that level vary greatly, mainly with the type of surface and the stability of the air. Stability is chiefly a function of the temperature structure in the boundary layer. One extreme of temperature structure is represented by a super adiabatic lapse rate in which temperature decreases rapidly with altitude so that air displaced upward will continue upward because it is warmer than its surroundings. The opposite extreme is a negative lapse rate or inversion, in which temperature increases with height, so that air displaced upward is cooler than its surroundings and tends to sink back to its original level. A neutral condition (adiabatic) exists when the temperature lapse rate is such that a parcel of air displaced vertically will experience no buoyant acceleration. In general, a neutral lapse rate is established by the turbulent mixing caused by strong winds at the surface.

A popular model for the shift of wind speed and direction with altitude in the boundary layer, originally developed for ocean depth, is termed the Ekman Spiral. Since its introduction, micrometeorologists have studied the energy trans-

CHAPTER 17

Table 17-1 Ratio of wind speed at height H to speed at 91 m over open prairie obtained during the Great Plains Turbulence Field program.

H (meters)	Lapse Rate ($dt/dz < 0$)	V_H/V_{91} Inversion ($dt/dz > 0$)	Isothermal ($dt/dz = 0$)
91	1.000	1.000	1.000
30	0.965	0.689	0.872
21	0.944	0.608	0.829
15	0.915	0.538	0.792
9	0.866	0.452	0.733
6	0.825	0.403	0.686
3	0.749	0.339	0.604
1.5	0.662	0.275	0.518
0.6	0.556	0.231	0.424
0.3	0.470	0.200	0.336
0.15	0.383	0.166	0.300

fer and diffusion phenomena in the boundary layer and have found improved empirical relationships to fit wind speed data gathered at various heights above ground. Table 17-1 shows the ratio of wind speed at various heights to that at 91 m over open prairies. Values are based on actual wind measurements taken during the Great Plains Turbulence Field Program conducted at O'Neill Nebraska [Lettau and Davidson, 1957]. The ratios are shown for typical daytime lapse rates ($dT/dz < 0$), night-time ($dT/dz > 0$) and for isothermal conditions ($dT/dz = 0$).

17.1.1 Variation of Wind Speed with Height (Lowest 100 m)

Two alternative classes of models have been used to estimate the increase of wind speed with height: *logarithmic* and *power* models.

In one logarithmic model

$$\frac{V}{V_1} = \frac{\ln\frac{(1 + H)}{H_o}}{\ln\frac{(1 + H_1)}{H_o}} \qquad (17.1)$$

V is the mean wind speed at height H, V_1 is the mean wind speed at the reference level H_1 (anemometer level), and H_o is the *roughness parameter* a length determined by the characteristic ground surface. The boundary condition at $H = 0$ is $V = 0$. This model has the advantage that the effect of terrain is included explicity. Typical values for H_o are given in Table 17-2. Table 17-3 lists V/V_1 for a variety of roughness parameters H_o, for two reference levels, and for various heights H.

In the simplest power model the mean wind speed V at height H is approximated by

Table 17-2. Typical values of the roughness parameter, H_o.

Type of Surface	H_o (cm)
Smooth (mud flats, ice)	0.0009
Lawn, grass up to 1 cm	0.09
Downland, thin grass up to 10 cm	0.61
Thick grass, up to 10 cm	2.25
Thin grass, up to 240 cm	4.9
Thick grass, up to 240 cm	9.1

$$\frac{V}{V_1} = \left(\frac{H}{H_1}\right)^p \qquad (17.2)$$

where the exponent p depends on the height, terrain, thermal stratification and speed of the overall airflow. The parameter p is larger for rough ground, for altitudes below H_1, and for relatively small V_1. It varies within limits, approximately 0.05 to 0.8, and averages between 0.1 and 0.3. Table 17-4 lists mean values of p, determined for several locations and two types of terrain. Table 17-5 lists mean values of p that were determined for typical daytime ($dT/dz < 0$) night-time ($dT/dz > 0$), and isothermal conditions in the lowest 9 m over open prairie country (Great Plains Turbulence Field Program). The exponent p is larger when there is a stabilizing inversion and smaller when there is a positive lapse rate. According to Sherlock [1952] a typical value for p is 1/7 (or 0.143). Early workers had recognized that his p-value was applicable to typical steady or mean winds but not applicable to gustiness. Sherlock [1952] noted that gusts were better described with a value of $p = 0.0625$. Shellard [1965], in reducing high wind speeds and gusts to a common height of 10 m, used Equation (17.2) with $p = 0.17$ for mean speeds and 0.085 for gusts.

A special study was made with data from the Windy Acres Project [Izumi, 1971] consisting of 39 h of 1-s wind

Table 17-3 Ratios of wind at height H to wind at height H_1 for two reference levels and three roughness parameters H_o, calculated by Equation (17.1)

H (m)	$H_o = 30$ cm	$H_o = 3$ cm	$H_o = 0.3$ cm
	V/V_1 for $H_1 = 6$ m		
100	1.91	1.53	1.37
10	1.16	1.10	1.07
1	0.48	0.67	0.76
0.5	0.32	0.54	0.67
	V/V_1 for $H_1 = 3$ m		
100	2.42	1.76	1.51
10	1.47	1.26	1.17
1	0.61	0.77	0.84
0.5	0.41	0.62	0.74

Table 17-4. Mean values for heights ranging from 10 to 100 m of exponent p in Equation (17.2).

Location and Terrain	p
Fairly Level Open Country	
Ann Arbor, Michigan	0.14
Sale, Victoria, Australia	0.16
Cardington, Bedfordshire, England	0.17
Leafield, Oxfordshire, England	0.17
Fairly Level Wooded Country	
Quickborn, Germany	0.23
Upton, Long Island, New York	0.26
Akron, Ohio	0.22

speeds, taken in the summer of 1968, at 8 heights on a 32 m tower in southeast Kansas. The terrain was very flat and partly covered with wheat stubble 16 to 22 cm tall. The 1-s wind was never more than 15.3 m/s. Figure 17-1 (right side) presents resulting isopleths of p as a function of windspeed, from 1-s winds including gusts to 1-h integrated wind speeds. The exponents p were found for wind speeds of the same probability, from level to level, as opposed to the mean wind speed. Nevertheless previous results were supported. There is a systematic decrease of p, from 0.7 down to 0.12, with increasing wind speed for either 1-s, 1-min, 5-min or 1-h winds, when the lower 90% of the wind speeds are considered. For winds equal to or greater than the 90th percentile, the exponent p is almost uniform at 0.12 except for gusts or short-duration winds of 1 min or less. The Windy Acres winds became turbulent above the 95% speeds. In gusts, the value of p varies from 0.11 down to 0.08, suggesting a tendency toward a common speed throughout the turbulent layer.

Equation (17.2) can be used to standardize the height of the wind data of individual stations to one level even though they have differing anemometer levels. In the publication "Climatic Extremes for Military Equipment" [MIL-STD-210B] the value of the exponent p was adopted at 0.125 when the wind is strong but steady, and at 0.08 when the wind is strong and gusty.

The validity of a wind extrapolation to another height is dependent on the representativeness of the wind measurement at the reference level. If the anemometer mast is in a poorly defined terrain, the use of the wind profile formula is questionable. Uniformity of the terrain would improve the result. Winds observed with a land-based anemometer cannot be used to estimate the wind speed over an adjacent water surface. In certain cases the wind speed over terrain may attain a maximum speed at a level significantly below 100 m. Such cases usually occur in cold air flow, for example, nocturnal down-slope winds or sea breezes.

Table 17-5. Mean values of exponent (p) for the lowest 9 m in Equation (17.2), determined from wind profile measurements.

H (m)	p		
	Lapse Rate ($dt/dz < 0$)	Inversion ($dt/dz > 0$)	Isothermal ($dt/dz = 0$)
9.0	0.11	0.38	0.14
4.6	0.13	0.31	0.16
0.9	0.18	0.23	0.21

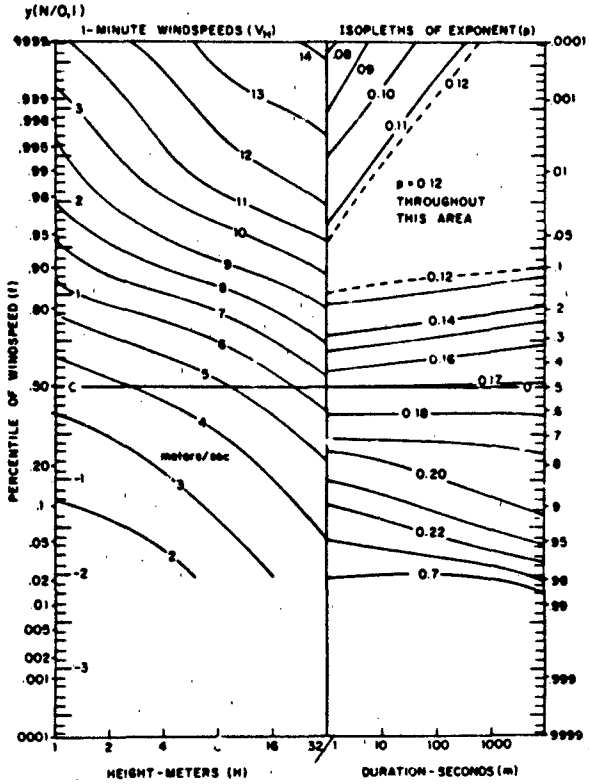


Figure 17-1. A nomogram to obtain windspeeds of duration 1 to 1000 s at a height of 1 to 32 m. (The right hand side of the Nomogram shows the exponent (p) in the wind power model (Equation 17.2) as a function of the duration of the wind, including 1-s, 1-min, 5-min, and 1-h averages, and as a function of the strength of the wind (in percentiles). Left-hand side shows percentiles of 1-min windspeed as a function of height above ground level.)

17.1.2 Wind Direction Shifts (Below 3000 m)

Normally wind direction changes with height. Changes in direction with height are termed veering if the wind turns clockwise and backing if the wind turns counter clockwise. Veering is usually expressed as the rate of turning in degree per height intervals (negative for backing).

CHAPTER 17

Wind direction shifts with height are caused by surface frictional effects and by height changes of the horizontal pressure patterns controlling the mean airflow. The following discussion excludes such phenomena as slope winds, land and sea breezes, nocturnal low-level airflow conditions in the equatorial zone, and rapid, small-scale wind fluctuations.

In the free atmosphere, the mean horizontal airflow is a gradient wind, approximately parallel to the isobars, the lower pressure being to the left. Surface friction reduces the speed, causing a component of the surface wind to blow across the isobars toward lower pressure. Thus the wind direction changes with height to align itself with the gradient wind in the free atmosphere.

To a first approximation, it can be said that under strong mean wind conditions in the Northern Hemisphere, the winds will veer with height in the lower 1000 m, the magnitude of the veering being determined by the intensity of thermal advection processes. Warm air advection increases veering, and cold air advection either decreases the magnitude of frictional veering or causes backing of the winds. The conditions for veering and backing are reversed in the Southern Hemisphere. North of approximately 20°N latitude, winds in the lowest 1000 m of the atmosphere will usually display varying degrees of veering with relatively few cases of backing. Southerly surface winds will veer more with height than northerly surface winds. Above approximately 1000 m, southerly winds will continue to veer with height, while northerly winds will begin to show backing with height.

Tables 17-6, 17-7 and 17-8 indicate the order of magnitude of veering with associated general meteorological conditions. Because no advection is permitted for the situation described in Table 17-6, the veering with height is almost constant and represents the gradual decrease of the surface frictional effect with height. Once the gradient level of the free atmosphere is reached, frictional effects and hence veering become negligible in the case of no advection. Strong surface winds generally make an angle with the gradient wind of 10° to 30°; this is the overall veering found in the frictional layer of the atmosphere when little or no advection exists. The direction of the surface wind is insignificant. Total veering over the entire layer of frictional influence depends primarily on the roughness characteristics of the earth surface: the average for oceans is 5° to 15°, for continents, 25° to 45°. The average veering is usually greater in winter than in summer, and greater at northern stations than at southern stations. The averaging process masks the variability of veering that would be encountered with isolated observations in time and space. In the first 1000 m of the atmosphere, however, the importance of the general direction of the surface wind in obtaining reasonable estimates of veering appears doubtful. Above this layer southerly winds veer with height, northerly winds back with height. Maximum values of veering in the frictional layer are near 3°/30 m; isolated cases of backing of the same order of magnitude are observed.

In summary, the average total veering (or backing) in the lower 1000 m is 20° to 40°, with isolated cases of 70° to 90°. To a first approximation, it may be assumed that this veering (or backing) is evenly distributed throughout the layer. Above this layer, primarily dependent on horizontal advection conditions, winds will show veering or backing with approximately the same average order of magnitude as in the frictional layer.

Table 17-6. Typical frictional veering of wind over plain land with moderately strong gradient winds (18 m/s) and no temperature advection; 51.3°N, 12.5°E, 20 October 1931. [H. Lettau, 1950]

Altitude (100m)	Average Speed (m/s)	Angle* (deg)	Veering (deg/100m)
6.0 to 9.0	18	7	2.3
3.0 to 6.0	16	15	2.6
1.5 to 3.0	13	21	2.6
0.0 to 1.5	9	25	2.3

*Angle between the wind and the gradient wind

Table 17-7. Average angle formed by the wind and the gradient wind, and average veering for weather stations in Germany, grouped according to topography. [H. Lettau, 1939]

Altitude (100m)	Coastal Plains	Rolling Country	Hilly Land
	Angle (deg)		
9.0 to 15.0	0	2	3
6.0 to 9.0	2	5	10
3.0 to 6.0	10	17	25
1.5 to 3.0	22	30	36
0.0 to 1.5	29	36	43
	Veering (deg/100m)		
9.0 to 15.0	0.3	0.7	1.0
6.0 to 9.0	0.7	2.0	3.9
3.0 to 6.0	4.3	5.9	5.2
1.5 to 3.0	4.9	4.6	4.3
0.0 to 1.5	4.9	3.0	4.3

17.1.3* Diurnal Variation and Low-Level Jet Streams (below 2000 m)

The mean diurnal variation of wind speed at various heights above any given site is caused by diurnal variations of both the horizontal pressure gradient force and the frictional force. The regular variations of the former are controlled by tidal effects (solar and lunar) which produce pre-

*This is based on the section by H. H. Lettau and D. A. Haugen in the *Handbook of Geophysics for Air Force Designers*, 1957.

Table 17-8. Average veering (deg 100 m) for various ranges of height; means (1918 to 1920) for three stations in U.S.A. in the latitude zone 31° to 36° (average 33°N), and three in latitude zone 40° to 45° (average 43°N). [Condensed from W. R. Gregg, 1922.]

Altitude (100 m)	Southerly Surface Winds				Northerly Surface Winds			
	Summer		Winter		Summer		Winter	
	33°N	43°N	33°N	43°N	33°N	43°N	33°N	43°N
27.0 to 36.0		2.3	3.2	2.3		−3.3	−3.0	−2.3
15.0 to 27.0	1.0	3.3	2.3	3.3	−1.0	−3.3	−1.0	−2.3
6.0 to 15.2	1.0	2.3	2.3	3.0	−1.6	−1.3	−1.0	−2.3
3.6 to 6.0	3.0	2.3	3.3	3.3	−1.0	−1.0	3.0	3.0
1.5 to 3.6	2.3	2.3	3.0	4.3	2.3	2.3	3.0	4.3
0.0 to 1.5	2.3	2.3	3.6	4.3	1.6	2.3	3.6	4.3

dominantly a 12-hourly wave, and by differential solar heating of the air over different locations and subsequent horizontal density gradients. In the troposphere the tidal motions of the atmosphere are small (range less than 0.5 m/s). The barometeric effects of differential solar heating produce marked diurnal wind variations only in special locations (along coast lines and the rims of extended high plateaus). Over most parts of the continents, the diurnal variation of wind speed is controlled by the horizontally uniform effects of the cycle of solar heating and nocturnal cooling of the earth's surface. Consequent changes in the vertical thermal stratification of the atmosphere at 1000 to 2000 m significantly influence the effective frictional force in large-scale air flow.

Over relatively smooth land the daytime thermal stratification intensifies the vertical mixing and the nocturnal thermal stratification weakens it. This causes a wind speed maximum near the ground at about mid-afternoon and a minimum in the early morning hours. As seen in Figure 17-2, the phase of diurnal wind variation is reversed approximately 100 m above the ground, a level that varies with climatic zone, season, and surface roughness from 100 to 200 m. The amplitude of the diurnal variation of wind speed normally has two maxima, at approximately 6 and 600 m. The vertical extent of such diurnal variation varies roughly as the vertical extent of convective activity (2000 m).

Over the midwestern United States the nocturnal maximum of wind speed above approximately 100 m frequently leads to a sharp peak of the velocity profile at heights of 300 to 900 m. The peak, usually at the top of the nocturnal inversion, is significantly stronger (supergeostropic) than explained by a balance between the horizontal pressure gradient and the Coriolis forces; it is often associated with extremely large values of wind shear. These peak winds are called low-level jet streams. The supergeostrophic wind speeds (peak speeds up to 26 m/s in a pressure field resulting in 10 to 15 m/s of geostrophic speed) suggest that an inertial oscillation of the air masses is induced when the constraint imposed by the daytime mixing is released by the initiation of stable thermal stratification near sunset. Average wind profiles showing the development of a low-level nocturnal jet stream over O'Neill, Nebraska, are given in Table 17-9. A particularly large value of wind speed variation with height, recorded during the development of a low-level jet stream, was obtained by kite observation at Drexel, Nebraska on the night of 18 March 1918; the surface wind speed was 26 m/s, but at an elevation of 238 m the reported wind speed was 36 m/s, yielding an average shear of 0.14 s^{-1}.

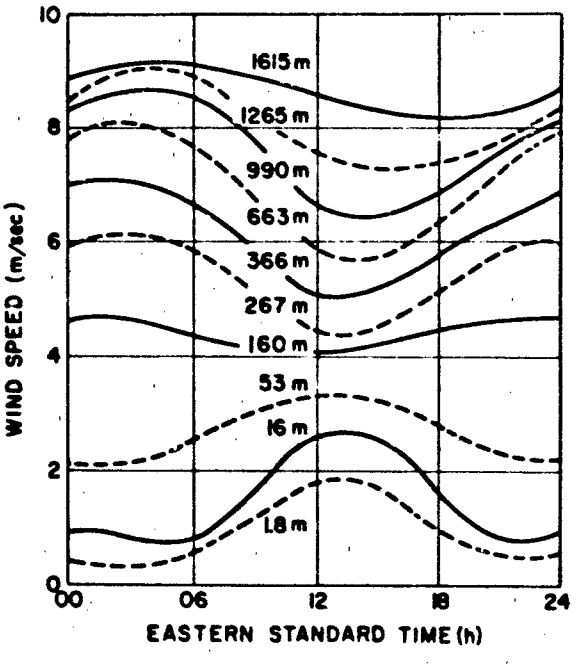


Figure 17-2. Diurnal curves (slightly smoothed) of annual mean wind speed at Oak Ridge, Tennessee, average for 1948 to 1952. (The wind speed data at heights of 1.8 and 16 m are from anemometer recordings; for other heights, data were obtained from pilot balloon soundings at 4 h intervals [data lines are for clarity of illustration].)

Table 17-9. Average wind profiles (speed and direction at various heights) showing development of nocturnal low-level jet stream (Average of five nights of observations, Great Plains Turbulence Field Program, O'Neill, Nebraska.)

Height (km)	Mean Local Time							
	1800		2000		2200		2400	
	Speed (m/s)	Dir. (deg.)	Speed (m/s)	Dir. (deg.)	Speed (m/s)	Dir. (deg.)	Speed (m/s)	Dir. (deg.)
2.0	6.1	223	4.3	225	3.5	223	3.1	231
1.8	7.2	219	5.0	226	4.0	227	3.4	233
1.5	9.4	213	8.3	225	5.6	229	5.0	232
1.2	10.9	207	12.3	217	10.8	225	8.9	223
0.9	11.6	194	14.7	202	13.9	211	13.1	219
0.8	11.7	190	15.5	193	15.9	204	16.2	213
0.6	11.6	182	15.8	190	17.2	196	19.6	203
0.3	10.8	177	14.7	178	18.0	183	18.9	189
0.2	9.7	174	13.6	171	14.9	175	15.1	183
0.1	9.2	173	12.2	167	12.8	170	12.9	170
	0200		0400		0600		0800	
	Speed (m/s)	Dir. (deg.)	Speed (m/s)	Dir. (deg.)	Speed (m/s)	Dir. (deg.)	Speed (m/s)	Dir. (deg.)
2.0	4.0	227	3.7	214	3.3	208	4.5	208
1.8	4.2	226	3.9	220	3.7	212	4.9	208
1.5	5.4	226	4.5	227	4.8	213	6.2	204
1.2	8.5	227	8.0	228	6.7	218	8.3	203
0.9	13.5	224	12.1	228	10.5	220	11.2	213
0.8	16.5	219	14.5	225	12.7	220	12.4	215
0.6	18.9	211	17.5	217	15.6	217	13.9	215
0.5	19.9	203	18.9	209	17.2	213	15.9	213
0.3	19.0	196	18.7	199	17.3	204	14.0	204
0.2	17.1	189	16.2	191	14.5	195	12.6	195
0.1	13.4	187	13.9	189	12.7	191	11.8	191

Over the ocean, the diurnal variation of wind speed is negligible because diurnal variations in thermal stratifications are extremely small. In coastal regions where land and sea breezes are experienced, the amplitude and phase of the diurnal variation of wind speed are comparable with those over land. A reversal of phase with height is questionable, however. The vertical extent of the sea breeze is roughly 1000 m, the land breeze roughly 500 m, the overall land-sea breeze system including return flow, about 3000 m.

In mountainous regions, the daytime, up-slope valley winds (1 to 3 m/s) are normally stronger than the night-time down-slope mountain winds (0.2 to 2 m/s). The mean amplitude of the variation decreases very slightly with height. A reversal of phase with height is occasionally found. The vertical extent of the valley breeze is about that of the height of the surrounding ridges, the mountain breeze, 60 to 100 m. The whole system extends about 300 m above the height of the ridges.

The various features of local wind variation discussed here occur quite frequently. In general, they are easily observed during conditions of clear skies and light to moderate intensity of the large-scale airflow. It should be remembered, however, that when the diurnal variations are not readily observed, it may be because they are superimposed on the large-scale airflow. Empirical wind structure models not accounting for possible diurnal variations could then be in error.

17.2 LARGE SCALE WIND STRUCTURE

This section provides information on the vertical and horizontal distribution of winds for altitudes up to 90 km. The wind data for altitudes up to about 26 km are based on routine rawinsonde observations taken by the National Weather Service (NWS). The root mean square (rms) observational errors in vector wind measurements using FPS-16, T-19, or similar tracking radar for altitudes up to 26

km are 1 m/s plus 2% the vector wind speed. The information presented on winds between 26 and 60 km was obtained from Meteorological Rocket Network (MRN) observations. Most of the available observations are for locations in the Northern Hemisphere. The estimated rms observational errors at these altitudes are 2 m/s plus 3% of the vector wind. The wind distribution at altitudes between 60 and 90 km are based on data obtained from grenade and inflatable falling sphere experiments. The estimated rms error for the falling sphere observations, which make up the largest portion of the data, is 2 to 3 m/s between 60 and 90 km.

17.2.1 Seasonal and Day-to-Day Variations

The broad features of the seasonal change in winds between the surface and 90 km are reasonably well established. Figure 17-3 is a meridional cross section of the observed mean monthly zonal wind components for January and July from the surface to 80 km. The seasonal change in the stratospheric and mesospheric wind fields differs from tropospheric seasons in timing and length of season. At middle and high latitudes, long periods of easterly and westerly flow are separated by shorter periods of transition as shown by the curves of the mean monthly 50-km zonal wind components in Figure 17-4 for White Sands, 32°N, Wallops Island, 38°N, and Ft. Churchill, 59°N. The mean monthly meridional components (Figure 17-4) are less variable and maintain a slight southerly (positive) component at most locations throughout the year.

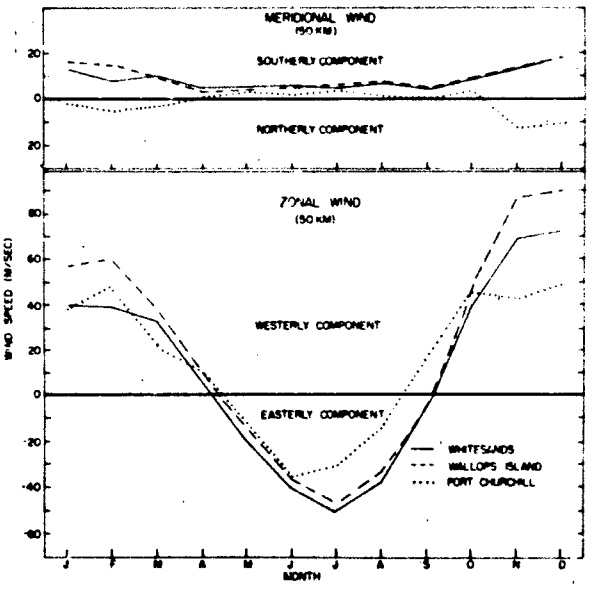


Figure 17-4. Annual variation of meridional and zonal wind at White Sands, Wallops Island and Ft. Churchill.

Profiles of mean monthly zonal wind speeds (m/s) for each of the midseason months at Ascension Island, Wallops

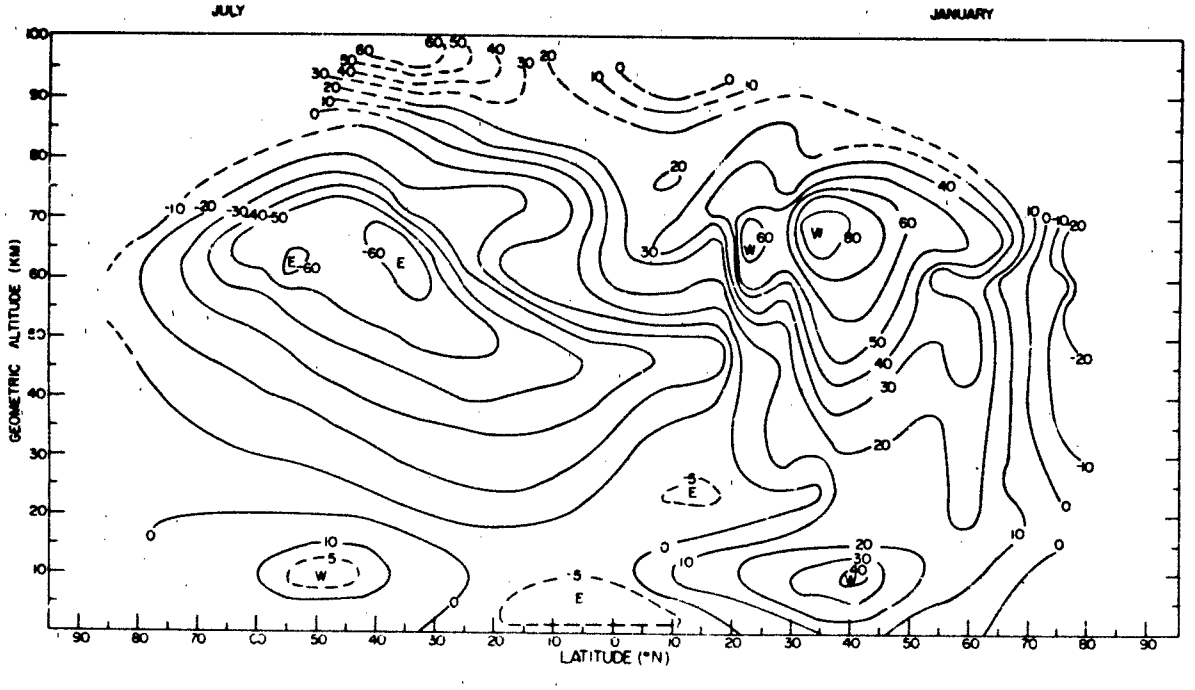


Figure 17-3. Meridional cross section of zonal wind speed (m/s).

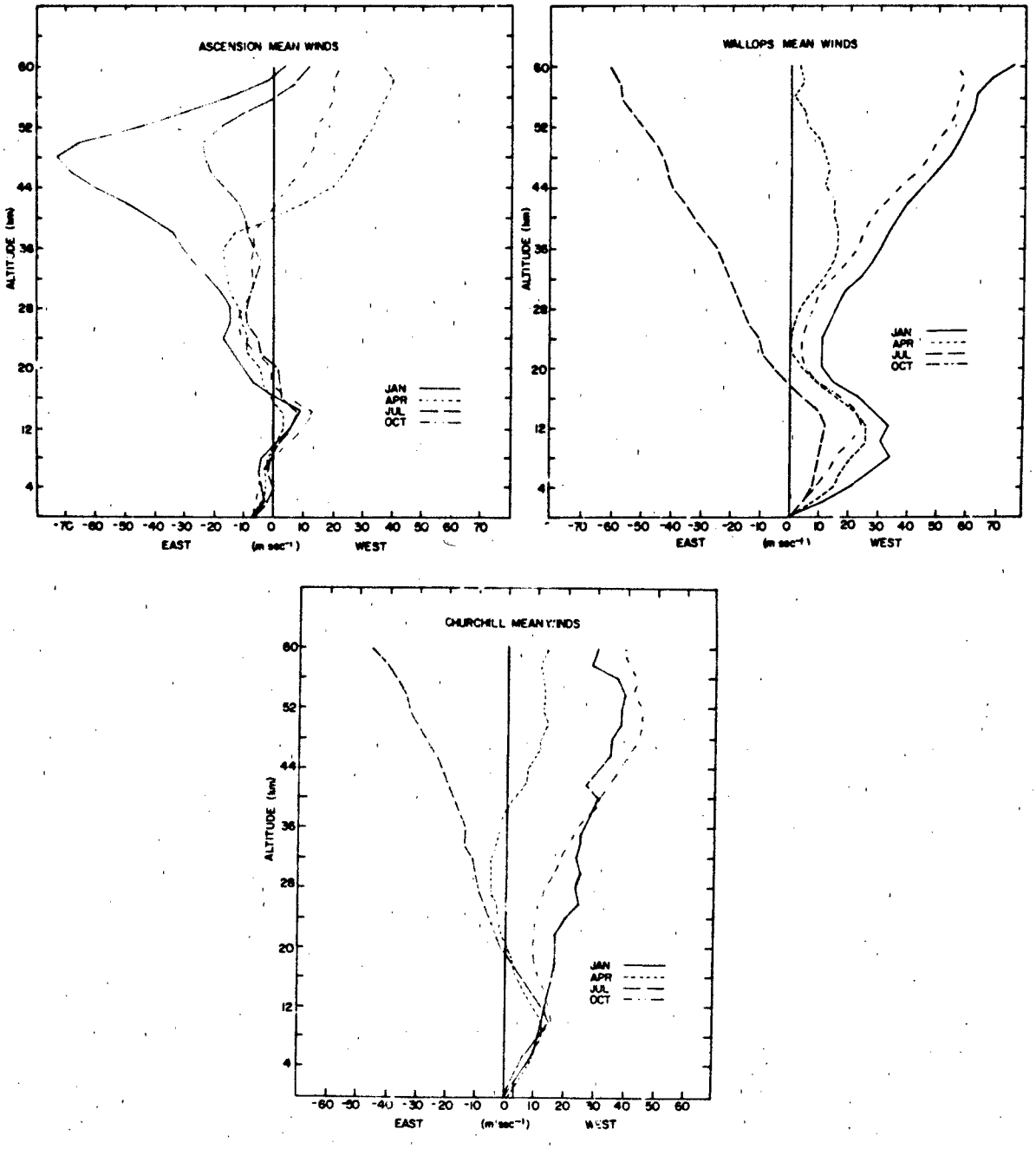


Figure 17-5a. Seasonal effects on the zonal wind profiles at Ascension Island, Wallops Island and Ft. Churchill.

Island and Fort Churchill, from a report by Kantor and Cole [1980] are plotted in Figure 17-5 for altitudes up to 60 km. The seasonal variations in the mean monthly zonal winds are largest at Wallops Island, a midlatitude location, and at altitudes above 30 km. The largest day-to-day variations around the mean monthly values (Figure 17-6) occur during the winter months at altitudes above 20 km at middle and high latitudes.

17.2.2 Time and Space Variations

The change and movement of pressure patterns in the atmosphere cause standard wind observations (for example, the mean wind over 1 min) to vary from time to time at a given place, and from place to place at a given time. In general, as shown in Figure 17-7, the amount of change in wind between two observations increases with the time in-

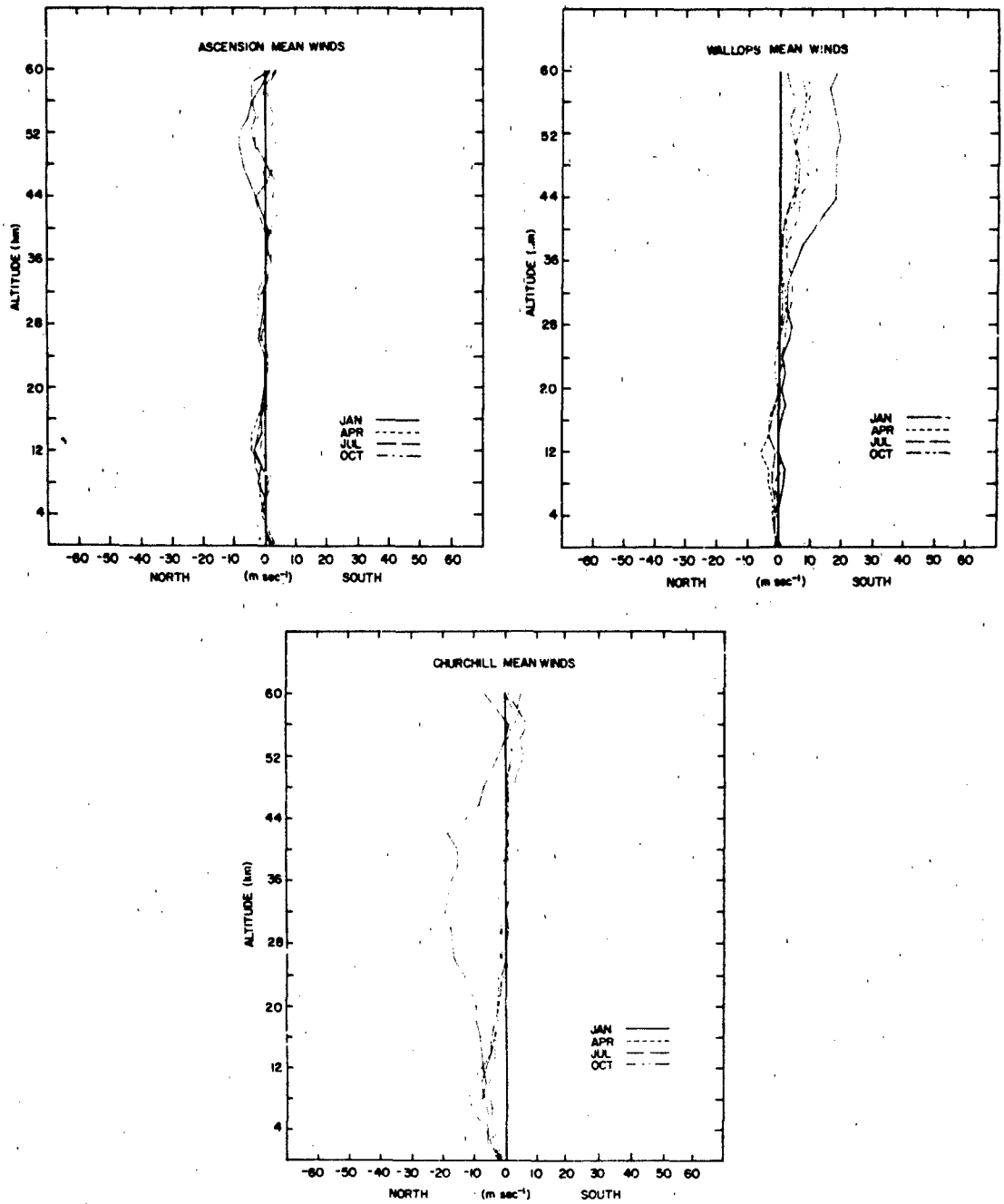


Figure 17-5b. Seasonal effects on the meridional wind profiles at Ascension Island, Wallops Island and Ft. Churchill.

terval between them and with the distance between observing points. The rate of increase in wind change with increasing time or space interval between observations is, in turn, a variable depending upon season, geographic location average wind speed, nature of the sample, and, to some extent, height above the ground.

Most of the wind variability data pertaining to the free atmosphere below 30 km are derived from standard pilot balloon or rawinsonde observations. Lower limits of resolution of these observations are such that minimum intervals for wind variability are about 15 min or 1.6 km. The small-scale fluctuations appear to be fairly random, and their combined effect on ballistic or synoptic-scale forecasting problems normally cancel out and, thus, are neglected.

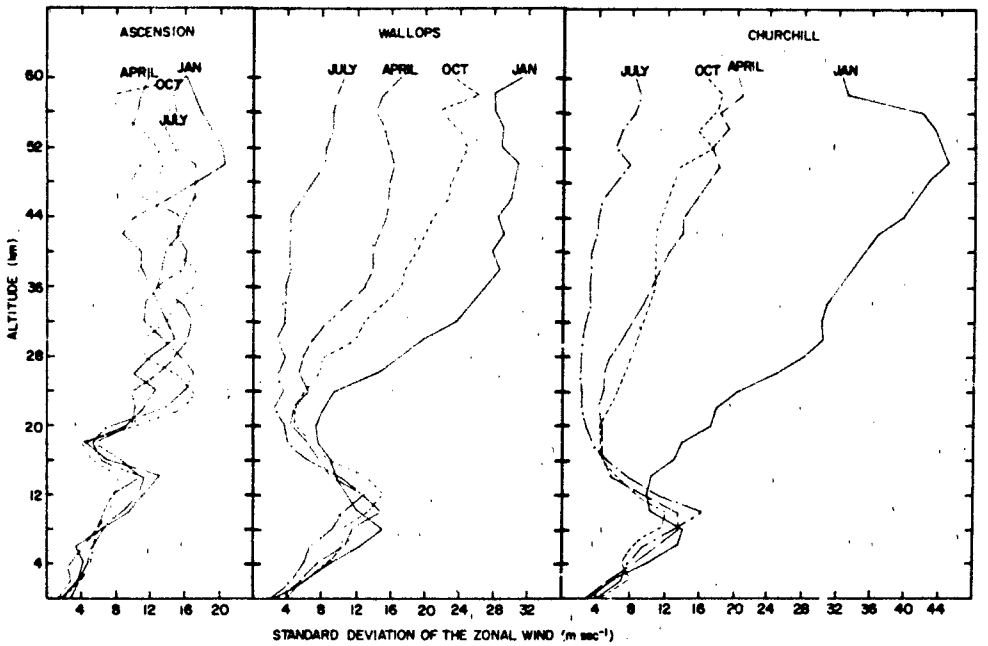


Figure 17-6a. Day-to-Day variability around mean monthly zonal winds for the midseason months at Ascension Island, Walleps Island, and Ft. Churchill.

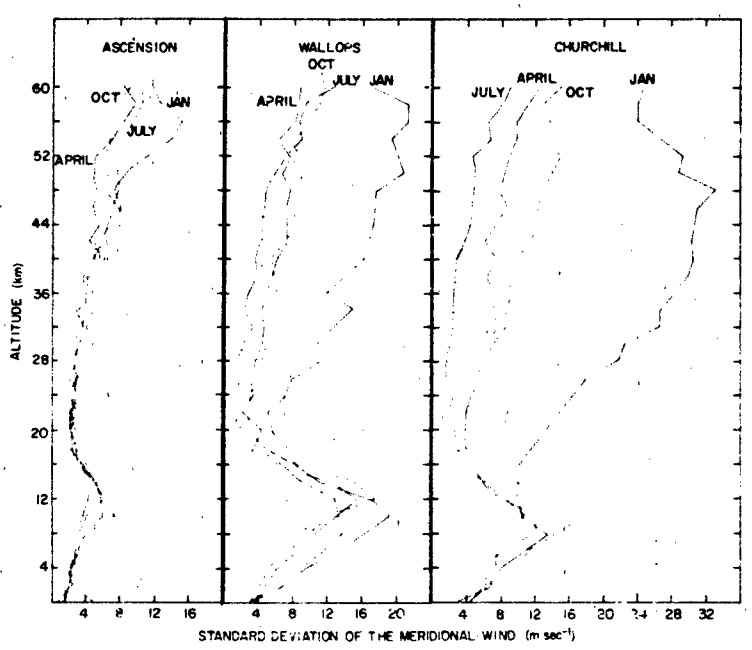


Figure 17-6b. Day-to-Day variability around mean monthly meridional winds for the midseason months at Ascension Island, Wallops Island and Ft. Churchill.

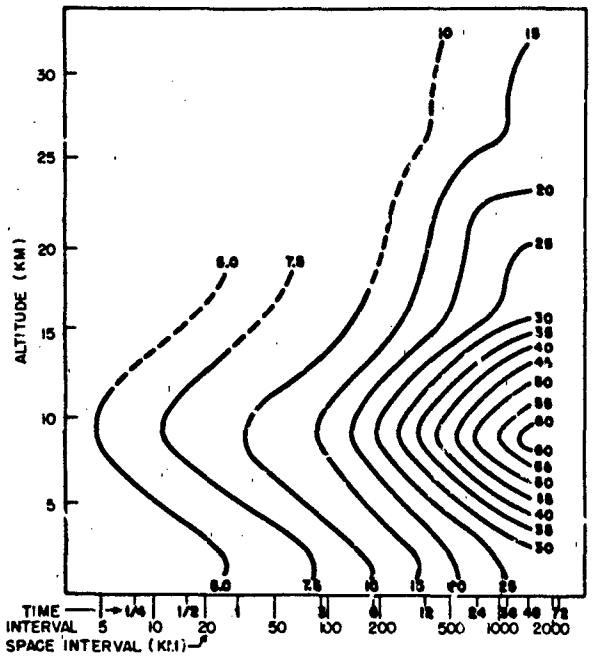


Figure 17-7. Isopleths of time and space variabilities of changes in the mean vector wind with altitude; the rms (63rd percentile) values, in kilometers, of observed changes for various time lags between observations, and of derived changes for various distances are given on each curve [Ellsaesser, 1960].

Observations of VHF and UHF radar backscatter from turbulence in the clear atmosphere, conducted at several locations during the past several years, permit much more detailed measurement of the time variation of vector winds than was previously attainable. Depending on the radar wavelength and the power-aperture product [Gage and Balsley, 1978], measurements are obtainable from near the surface to an altitude of about 100 km. Several radars which were originally designed for ionospheric research have been used for observations of the neutral (non-ionized) atmosphere, and a few radars have been built specifically for tropospheric and stratospheric wind measurements. The latter include NOAA radars in Colorado and Alaska and the SOUSY-VHF-Radar operated by the Max-Planck-Institut fur Aeronomie in the Federal Republic of Germany. Examples of data are presented by Green et al. [1979], Balsley and Gage [1980], Rottger [1979], and in numerous other publications. A radar of this type has been incorporated into the NOAA Prototype Regional Observation and Forecasting System (PROFS).

Several measures of wind variability are possible. The most useful measure is the (RMS) of the vector change in wind; others commonly used are the mean and the median absolute vector differences. All of these are scalar measures computed from the magnitudes of the difference vectors. They are related in a circular normal distribution which is usually a good approximation to the frequency distribution of wind changes.

17.2.2.1 Time Variability up to 30 km. The time rate of change of wind in the frictional layer is affected by the topography and the thermal structure. The results in Table 17-10 are for steady southerly flow conditions from seven observational periods, all but one of which extended over 24 h (pilot balloon, rawinsonde, and smoke-puff observations are combined). The effect of thermal stratification is indicated by comparision of the day and night values for time differences up to 8 h.

The change in wind variability above the frictional layer with increasing time between observations and with altitude is illustrated in Figure 17-7. Figure 17-8 shows the effect of latitude and season for a 24-h lag between wind observations.

For relatively short periods during which the pattern of the winds is fairly stable, the variability of the wind is given directly by

$$S_t = Kt^p, \qquad (17.3)$$

where S_t is the rms change in wind during the time interval t, and K is a constant. The exponent p depends on r_t, the correlation coefficient between winds separated by the time interval t. At short lags where r_t is 1, p is 1; at greater lags where r_t is 0, p is 0.5. For t in hours and S_t in m/s

$$S_t = 1.8\, t^{0.5} \qquad (17.4)$$

is a suitable generalization for middle latitudes and for lag intervals of 30 min to about 12 h. Although this empirical relation is an acceptable average, K actually depends on the mean wind speed. The mean wind varies with season, altitude, and geographic location (see Figure 17-5); hence, values of K other than 1.8 will on some occasions be more applicable to engineering problems. Values of K from 1.5 to 6.3 are tabulated by Arnold and Bellucci [1957]. An analysis of the relationship of K to the mean wind in a stable flow pattern shows that K increases from about 0.5 at speeds of 2 m/s to perhaps 2 or 3 at speeds of 16 m/s and higher.

An indirect model relates the variability in time to r and to the climatological dispersion of the winds. This relationship is given by

$$S_t = \sigma_t\sqrt{2(1 - r_t)}, \qquad (17.5)$$

where σ_t is the standard vector deviation of the winds (rms deviation from the vector resultant wind). The vector stretch correlation coefficient between the initial wind having components u and v and the wind after a time interval t, having components x and y, is given by

Table 17-10. Mean and standard deviation of absolute value of vector velocity differences at various time intervals, Δ t, in the lower 1829 m over smooth open terrain (Great Plains Turbulence Field Program).

Height (1000 m)	Velocity Differences (m/s)					
	Δt = 2 hours		Δt = 4 hours		Δt = 8 hours	
	Day	Night	Day	Night	Day	Night
1.8	3.0 ± 1.7	3.3 ± 2.4	5.0 ± 2.6	4.5 ± 3.2	7.4 ± 2.4	4.0 ± 2.9
1.4	2.6 ± 1.5	2.8 ± 1.9	4.2 ± 2.5	4.2 ± 2.6	5.4 ± 2.8	5.1 ± 3.3
0.9	2.6 ± 1.7	3.9 ± 2.5	4.1 ± 2.2	5.6 ± 2.6	5.4 ± 1.9	7.5 ± 3.1
0.5	3.1 ± 2.2	4.0 ± 2.8	4.8 ± 3.1	6.5 ± 3.5	6.8 ± 2.8	9.6 ± 3.1
0.4	3.0 ± 1.9	4.0 ± 3.0	4.8 ± 3.1	6.9 ± 4.5	7.1 ± 3.5	6.2 ± 3.5
0.2	2.6 ± 2.4	3.4 ± 3.1	4.3 ± 2.9	6.0 ± 5.4	6.2 ± 3.1	8.9 ± 4.9
	Δt = 12 hours		Δt = 16 hours	Δt = 20 hours	Δt = 24 hours	
1.8	5.6 ± 2.9		6.4 ± 3.1	7.5 ± 3.5	6.5 ± 3.3	
1.4	6.3 ± 3.1		6.6 ± 3.0	5.9 ± 2.1	4.8 ± 2.6	
0.9	7.8 ± 3.0		7.7 ± 3.5	5.9 ± 2.8	4.4 ± 2.6	
0.5	10.2 ± 3.9		8.8 ± 4.1	6.9 ± 4.0	4.9 ± 3.3	
0.4	10.4 ± 4.1		8.8 ± 3.6	7.1 ± 4.9	4.6 ± 2.9	
0.2	8.3 ± 3.7		7.2 ± 3.8	6.0 ± 4.9	4.3 ± 3.9	

$$r_t = \frac{\sum(ux + vy)}{\sqrt{\sum(u^2 + v^2)\sum(x^2 + y^2)}} \qquad (17.6)$$

This parameter undergoes an exponential decay with increasing lag for intervals up to 24 h or more, and then appears to oscillate about zero. The relation,

$$r_t = \exp(-0.0248t), \qquad (17.7)$$

with t in h, is widely used. This equation, in conjunction with Equation (17.5) and values of σ_t allows an estimate of the wind variability for a desired lag interval that pertains specifically to the place, season, and altitude of interest. This model has two serious limitations; it will not permit r_t to become negative, nor does the constant coefficient of t allow for variations in the rate of decay of the correlation. At sufficiently large time intervals, r_t does become negative, and r_t is so close to zero for lags in excess of 72 hours that the model becomes unreliable. In some cases r_t becomes negative at shorter lags, and the lag at which this occurs varies from place to place. Investigations of the rate of decay of correlation show that it varies geographically, seasonally, and probably with altitude. The variation has been mapped only for the United States at 5.5 km [Ellsaesser, 1960].

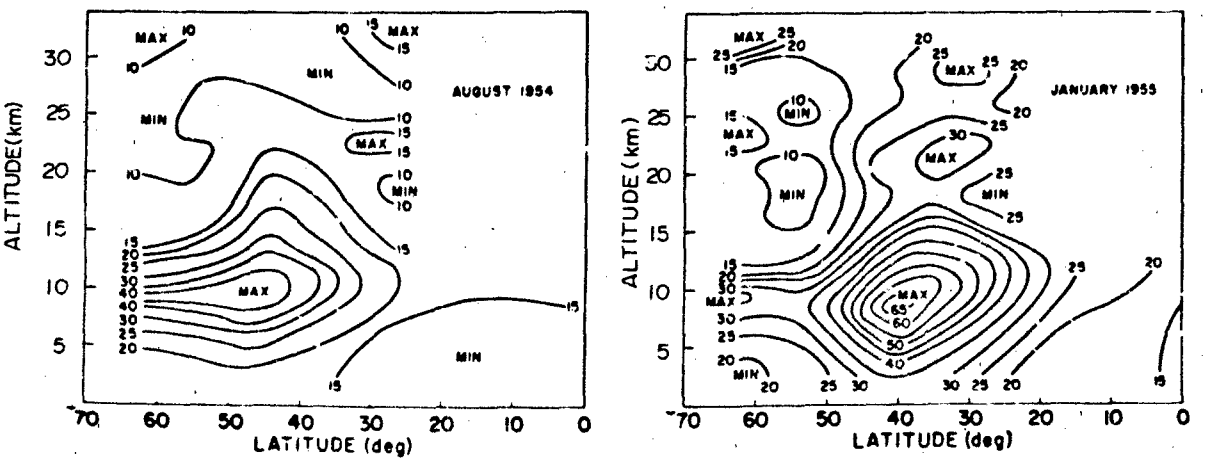


Figure 17-8. Distribution in latitude and altitude of the vector wind variability (St) for a 24-h time lag between observations along the meridian 75 W in summer and winter (August 1954 data are for East Africa); rms values in knots are shown on curves [Ellsaesser, 1960].

Attempts to develop a more precise model of the variability of winds with time have resulted either in only moderate improvement or extremely complex models. Thus Equation (17.7) is considered the most useful approximation. When precision in estimating the wind variability with time is required, a special climatological study must be made.

17.2.2.2 Spatial Variability up to 30 Km. In general, varability increases with increasing distance between observation points, and the rate of increase with distance depends on geographic location, season, and altitude. A change in the wind with time can be thought of as resulting in part from the movement of wind-field patterns over the observing point, and thus as analogous to the spatial variability.

Extending this analogy to the models, the space variability of wind S_d is given by $K'd^P$, where d is the distance between observing points. The parameter K′ varies with season, geographic location, and altitude, but no detailed examination has been made of the way in which these factors act. Arnold and Bellucci [1957] tabulated values of K′ from 1.1 to 6.1. They consider the expression, representative for middle latitudes.

$$S_d = 1.5d^{0.5}, \qquad (17.8a)$$

where S_d is in mph and d in miles, or converted to metric units

$$S_d = 0.53d^{0.5} \qquad (17.8b)$$

where S_d is in m/s and d in km.

The indirect model provides several empirical curves for the decay of the correlation coefficient of winds with increasing separation between observing points. These curves indicate differing rates of decay depending upon latitude and upon the orientation of the line connecting the observing points. The analogy between time and space variability of the winds extends to these curves. Figure 17-7 indicates that, for temperate latitudes, a general approximation to the space variability of winds can be obtained by taking 3 hours as equivalent to 92.6 km. The space variability is then estimated in a manner similar to time variability from $\sigma_d \sqrt{2(1 - r_d)}$; where r_d is the correlation coefficient between winds separated by the distance interval d.

17.2.2.3 Time and Space Variations—30 to 60 km. Observations show that pronounced diurnal and semi-diurnal oscillations exist in the winds at altitudes above 30 km. Sufficient data, however, are not available to permit the development of a satisfactory model for all latitudes and seasons. The results of an analysis of a series of wind observations taken at Wallops Island during a 48-h period in May 1977, shown in Figure 17-9, provide an indication of the magnitude of the diurnal oscillation at a midlatitude location. The X's in Figure 17-9 are average values of the north-south and east-west wind components at the indicated local times. The curves are based on a harmonic analysis of the meridional (N-S) and zonal (E-W) wind components. The amplitude of the combined diurnal and semidiurnal oscillations of the N-S component, increases with altitude above 30 km, reaching a maximum of 10 to 11 m/s between 50 and 55 km. The diurnal variations of the E-W component

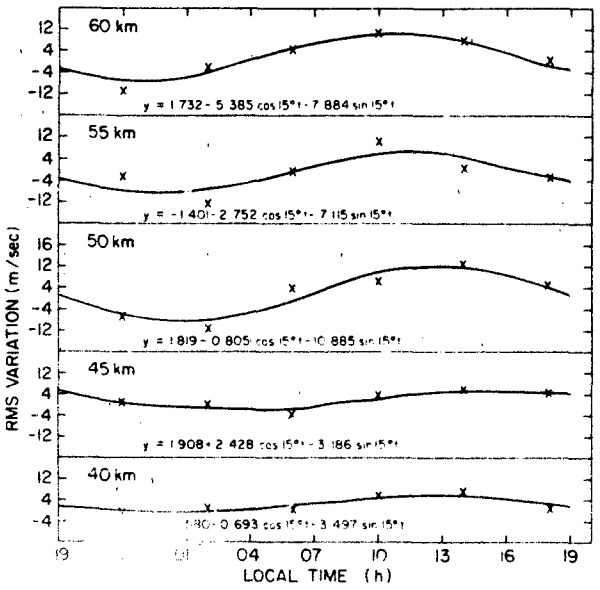


Figure 17-9a. Diurnal wind variations at Wallops Island in May (N-S).

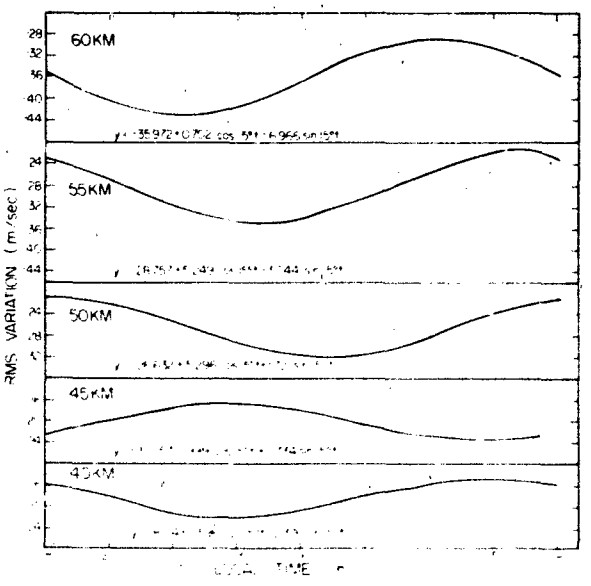


Figure 17-9b. Diurnal wind variations at Wallops Island in May (E-W).

CHAPTER 17

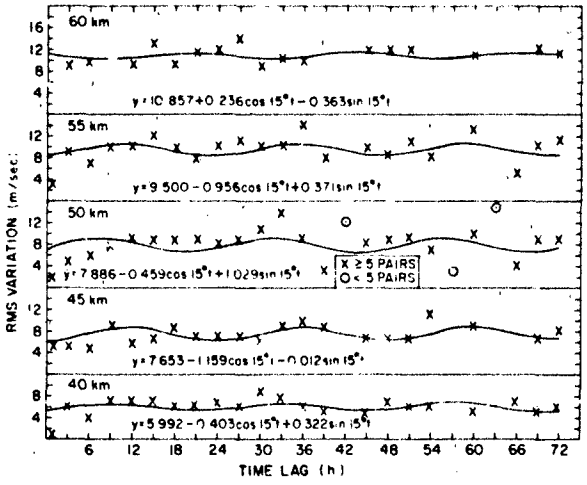


Figure 17-10. Rms differences between north-south winds observed 1 to 72 h apart in Ascension, Ft. Sherman and Kwajalein.

are not as well defined at 40 and 45 km. At altitudes of 50 km and above, however, the amplitudes are between 12 and 14 m/s. The magnitude of the diurnal oscillations are somewhat smaller in tropical and arctic regions.

Rms differences between wind components observed from 1 to 72 h apart have been analyzed at altitudes between 30 and 60 km to obtain estimates of the variations in these components with time. Rms differences between the N-S wind components observed from 1 to 72 h apart at Ascension, Ft. Sherman and Kwajalein are shown in Figure 17-10. Data for all seasons for these three stations have been combined. The values display a relatively stable rms variation of roughly 10 to 12 m/s for periods up to 72 h at 55 and 60 km and about 6 or 7 m/s at altitudes of 40 and 45 km. There is an indication of a diurnal cycle with an amplitude of 1 to 2 m/s at these altitudes as shown by the 24-h harmonic curves which have been fitted to the individual values. Most of the observed variability can be attributed to random measurement errors as there is little difference in the rms variations between observations taken 6 h apart and those taken 72 h apart.

In summer conditions are approximately the same at all latitudes, with little change in rms values with time. In winter, however, the rate at which the rms values increase with time varies with latitude. The larger day-to-day changes in the synoptic patterns at middle and high latitudes are reflected in well-defined increases in the rms variability with time. At 50 km, for example, rms values for the N-S components at White Sands Missile Range (32°N) increase with time from 6 to 8 m/s to 16 or 18 m/s in 72 h. At Poker Flats (64°N) and Ft. Churchill (59°N), they increase from 6 or 8 m/s to 25 m/s. The diurnal cycle in observed data is masked by the synoptic changes and instrumentation errors during the winter months at middle and high latitudes.

17.3 WIND PROFILES

17.3.1 Wind Shear

Wind shear is the derivative of the wind vector with respect to distance and is itself a vector. The shear of the horizontal wind is of primary interest and is the one discussed in this section. The terms vertical wind shear and horizontal wind shear are commonly used in referring to the shear of the horizontal wind in the vertical and horizontal directions, respectively. Horizontal wind shear is the derivative of the horizontal wind with respect to an axis parallel to the earth's surface. Its applications are restricted largely to meteorological analysis. Vertical wind shear is expressed as $\Delta W/\Delta y$, where Δ is the change of the horizontal wind in the altitude interval Δy; the unit of shear is s^{-1}. Although direction is also necessary to specify the shear vector, it is usually ignored; for design purposes shear is normally applied in the most adverse possible direction.

The climatology of vertical wind shear is applicable to problems dealing with design and launch of vertically rising vehicles and jet aircraft, radioactive fallout investigations, and many phases of high-altitude research. Most investigations of shear climatology are for specific locations in order to satisfy design and operational requirements of missiles and other vehicles at or near launch sites. For vehicles with other than a vertical flight path, vertical wind shear can be determined by multiplying the shear by the cosine of the angle between the vertical axis and the vehicle trajectory.

Measurements of vertical wind shear indicate that average shear behaves inversely with layer thickness (scale-of-distance). This is illustrated in Figure 17-11, which is based on a relatively large number of observations during

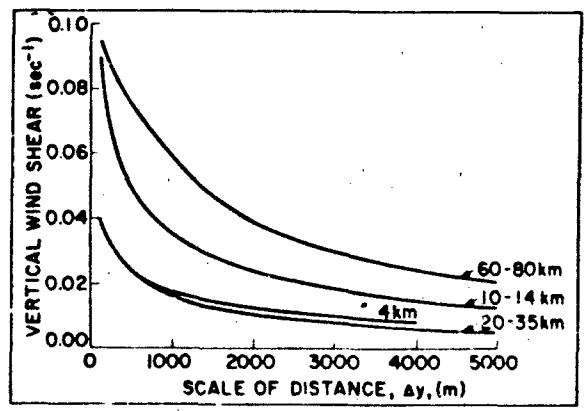


Figure 17-11. Selected vertical wind shear spectrums (4, 10, 14, 20, 35, and 60-80 km altitude) for use with 5% and 1% probability level wind profile envelope, Cape Canaveral, Florida [Scoggins and Vaughan, 1962].

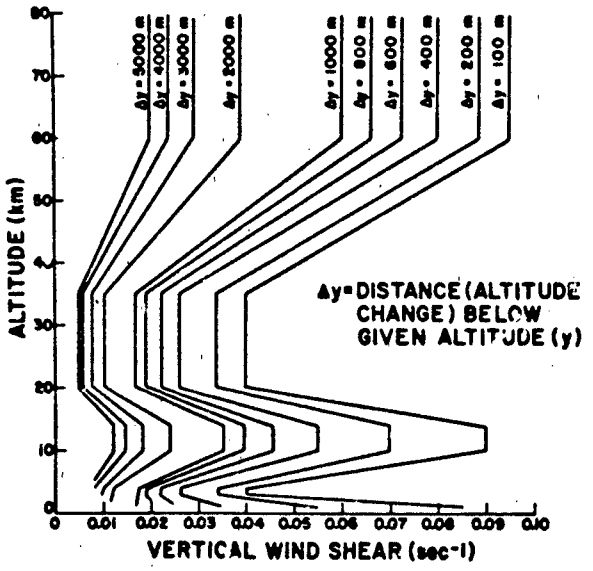


Figure 17-12. 1% probability-of-occurrence vertical wind shear spectrum as function of altitude and scale-of-distance for association with the 5% and 1% wind speed profile envelope for Cape Canaveral, Florida [Scoggins and Vaughan, 1962].

the windiest months over Cape Canaveral, Florida. Figure 17-12, based on the same data, shows the variation of shear with altitude and with layer thickness; it provides, as an example, the vertical wind shear spectrum at Cape Canaveral with a 1% probability of occurrence.

Wind shears for scales-of-distance $\Delta y \geq 1000$ m in thickness are computed directly from rawinsonde and rocketsonde observations, whereas smaller scale shears can be calculated directly only from special fine-scale observations. However, shears associated with scales-of-distance $\Delta y < 1000$ m can be estimated from the following relationship [Fichtl, 1972]:

$$\Delta u = \Delta u_{(1000)} \left(\frac{1000}{\Delta y} \right)^{0.3} , \qquad (17.9)$$

where Δu is the shear, $\Delta u_{(1000)}$ is the 1000 m shear, and Δy is the scale-of-distance in m for thickness < 1000 m.

Wind shear statistics for various locations differ primarily because of prevailing meteorological conditions, orographic features, and data sample size. As a result, significant differences exist in the shear structure for different locations. Consistent shear data for five vehicle launch and/or landing sites are presented in Tables 17-11 through 17-15 for the Eastern Test Range, Florida; Vandenberg AFB, California; Wallops Island, Virginia; White Sands Missile Range, New Mexico; and Edwards AFB, California. To get actual shear (s^{-1}) from the indicated wind speed changes, divide by the appropriate scale-of-distance. Table 17-16 gives envelopes of the 99% wind speed change for the five locations combined. The data contained in Table 17-16 are applicable when design or operational capability is not restricted to a specified launch site or may involve several geographical locations. Equation (17.9) was used to construct Tables 17-11 through 17-16 for scales-of-distance < 1000 m.

17.3.2 Interlevel Correlations

Deviations in the assumed vertical wind profile over a target or reentry point affect the range and cross-range of a ballistic missile. These effects must be considered in the design of guidance systems for entry vehicles and for targeting ballistic missiles.

The mean effect E of mean monthly or mean seasonal winds on the range and cross range of a missile can be determined for a particular location by computer-simulated flights through mean monthly component wind profiles if

Table 17-11. Envelopes of 99% wind speed change (m/s), 1-80 km altitude region, Eastern Test Range [Kaufman, 1977].

Wind Speed at Reference Altitude (m/s)	Scales of Distance (m) Thickness									
	5000	4000	3000	2000	1000	800	600	400	200	100
≥ 90	77.5	74.4	68.0	59.3	42.6	36.4	29.7	22.4	13.8	8.5
= 80	71.0	68.0	63.8	56.0	40.5	34.7	28.5	21.4	13.2	8.1
= 70	63.5	61.0	57.9	52.0	38.8	33.1	27.0	20.3	12.5	7.7
= 60	56.0	54.7	52.3	47.4	36.0	31.0	25.3	18.9	11.7	7.2
= 50	47.5	47.0	46.2	43.8	33.0	28.3	23.2	17.5	10.7	6.6
= 40	39.0	38.0	37.0	35.3	29.5	25.3	20.6	15.5	9.6	5.9
= 30	30.0	30.0	29.4	26.9	22.6	19.4	15.8	11.9	7.3	4.5
= 20	18.0	17.5	16.7	15.8	14.6	12.5	10.2	7.5	4.7	2.9

Table 17-12. Envelopes of 99% wind speed change (m/s), 1–80 km altitude region, Vandenberg AFB [Kaufman, 1977].

Wind Speed at Reference Altitude (m/s)	Scales of Distance (m) Thickness									
	5000	4000	3000	2000	1000	800	600	400	200	100
≥ 90	66.9	62.5	57.8	51.5	37.5	32.1	26.1	19.7	12.0	7.4
= 80	64.1	60.8	56.6	48.8	36.9	31.5	25.6	19.1	11.6	6.8
= 70	62.0	59.2	54.8	48.1	36.0	31.0	25.0	18.6	11.2	6.5
= 60	57.1	54.5	51.3	45.4	32.7	28.5	23.0	17.1	10.2	5.3
= 50	49.6	47.8	45.7	42.1	30.1	25.9	21.8	15.6	9.2	5.0
= 40	39.4	38.8	37.9	35.5	25.9	23.5	19.6	14.9	8.8	4.8
= 30	30.0	29.4	28.3	26.3	20.5	18.6	15.8	12.2	8.0	4.6
= 20	20.0	19.8	19.5	18.4	15.0	13.1	10.9	9.0	6.3	4.3

Table 17-13. Envelopes of 99% wind speed change (m/s), 1–80 km altitude region, Wallops Island [Kaufman, 1977].

Wind Speed at Reference Altitude (m/s)	Scales of Distance (m) Thickness									
	5000	4000	3000	2000	1000	800	600	400	200	100
≥ 90	72.5	67.0	60.2	50.5	37.6	32.3	26.3	19.8	12.2	7.5
= 80	66.5	62.5	57.5	48.8	37.0	31.7	25.9	19.5	12.0	7.4
= 70	61.2	58.5	53.8	46.5	35.8	30.7	25.1	18.9	11.6	7.1
= 60	54.4	52.5	50.0	44.4	34.5	29.6	24.2	18.2	11.2	6.9
= 50	46.2	44.2	42.3	38.8	33.0	28.3	23.2	17.4	10.7	6.6
= 40	36.7	35.6	34.5	32.3	27.6	23.7	19.3	14.5	8.9	5.5
= 30	27.2	26.3	25.3	24.2	20.6	17.7	14.7	10.8	6.7	4.1
= 20	17.8	17.3	16.8	16.4	15.2	13.0	10.6	8.0	4.9	3.0

Table 17-14. Envelopes of 99% wind speed change (m/s), 1–80 km altitude region, White Sands Missile Range [Kaufman, 1977].

Wind Speed at Reference Altitude (m/s)	Scales of Distance (m) Thickness									
	5000	4000	3000	2000	1000	800	600	400	200	100
≥ 90	70.7	67.0	61.2	52.4	42.0	36.0	29.4	22.1	13.6	8.4
= 80	66.0	63.0	57.7	50.0	40.2	34.5	28.1	21.2	13.0	8.0
= 70	60.2	57.0	53.0	46.5	38.0	32.6	26.6	20.0	12.3	7.6
= 60	52.6	50.0	46.5	42.3	35.5	30.5	24.9	18.7	11.5	7.1
= 50	45.0	43.0	40.2	37.0	32.0	28.3	23.1	17.4	10.7	6.6
= 40	36.5	35.5	34.8	33.5	29.3	25.1	20.5	14.5	9.5	5.5
= 30	27.4	27.0	26.4	24.8	22.0	19.3	15.8	11.8	7.3	4.5
= 20	18.4	17.7	17.3	16.5	15.0	12.9	10.5	7.9	4.9	3.0

Table 17-15. Envelopes of 99% wind speed change (m/s), 1–80 km altitude region, Edwards AFB [Kaufman, 1977].

Wind Speed at Reference Altitude (m/s)	Scales of Distance (m) Thickness									
	5000	4000	3000	2000	1000	800	600	400	200	100
≥ 90	75.2	72.0	67.3	59.0	42.8	36.7	30.2	22.5	13.9	8.5
= 80	68.0	66.3	62.5	55.5	40.8	35.0	28.6	21.5	13.2	8.1
= 70	60.4	59.0	56.8	51.4	38.7	33.2	27.0	20.4	12.5	7.7
= 60	53.0	51.8	49.3	45.0	36.0	30.9	25.2	19.0	11.7	7.2
= 50	44.8	43.6	41.5	38.4	32.0	27.5	22.4	16.9	10.4	6.4
= 40	36.5	35.5	34.5	33.0	27.0	23.2	18.9	14.2	8.8	5.4
= 30	28.0	27.3	26.9	26.3	21.4	18.4	15.0	11.3	6.9	4.3
= 20	18.0	17.7	17.4	16.7	15.2	13.0	10.6	8.0	4.9	3.0

the appropriate influence coefficients for the missile at various levels are given:

$$E = \sum Ci\, V_i \qquad (17.10)$$

where Ci is the influence coefficient at the ith level that describes the portion of the total response of a missile assignable to that level, and V_i represents the mean of the component wind speed at that level. The variation around this average effect caused by day-to-day fluctuations in the winds is obtained from

$$\sigma^2_{(S)} = \sum_{i,j} C_i C_j r_{ij} \sigma_i \sigma_j \qquad (17.11)$$

where $\sigma^2_{(S)}$ is the integrated variance for all levels considered. C_i and C_j are the influence coefficients at the ith and jth levels, σ_i and σ_j are the standard deviations of the component wind at these levels, and r_{ij} is the correlation between the component wind at the ith level and that at the jth level. The square root of the solution to this equation yields the standard deviation for each component of the ballistic wind.

The two components can be combined and used to determine the probability of occurrence of deviations of any desired magnitude from a trajectory or impact point based on mean monthly winds. In these computations, it is assumed that the cross-component correlations are zero at and between levels, and the wind frequency distributions are essentially circular normal.

The wind climatology necessary for determination of ballistic effects consists of statistical arrays of mean monthly north-south and east-west wind components, their standard deviations, and interlevel correlations between the same component at different levels. This information has been prepared in matrix form for a relatively large number of locations. Examples for levels between the surface and 60 km are shown in Table 17-17 for the month of January at Ft. Churchill, Wallops Island and Ascension Island. Winds are lighter in summer at middle and high latitudes. For targeting purposes, the wind data must pertain to the locations of interest. For design purposes, however, a representative sample of data from the various climatic regions should be used.

Table 17-16. Envelopes of 99% wind speed change (m/s), 1–80 km altitude region, for All five locations [Kaufman, 1977].

Wind Speed at Reference Altitude (m/s)	Scales of Distance (m) Thickness									
	5000	4000	3000	2000	1000	800	600	400	200	100
≥ 90	77.5	74.4	68.0	59.3	42.8	36.7	30.2	22.5	13.9	8.5
= 80	71.0	68.0	63.8	56.0	40.8	35.0	28.6	21.5	13.2	8.1
= 70	63.5	61.0	57.9	52.0	38.8	33.2	27.0	20.4	12.5	7.7
= 60	57.1	54.7	52.3	47.4	36.0	31.0	25.3	19.0	11.7	7.2
= 50	49.6	47.8	46.2	43.8	33.0	28.3	23.2	17.5	10.7	6.6
= 40	39.4	38.8	37.9	35.5	29.5	25.3	20.6	15.5	9.6	5.9
= 30	30.0	30.0	29.4	26.9	22.6	19.4	15.8	12.2	7.3	4.6
= 20	20.0	19.8	19.5	18.4	15.2	13.1	10.9	9.0	6.3	4.3

Table 17.17a Zonal Winds from the surface to 60 km at Ft. Churchill.

Correlation at Pairs of Levels for Jan 1969–1976
Fort Churchill, Manitoba
East West Wind (+ W, - E)

KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	1035	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	4	4	7	10	12	13	14	15	16	17	17	17	20	25	24	25	24	25	25	28	30	26	31	35	35	38	39	39	37	27	30
STDV	32	64	98	132	130	104	99	108	129	137	170	199	202	248	281	305	301	308	324	337	353	366	395	410	427	447	440	435	418	330	324
N	50	50	51	52	52	52	51	50	49	46	44	43	38	52	53	52	53	53	53	53	53	49	52	53	52	53	52	53	49	44	43
2	54	77																													
4	33	81																													
6	23	62	95																												
8	18	59	87	96																											
10	17	45	71	82	92																										
12	15	36	55	61	76	91																									
14	5	15	34	33	60	80	91																								
16	0	6	23	30	38	68	82	96																							
18	4	10	18	16	26	51	72	86	92																						
20	3	5	15	15	24	50	66	80	82	97																					
22	4	6	12	11	21	46	62	74	82	95	98																				
24	1	6	5	7	5	36	55	66	76	90	94	98																			
26	1	6	13	17	30	46	59	68	74	82	87	90	92																		
28	2	5	10	13	28	44	59	65	73	81	86	89	90	99																	
30	2	4	9	12	25	33	50	54	64	76	81	84	84	96	98																
32	6	10	9	10	22	38	49	55	63	70	75	77	79	92	94	97															
34	5	12	11	12	23	38	46	52	58	66	71	74	76	88	91	94	98														
36	6	12	9	10	20	35	40	45	51	55	60	63	64	79	83	88	93	96													
38	5	10	4	7	16	29	32	37	42	44	50	54	57	73	76	81	89	92	98												
40	3	15	6	9	16	25	23	23	26	27	34	37	41	60	63	70	78	83	91	95											
42	1	12	1	4	8	9	1	8	11	10	19	23	26	42	46	54	66	70	81	88	95										
44	10	17	6	7	10	14	6	6	7	8	16	20	26	37	40	49	62	67	75	82	90	96									
46	6	7	1	0	3	7	1	2	0	2	9	13	21	30	31	40	53	58	67	74	83	90	96								
48	15	9	0	1	0	3	8	7	8	6	0	5	15	23	23	32	46	52	61	68	78	85	92	97							
50	13	9	1	4	6	0	11	9	10	7	3	1	14	19	17	25	41	46	55	62	71	78	86	91	96						
52	11	4	5	7	8	1	10	7	9	7	3	4	13	14	13	22	37	43	53	59	67	76	82	88	93	97					
54	14	1	7	9	9	4	10	8	8	9	5	4	12	7	6	14	27	31	41	49	56	68	75	82	88	92	95				
56	12	2	14	13	13	13	26	26	22	25	20	14	1	2	0	9	23	28	39	48	57	73	76	80	86	90	90	95			
58	15	2	18	18	16	10	22	11	11	21	21	16	15	6	2	5	18	20	27	38	42	53	58	65	75	80	81	86	93		
60	15	5	23	22	19	13	14	13	12	17	14	7	28	12	5	10	23	26	32	43	44	51	58	68	75	79	79	85	91	94	

**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 17-17b Zonal Winds from the surface to 60 km at Wallops Island.

Correlation at Pairs of Levels for Jan 1969–1976
Wallops Island, VA
East-West Wind (E is West +)
KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	0.15	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	1	12	21	28	34	31	33	28	23	15	11	11	11	14	16	19	24	28	31	33	37	40	45	50	54	57	59	62	63	68	76
STDV	30	61	90	123	150	120	109	96	88	76	74	81	95	140	170	199	236	254	269	286	279	291	287	300	303	308	290	291	282	282	313
N	44	44	44	44	41	29	25	22	21	20	21	19	22	42	43	44	44	44	44	44	44	44	44	44	44	44	44	44	40	34	25
2	25	2																													
4	31	55																													
6	30	68	94																												
8	34	50	85	91																											
10	4	31	60	77	91																										
12	21	18	32	44	68	77																									
14	21	17	40	62	83	89	87																								
16	13	4	32	51	67	76	86	88																							
18	13	29	5	31	53	72	82	75	75																						
20	18	34	8	32	52	69	82	76	85	93																					
22	11	17	3	17	34	55	57	81	89	93	98																				
24	3	29	11	30	21	57	63	72	73	85	92	95																			
26	7	16	13	18	20	39	46	48	42	67	80	38	88																		
28	12	17	19	26	30	45	43	52	24	61	21	32	83	95																	
30	10	18	18	26	32	44	39	52	20	51	17	30	72	89	96																
32	14	18	23	37	41	48	38	25	20	53	18	28	66	82	91	97															
34	11	18	24	38	42	46	36	21	19	52	18	23	50	75	83	90	97														
36	8	16	26	37	40	41	32	29	23	55	22	33	53	63	75	83	90	96													
38	3	21	22	30	31	38	46	32	25	52	20	29	48	55	66	72	79	86	95												
40	1	18	29	33	30	28	36	25	12	36	2	4	33	46	57	60	68	75	84	93											
42	2	13	23	23	20	9	16	8	3	23	19	17	18	32	40	41	50	58	68	81	95										
44	3	9	18	18	13	1	14	9	5	15	21	21	13	18	22	23	33	41	53	69	86	96									
46	2	10	18	17	14	13	5	4	2	5	25	32	8	3	4	5	15	24	35	52	72	85	95								
48	2	6	14	13	8	16	0	7	9	7	16	22	1	5	5	5	4	12	23	41	61	76	88	97							
50	5	0	8	6	1	20	7	7	9	14	2	11	1	8	10	11	4	4	13	30	50	67	80	89	96						
52	3	7	15	12	5	13	3	13	13	13	6	12	3	10	10	11	4	4	14	31	51	67	79	87	94	98					
54	1	13	19	13	6	16	9	5	5	0	18	21	11	10	9	9	3	5	15	30	52	66	76	84	91	92	97				
56	1	20	20	12	5	16	10	5	10	11	11	17	6	2	0	3	1	6	11	24	47	60	70	76	85	87	93	96			
58	9	23	19	13	15	9	0	3	3	9	6	9	18	18	19	17	17	21	21	32	52	59	63	66	76	78	84	89	93		
60	6	28	31	26	32	4	1	9	2	8	6	13	9	9	12	8	10	16	12	18	39	45	52	60	73	78	84	87	90	96	

**Multiply tabular values by 0.01 to obtain correlation coefficients

Table 17.17c. Zonal Winds from the surface to 60 km at Ascension Island.

Correlation at Pairs of Levels for Jan 1969-1976
Ascension Island
East-West Wind M/S West +
KM Kilometers Above Sea Level
MEAN Average of Observed Values
STDV Standard Deviation of Values Times 10
N Number of Values at Each Altitude

KM	.079	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
MEAN	-6	-3	4	-5	4	1	6	9	0	7	-10	-14	-17	-15	-15	-18	-22	-27	-31	-34	-41	-50	-60	-68	-73	-65	-45	-29	-15	-2	4
STDV	16	37	42	44	67	98	113	130	70	45	92	113	126	102	120	147	139	134	118	110	111	88	100	139	177	206	204	189	178	173	161
N	42	43	43	43	43	43	43	43	43	43	43	43	41	43	42	43	43	43	43	43	43	43	43	43	43	43	43	43	39	30	26
2	12	**																													
4	33	30																													
6	23	33	28																												
8	23	28	7	55																											
10	15	8	9	30	77																										
12	11	-3	-5	37	58	86																									
14	-4	-22	23	14	40	68	82																								
16	-8	0	16	22	34	59	67	82																							
18	-12	19	-24	12	7	10	5	8	20																						
20	-2	18	14	10	-18	-10	1	0	4	48																					
22	2	22	2	4	8	7	14	11	11	38	85																				
24	9	28	21	17	40	26	25	16	13	27	47	84																			
26	12	14	22	-10	19	29	27	25	-6	-24	-7	18	42																		
28	15	18	11	-1	16	21	12	24	1	30	-61	-39	-4	61																	
30	1	-23	4	-5	11	19	9	27	13	35	-72	-65	-37	34	89																
32	3	-22	11	4	18	28	15	30	22	20	-71	-68	-40	26	81	95															
34	3	-24	16	5	14	20	5	12	5	-17	-75	-77	-53	7	65	76	86														
36	3	-26	5	9	7	7	-4	4	0	-18	-74	-78	-60	-12	50	59	68	91													
38	9	-33	14	5	7	2	-6	2	-6	-20	-59	-71	-64	-24	30	39	45	73	87												
40	-14	9	-22	-14	-11	-14	-12	-16	-19	-17	-20	-15	-40	-17	2	-1	13	45	58	76											
42	-13	23	3	7	-16	-10	-5	-14	-3	42	22	5	-7	34	-27	-27	-15	4	5	19	57										
44	-5	21	16	3	0	-2	8	7	1	25	-34	-28	-17	-21	-41	-48	-40	-19	-9	-3	26	53									
46	-2	2	19	12	1	-12	-4	7	-7	12	21	24	18	-23	-34	-35	-36	-34	-21	-18	-16	-2	51								
48	3	6	3	8	-1	-24	13	-16	-11	9	33	35	19	-31	-48	-50	-51	-41	-27	-22	-15	0	46	75							
50	4	20	8	0	9	21	18	22	-19	3	22	23	10	-34	-36	-36	-36	-26	-17	-13	2	9	37	44	82						
52	1	31	16	10	13	-17	14	19	-17	4	7	5	9	-36	-31	-26	-23	-10	-3	4	26	24	32	16	51	78					
54	11	40	13	16	6	14	10	20	11	7	5	2	17	-47	-42	-31	-27	-15	-6	6	33	31	27	0	23	44	82				
56	24	42	3	10	-5	23	15	16	5	13	1	-7	-14	-45	-45	-23	-27	-27	-20	-8	1	14	5	-2	22	34	50	72			
58	14	14	5	13	22	1	9	2	-10	13	1	2	7	-10	34	-30	-33	-28	-21	-8	-5	-11	28	6	34	41	34	35	75		
60	5	1	0	8	19	9	1	2	15	11	4	2	7	-4	28	-28	30	-19	-8	-6	-12	-20	28	13	40	42	21	10	46	83	

**Multiply tabular values by 0.01 to obtain correlation coefficients

17.4 DESIGN DATA ON WINDS

Wind statistics are presented in a variety of ways, each of which is intended for maximum usefulness for particular aspects of design and operational problems. In "Upper Wind Statistics," Crutcher [1959] presents northern hemisphere charts of some 15 wind variables or statistics. For surface winds, in the "Climatic Atlas of the United States," the National Weather Service [1968] presents monthly and annual charts of the prevailing direction, mean wind speed, the fastest wind on record and its direction, and wind roses that give the frequency distribution of the wind by direction including frequency of calm, the mean wind vector and direction. There is also a table of the frequencies of 9 categories (Beaufort) of wind speeds for some 120 U.S. stations.

Generally speaking, any data source on surface wind speed is limited to providing several key parameters from which the general frequency distribution of the wind, in speed and direction, can be estimated or reconstructed. But to do so, a practical model of the distribution of the wind speed must be adopted. Three such models are the gamma distribution, Weibull distribution, and the circular normal bivariate distribution. Each has its merits and drawbacks.

17.4.1 Hourly Surface Wind Speeds

The distributions of surface wind speeds, observed every hour on the hour, have been studied by many climatologists and statisticians whose conclusions and consequent models differ. Some favor the distribution of wind speeds as given by the circular bivariate normal distribution, others the log-normal, still others the gamma distribution. A historical record of at least 5 years should be used to obtain good estimates of wind speed distributions, especially small probabilities (for example 1%). Although there are about 600 city and airport stations in the United States where hourly records are kept, they generally represent wind fields inadequately. A location that is close to a weather reporting station should have similar wind characteristics, but terrain effects including manmade effects and proximity to large bodies of water superimpose a spatial variability that is difficult to generalize. Fairbanks, Alaska, for example, has a mean wind speed in January of 0.9 m/s with standard deviation 1.7 m/s. Yet Big Delta, only 121 km away, has a mean wind speed of 5.4 m/s and standard deviation 4.9 m/s.

The following are alternative models of windspeed frequency distribution:

Model Alternative 1: A gamma distribution. In terms of the mean wind speed $\bar{V}_s$ and standard deviation s_s, the probability density function of the wind speed V_s is given in terms of a transformed variable y by

$$f(y) = 0.5\, y^2 e^{-y} \qquad (17.12)$$

$$\text{where } \frac{(y-3)}{\sqrt{3}} = \frac{V_s - \bar{V}_s}{s_s}. \qquad (17.13)$$

The cumulative probability of wind speed, $P(V_s)$, equal to or less than V_s is given by

$$P(V_s) = 1 - e^{-y}\left(1 + y + \frac{y^2}{2}\right). \qquad (17.14)$$

As an example, in the month of January, the noontime Bedford, Mass., wind has a mean hourly speed $\bar{V}_s$ of 5.1 m/s and standard deviation s_s of 3.5 m/s. Equation (17.14) gives the cumulative probability for several values of the wind speed as follows.

V_s, m/s	y	Estimated $P(V_s)$	Observed Distribution
calm	0.445	0.011	0.115
1.5	1.219	0.125	0.168
3.1	1.993	0.321	0.322
5.2	3.03	0.583	0.592
8.2	4.57	0.834	0.838
10.8	5.89	0.932	0.946
13.9	7.41	0.978	0.989
17.0	8.96	0.9936	0.999

This kind of approximation is good for moderate to strong wind speeds and is generally useful except for very light winds or calm conditions.

Model Alternative 2: A Weibull Distribution. A recent study [Bean and Somerville, 1979] has produced the model

$$P(V_s) = c + (1 - c)[1 - \exp(-\alpha v^{\beta})] \qquad (17.15)$$

where c is the probability of calm, and α and β are parameters of the Weibull distribution that are determined either to make the model fit the observed frequencies of wind speed (in a least squares sense), or are estimated in other effective ways. The records of many stations indicate a high probability of "calm," which makes this formulation desirable.

Records like the "Revised Uniform Summaries of Surface Weather Observations" (RUSSWO) (Compiled by the Environmental Technical Applications Center, Scott AFB, Ill.) contain the relative frequencies f_i of each of 11 or fewer winds speeds V_i. Where x_i, the middle value in the category, is used for the ith category of V_s, formulas have been found for α and β as follows:

$$\hat{\alpha} = \frac{\sum f_i}{\sum f_i x_i^{\beta}} \qquad (17.16)$$

$$\hat{\beta} = \left[\frac{\sum f_i \cdot x_i^{\beta} \ln x_i \quad \sum f_i \ln x}{\sum \quad \sum x_i^{\beta}}\right]$$

To solve for α and β, an initial guess is entered for β in the right-hand side of Equation (17.16), and the equation is solved for a first estimate of β. With this revised estimate, the equation is solved again for a second and better estimate, and so on, until the value of β is stabilized.

The RUSSWO for Bedford, Mass., January at noontime, gives f, for categories of V_c (or x_i) which, when entered in Equation (17.16) produced the following values after several iterations:

$$\hat{\beta} = 1.961$$

$$\hat{\alpha} = 0.02512$$

When these values for α and β are entered in Equation (17.15) together with $c = 0.115$, the probabilities $P(V_c)$ become

V_c (m/s)	$P(V_c)$	Observed Distribution
calm	0.115	0.115
1.5	0.166	0.168
3.1	0.297	0.322
5.2	0.527	0.592
8.2	0.816	0.838
10.8	0.9395	0.946
13.9	0.989	0.989
17.0	0.9987	0.999

Model Alternative 3: The circular normal distribution. The assumption of the circular normal distribution for the wind implies that the zonal u and meridional v components each have a normal Gaussian distribution with individual means $\bar{u}$ and $\bar{v}$ but with the same standard deviation s. The two components are independently distributed. As a consequence the probability density function of each is

$$f(u) = \frac{\{\exp[-(u-\bar{u})^2/2]\}}{\sqrt{2\pi}}$$

$$f(v) = \frac{\{\exp[-(v-\bar{v})^2/2]\}}{\sqrt{2\pi}} \qquad (17.17)$$

The magnitude of the mean wind vector $(\bar{V})$ is given by

$$|\bar{V}| = \sqrt{\bar{v}^2 + \bar{u}^2}. \qquad (17.18)$$

The standard vector deviation (s_c) is simply

$$s_c = s\sqrt{2} \qquad (17.19)$$

The mean wind speed $\bar{V}$, and the standard deviation of the wind speed s, are not easily derived from the components. The cumulative probability $P(V_c)$ of the wind speed, however, can be given in terms of $\bar{V}$ and s or s_c. One of the most applicable formulas is

$$P(V_c) = e^{-y} \sum_{n=0}^{10} \frac{y^n}{n!}\left[1 - e^{-x} \sum_{m=0}^{n} \frac{x^m}{m!}\right], \qquad (17.20)$$

where $x = V_c^2/s_c^2$,

$y = |\bar{V}|^2/s_c^2$.

In the example of Bedford, Mass. January noontime winds, the RUSSWO table of wind speed versus wind direction yields

$|\bar{V}| = 2.9$ m/s
$s_c = 5.1$ m/s
or $s = 3.6$ m/s
as opposed to $V_c = 5.1$ m/s
$s_c = 3.7$ m/s.

When the values for $|\bar{V}|$ and s_c are used in Equation (17.20) the probabilities $P(V_c)$ become

V_c m/s	$P(V_c)$	Observed
1.5	0.063	0.168
3.1	0.231	0.322
5.2	0.521	0.592
8.2	0.853	0.838
10.8	0.964	0.946
13.9	0.9936	0.989
17.0	0.9965	0.999

The distribution of the wind V around the mean wind vector $(\bar{V})$ is sometimes of prime interest as it would be in targeting. It is the difference between the actual vector wind and the mean vector wind and has the distribution

$$P(V) = 1 - \exp\left(-\frac{v^2}{s_c^2}\right) \qquad (17.21)$$

Thus a circle of radius equal to one standard vector deviation, drawn about the tip of the mean wind vector, would encompass 63% of the population of winds. Half of the wind vectors will fall within a circle of radius 0.83 times the standard vector deviation.

17.4.2 Surface Wind Direction

Winds tend to have a modal direction, like the trade winds or the prevailing westerlies. Using the model of the bivariate normal distribution, the estimate of the frequency

$P(\theta_1 \leq \lambda < \theta_2$ of the wind direction λ between directions θ_1 and θ_2 is given by

$$P(\theta_1 \leq \lambda < \theta_2) = (1/\sqrt{2\pi}) \cdot \sum_{u=0}^{u=\infty} [\exp - \frac{(\bar{u} - u)^2}{2} \delta u][F(x_2)) - F(x_1)] \quad (17.22)$$

where $u = V/s$.

V is the wind component in the direction (θ_0) of the mean wind vector $\bar{V}$, and s is the standard deviation of the appropriate wind component.

$\bar{u} = |\bar{V}|/s$.
δu is a small increment of u.
$x_2 = \tan(\theta_2 - \theta_0)$.
$x_1 = \tan(\theta_1 - \theta_0)$.

and F(x) is approximated by

$$F(x) = 1 - 0.5(1 + c_1x + c_2x^2 + c_3x^3 + c_4x^4)^{-4}$$

where

$c_1 = 0.196854$
$c_2 = 0.115194$
$c_3 = 0.000344$
$c_4 = 0.019527$

In the example for Bedford, Mass., January noontime wind $\theta_0 = 300°$. After correcting for calm frequency of 0.115, the computed frequencies of wind directions were as follows:

Direction	$P(\theta_1 \leq \lambda < \theta_2)$	Observed
N	0.082	0.067
NNE	0.056	0.041
NE	0.027	0.030
ENE	0.027	0.013
E	0.022	0.019
ESE	0.019	0.010
SE	0.018	0.014
SSE	0.020	0.009
S	0.025	0.022
SSW	0.032	0.051
SW	0.036	0.067
WSW	0.070	0.063
W	0.098	0.108
WNW	0.122	0.143
NW	0.126	0.110
NNW	0.110	0.106

17.4.3 Surface Wind Gusts

Studies on the relationship of gusts to the steady wind speed are in general agreement; however, quantitative results have varied depending on the data and the analytical methods used. Sissenwine et al. [1973] reviewed the results of many of these; he also presented a comprehensive analysis of gustiness based on actual wind records during periods of strong winds, observed at 14 airfields in the Northern Hemisphere between 14° and 77° latitude. Tattelman [1975] analyzed the same data to develop 50%, 75%, 90%, and 98% gust factor curves for 5-min, 1-min, and 30-s average (or steady) wind speeds. Gust speeds reported in weather observations are normally considered to be about 2-s averages, but for designers of various sized equipment, other short duration gusts may be applicable.

Since resolution of the wind records was approximately 2 s the shortest period gusts were considered to be 2-s wind speeds. Gust factor (GF) curves were fitted using the equation

$$GF = 1 + Ae^{-BV}, \quad (17.23)$$

where A and B are constants, and V is the speed of the steady wind. The values for B were dertermined by a least squares fit of all the data to the 5-min, 1-min, and 30-s speeds. Using these values of B, the value for A was determined for each percentile curve by a least squares fit to the mid-class wind speed values (for example, 12.6 m/s for V = 10.3–14.9 m/s) and weighted for the number of observations at each point. Figure 17-13 shows curves of the 30-s gust factor as related to the 5-min steady speed, and the 1-min gust factor related to the 5-min speed using all the data. Figure 17-14 shows the relationship between the 2-s gust factor and the 5-min steady speed. Curves in both

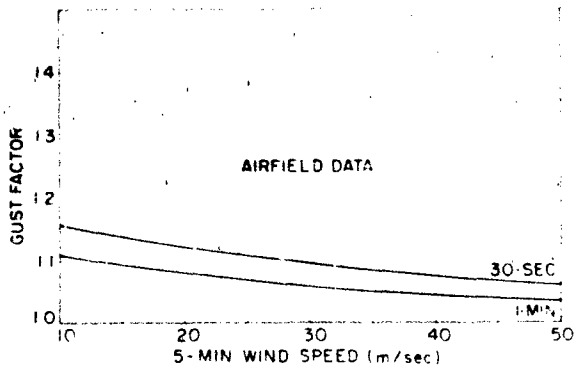


Figure 17-13 Curves of the 30-s and 1-min gust factor to 5-min wind speed for the airfield data fitted to $GF = 1 + Ae^{-BV}$ [Tattelman, 1975]

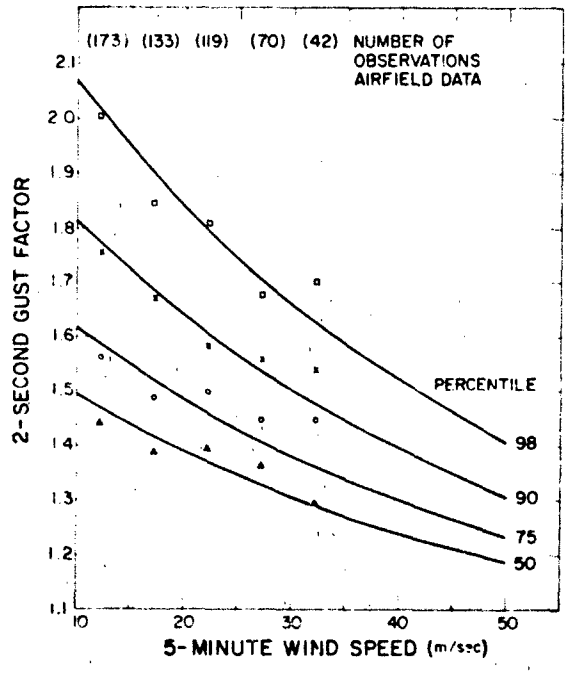


Figure 17-14 Relationship between the 2-s gust factor and the 5-min wind speed at indicated percentiles for the airfield data (Data were fitted to GF = 1 + Ae^{-B} and weighted for the number of observations at each mid-class value).

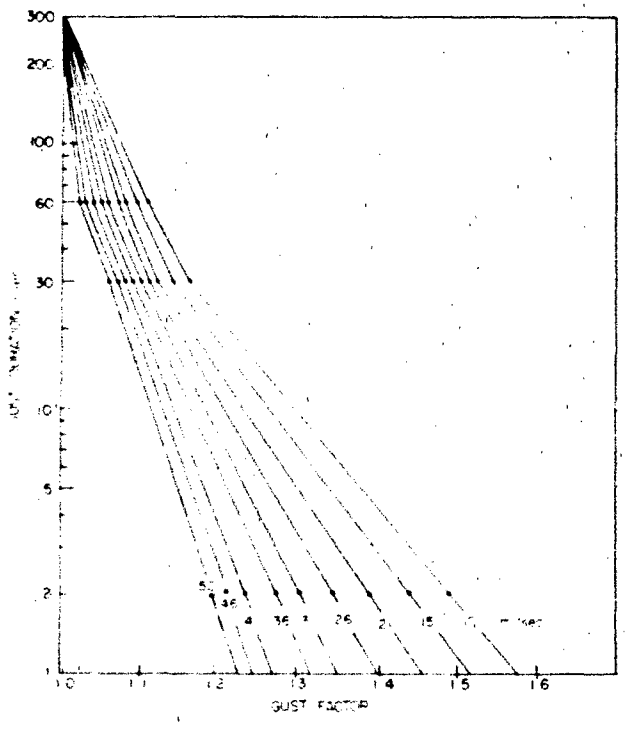


Figure 17-15 Expected (50%) gust factors versus gust duration and 5-min steady speed.

figures extend beyond the limit of the wind data, for which the highest steady speed is 35 m/s.

Sherlock [1947] found that a gust must have a duration such that its size is about 8 times the downwind dimension of a structure in order to build a force on the structure commensurate with the gust speed. Because the orientation of most equipment will not be taken into consideration, the shortest horizontal dimension of the equipment should be considered as the downwind dimension. For example, if a structure has a 5-m downwind dimension, a gust must be 40 m long to build up full dynamic pressure. Therefore, a speed of 40 m/s would build up full force on such a structure in only 1 s; a speed of 20 m/s would require 2 s. Figure 17-15 shows the expected (50%) gust factors versus gust duration and 5-min steady speed to aid in calculating the design windspeed for structures of differing downwind dimensions. The curves were drawn to four points: a gust factor of 1.0 at 300 s, the 60-s and 30-s gust factors for each speed from Figure 17-13, and the 2-s, 50% gust factors from Figure 17-14. Lines on Figure 17-15 were extended from the 2-s gust factors to include 1-s gust factors.

Tattelman [1975] also presented curves of the 50%, 75%, 90%, and 98% wind speed range (the difference between the maximum and minimum 2-s speed) as a function of time interval and mean 5-min wind speed. The 90% wind speed range is shown in Figure 17-16. The dashed lines indicate extrapolation beyond the limits of the data. The gust factor and wind speed range curves in this section are considered applicable to most airport locations at the average height of the data used, 15 m.

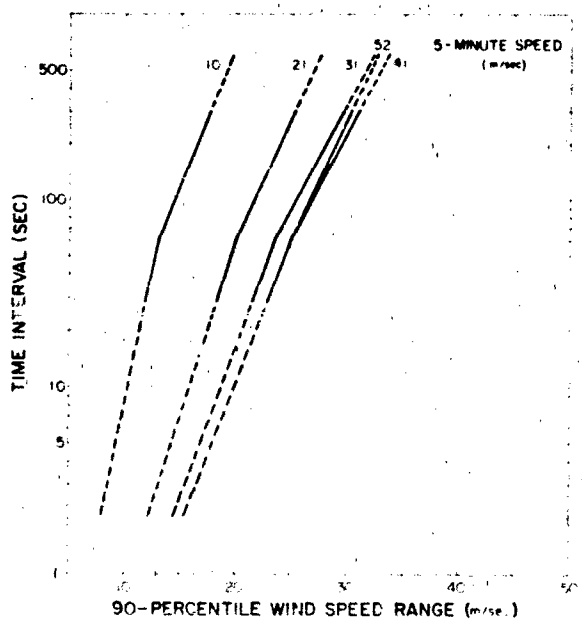


Figure 17-16 90% wind speed range versus time interval and 5-min speed.

17.4.4 Extreme Surface Wind Speeds

Various statistics are used for the daily maximum wind speed. One statistic is the maximum 5-min wind of the 288 5-min intervals of one day. Another popular measure is the *fastest mile*, which is the reciprocal of the shortest interval (in 24 h) that it takes one mile of air to pass a given point. The *strongest gust* is the highest reading of an instantaneous recording anemometer. There is no unique relation between these units. The Washington, D.C. records show that, for speeds greater than 15 m/s, the maximum 5-min wind speed is roughly 0.93 times the fastest mile and roughly 0.67 times the strongest gust. The maximum 1-min wind speed is roughly 0.71 times the strongest gust. Canadian and British records show the maximum 1-h wind speed to be 0.62 times the strongest gust.

A model distribution of the annual extreme has been the subject of conflicting viewpoints. The most favored distribution is the Fisher-Tippett Type I, now generally known as the Gumbel Distribution. If x is the symbol for the annual extreme, then the cumulative probability of x is given by

$$P(x) = \exp[-\exp(-y)] \qquad (17.24)$$

where y is the *reduced variate*, whose mean is $\bar{y} = 0.5772$ and whose standard deviation is $\sigma_y = 1.28255$.

A relatively simple equation can be used to relate x to y:

$$\frac{(y - \bar{y})}{\sigma_y} = \frac{(x - \bar{x})}{s_x} \qquad (17.25)$$

where $\bar{x}$ and s_x are the mean and standard deviation, respectively, of the annual extremes (x).

The work that is implied to make adequate estimates of percentiles of x is to find the sample mean and standard deviation of a set of N extremes for N years. Other approaches, although more rigorous, require ordering the N extremes by magnitude and/or estimating parameter values by successive approximations. Estimates also might depend directly on sample size.

To judge the risk of an extreme event in n years, then Equation (17.24) above is replaced by

$$P(x,n) = \exp[-n\exp(-y)]. \qquad (17.26)$$

To estimate x for probability P in n years, or an n-year risk (1-P), and knowing the annual extreme mean $\bar{x}$ and standard deviation (s_x), Equation (17.25) then gives

$$\hat{x} = \bar{x} + s_x \frac{(y - \bar{y})}{\sigma_y}, \qquad (17.27)$$

where $y = \ln(n) - \ln(-\ln(P))$. (17.28)

As an example, a 19-yr record at Denver, Colorado shows an average annual 5-min wind speed maximum 15 m above ground level of 19.6 m/s and standard deviation 2.5 m/s. If a tower is to be built to last for 25 years, with 99% certainty of survival throughout the 25 years, then the critical wind speed, given by Gumbel model (Equations (17.28), (17.27)), is estimated at 33.7 m/s.

An estimate is subject to sampling variations, in the above example a 19-yr sample. One standard error of this estimate s(x) is approximately

$$s(\hat{x}) = s_x\sqrt{\alpha/N}$$

where N is the number of years, and α is given by

$$\alpha = 0.710 + 0.116y + 0.669\,y^2.$$

In the Denver example, the standard error of the estimate (33.9 m/s) is approximately s(x) = 7.8 m/s, or 13% of the estimate itself. A brief meteorological record of annual extremes permits estimates only within confidence bands that vary with the size of the record.

Publications by climatic agencies such as NOAA Environmental Data Service and the Canadian Meteorological Service give extreme values. Yet each permanent installation might require a special study for a particular site and/or for a particular design problem. For example, the extreme annual wind speed at Logan airport (Boston) would be an underestimate, by about 33%, of the extreme on Blue Hill, 19 km to the south-southwest. Even the record at the exact location of interest may not be representative of the wind effects on a tall (or short) structure.

Table 17-18 gives means and standard deviations of annual extreme wind speeds and peak gusts at some representative stations around the world including stations that are typically in the paths of hurricanes. The 1% risk in 10 yrs was computed by the above formula, assuming the Gumbel distribution. This column is virtually the same as for the 1000-yr return period, and thus is the speed that has only 1/10 of 1% chance of being exceeded each year. There is a noticeable difference between the extreme speeds at airports and at city stations, attesting to the effects of tall buildings in the cities. The reduction of speed might be as high as 53%, but averages roughly 25%.

Tornadoes are severe whirlwinds, typically a hundred meters in diameter, which fortunately are rare accompaniments of severe thunderstorms. They are most common in Australia and the United States where they number some 140 to 150 per year, most frequent in spring in the plains area of the United States between the Rockies and the Appalachians. The wind speeds are generally less than 110 m/s in the lowest 10–20 m above the ground. Speeds up to 140 m/s occur occasionally in very severe tornadoes; higher wind speed estimates are generally suspect [Golden, 1976]. These high wind speeds place them outside the distributions

Table 17-18. Extreme annual wind speed (fastest m/s) at 15.2 m above ground at the given stations; (A) denotes airport station.

Station	Years of Record	Mean (m/s)	S.D. (m/s)	1% Risk in 10 yr (m/s)
Tampa, Fla (A)	1941–56	23	3.9	42
Miami, Fla	1943–58	24	8.0	64
Wilmington, N.C. (A)	1951–58	30	7.1	65
Hatteras, N.C.	1912–57	28	6.0	58
Dallas, Tex (A)	1941–58	23	2.9	38
Washington, D.C. (A)	1949–58	22	3.8	41
Dayton, Ohio (A)	1944–58	27	3.8	46
Atlanta, Ga. (A)	1933–58	22	3.3	39
Abilene, Tex (A)	1945–58	28	6.1	59
Columbia, Mo (A)	1949–58	25	2.8	39
Kansas City, Mo	1934–58	25	3.1	40
Buffalo, N.Y. (A)	1944–58	26	3.7	44
Albany, N.Y. (A)	1938–58	23	3.8	42
Boston, Mass (A)	1936–50	26	5.4	53
Chicago, Ill. (A)	1943–58	23	2.5	35
Cleveland, Ohio (A)	1941–58	26	2.6	39
Detroit, Mich (A)	1934–58	22	2.6	34
Minneapolis, Minn (A)	1938–58	23	5.0	48
Omaha, Nebraska (A)	1936–58	26	5.9	55
El Paso, Tex (A)	1943–58	26	2.0	36
Albuquerque, N.M. (A)	1933–58	27	4.6	50
Tucson, Ariz	1948–58	22	3.2	38
San Diego, Calif	1940–58	16	2.7	30
Cheyenne, Wyo	1935–58	28	3.1	43
Rapid City, S.D.	1942–58	30	3.0	44
Bismarck, N.D.	1940–58	30	2.3	41
Great Fall, Mont.	1944–54	29	1.6	37
Portland, Ore	1950–58	25	3.0	41
New York, N.Y.	1949–58	26	2.1	37
Pittsburgh, Pa.	1935–52	23	2.8	37
	Number of Years of Data			
Fairbanks, Alaska	9	17	3.7	35
Nome, Alaska	11	27	4.1	47
Elmendorf AFB, Alaska	14	20	3.2	36
Shemya Island, Alaska	10	31	7.8	45
Hickam AFB, Hawaii	17	20	3.8	38
Clark AB, Philippines	13	17	5.5	45
Lajes Field, Azores	13	28	7.6	65
Albrook AFB, Canal Zone	18	12	1.8	21
San Pablo, Spain	11	34	6.8	68
Wheelus AB, Libya	14	22	5.3	48
Stuttgart, Germany	15	18	2.7	29
Keflavik, Iceland	9	38	4.8	62
Thule, Greenland	14	36	5.5	63
Tainan, Formosa	39	24	9.5	71
Taipei, Formosa	39	26	9.8	75

Table 17-18. *(Continued)*

Station	Years of Record	Mean (m/s)	S.D. (m/s)	1% Risk in 10 yr (m/s)
Itazuke AB, Japan	14	19	4.5	42
Misawa AB, Japan	11	21	3.2	37
Tokyo Intl. Airport, Japan	15	23	5.5	46
Kimpo AB, Korea	8	19	3.6	37
Bombay, India	6	22	6.3	54
Calcutta, India	6	25	3.3	42
Gaya, India	6	23	3.0	38
Madras, India	6	20	3.4	37
New Delhi, India	6	23	1.7	31
Poona, India	6	17	2.7	31
Central AB, Iwo Jima	17	35	16.9	119
Kadena AB, Okinawa	14	37	11.3	93

discussed previously. Their small dimensions and infrequent occurrence make it impossible to state precise probabilities of occurrence at specific locations. Statistics of occurrence on an areal basis have been developed by Abbey [1976] and Kelly et al. [1978].

Hurricanes, while much more frequent, larger, and more predictable, still cause problems for the statistical analyst. Recent studies at the National Bureau of Standards suggest that the Gumbel (Type I) distribution of extreme winds does not provide as good a description of the extreme wind distributions when they must include hurricane winds. The departures of the winds from the Gumbel estimates tend to be on the weaker side. This means that the estimates of Table 17-18 should be used with caution. Hurricanes may strike at some stations while missing others, so that the record extreme wind speeds of one station may not represent the potential risk of still stronger winds at another nearby station or even at the same station. We conclude, however, that the Gumbel distribution is the best single model for all extreme winds data including hurricane winds.

Directly recorded data for record wind extremes are rare due to damage or destruction of the wind measuring instruments and power outages, during an event such as a tornado or hurricane. Newspaper reports give estimates whose reliability is unknown and variable. Some might be from a weather station where visually observed wind-speed dials were still in operation even though the recorder lost power. Others might be calculated from the amount of force required to blow over a building or tree or to drive a metal rod through a wooden post.

The recognized worldwide maximum wind speed measured at a surface station is a 5-min speed of 91 m/s and a 1-s gust of 101 m/s measured at Mt. Washington, New Hampshire Observatory on 12 April 1934. Mt. Washington is 1915 m above mean sea level and the anemometer was mounted at 11.6 m.

Operationally speaking, the greatest wind extremes are typically in northern Scotland. In the windiest month the 1% extreme is set at 22 m/s (a 1-min wind at 3 m above ground level). Gusts accompanying this extreme are estimated, by the shortest dimension of equipment, at

Dimension (m)	Speed (m/s)
<0.6	32
1.5	30
3.0	29
7.6	27
15.2	26
30.5	25

17.4.5 Structure of Jet Streams

Belts of exceptionally strong winds are jet streams. (The term as commonly used applies to the strong westerly winds found at the base of the stratosphere.) Because the meteorological disturbances vary in size and intensity and are moving, the jet streams vary in extent, severity, and location. Geographically-oriented descriptions obscure significant features. As Figure 17-17 shows, this is overcome by relating the features to the position of the core. In winter, the centers (or cores) of jet streams are found at altitudes between 9 and 12 km in latitudes 25° to 70° in each hemisphere. In summer, jet streams are weaker, higher (11 to 14 km), and farther poleward (35° to 75° latitude). Generally speaking, a jet stream is several thousand kilometers long,

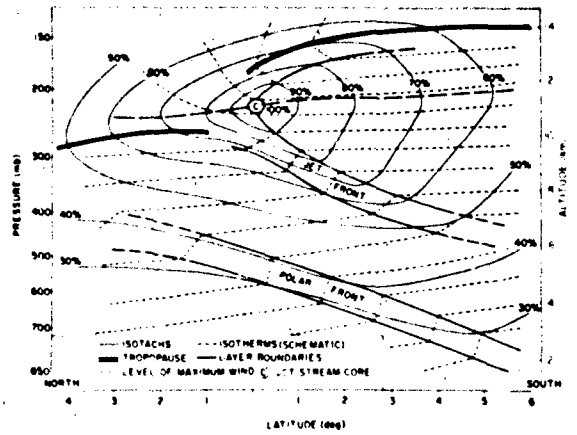


Figure 17-17. Idealized model of the jet stream, average structure in a cross section perpendicular to the flow. (Percent of the core speed is given for each isotach; wind direction is into the page.)

several hundred kilometers wide, and several kilometers deep.

The typical structure of jet streams is shown in Figure 17-17; this is a cross section showing isotachs (lines of equal wind speed blowing "into the page") labeled in percent of the wind speed at the jet core. Core wind speeds range from 50 m/s to as much as 130 m/s on rare occasions. Speeds of 75 m/s are common in winter. Figure 17-17 also shows the thermal field, the tropopause, and the level of maximum wind. Typical wind and temperature variations across jet streams at a constant altitude or vertically through jet streams can be determined from this figure.

Individual jet streams are variable and may differ considerably from the typical. In some cases, two or even three jet cores may be found paralleling each other, separated by distances of several hundred km. In addition to these deviations from the average, eddies of various sizes are present in the wind and temperature fields, as shown in Figure 17-18. Extreme horizontal changes of wind and temperature are of the order of 20 m/s and 5 K in a 16-km distance measured perpendicular to the jet stream. Variations parallel to jet streams are an order of magnitude smaller.

Clear air turbulence is closely related to jet stream structure; measurements show that it is more likely in certain regions than in others [McLean, 1962]. At the level of maximum wind (where vertical shear changes sign), turbulence was observed about 14% of the time. From Figure 17-19 it can be seen that turbulence is most likely to be found north of the jet stream core near the polar tropopause and above the core along the tropical tropopause. Indication of a maximum is also found in the jet stream front.

Other wind currents which exhibit properties similar to those shown in Figure 17-17 are the low-level jet and the tropical easterly jet stream which is observed in certain areas

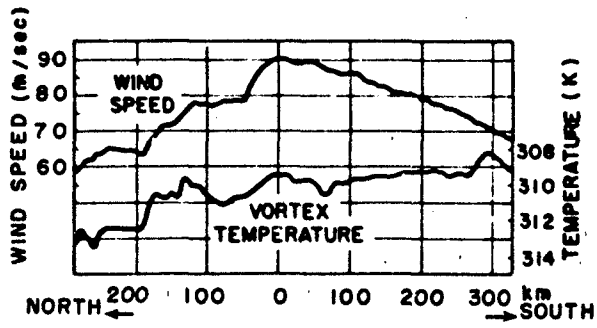


Figure 17-18. Example of wind and temperature fields near the jet stream.

(for example, southern India) at an altitude of about 14 km in summer.

Frequently it is too expensive to design equipment to operate under extreme conditions. In such cases, calculated risks are assumed or alternate methods of operation are devised for periods when such conditions exist. If high wind speeds or strong shears normally associated with them adversely affect the operation of a particular vehicle that is released vertically into the atmosphere, it may be desirable to select several release sites spaced far enough apart to ensure that operational conditions will exist above one of the stations when the critical wind speed or shear is exceeded above the others. In such cases, Figure 17-17 will be useful in selecting alternate sites; it indicates the average rate with which wind speed decreases with distance on both sides of the jet stream core. For example, with a 100 m/s wind at the jet stream core, Figure 17-17 indicates that, on the average, a belt of winds exceeding 70 m/s can be expected to extend more than 300 km south and nearly 200 km north of the core. Consequently, if the maximum wind likely to occur over a region is 100 m/s and the vehicle is only designed to withstand speeds up to 70 m/s alternate release

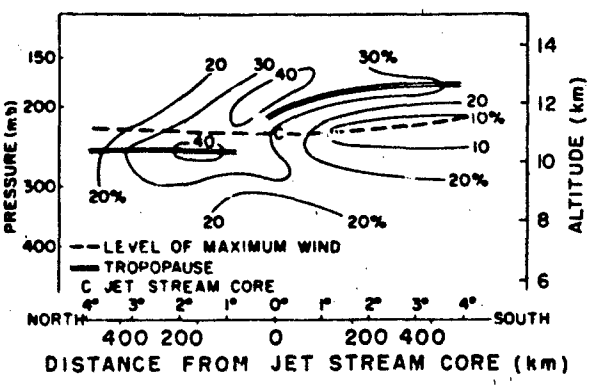


Figure 17-19. Turbulence in various sectors of a typical jet stream cross section. (Frequency (%) of occurrence is shown for each contour.)

sites should be selected at least 500 km apart. As the jet stream axis occasionally approaches a north-south orientation, selected sites should be more than 500 km apart in longitude as well as in latitude.

17.5 THERMOSPHERIC WINDS

Neutral air motions are a major key to our understanding of the upper atmosphere at thermospheric altitudes. A variety of phenomena associated with the neutral and ionized atmosphere need to be explained in terms of a unifying description of thermospheric dynamics. Neutral density, composition, and temperature characteristics are related to geomagnetic storm-induced circulation processes and seasonal and diurnal tides. Aspects of ionospheric structure depending directly on winds and compositional transport include maintenance of the nighttime ionosphere, spread F-layers, scintillations, and geomagnetic storm effects.

The dynamics of the upper atmosphere are controlled mainly by solar EUV and UV heating, waves and tides propagating upward from the mesosphere, ion drag interactions with the ionosphere, and energy and momentum sources associated with auroral region processes (magnetospheric convection and particle precipitation). The major heat source, due to solar electromagnetic radiation, is deposited mainly at low and midlatitudes and is reasonably well understood. At high latitudes, magnetospheric convection effects are important even during geomagnetically quiet times. They dominate the high latitude circulation during geomagnetic storms. Magnetospheric convection processes are extremely variable and are not well understood (see Chapter 8). The phenomenon involves ionospheric plasma driven into motion by electric fields interacting with the neutral atmosphere via collisions. The neutral constituents gain energy through Joule dissipation of currents and gain momentum through ion-neutral drag [Straus, 1978; Mayr et al., 1978].

Most of our present knowledge of thermospheric neutral winds is based on calculations made by using semiempirical dynamic models of the neutral atmosphere. The pressure gradients inherent in these density models provide the pressure forces that drive the thermospheric circulation. At thermospheric altitudes the collision frequency is sufficiently high that the air may be regarded as a fluid subject to hydrodynamic equations of motion. Therefore, the equations of fluid motion can be integrated to determine the wind system that should result from the pressure distribution. These equations are the same as those used by meteorologists to study the lower atmosphere, but are modified to include two additional terms that are important in the thermosphere: a viscous force and an ion drag force. Kinematic molecular viscosity transfers momentum between various altitude regions and thus smoothes out vertical gradients in wind velocity. It is inversely proportional to density and is large enough above 300 km to prevent development of large vertical and horizontal shears. Ion drag is a collisional interaction between charged particles and neutral particles. Outside the auroral zone, where electric fields are small, the ions above 120 km can be assumed to corotate with the earth. A neutral wind flowing through the ions experiences a collisional drag that is a maximum at the peak of the ionospheric layer. A global distribution of electron density [for example Ching and Chiu, 1973] is used to determine the ion drag in various model calculations.

The important terms in the equations of motion of the neutral air velocity U are

$$\frac{d\mathbf{U}}{dt} = \mathbf{F} - 2\,\boldsymbol{\Omega} \times \mathbf{U} - \nu_{ni}(\mathbf{U} - \mathbf{V}) + \frac{\mu}{\rho}\nabla^2\mathbf{U} + \mathbf{g}$$

(Acceleration) = (Pressure gradient force) − (Coriolis term) − (Ion Drag) + (Viscous Drag) + (Gravity)
where

$$\frac{d\mathbf{U}}{dt} = \frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U} \cdot \nabla)\mathbf{U}$$

and **V** is the ion drift velocity, $\boldsymbol{\Omega}$ is earth's angular velocity, **F** is the driving force per unit mass due to pressure gradients determined from a model, ν_{ni} is the appropriate neutral-ion collision frequency, μ is the coefficient of kinematic viscosity, ρ is the atmospheric density, and **g** is the acceleration due to gravity. Since this equation contains the ion velocity and the collision frequency that is proportional to the ion concentration, it is necessary to take into consideration the equation of motion and continuity equation for the ions. Methods of solutions to this set of equations for thermospheric altitudes have been extensively described by Rishbeth [1972]. (Note that the definition of winds does not include the corotation of the atmosphere. Results in this chapter are presented in a coordinate system that rotates with the earth). The accuracy of all models of global winds systems is limited by lack of experimental data and self-consistent model solutions. Complete specification of the wind fields requires solving a large set of non-linear partial differential equations in three dimensions that relate the hydrodynamic and electrodynamic properties of the neutral and ionized components in the atmosphere to the energy, mass, and momentum sources of the magnetosphere-ionosphere-thermosphere system.

A brief review of thermospheric wind properties deduced from experimental data is given in Section 17.5.1. The remainder of this chapter describes winds data derived from various models, which are only briefly described here. References provided should be consulted for the theoretical aspects and for detailed descriptions of the results. Numerous plots excerpted from the relevant journal articles, summarize these data. The development over the past decade

of a general circulation model at the National Center for Atmospheric Research (NCAR) is traced historically. Variations in circulation as a function of season, solar cycle, altitude, and latitude are given in Section 17.5.2. Local time variations are shown in Section 17.5.3. Results of including the effects of magnetospheric convection are shown in Section 17.5.4. This section also includes the recent work, described in a series of papers, of the University College of London group in determining the influence of magnetospheric convection. At present, the effects of heating by particle precipitation at high latitudes and by upward propagating atmospheric tides and momentum deposition by tides at lower latitudes are not included in the above models. There are insufficient data on which to build an adequate model of particle precipitation. Tidal phenomena have been extensively described by Forbes and Garrett [1979] and by Forbes [1982a,b] and should eventually be incorporated as part of thermospheric dynamic models.

17.5.1 Observational Summary and Limitations

Direct measurements of winds in the thermosphere are difficult to make and there have been insufficient observations for development of empirical models. Techniques include ground-based Fabry-Perot spectrometer measurements of the 6300 Å line in the airglow [for example, Hays and Roble, 1971; Sipler et al., 1982] and incoherent scatter radar techniques [for example, Evans, 1975], rocketborne chemical release experiments [for example, Kelley et al., 1977], and satellite data from mass spectrometers [for example, Spencer et al., 1976] and accelerometers [DeVries, 1972; Marcos and Swift, 1982]. These limitations are being ameliorated at F region altitudes with Dynamics Explorer satellite data from the Fabry-Perot Interferometer [Hays et al., 1981] and Winds and Temperature Spectrometer [Spencer et al., 1981]. An extensive review of recent thermospheric wind measurements has been given by Roble [1983]. General features of wind patterns deduced from these experimental data in this reference are summarized below for different latitude regions. For more detailed information on various experimental techniques and results obtained, the reader should consult Roble [1983] and references therein.

The general circulation at midlatitudes at F region heights is controlled mainly by solar heating during quiet geomagnetic conditions and high latitude heat sources during storms. Experimental data obtained when geomagnetic activity is low indicate a flow from the high temperature dayside of the thermosphere to the low temperature nightside. Daytime poleward meridional winds are 50–100 m/s and nighttime equatorward winds are about 100–200 m/s. These values depend on ion drag and season. Zonal winds are generally found to be westward before local noon and eastward in the afternoon. A nighttime transition occurs in early morning (winter) or near midnight (summer). During geomagnetic storms, data show the influence of high latitude heat sources. Observed effects include enhanced nighttime winds, earlier nighttime transition of zonal winds from eastward to westward, a tendency for a midnight "surge" of winds, and occasionally enhanced poleward daytime winds. Fabry-Perot data at Fritz Peak, Colorado, showed nighttime peak equatorial winds of 650 m/s were linearly related to the time rate of change of the equatorial ring current, and peak equatorial meridional and westward zonal winds were related to the square of the auroral electrojet index.

At near-equatorial latitudes the circulation is primarily controlled by solar heating, upward-propagating tides from below 100 km, large-scale neutral gas-plasma interactions (such as $\mathbf{E} \times \mathbf{B}$ drifts), and effects related to geomagnetic storms. A unique feature in this region is the equatorial midnight temperature bulge discovered in incoherent scatter and satellite data. This persistent but variable feature has been attributed to a complex interaction between the upward-propagating semidiurnal tide and ion-neutral momentum coupling associated with the diurnal variation of ion density. A general trend observed is that nighttime equatorward winds weaken and sometimes reverse after midnight in response to this bulge. Other features of the low latitude circulation include seasonal variations with stronger equatorward winds in summer than in winter, and zonal winds that are eastward throughout the night during winter and spring but have a westward reversal during early morning hours in the summer.

At high latitudes, F region data indicate magnetospheric convection is the dominant factor controlling wind systems. Fabry-Perot measurements [Hays et al., 1979] show early evening westward winds in the direction of magnetospheric convection and give evidence of a heat source located near local midnight. Heppner and Miller [1982], analyzing 39 chemical-release rocket experiments at high latitudes, found that winds in the magnetic dusk sector above 160 km follow the magnetospheric convection pattern with a time lag of about two hours. They also observed a 140–160 km altitude transition region in the wind pattern. Meriwether et al. [1982] used airglow data to show strong magnetospheric convection-driven winds with superimposed variability probably due to particle precipitation. Dynamics Explorer (DE) satellite data analyzed by Killeen et al. [1982] show wind velocities of 600–800 m/s. They further show that the DE data can be described as a superposition of a "background" solar-driven wind field and a wind field driven by high latitude energy and momentum sources. DE data of Spencer et al. [1982] showed zonal winds of about 1 km/s related to convecting ions. These strong zonal winds frequently correlated with vertical motions having velocities of 100–250 m/s. Knutson et al. [1977] also found evidence of large vertical winds, up to 65 m/s, in the early morning sector of the auroral zone, with no simple correlation with Kp.

17.5.2 Diurnally Averaged Winds

17.5.2.1 Seasonal Variations at Solar Maximum. The zonal mean (diurnally averaged) thermospheric circulation has been calculated for solar maximum equinox and solstice conditions respectively by Dickinson et al. [1975, 1977] using a two-dimensional (time-dependent) model run under steady state conditions. Their model utilized a zonally averaged version of lower atmospheric dynamic meteorology equations. Heat and momentum sources used were solar EUV and UV heating, high latitude heating due to auroral processes (assumed distributed uniformly in longitude) and a momentum source resulting from correlation of the diurnal variation of winds and ion drag. This model uses a 5° latitude-by-longitude grid with 24 constant-pressure surface layers extending in altitude from about 100 to 500 km. (Note that the altitude of a constant pressure surface can vary.) Their results showed that while the winds are driven primarily by heating due to solar EUV and UV absorption, a high latitude heat source was also required to give an equatorward mean meridional circulation in agreement with incoherent scatter observations [Evans, 1972]. Equinoctial circulation properties are schematically represented in Figure 17-20. This figure shows the mass flow of the cellular motion above about 80 km. Figure 17-20a shows the circulation during a very quiet geomagnetic period. Upward motion occurs over the equatorial subsolar point and flows poleward to high latitudes where sinking motion occurs. The small reverse circulation at high latitudes is due to a small high-latitude heat source as well as persistent solar heating over the polar regions. During average geomagnetic activity (Figure 17-20b) the equatorward circulation is larger. Below ~150 km the strong solar UV heating still maintains a poleward flow. At higher altitudes the flow is equatorward to about 20° latitude. During large geomagnetic storms (Figure 17-20c) the equatorward circulation extends to the equator and circulation below 150 km is also modified. The situation is similar during solstice. Under geomagnetically quiet conditions (Figure 17-21a) the asymmetry in solar heating drives a summer-to-winter circulation. For average geomagnetic conditions (Figure 17-21b), the high-latitude heat source reinforces the summer-to-winter circulation in the summer hemisphere but forces a reverse circulation in the high latitude winter hemisphere, with transition from

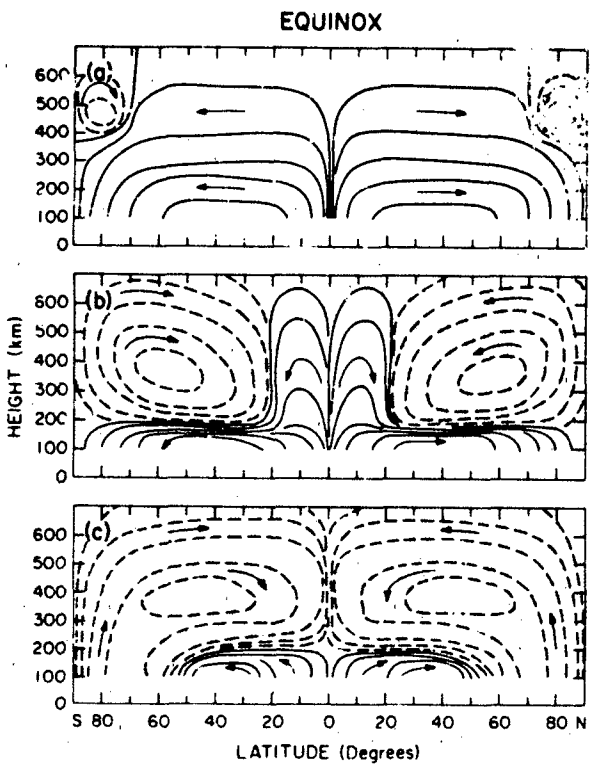


Figure 17-20. Schematic diagram of the zonal mean meridional circulation in the earth's thermosphere during equinox for various levels of auroral activity (a) extremely quiet geomagnetic activity, (b) average activity (10^{15} erg/s), and (c) geomagnetic substorm (10^{16} erg/s). The contours schematically illustrate the mass flow, and the arrows indicate the direction of the motion [Roble, 1977].

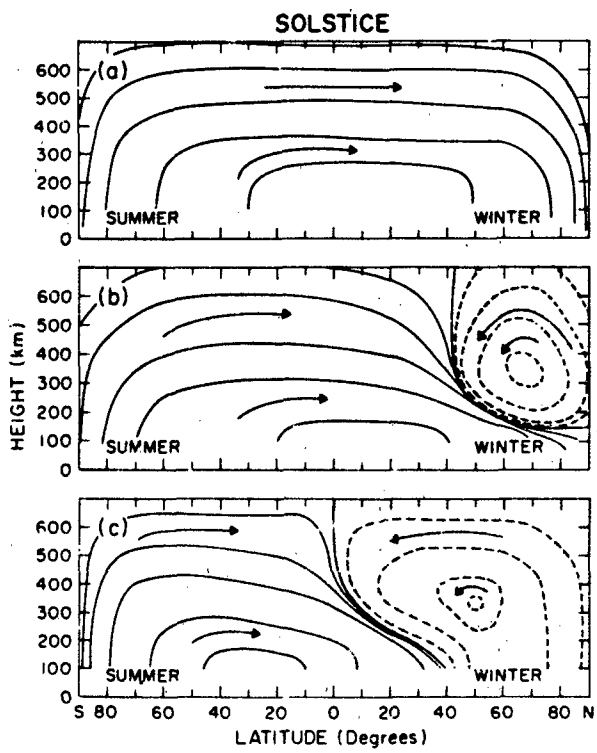


Figure 17-21. Schematic diagram of the zonal mean meridional circulation in the earth's thermosphere during solstice for various levels of auroral activity (a) extremely quiet geomagnetic activity, (b) average activity (10^{15} erg/s), and (c) geomagnetic substorm (10^{16} erg/s). The contours schematically illustrate the mass flow, and the arrows indicate the direction of the motion [Roble, 1977].

one cell to another in the midlatitude winter hemisphere. Below ~150 km the summer-to-winter circulation is maintained at all latitudes. During large geomagnetic storms (Figure 17-21c) the circulation is equatorward above 300 km in both hemispheres, with an asymmetry at lower altitudes.

Roble et al. [1977] have extended this analysis to determine seasonal circulation patterns for solar maximum (F = 165 units) and solar minimum conditions (F = 80 units). A solstice-type circulation is maintained except for about a week either side of equinox. This circulation is most intense at solstice and weakens as equinox is approached. The transition period is controlled in part by the high-latitude heat source which in turn is related to geomagnetic activity. For higher geomagnetic activity the transition from equinox to solstice circulation occurs more rapidly. The derived diurnally average meridional, zonal, and vertical winds for solar maximum conditions are shown in Figures 17-22–17-24 respectively. These results correspond to average geomagnetic conditions.

During equinox the meridional flow (Figure 17-22a) below 150 km is directed from the equator to both poles at about 5 m/s. At higher altitudes the pole-to-equator flow has a maximum velocity of about 50 m/s near 60° latitude in both hemispheres. By October 6 (Figure 17-22b) the circulation below 150 km is almost completely reversed in the summer hemisphere. The summer-to-winter circulation intensifies from 50 to 200 m/s as the season progresses toward the solstice conditions of Figure 17-22c and d. Above ~200 km an intensification of the summer hemisphere equatorward circulation penetrates into the winter hemisphere and forces the boundary between the two oppositely directed winds poleward as the season progresses toward solstice. This boundary is near the equator at equinox and at ~60° latitude for solstice. Also, the strongest equatorward winds in the high latitude winter hemisphere decrease from about 50 m/s at equinox to about 30 m/s at solstice.

The derived zonal winds during equinox (Figure 17-23a) are relatively weak above 200 km. As the season progresses (Figure 17-23b and c) there is a gradual increase in wind speeds above ~150 km in the summer hemisphere and below 150 km in the winter hemisphere. At solstice (Figure 17-23c) there are generally westward winds in the summer hemisphere and eastward winds in the winter hemisphere. In the winter hemisphere the velocity at ~120 km increases to 120 m/s. Above 200 km the maximum solstitial velocities increase to 60 m/s in the summer hemisphere and 40 m/s in the winter hemisphere.

Vertical velocities shown in Figure 17-24 are given in units of cm/s. Below ~120 km the vertical velocity at equinox (Figure 17-24a) is upward in the equatorial region and downward at middle and high latitudes. By October 6 (Fig-

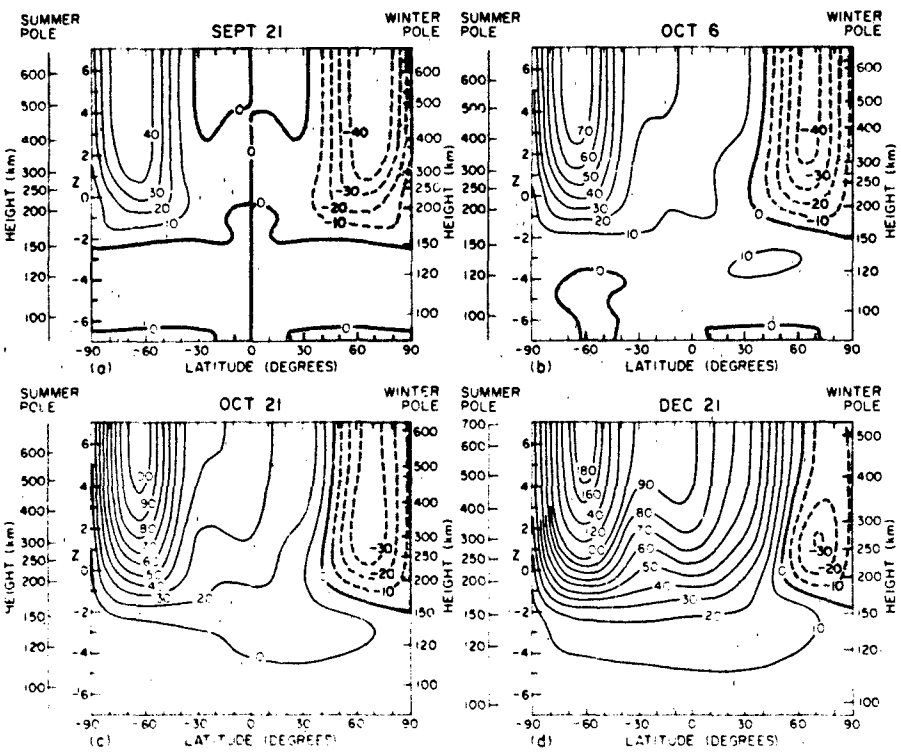


Figure 17-22 The calculated zonal mean contours of meridional wind component (in m/s) for 4 days representing the seasonal variation during solar maximum (a) equinox, September 21, (b) October 6, (c) October 21, and (d) winter solstice, December 21 [Roble et al. 1977].

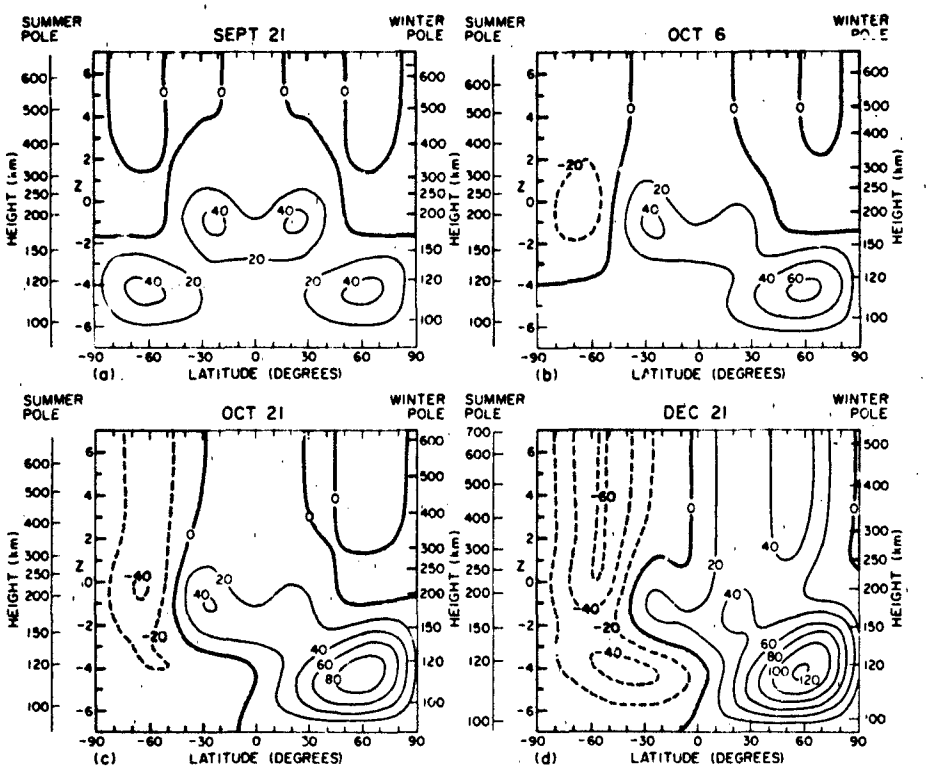


Figure 17-23. The calculated zonal mean contours of zonal wind component (in m/s) for 4 days representing the seasonal variation during solar maximum (a) equinox, September 21, (b) October 6, (c) October 21, and (d) winter solstice, December 21 [Roble et al, 1977].

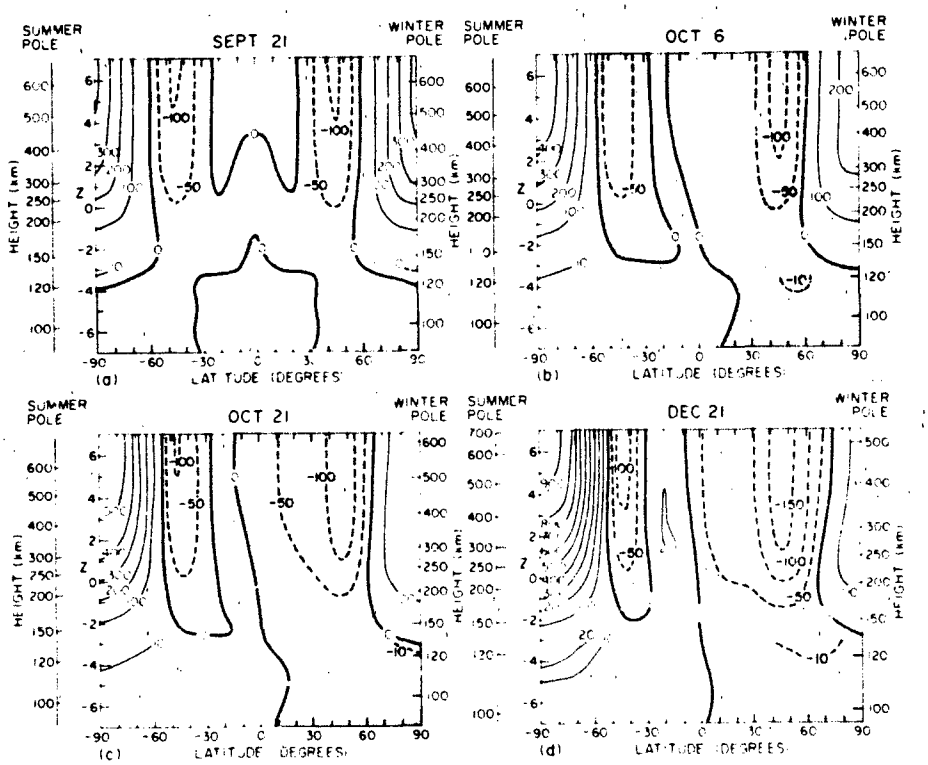


Figure 17-24. The calculated zonal mean contours of vertical wind component (in cm/s) for 4 days representing the seasonal variation during solar maximum (a) equinox, September 21, (b) October 6, (c) October 21, and (d) winter solstice, December 21 [Roble et al, 1977].

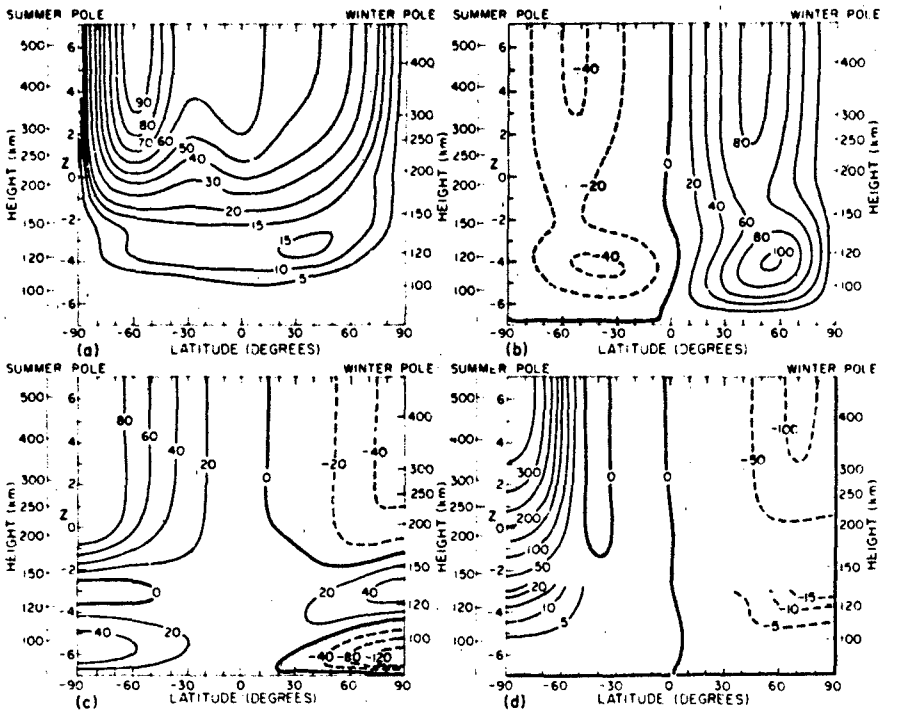


Figure 17-25. Calculated contours of the perturbation zonal mean temperature and winds for December solstice during solar minimum: (a) meridional winds (in m/s), (b) zonal winds (in m/s), (c) temperature (in K) and (d) vertical winds (in cm/s) [Roble et al. 1977].

ure 17-24b) these motions become upward in the summer hemisphere and downward in the winter hemisphere. This pattern intensifies as solstice is approached (Figure 17-24c and d). Maximum upward velocities occur at high latitudes and maximum downward velocities occur at mid-latitudes. The main seasonal effect above ~200 km is the increasing asymmetry between hemispheres as the season progresses. At high latitudes the summer hemisphere upward velocities are about six times greater than those in the winter hemisphere.

17.5.2.2 Solstice Variations at Solar Minimum. Results of calculations for solar minimum (F = 80) December solstice conditions are presented as Figures 17-25a–d for the meridional wind, zonal wind, perturbation temperature, and vertical wind, respectively. The high latitude heat source required to obtain reasonable agreement between the model and experimental data was 4.4 times smaller than that required for solar maximum conditions.

The meridional winds (Figure 17-25a) are directed from the summer to winter hemisphere at all altitudes and latitudes. A maximum velocity of 90 m/s occurs near 330 km at 60° latitude in the summer hemisphere. The magnitude of the winds in the summer hemisphere are lower than those at solar maximum by about a factor of two. The derived high latitude heat source in the winter hemisphere is not sufficiently large to drive a reverse circulation. Poleward winds are encountered at all latitudes in the winter hemisphere. The zonal winds (Figure 17-25b) are westward in the summer hemisphere and eastward in the winter hemisphere. These winds are similar in magnitude to those of solar cycle maximum. In general, the zonal jets near 120 km are slightly greater during solar maximum than during solar minimum. At higher altitudes the converse is true. Perturbation temperature data are shown in Figure 17-25c but are not discussed here. Vertical velocity results are shown in Figure 17-25d. In the summer hemisphere the winds are generally directed upward except for a small region of downward velocities near 40° above ~175 km. Downward winds occur at all latitudes in the winter hemisphere. Vertical velocities are smaller during solar minimum than during solar maximum.

17.5.3 Local Time Variations

A time-dependent "Three-dimensional General Circulation Model" (TGCM) developed by Dickinson et al. [1981] permits specification of winds in latitude-local time coordinates. This model uses the same geophysical inputs as in Dickinson et al. [1975, 1977] and Roble et al. [1977]. Results are presented in Figures 17-26 and 17-27 for solar

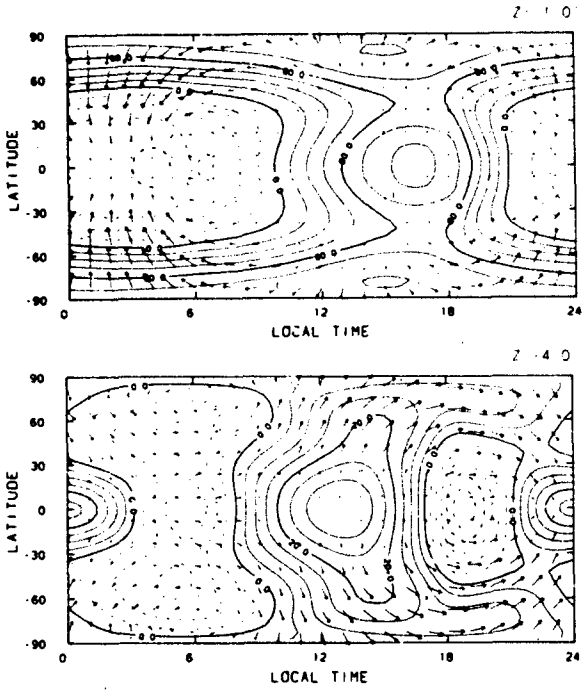


Figure 17-26. Calculated global distribution of winds and perturbation temperature (K) along two constant pressure surfaces: (a) z = 1 at approximately 300 km and (b) z = −4 at approximately 120 km for equinox conditions during solar cycle maximum. Both solar heating and a high-latitude heat source are included. The maximum wind arrow is 130 m/s in (a) and 90 m/s in (b). [Dickinson et al., 1981].

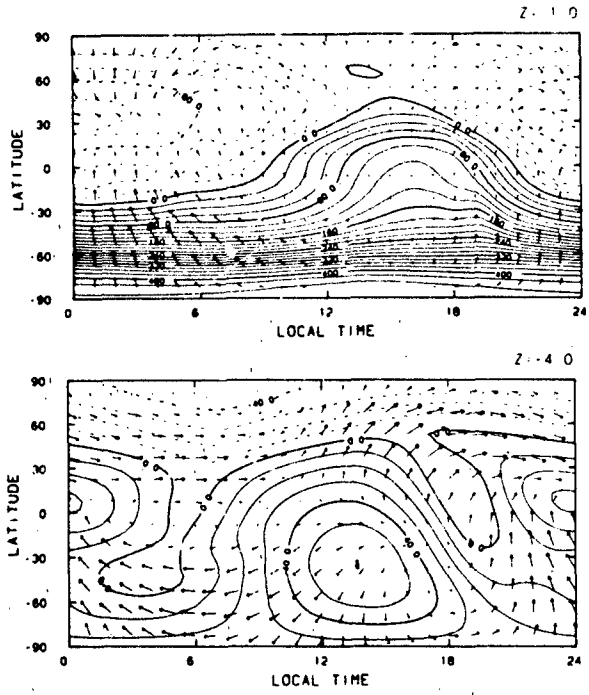


Figure 17-27. Calculated global distribution of winds and perturbation temperature (K) along two constant pressure surfaces: (a) z = 1 at approximately 300 km and (b) z = −4 at approximately 120 km for solstice conditions during solar cycle maximum. Both solar heating and a high-latitude heat source are included. The maximum wind arrow is 180 m/s in (a) and 75 m/s in (b) [Dickinson et al., 1981].

maximum equinox and solstice conditions, respectively, at altitudes of 120 km and 300 km. Solid lines shown in these figures represent a calculated perturbation temperature from global mean values derived from the model and are fully described in Dickinson et al. [1981]. The present discussion considers only the wind results, depicted by arrows.

At 300 km the equinoctial winds blow from the daytime high pressure region near 1700 LT into the nighttime low pressure region near 0400 LT as shown in Figure 17-26a. Maximum winds of 130 m/s occur at night when ion drag is smallest. At 120 km (Figure 17-26b) the maximum arrow corresponds to 90 m/s. The low latitude dynamic structure exhibits a semi-diurnal variation. At high latitudes the circulation remains nearly diurnal.

Results for solstice (southern hemisphere summer) at 300 km are shown in Figure 17-27a. High latitude heating greatly enhances the flow toward the winter hemisphere. The maximum wind vector is in the nighttime summer hemisphere with an equatorward velocity of 180 m/s. At 120 km (Figure 17-27b) the complex circulation pattern again shows a pronounced semi-diurnal oscillation at low latitudes and a dominent diurnal variation at high latitudes. The maximum velocity of 90 m/s occurs for high latitude night conditions.

17.5.4 Magnetospheric Convection Effects on Neutral Winds

More detailed and higher resolution models of thermospheric dynamics have resulted from incorporating electric field models to provide both an energy and momentum source at high latitudes. This permits including the important influence of high-latitude plasma convection. The magnitude of the magnetospheric convection is defined by the electric potential drop across earth's polar cap. Values to 150 kV are possible during strong geomagnetic storms. Correlation of polar cap convection potential drop with more conventional indicators of geomagnetic activity including the Kp and AE (auroral electrojet) indices has been studied by Reiff et al. [1981]. They found that AE had a correlation coefficient of 0.71 with the electric field data while Kp had a relatively weak correlation coefficient of 0.55.

Roble et al. [1982] have used the electric field model of Volland [1975, 1979a,b] in their Thermospheric General Circulation Model. Results have been initially presented to compare two cases: (1) solar EUV and UV heating only and (2) solar heating plus a 26 kV cross-tail potential to simulate low geomagnetic conditions. These runs assume equinox

and solar cycle maximum conditions and coincident geographic and geomagnetic poles. Circulation and temperature data are shown on polar plots for 50°N to 90°N geographic latitude. As in Figures 17-26 and 17-27, arrows represent wind magnitude and direction, and heavy solid lines represent perturbation temperatures.

Figures 17-28 a and b show the circulation pattern predicted at 130 km and 300 km, respectively, in response to solar heating only. These results are the same as obtained for Figures 17-26 and 17-27, but in polar plot format. The characteristic dayside to nightside flow is found at the 130 km level (Figure 17-28a). At high latitudes the daily variation pattern is mainly diurnal while a semidiurnal component is indicated at low latitudes. A counter-clockwise circulation is centered near 0600 LT and 75°N latitude. Maximum winds of 75 m/s occur in the polar region. At the 300 km level (Figure 17-28b) the strongest winds are again in the polar region. However, the day-to-night flow is now rotated to the left of that for the 130 km case.

The effect of including a 20 kV cross-tail potential is shown in Figure 17-29. In the lower thermosphere (Figure 17-29a) the dawn-dusk temperature contrast is enhanced in the polar region but the winds are only slightly changed (from a maximum of 75 m/s to 80 m/s). However, in the upper thermosphere (Figure 17-29b) the overall temperature contours are not significantly modified, but the polar region winds are greatly affected in magnitude and direction. Maximum winds increase from 120 m/s to 200 m/s over the poles. Also, the pattern of circulation in the polar cap is in the direction of the plasma drift associated with magnetospheric convection. At lower latitudes the influence of magnetospheric convection is much smaller.

The TGCM [Roble et al., 1982] was then modified to include displaced geographic and geomagnetic poles. Data are given below for a 20 kV and 60 kV (moderate geomagnetic activity) case. Figure 17-30 shows circulation patterns for the 20 kV case at 130 km and 300 km for 0600 UT and 1800 UT. At 130 km (Figures 17-30a and b) the counter-clockwise circulation on the morningside is near 80° latitude and 0° longitude at 0600 UT. Figure 17-30b (1800 UT) shows it has a UT dependence due to the displaced poles. The 300 km circulation patterns are shown in Figures 17-30c and d. These are analyzed as zonal and meridional components in Figure 17-31. Maximum zonal winds (Figures 17-31a and b) are about 200 m/s in the polar region at 1800 UT and are about 150 m/s at 0600 UT. Maximum meridional winds (Figures 17-31c and d) at high latitudes are also generally about 50 m/s greater at 1800 UT than at 0600 UT. Model results obtained by increasing the cross-tail potential to 60 kV (moderate geomagnetic activity) are shown in Figure 17-32. At 120 km (Figures 17-32a and b) the maximum wind velocity increases to 102 m/s. The counterclockwise wind rotation around the low temperature vortex also increases in magnitude. At 300 km (Figures

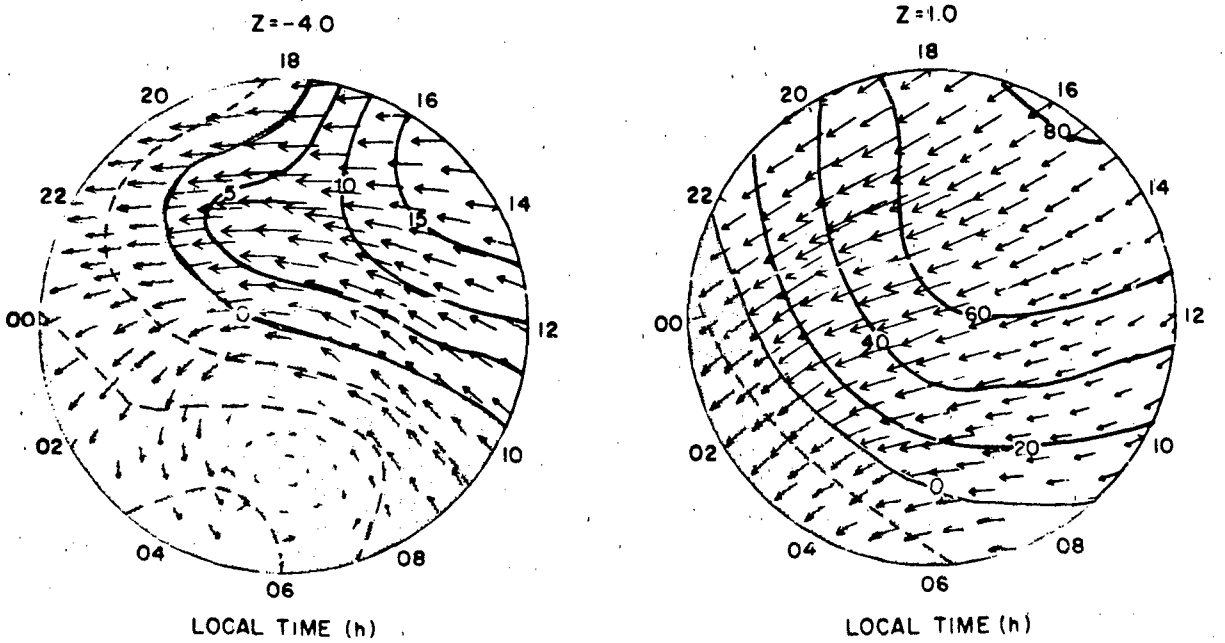


Figure 17-28 Polar plots giving the direction and magnitude of the calculated northern hemisphere high-latitude circulation and contours of perturbation temperature (K) along (a) the $z = -4$ (130 km) and (b) $z = +1$ (300 km) constant-pressure surfaces for the case of solar heating alone. Coincident poles are assumed. The wind speed associated with the maximum arrow is 75 m/s in (a) and 150 m/s in (b) [Roble et al., 1982].

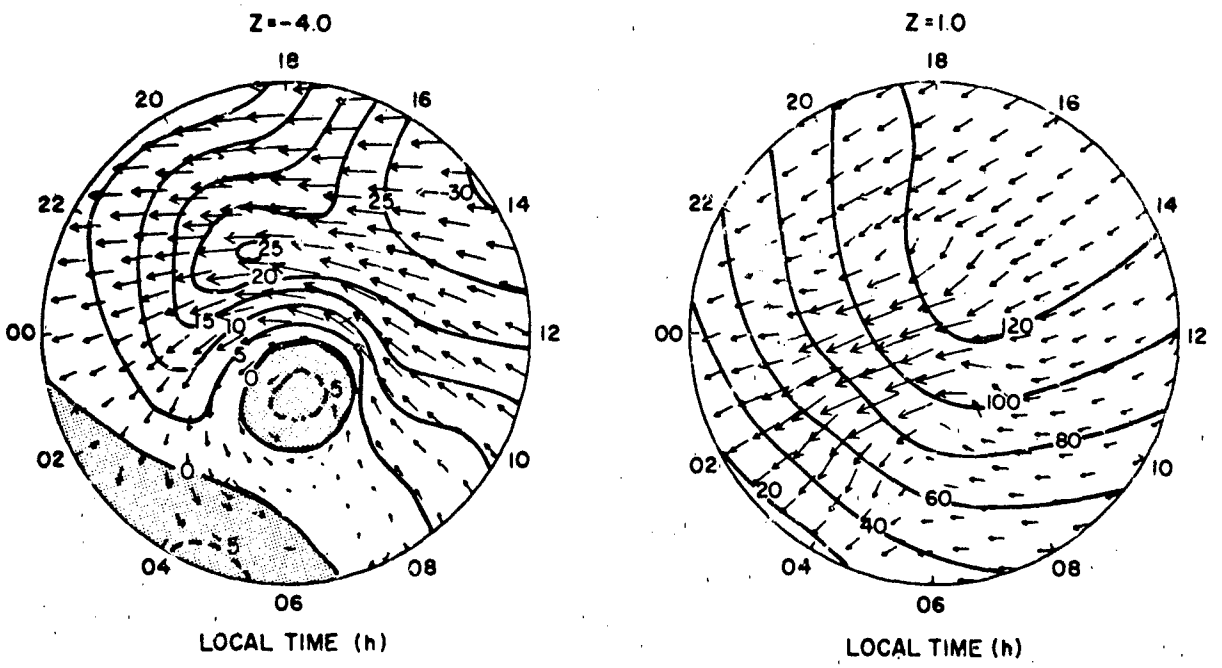


Figure 17-29. Polar plots giving the direction and magnitude of the calculated northern hemisphere high-latitude circulation and contours of perturbation temperature (K) including magnetospheric convection with a cross-tail potential of 20 kV along (a) the $z = -4$ (130 km) and (b) $z = +1$ (300 km) constant-pressure surfaces for the case of solar heating alone. The wind speed associated with the maximum arrow is 80 m/s in (a) and 225 m/s in (b) [Roble et al., 1982].

17-32c and d) the winds in the magnetic polar cap increase to 380 m/s. These winds are enhanced by ion drag with the incremental winds tending to follow the twin ion drift vortices associated with magnetospheric convection. The double vortex wind system at F-region heights extends downward into the lower thermosphere. However, there is a significant shift in the pattern with the main features at 120 km associated with the cold low-pressure and warm high-pressure regions.

Roble et al. [1983] have extended the TGCM to examine seasonal variations for solar cycle maximum conditions. This version of the model assumes coincident geographic and geomagnetic poles; universal time effects are not included. Results are shown for the three cases: (1) solar EUV and UV heating only, (2) solar heating plus a 20 kV cross-tail potential, and (3) solar heating plus a 60 kV cross-tail potential.

Results for southern hemisphere summer at the altitude of approximately 130 km are given in Figure 17-33. With solar heating only (Figure 17-33a) the circulation is again from the high temperature noon sector toward the midnight sector where maximum wind velocities of about 75 m/s occur. With a 20 kV cross-tail potential (Figure 17-33b) there is very little change in the circulation pattern. The maximum wind speed is 77 m/s. For the 60 kV case (Figure 17-33c) the circulation pattern changes, reflecting the influence of the larger magnetospheric convection ion velocities. The maximum wind speed of 89 m/s occurs directly over the polar region. Vortex development occurs on the morningside of the polar cap, obstructing the seasonal flow from the high-to-low temperature regions. The effect of magnetospheric convection is to redirect the overall flow pattern, resulting in a jet over the polar cap.

Results for winter conditions in the northern hemisphere at 120 km are given in Figure 17-34. In each case, a counterclockwise circulation is driven throughout the high latitude region by a low-pressure system over the pole. Maximum wind speeds increase from 108 m/s to 117 m/s to 120 m/s as the cross-tail potential increases from zero kV to 20 kV to 60 kV. Magnetospheric convection effects are less than in summer because of the lower electron density and consequently decreased ion drag momentum source.

Results for the altitude of approximately 300 km for southern hemisphere summer winds are given in Figure 17-35. Maximum winds increase from 110 m/s for the zero kV case (Figure 17-35a) to 200 m/s for the 20 kV case (Figure 17-35b). For both cases the neutral winds tend to follow the two-cell pattern. For the 60 kV case (Figure 17-35c) anti-sunward winds increase to 380 m/s. Further, at the boundary of the polar cap, the wind direction is sunward, following the pattern of magnetospheric convection. An equatorward surge in wind velocities is also observed near local midnight.

Circulation patterns for the 300 km winter case are given

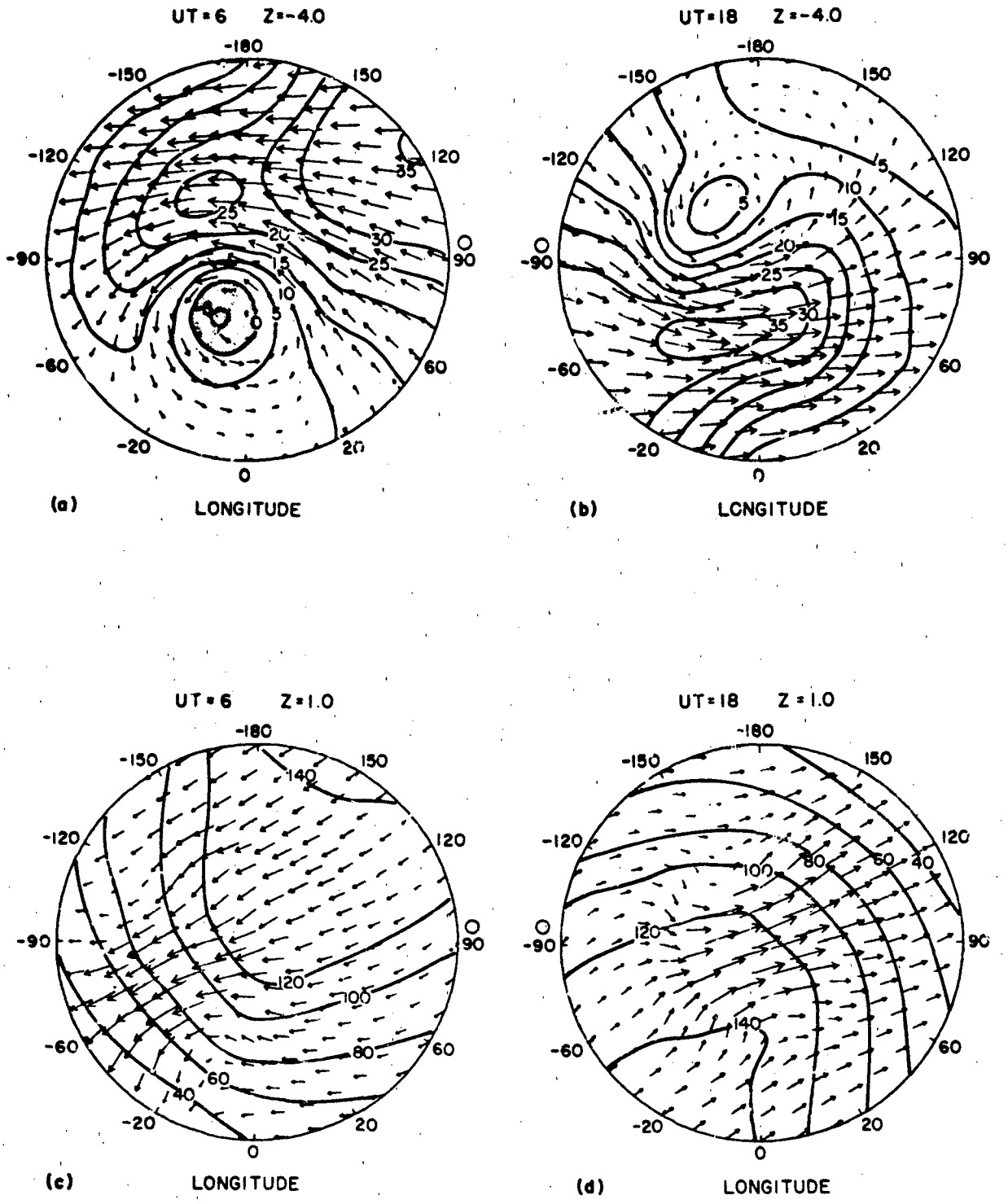


Figure 17-30. Winds and contours of perturbation temperature (K) along the $z = -4$ (130 km) surface for two universal times, 0600 UT (a) and 1800 UT (b), and along the $z = +1$ (300 km) surface at 0600 UT (c) and 1800 UT (d) for the case where the geographic and geomagnetic poles are displaced and the cross-tail potential for the magnetospheric convection model is 20 kV. The circulation is indicated by the arrows giving direction, and the length is the magnitude of the winds. The length of the maximum arrow represents 76 m/s in (a), 87 m/s in (b), 219 m/s in (c), and 222 m/s in (d), respectively. Local noon is indicated by an open circle on the boundary [Roble et al., 1982].

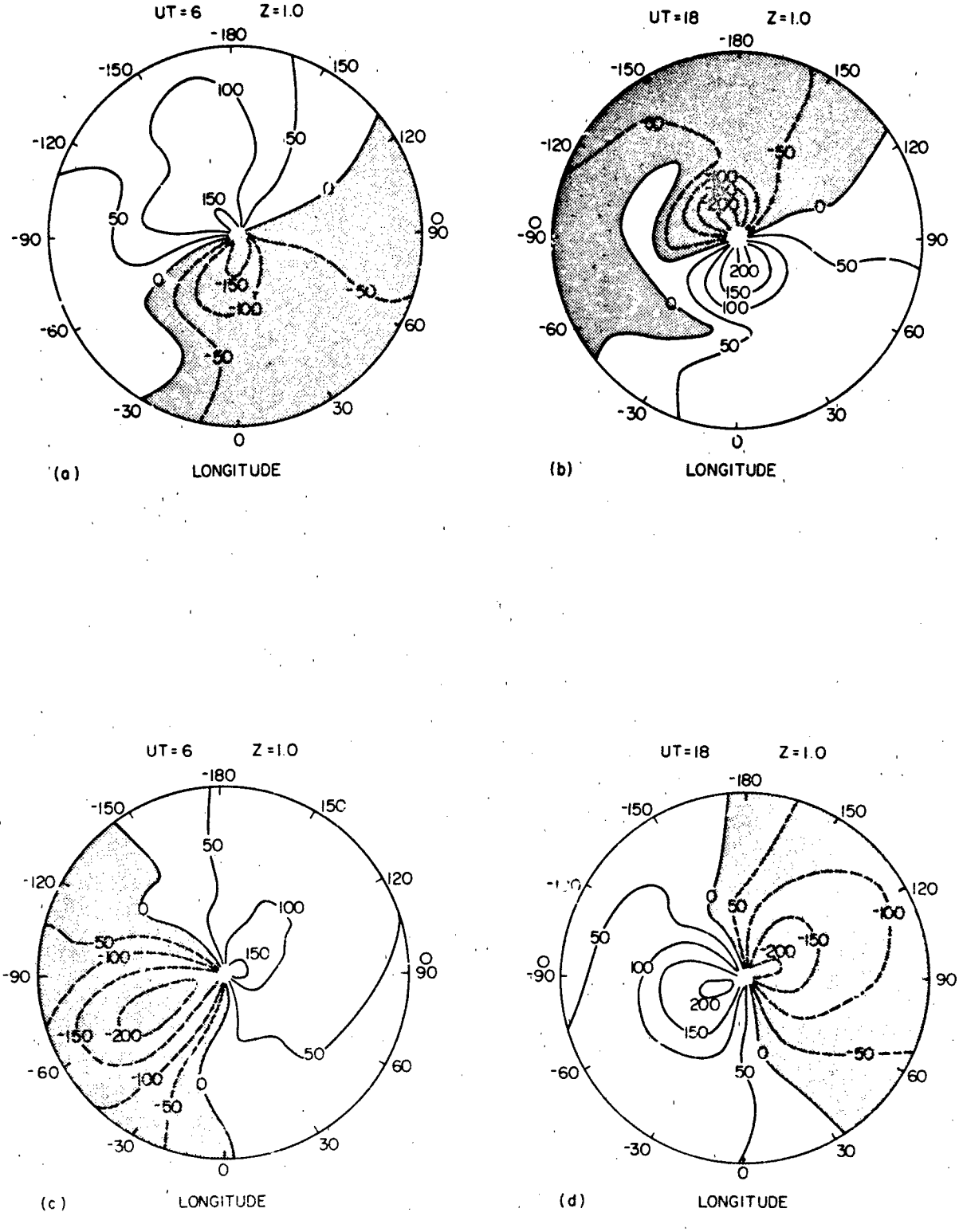


Figure 17-31. Contours of the zonal and meridional wind components at 0600 UT, (a) and (c), and 1800 UT, (b) and (d), respectively, along the $z = +1$ (300 km) surface for the case where the poles are displaced and the cross-tail potential is 20 kV. Local noon is indicated by an open circle on the boundary [Roble et al., 1982].

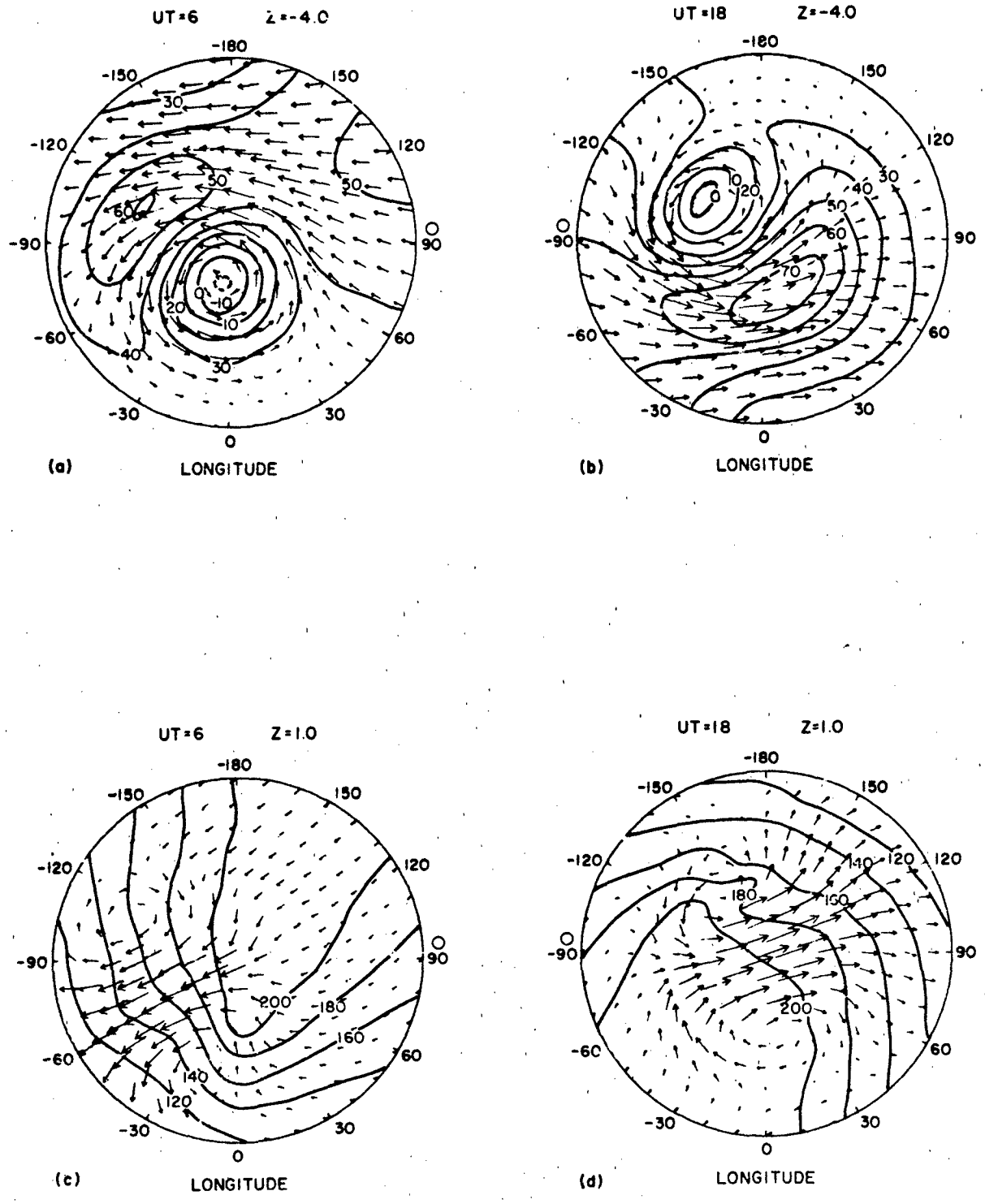


Figure 17-32. Winds and contours of perturbation temperature (K) along the $z = -4$ (130 km) surface for two universal times, 0600 UT (a) and 1800 UT (b), and along the $z = +1$ (300 km) surface at 0600 UT (c) and 1800 UT (d) for the case where the geographic and geomagnetic poles are displaced and the cross-tail potential for the magnetospheric convection is 60 kV. The circulation is indicated by the arrows giving direction, and the length is the magnitude of the winds. The length of the maximum arrow represents 100 m/s in (a), 109 m/s in (b), 395 m/s in (c), and 373 m/s in (d). Local noon is indicated by an open circle on the boundary [Roble et al., 1982].

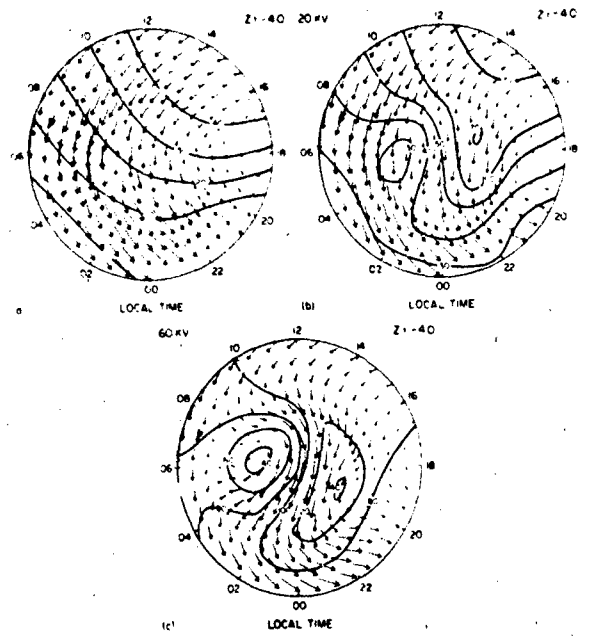


Figure 17-33. Polar plots giving the direction and magnitude of the calculated southern hemisphere (summer) high-latitude circulation and contours of perturbation temperature (K) along the z = −4 (130 km) constant-pressure surface at December solstice for (a) solar heating only, (b) solar heating plus magnetospheric convection with a cross-tail potential of 20 kV, and (c) solar heating plus magnetospheric convection with a cross-tail potential of 60 kV. The wind speed associated with the maximum arrow is 75 m/s in (a), 77 m/s in (b) and 89 m/s in (c) [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

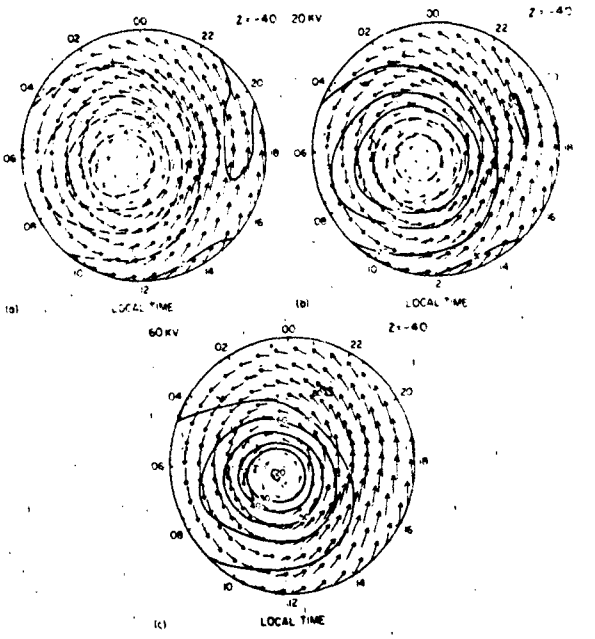


Figure 17-34. Polar plots giving the direction and magnitude of the calculated northern hemisphere (winter) high-latitude circulation and contours of perturbation temperature (K) along the z = −4 (130 km) constant-pressure surface at December solstice for (a) solar heating only, (b) solar heating plus magnetospheric convection with a cross-tail potential of 20 kV, and (c) solar heating plus magnetospheric convection with a cross-tail potential of 60 kV. The wind speed associated with the maximum arrow is 108 m/s in (a), 117 m/s in (b) and 120 m/s in (c) [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd © 1983.)

in Figure 17-36. As for the corresponding case at 120 km, the circulation patterns for the zero and 20 kV cases are similar, with the maximum wind increasing from 181 (Figure 17-36a) to 200 m/s (Figure 17-36b). For the 60 kV case (Figure 17-36c) polar winds increase to 330 m/s. These winds also show a surge near 0200 LT. There is again a tendency to follow the two-cell pattern of magnetospheric convection with a reduced effect due to the lower winter electron densities.

Pole-to-pole cylindrical equidistant projections of winds and temperatures as a function of local time as given by Roble et al. [1983] are shown in Figure 17-37 for 120 km and Figure 17-38 for 300 km. Figure 17-37 shows that the major effect of magnetospheric convection in the lower thermosphere occurs at high latitudes in the summer with a 60 kV potential as discussed previously. In the upper thermosphere (Figure 17-38) the 60 kV potential has a much more pronounced influence on the high-latitude circulation. Maximum winds of about 400 m/s occur at high latitudes near local midnight and are equatorward. Changes are observed even at lower latitudes, indicating that the high-latitude energy input is redistributed globally by the thermospheric wind system.

Roble et al. [1983] also give meridional winds and temperature data in altitude-latitude coordinates. Results are given for the cases of a zero and a 60 kV potential for local times of 1200 h and 0000 h. In Figures 17-39 (1200 LT) and 17-40 (0000 LT) the top two plots show temperatures for the two heating cases and the bottom two plots show the corresponding wind fields. At 1200 LT, maximum velocities for solar heating only are 75 m/s and 125 m/s over the summer and winter poles respectively. The inclusion of magnetospheric convection increases the flow over the pole to 150 m/s and 300 m/s for the summer and winter cases respectively. Also, the higher summer electron densities permit the ion drag acceleration source to accelerate the winds down to about 130 km in the summer hemisphere compared to about 180 km in the winter hemisphere. Along the 0000 LT meridian, comparison of Figures 17-40a and b shows the magnetospheric convection greatly enhances equatorward flow. As the cross-tail potential is increased from zero to 60 kV, winds over the summer pole increase from 100 m/s to 400 m/s while over the winter pole they increase from 100 m/s to 300 m/s. The sharp wind gradient near 60° in both hemispheres indicates the equatorward limit of magnetospheric convection forcing for this local time.

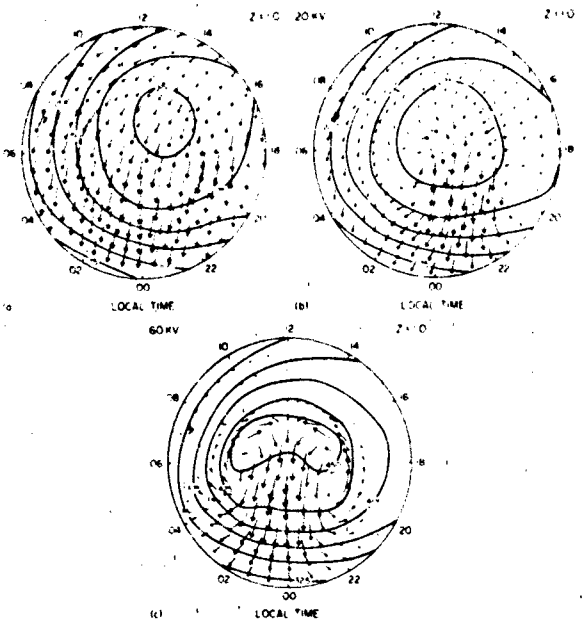


Figure 17-35. Polar plots giving the direction and magnitude of the calculated southern hemisphere (summer) high-latitude circulation and contours of perturbation temperature (K) along the $z = +1$ (~300 km) constant-pressure surface at December solstice for (a) solar heating only, (b) solar heating plus magnetospheric convection with a cross-tail potential of 20 kV, and (c) solar heating plus magnetospheric convection with a cross-tail potential of 60 kV. The wind speed associated with the maximum arrow is 110 m/s in (a), 200 m/s in (b), and 380 m/s in (c) [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

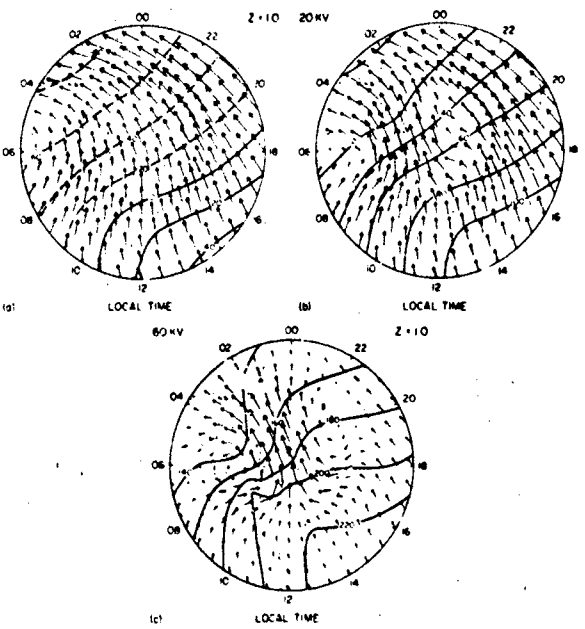


Figure 17-36. Polar plots giving the direction and magnitude of the calculated northern hemisphere (winter) high-latitude circulation and contours of perturbation temperature (K) along the $z = +1$ (~300 km) constant-pressure surface at December solstice for (a) solar heating only, (b) solar heating plus magnetospheric convection with a cross-tail potential of 20 kV, and (c) solar heating plus magnetospheric convection with a cross-tail potential of 60 kV. The wind speed associated with the maximum arrow is 181 m/s in (a), 200 m/s in (b), and 339 m/s in (c) [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

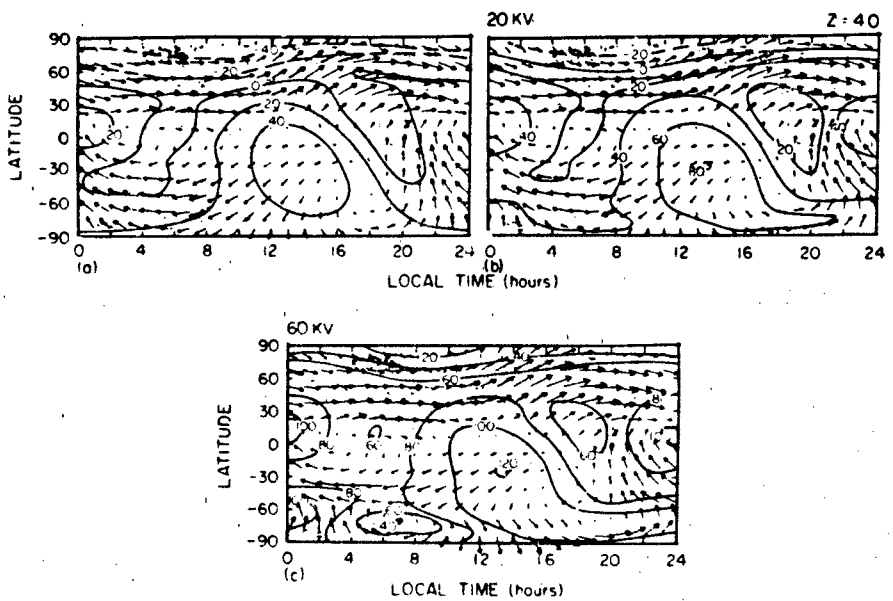


Figure 17-37. Cylindrical equidistant projections giving the direction and magnitude of the calculated circulation and contours of perturbation temperature (K) along the $z = +4$ (130 km) constant-pressure surface for December solstice for (a) solar heating only, (b) solar heating plus magnetospheric convection with a cross-tail potential of 20 kV, and (c) solar heating plus magnetospheric convection with a cross-tail potential of 60 kV. The wind speed associated with the maximum arrow is 108 m/s in (a), 117 m/s in (b) and 120 m/s in (c) [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

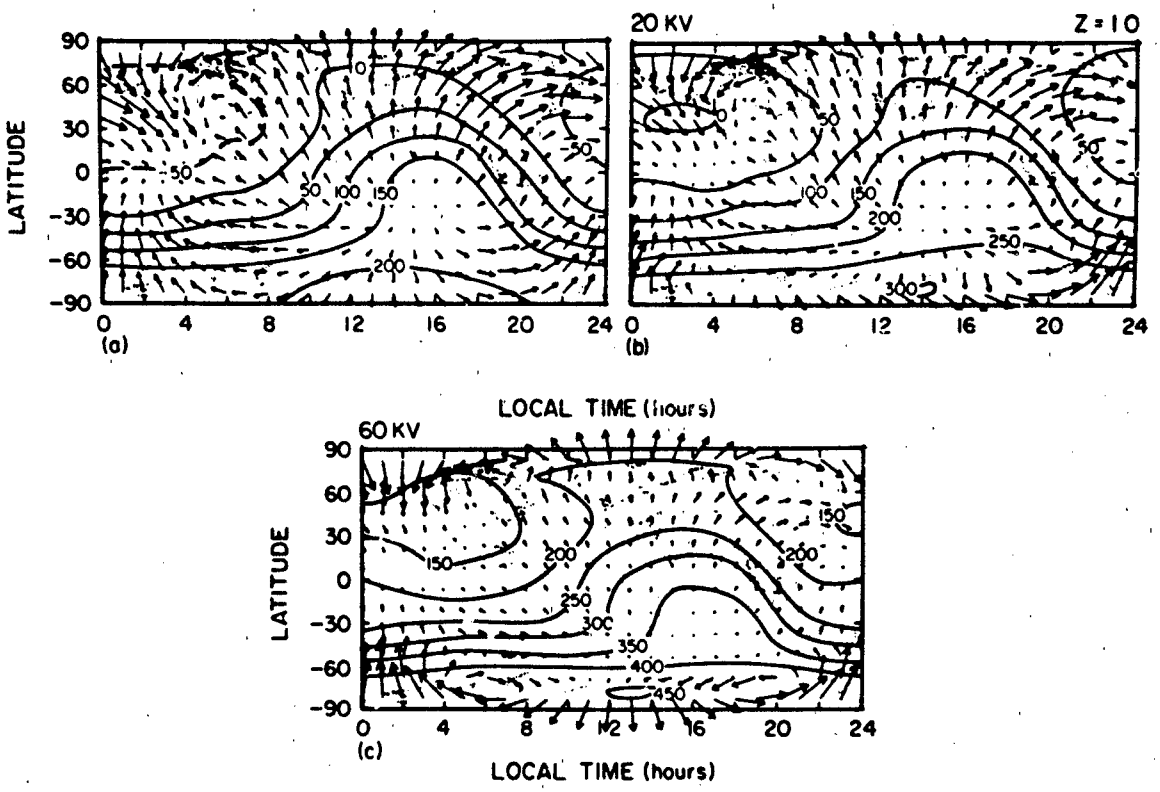


Figure 17-38. Cylindrical equidistant projections giving the direction and magnitude of the calculated circulation and contours of perturbation temperature (K) along the $z = +1$ (300 km) constant-pressure surface for December solstice for (a) solar heating only, (b) solar heating plus magnetospheric convection with a cross-tail potential of 20 kV, and (c) solar heating plus magnetospheric convection with a cross-tail potential of 60 kV. The wind speed associated with the maximum arrow is 181 m/s in (a), 220 m/s in (b) and 380 m/s in (c) [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

The TGCM model has been run to simulate conditions encountered by the Dynamics Explorer (DE-2) satellite. Results from two on-board wind sensors [Killeen et al., 1982; Spencer et al., 1982] were combined to provide thermospheric winds along the spacecraft path, at an altitude of about 300 km. These data were then compared to the TGCM predictions. Experimental data obtained on the four orbits over high southern latitudes, shown in Figure 17-41a are considered representative of summer conditions. Local times are between night and early morning hours. Data are shown in polar plots and in geomagnetic coordinates. Further details of these data are in the paper by Roble et al. [1983]. This figure shows strong control of the winds by magnetospheric convection and the characteristic two-cell convection pattern. Maximum winds are over the polar cap and have velocities between 500 and 750 m/s. Winter (northern hemisphere) polar data for three passes are shown in Figure 17-41b. Local times are during twilight and daytime hours. Magnetospheric convection control of the winds is evidenced by their similarity with those for summer conditions. Maximum speeds of 500–800 m/s are found in the polar cap. In the auroral oval, the winds are sunward with speeds of about 300 m/s. The TGCM simulations used a 60 kV potential and displaced geographic and geomagnetic poles. Results for the summer and winter cases are given in Figures 17-42a and b respectively. The model data generally agree with the direction of the experimental data. However, the model wind magnitudes underestimate the measured winds. Maximum model values are 400–500 m/s. Roble et al. [1983] note that better agreement could be obtained by increasing the cross-tail potential and ion-drift velocities in the magnetospheric convection model, but that such a modification is not warranted until the ion-drag momentum source is better defined (from DE-2 satellite plasma measurements).

Roble et al. [1983] have also examined whether a linear combination of solar-driven and high-latitude auroral-driven circulation and temperature structures reproduce the total structure. Such a linear combination would allow high-latitude plasma-dynamic interactions with neutral gas to be analyzed independently, with the results superimposed linearly on the solar-driven results. This would have the advantages of simplifying analyses of Dynamics Explorer data and placing more confidence in linear perturbation model approaches to thermospheric dynamic studies. Their finding

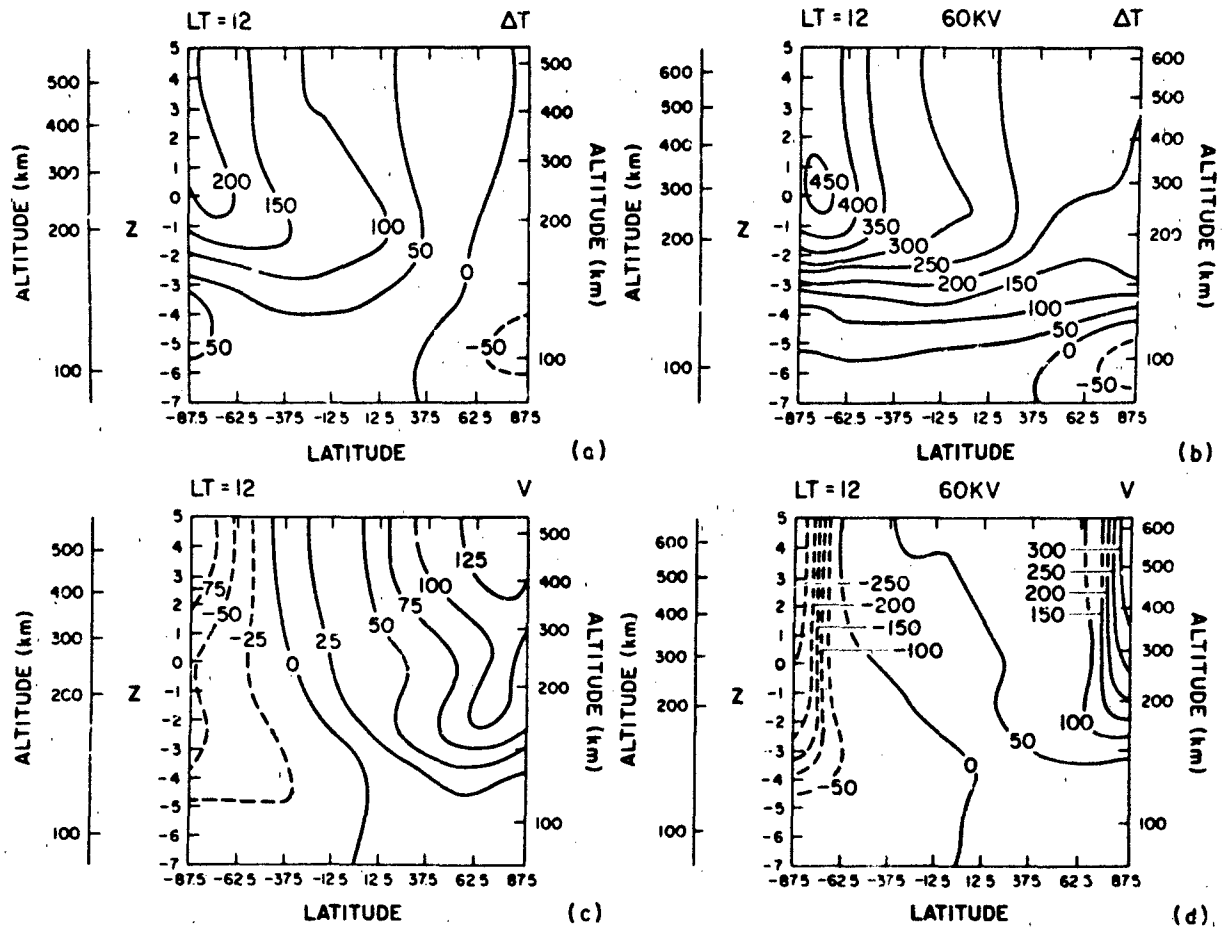


Figure 17-39. Meridional cross sections giving contours of the calculated perturbation temperature (K) [(a) and (b)] and meridional wind speed (m/s, positive northward) [(c) and (d)], respectively. The local time of the meridional slice is 1200 LT. (a) and (c) are the perturbation temperature and meridional wind for the case of solar heating only, and (b) and (d) are the perturbation temperature and meridional wind for the case of solar heating plus magnetospheric convection with a cross-tail potential of 60 kV [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

was that a linear superposition is accurate to within 10%–20% at F region heights, where motion is mainly controlled by the ion drag momentum source. In the lower thermosphere the linear combination is not a good approximation, apparently because of the non-linear nature of energy transport by winds and of Joule heating which is a function of the square of the difference between neutral and ion velocities. It therefore appears that non-linear models are required for lower thermospheric studies.

Thermospheric wind variations as a function of time due to a substorm corresponding to moderate geomagnetic activity have been calculated by Fuller-Rowell and Rees [1981]. These results were derived using the global three-dimensional, time-dependent numerical model of Rees et al. [1980]. Fuller-Rowell and Rees [1981] estimate that their input cross-tail potentials for the example given here represent a variation in Kp from about one to about six and back to about one in successive 3-h intervals. Equinoctial solar maximum conditions and coincident geographic and geomagnetic poles are assumed. Winds are calculated for altitudes 120 and 240 km from onset of the storm to 4.5 h after onset at 20-min intervals. Results are shown here at t = 0, 1, and 2 h to indicate the complex nature of the time-varying response. Longer term effects are also noted in the text. Data are again presented as polar plots from 50°–90° north geographic latitude. Arrows representing winds are scaled to 40 m/s = 2° latitude at 120 km and 160 m/s = 2° latitude at 240 km.

Pre-storm conditions, t = 0, at 120 km (Figure 17-43a) show a strong anti-solar flow in the evening polar cap which partially returns via a ~40 m/s sunward flow in the morning auroral oval. It is partially deflected in the midnight region into the westward flow of the evening auroral oval (2000–0200 LT) to eventually rejoin the general anti-sunward flow. These results are in general agreement with the 20 kV data of Roble et al. [1982] shown in Figure 17-29a. At t = 1 h (Figure 17-43b) peak wind amplitudes of ~180 m/s occur.

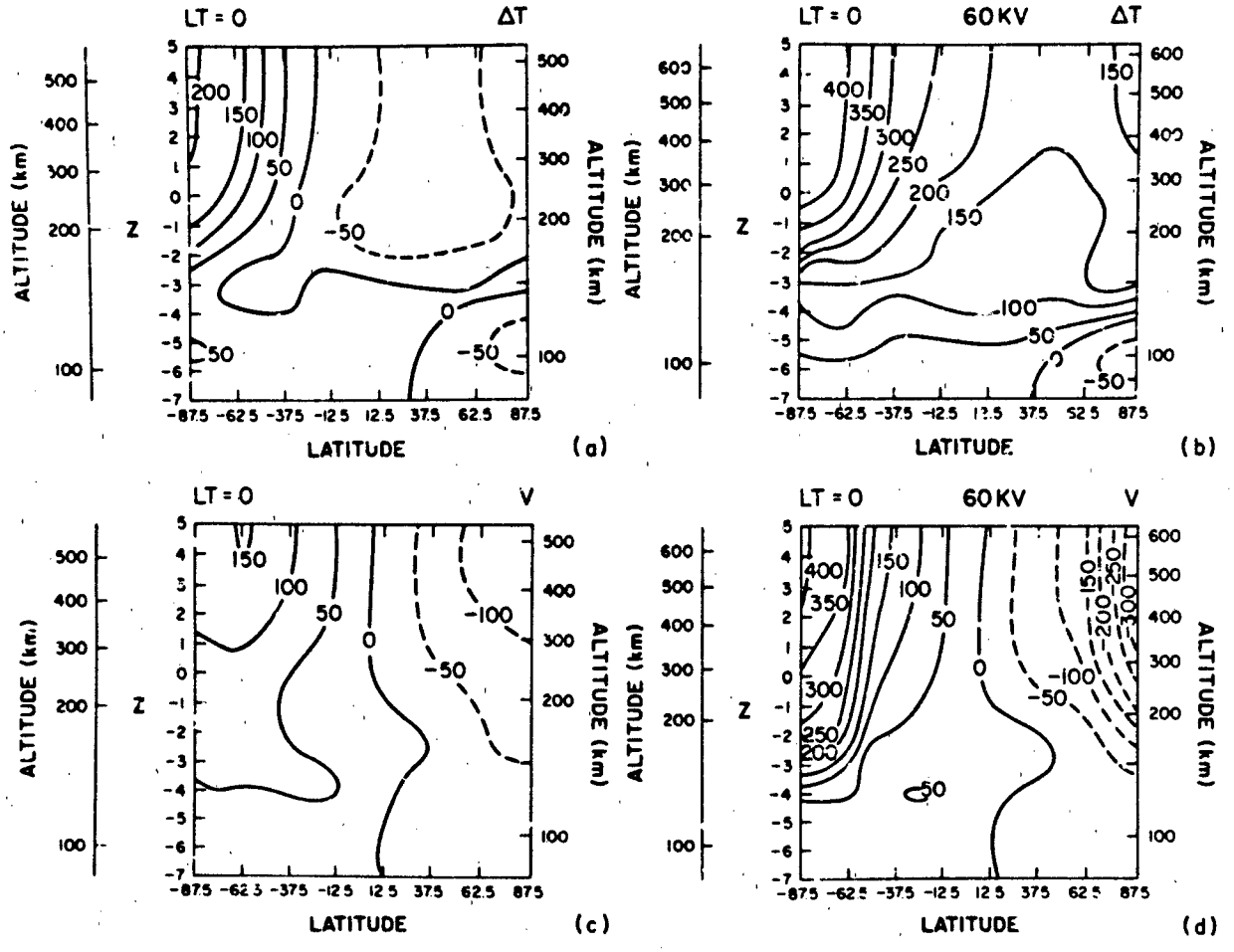


Figure 17-40. Meridional cross sections giving contours of the calculated perturbation temperature (K) [(a) and (b)] and meridional wind speed (m/s, positive northward) [(c) and (d)], respectively. The local time of the meridional slice is 0000 LT. (a) and (c) are the perturbation temperature and meridional wind for the case of solar heating only, and (b) and (d) are the perturbation temperature and meridional wind of the case of solar heating plus magnetospheric convection with a cross-tail potential of 60 kV [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

Other significant features are poleward winds in the afternoon sector south of 70°N latitude, equatorward winds from midnight to midday in the morning region south of 65°N latitude and a steadily increasing surge from dusk to dawn over the polar cap. At t = 2 h (Figure 17-43c) velocities have fallen to about half their maximum values. Also, an anti-cyclone circulation has been established with center near 2200 LT and 74° latitude. There is also a complimentary cyclonic cell centered near 1000 LT and 74° latitude. These cells have velocities of 160 m/s and decrease very slowly with time. At 4.5 h after onset the winds south of 65°N latitude revert to essentially pre-storm conditions in the afternoon and evening. In the morning sector equatorward winds of 60 m/s persist. Poleward of 65° latitude the anticyclonic nighttime and weaker cyclonic daytime vortices continue, along with the connecting dusk to dawn flow. Fuller-Rowell and Rees [1981] indicate that these vortices may persist for several more hours.

Greater wind magnitudes are observed at 240 km. The anti-sunward flow over the polar cap at t = 0 (Figure 17-44a) is about 300 m/s (compared with 150 m/s when solar heating is the only source in the thermosphere). Also, in the late evening auroral oval (1600-2400 LT) there is a significant sunward flow which forms a cell centered about 72° latitude with closure over the polar cap. South of 60°N latitude the wind circulation is only slightly modified by the electric field. One hour into the substorm (Figure 17-44b), winds in the evening auroral oval and over much of the polar cap exceed 300 m/s. Near 1500 LT and 73° latitude the velocity is 450 m/s. Horizontal winds converge near 70°–75° latitude and 1200 LT due to the auroral oval jets. While not evident in the limited data shown here, there are strong oscillations due to gravity waves which have already propagated to 50° latitude in the early morning hours, affecting both zonal and meridional winds. Wind magnitudes in the midday region have been significantly decreased but

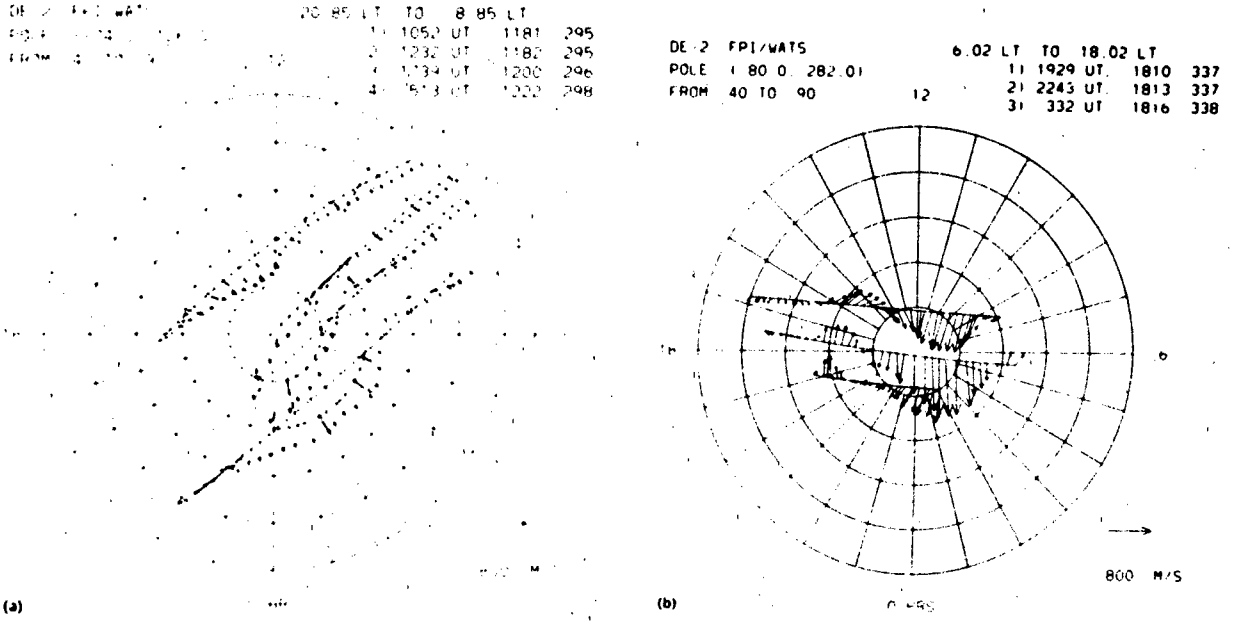


Figure 17-41. Polar plots of the measured wind vector at F-region heights along the DE-2 satellite path: (a) winds on four orbits measured over the southern magnetic polar cap on 22-23 Oct 1981 and (b) winds measured on three orbits over the northern magnetic polar cap on 3-4 Dec 1981 [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

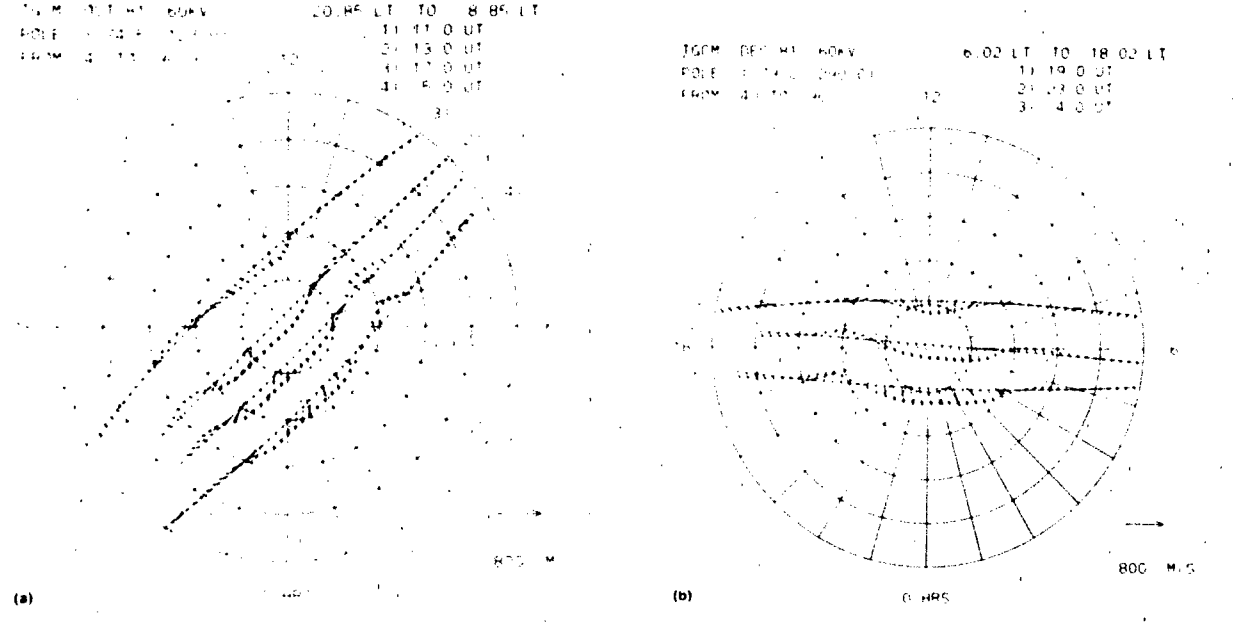


Figure 17-42. Polar plots of the calculated wind vector at F region heights along the same DE-2 satellite paths shown in Figure 17-31: (a) winds for the same four orbits over the southern hemisphere magnetic polar cap in October and (b) winds for the same three orbits over the northern magnetic polar cap in December [Roble et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

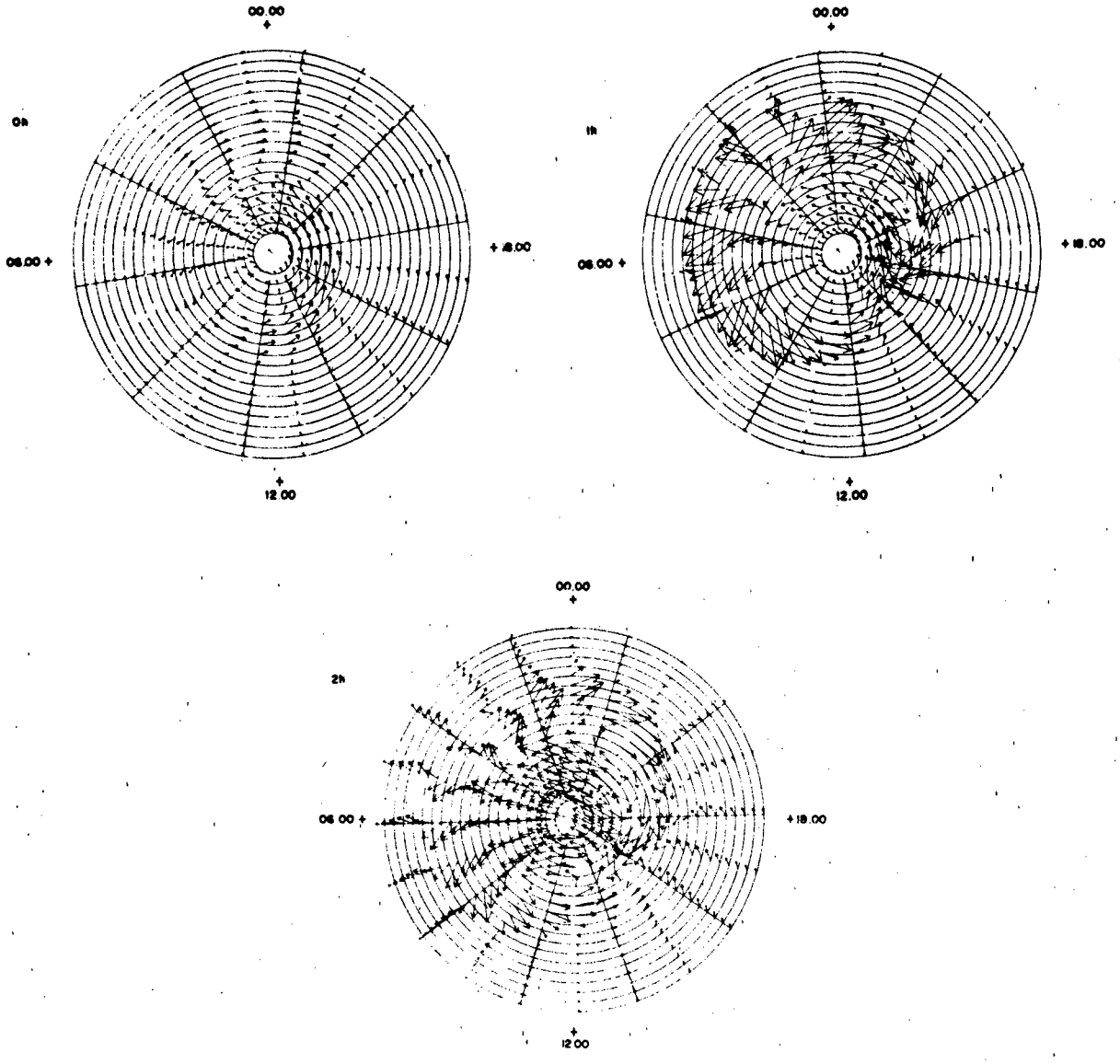


Figure 17-43. Polar plot of dynamical response of winds at 120 km to a large geomagnetic storm at (a) t = 0 h, (b) t = 1 h, and (c) t = 2 h. Winds are scaled to 40 m/s = 2° latitude [Fuller-Rowell and Rees, 1981]. (Reprinted with permission from Pergamon Press Ltd. © 1981.)

not reversed. Wind vectors over the polar cap appear to converge toward the midnight part of the auroral oval. This is due to the poleward gravity wave launched by the dayside auroral oval. The high velocity ion flow over the polar cap also provides momentum to drive surges of equatorward winds at midlatitude, particularly between midnight and dawn. At t = 2 h (Figure 17-44c) the morning circulation cell is tending to collapse, but the evening cell is hardly reduced. Polar cap winds have steadily increased to maximum values greater than 600 m/s in the night sector between 64°–70° latitude. Evidence of gravity wave characteristics are again observed. On the nightside (2200 LT to 0900 LT) strong oscillations of both meridional and zonal winds occur with latitude and longitude. Also, the dayside anti-solar circulation has been stopped, or "deflected" since the late morning winds are strongly enhanced relative to the case for solar heating only. At t = 4 h, the model runs show that the late morning circulation is still enhanced and that gravity waves reach mid to low latitudes. Simulations run out to 12 h after onset show that the wind circulation has not yet returned to prestorm conditions.

An improved version of this model has been developed

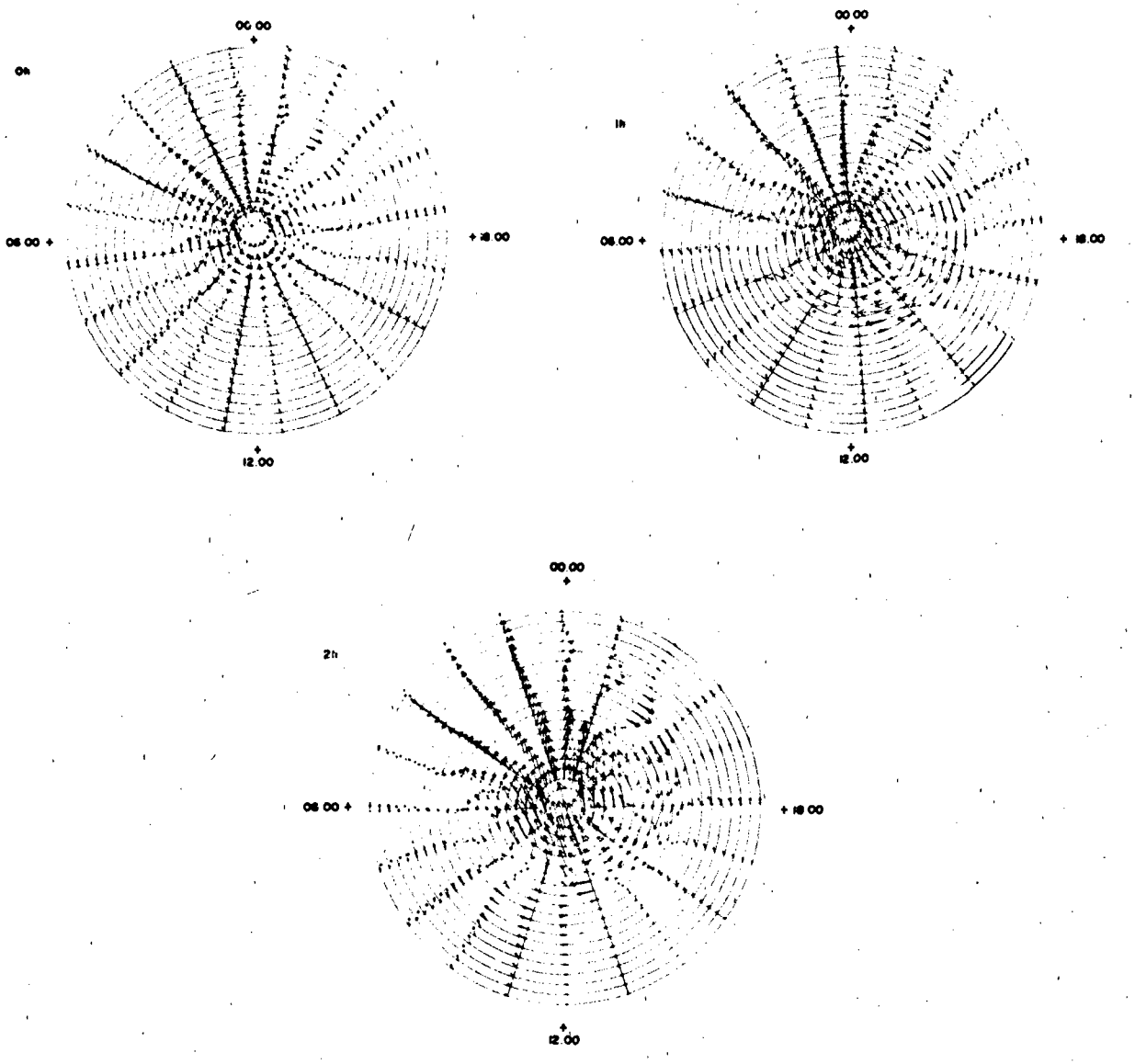


Figure 17-44. Polar plot of dynamical response of winds at 240 km to a large geomagnetic storm at (a) t = 0 h, (b) t = 1 h, and (c) t = 2 h. Winds are scaled to 166 m/s = 2° latitude [Fuller-Rowell and Rees, 1981]. (Reprinted with permission from Pergamon Press Ltd. © 1981.)

[Rees et al., 1983]. It incorporates the theoretical treatment by Quegan et al. [1982] of the polar and auroral ionosphere. This results in a much more realistic electron density distribution at high latitudes during other than low geomagnetic conditions since ionization enhancements due to magnetospheric phenomena are reflected. This second dynamical model and the earlier version [Rees et al., 1980] have been compared to the same DE-2 Oct 81 and Dec 81 data shown in Figure 17-41. Only a brief description of the extremely comprehensive analyses given by Rees et al. [1983] is presented below.

The DE-2 wind data, denoted "KVEC", are shown for each orbit in Figures 17-45 and 17-46 as a function of geographic latitude. The left side of the graph shows evening data and the right side of the graph shows morning data for the October data (Figure 17-45) while the reverse occurs for the December data (Figure 17-46). Model data for quiet geomagnetic conditions (labeled "CHIU") and moderately disturbed conditions (labeled "SHEFFIELD") are shown respectively above and below the experimental data for each orbit. The "SHEFFIELD" results are from the new dynamical model using variable electron densities. The following

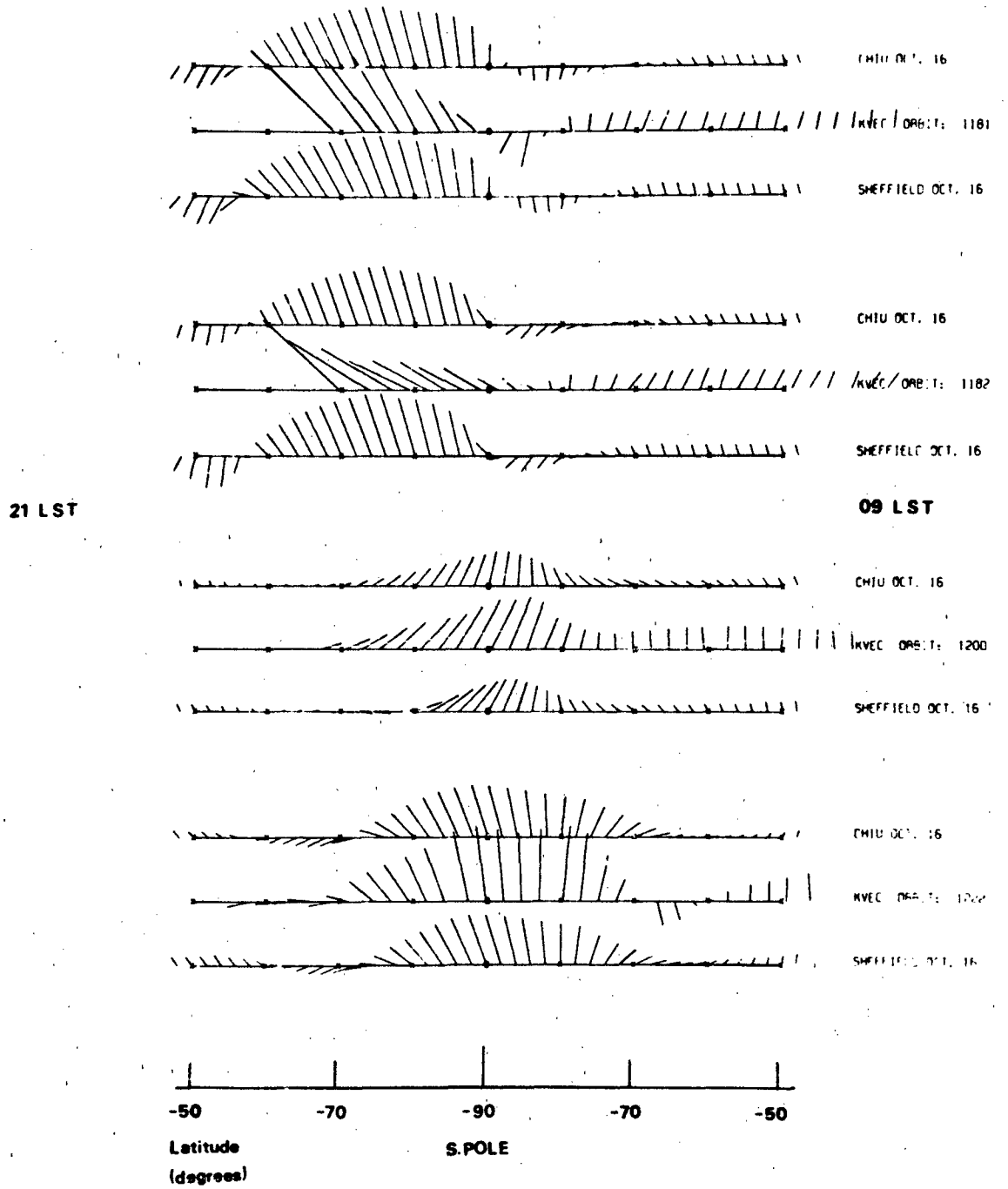


Figure 17-45. Comparison between each individual southern (Oct 1981) polar pass (marked KVEC orbit No.) and the corresponding polar slices from two theoretical global (three-dimensional time-dependent) model simulations at 320 km. Chiu—Quiet geomagnetic conditions. Sheffield—Moderately-disturbed geomagnetic conditions. Scale: 10° latitude ≡ 667 m s [Rees et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

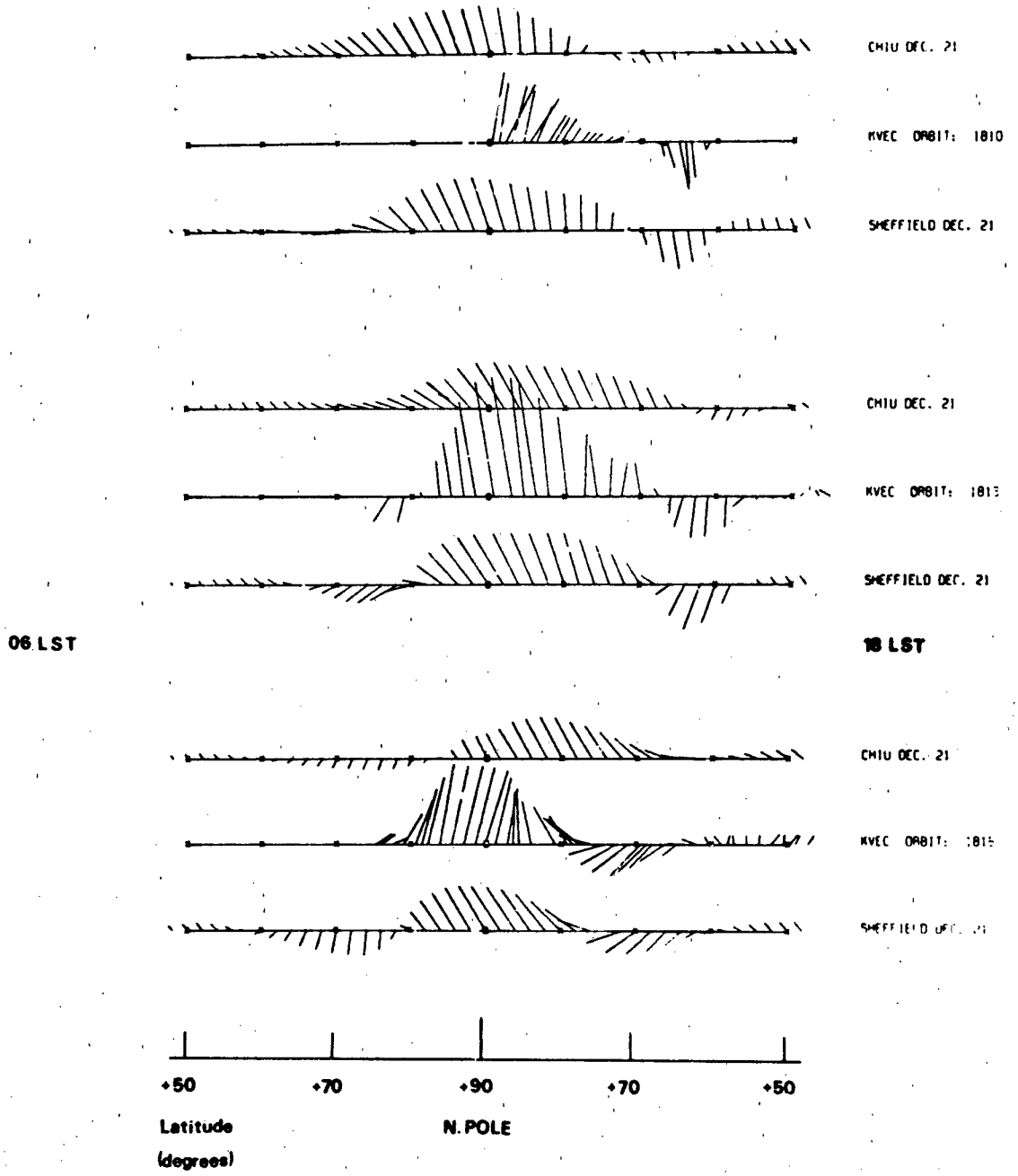


Figure 17-46. Comparison between each individual northern (Dec 1981) polar pass (marked KVEC orbit No.) and the corresponding polar slices from two theoretical global (three-dimensional time-dependent) model simulations at 320 km. Chiu—Quiet geomagnetic conditions. Sheffield—Moderately-disturbed geomagnetic conditions. Scale: 10° latitude = 667 m s [Rees et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

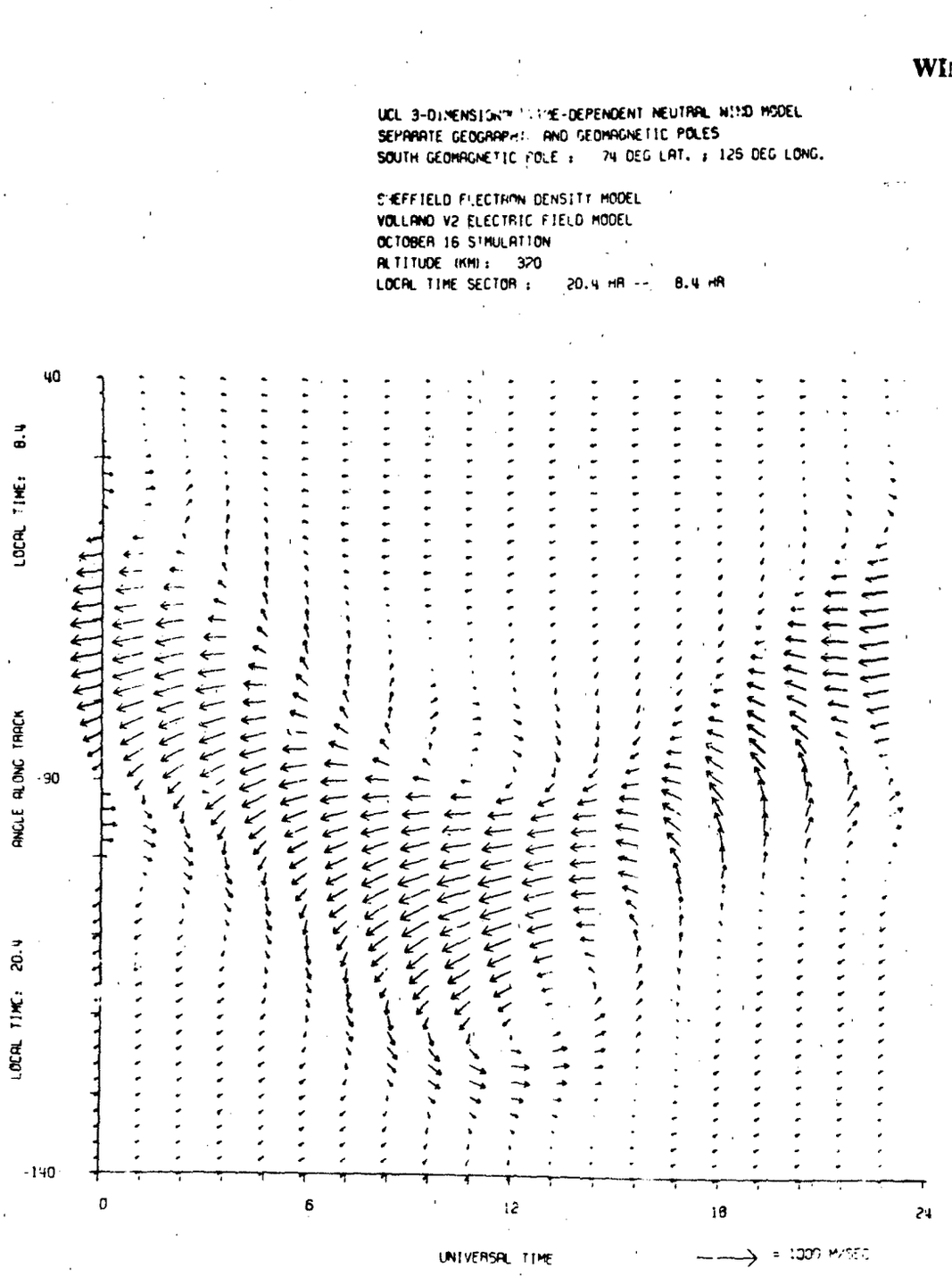


Figure 17-47. University College London three-dimensional, time-dependent neutral wind model. Separate geographic and geomagnetic poles. South geomagnetic pole: 74° Lat., 126° Long. Sheffield Electron Density Model plus Volland V2 Electric Field Model. 16 October simulation. Altitude 320 km; local time sector: 2040-0840 h. This presentation allows the UT geomagnetic dependence of vector winds from DE-2 to be compared orbit by orbit. Note this is not a flow pattern. Each slice is taken from the model at the same local time at 1.2 (U.T.) intervals as the earth rotates under the plane of the satelite orbit [Rees et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

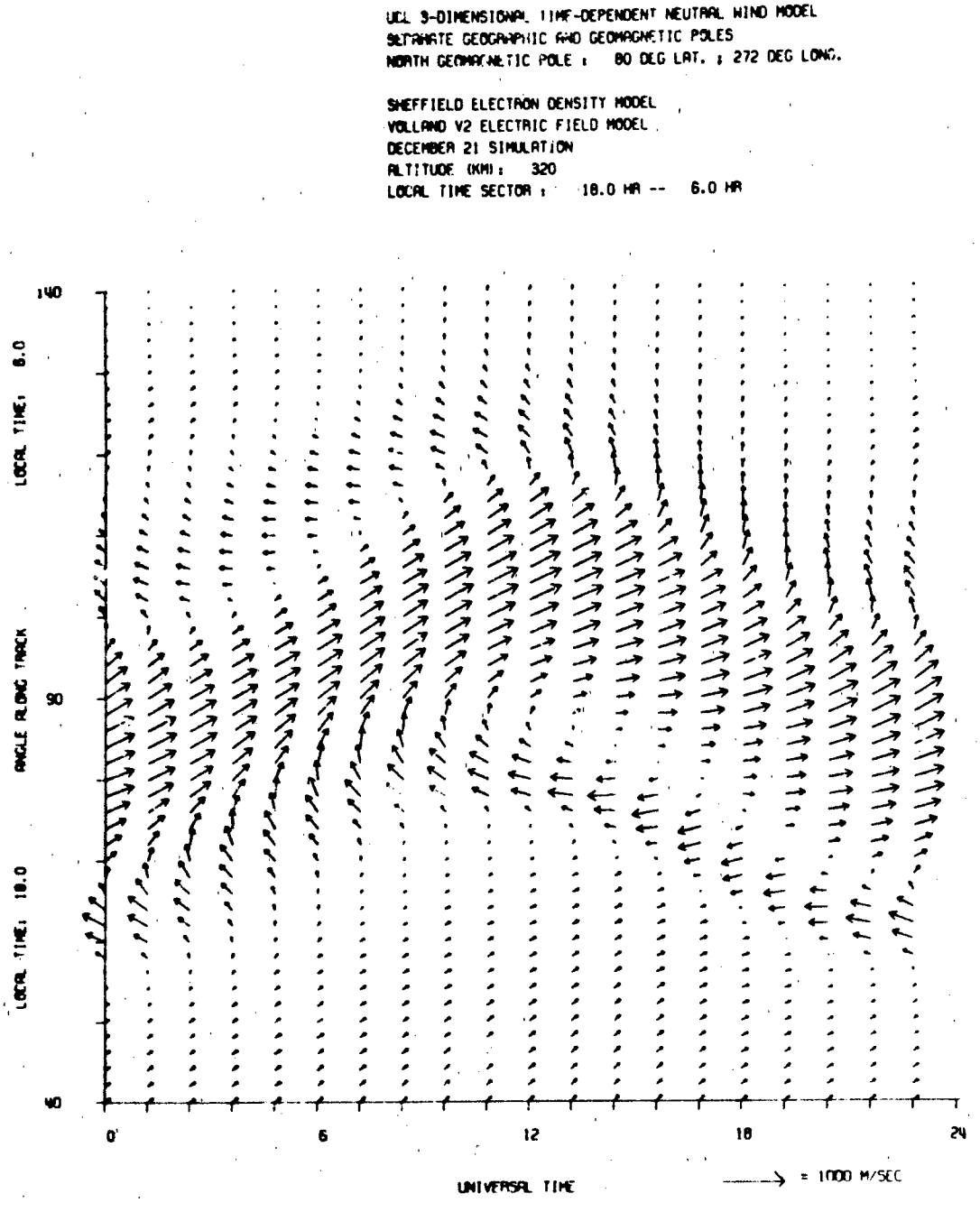


Figure 17-48 University College London three-dimensional, time-dependent neutral wind model. Separate geographic and geomagnetic poles. North geomagnetic pole: 80° Lat., 272° Long. Sheffield Electron Density Model plus Volland V2 Electric Field Model. 21 December simulation allows the UT geomagnetic dependence of vector winds from DE-2 to be compared orbit by orbit. Note this is not a flow pattern. Each slice is taken from the model at the same local time at 1.2 (UT) intervals as the earth rotates under the plane of the satellite orbit [Rees et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

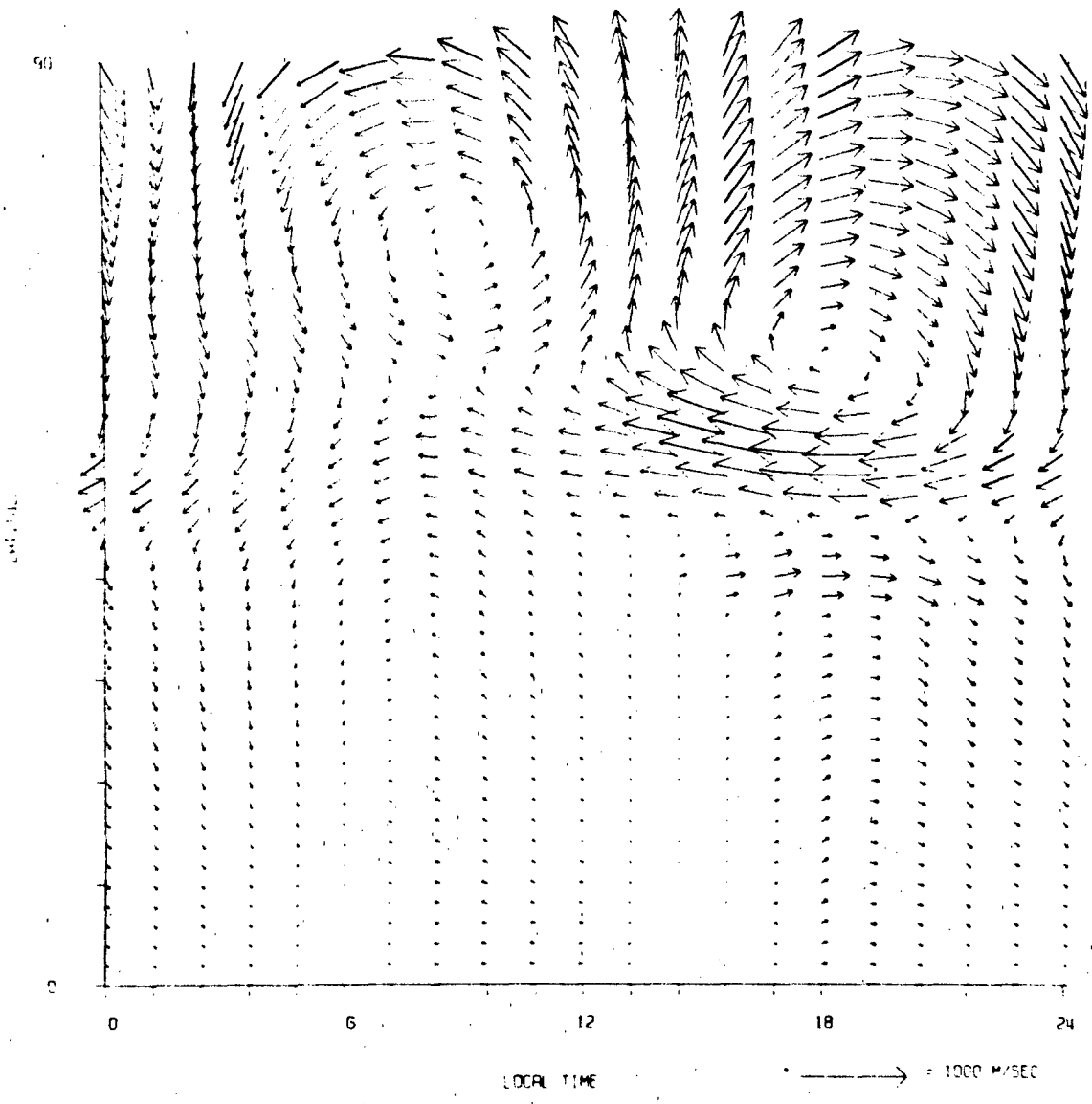


Figure 17-49. Storm-time simulation. University College London three-dimensional time-dependent neutral wind model. Separate geographic and geomagnetic poles. North geomagnetic pole; 80° Lat.; 272° Long. Sheffield Electron Density Model; LEF2 Large Electric Field Model. Expanded auroral oval and cross-cap potential. Altitude 320 km; U.T. 22.80. Note: This diagram shows the actual thermospheric flow pattern at 22.80 U.T. In December 1981, the DE-2 satellite was (approx.) in the 0600-1800 Local Solar Time plane and thus crossing the region of peak westward winds (~ 50°–60° geographical latitude at 22.80 U.T., over North America) [Rees et al., 1983]. (Reprinted with permission from Pergamon Press Ltd. © 1983.)

wind features were identified for each polar pass: (a) a region of strong antisunward flow over the geomagnetic polar cap with considerable variability within an orbit and from orbit to orbit, (b) sunward winds (or strongly reduced antisunward winds) in the dawn and dusk auroral ovals, (c) antisunward winds below the latitude of the auroral oval, as expected from a solar-driven source, and (d) highly variable high latitude winds with only a general correlation of velocity with geomagnetic activity.

The wind velocities in the different flow regimes and their dependence on Universal Time and geomagnetic events were studied further. The relation between ion drifts and neutral velocities showed that wind velocities are 30%–60% of ion velocities in the afternoon/evening auroral oval, greater than ion velocities over the polar cap, and not well correlated with ion velocities in the dawn/morning auroral oval. More detailed studies of the relationship between ion convection and neutral winds are planned using simultaneous electric field and ion drift data from DE-2. The UT variations in neutral winds as calculated from the moderate activity model for the local time planes of the Oct 81 and Dec 81 data are given in Figures 17-47 and 17-48 respectively. These plots are necessary to explain many features of the wind data related to locations of auroral oval boundaries. There is a modulation of about 45° in the geographic location of these boundaries as a function of Universal Time. Results of a simulated polar electric field enhanced to 130 kV, corresponding to a Kp value between ~5 and 7, are given in Figure 17-49. The predicted winds have been compared to those obtained during the disturbed Dec 81 period. At auroral oval latitudes the predicted winds exceed the measured data, but in the polar cap the maximum DE-2 winds (orbit 1813) are greater than the model values. The figure illustrates that during very disturbed periods the wind patterns and magnitudes accommodate to the expansion of the auroral oval and the increasing ion convection. Future studies are planned to examine these affects and to attempt to distinguish between effects of high latitude momentum and heating (Joule and particle heating) processes.

ACKNOWLEDGEMENTS

The critical comments of R. G. Roble, National Center for Atmospheric Research, Capt. D. V. Ridge, AFGL, and J. M. Forbes, Boston University, are greatly appreciated.

REFERENCES—CHAPTER 17

Abbey, R.F., Jr., "Risk Probabilities Associated with Tornado Windspeeds," *Proc. Symp. on Tornadoes*, Institute for Disaster Research, Texas Tech. Univ., Lubbock, Texas, 1976.

Amayenc, P., "Tidal Oscillations of the Meridional Neutral Wind at Midlatitudes," *Radio Sci.*, **9:** 281–294, 1974.

Arnold, A., and R. Bellucci, "Variability of Ballistic Meteorological Parameters," Tech Memo. No. M-1913, Ft. Monmouth, U.S. Army Sig. Eng. Lab, 1957.

Balsley, B.B., and K.S. Gage, "The MST Radar Technique: Potential for Middle Atmospheric Studies," *Pure Appl. Geophys.*, **118:** 452–493, 1980.

Bean, S.J., and P.N. Somerville, "Some Models for Windspeed," 4, AFGL-TR-79-0180, AD A077048, 1979.

Brekke, A.J., J.R. Doupnik, and P.M. Banks, "Observations of Neutral Winds in the Auroral-E Region During the Magnetospheric Storm of Aug 3–9, 1972," *J. Geophys. Res.*, **79:** 2248–2256, 1974.

Ching, B.K., and Y.T. Chiu, "A Phenomenological Model of Global Ionospheric Electron Density in the E, F1, and F2 Regions," *J. Atmos. Terr. Phys.*, **35:** 1616–1630, 1973.

Crutcher, H.L., "Upper Wind Statistics Charts of the Northern Hemisphere," NAVAIR 50-1C-535, 1959.

DeVries, L.L., "Analysis and Interpretation of Density Data from the Low-G Calibration System (Logacs)," *Space Res.*, **12:** 777–789, 1972.

Dickinson, R.E., E.C. Ridley, and R.G. Roble, "Medidional Circulation in the Thermosphere. I, Equinox Conditions," *J. Atmos. Sci.*, **2:** 1737–1754, 1975.

Dickinson, R.E., E.C. Ridley, and R.G. Roble, "Medidional Circulation in the Thermosphere. II, Solstice Conditions," *J. Atmos. Sci.*, **34:** 178–192, 1977.

Dickinson, R.E., E.C. Ridley, and R.G. Roble, "A Three-Dimensional General Circulation Model of the Thermosphere," *J. Geophys. Res.*, **86:** 1499–1512, 1981.

Ellsaesser, H.W., "Wind Variability," AWS Tech. Rept. No. 105-2, HQ Air Weather Service, Scott AFB, Ill., 1960.

Evans, J.V., "Ionospheric Movements Measured by Incoherent Scatter: A Review," *J. Atmos. Terr. Phys.*, **34:** 175, 1972.

Evans, J.V., "A Review of F Region Dynamics," *Rev. Geophys. Space Phys.*, **13:** 887, 1975.

Evans, J.V., "Incoherent Scatter Contributions to Studies of the Dynamics of the Lower Thermosphere," *Rev. Geophys. Space Phys.*, **16:** 195, 1978.

Fichtl, G.H., "Small-Scale Wind Shear Definition for Aerospace Vehicle Design," *J. Spacecraft Rockets*, **9:** 79, 1972.

Forbes, J.M., "Atmospheric Tides 1. Model Descriptions and Results for the Solar Diurnal Components," *J. Geophys. Res.*, **87:** 5222–5240, 1982a.

Forbes, J.M., "Atmospheric Tides 2. The Solar and Lunar Semidiurnal Components," *J. Geophys. Res.*, **87:** 5241–5252, 1982b.

Forbes, J.M., and H.B. Garrett, "Theoretical Studies of Atmospheric Tides," *Rev. Geophys. Space Phys.*, **17:** 1951–1981, 1979.

Fuller-Rowell, T.J., and D. Rees, "A Three-Dimensional Time-Dependent Global Model of the Thermosphere," *J. Atmos. Sci.*, **37:** 2545–2567, 1980.

Fuller-Rowell, T.J., and D. Rees, "A Three-Dimensional Time-Dependent Simulation of the Global Dynamical Response of the Thermosphere to a Geomagnetic Substorm," *J. Atmos. Terr. Phys.*, **43:** 701, 1981.

Gage, K.S., and B.B. Balsley, "Doppler Radar Probing of the Clear Atmosphere," *Bull. Amer. Meteor. Soc.*, **59:** 1074–1092, 1978.

Golden, J.H., "An Assessment of Windspeeds in Tornadoes," *Proc. Symp. on Tornadoes*, Institute for Disaster Research, Texas Tech. Univ., Lubbock, Texas, 5–42, 1976.

Green, J.L., K.S. Gage, and T.E. Van Zandt, "Atmospheric Measurements by VHF Pulsed Doppler Radar," *IEEE Trans. Geosci. Elec.*, **GE-17:** 262–280, 1979.

Gregg, W.R., "An Aerological Survey of the United States," *Monthly Weather Review*, Supplement No 20: 1922.

Hays, P.B., and R.G. Roble, "Direct Observations of Thermospheric Winds during Geomagnetic Storms," *J. Geophys. Res.*, **76:** 5316–5321, 1971.

Hays, P.B., J.W. Meriwether, and R.G. Roble, "Nighttime Thermospheric Winds at High Latitudes," *J. Geophys. Res.*, **84:** 1905–1913, 1979.

Hays, P.B., T.L. Killeen, and B.C. Kennedy, "The Fabry-Perot Interferometer on Dynamics Explorer," *Space Sci. Instr.*, **5:** 395–416, 1981.

Heppner, J.F., "Empirical Models of High-Latitude Electric Fields," *J. Geophys. Res.*, **82:** 1115, 1977.

Heppner, J.F., and M.L. Miller, "Thermospheric Winds at High Latitudes from Chemical Release Observations," *J. Geophys. Res.*, **87:** 1633–1647, 1982.

Hernandez, G., and R.G. Roble, "Direct Measurements of Nighttime Thermospheric Winds and Temperatures. 1. Seasonal Variations during Geomagnetic Quiet Periods," *J. Geophys. Res.*, **81:** 2065–2074, 1976a.

Izumi, Y., "Kansas 1968 Field Program," AFCRL-72-0041, AD739165, 1971.

Kantor, A.J., and A.E. Cole, "Wind Distributions and Interlevel Correlations, Surface to 60 Km," AFGL-TR-80-0242, ADA092670, 1980.

Kaufman, J.W. (ed.), "Terrestrial Environment (Climatic) Criteria Guidelines for Use in Aerospace Vehicle Design," NASA Tech. Memor. 78118, MSFC, 1977.

Kelley, M.C., T.S. Jorgensen, and I.S. Mikkelsen, "Thermospheric Wind Measurements in the Polar Region," *J. Atmos. Terr. Phys.*, **39:** 211–219, 1977.

Kelly, D.L., J.T. Schaeffer, R.P. McNulty, C.A. Doswell III, and R.F. Abbey Jr., "An Augmented Tornado Climatology," *Mon. Wea. Rev.*, **106:** 1172–1183, 1978.

Killeen, T.L., P.B. Hays, N.W. Spencer, and L.E. Wharton, "Neutral Winds in the Polar Thermosphere as Measured from Dynamics Explorer," *Geophys. Res. Lett.*, **9:** 957–960, 1982.

Knutson, J.R., D.C. Kayser, and W.E. Potter, "Mass Spectrometric Measurement of Thermospheric Wind," *J. Geophys. Res.*, **82:** 5253, 1977.

Lettau, H., *Atmosphärische Turbulenz*, Akademische Verlagsgesellschaft M.B.H., Leipzig, 114, 1939.

Lettau, H., "A Re-examination of the 'Leipsig Wind Profile' Considering Some Relations Between Wind and Turbulence in the Frictional Layer," *Tellus*, **2:** 125, 1950.

Lettau, H.H., and B. Davidson (eds.), *Exploring the Atmosphere's First Mile*, Pergamon Press, New York, 1957.

Marcos, F.A., and E.R. Swift, "Application of the Satellite Triaxial Accelerometer to Atmospheric Density and Wind Studies," AFGL-TR-82-0091, ADA120852, 1982.

Mayr, H.G., I. Harris, and N.W. Spencer, "Some Properties of Upper Atmosphere Dynamics," *Rev. Geophys. Space Phys.*, **16:** 539–566, 1978.

McLean, G.S., "The Jet Stream and Associated Turbulence," in *Winds for Aerospace Vehicle Design*, Sissenwine and Kasten, editors, AFCRL-62-273(I), 1962.

Meriwether, J.W., C.A. Tepley, F.A. Price, P.B. Hays, and L.L. Cogger, "Remote Ground-based Observations of Terrestrial Airglow Emissions and Thermospheric Dynamics at Calgary, Alberta," *Opt. Engr.*, 1982.

NOCD (Naval Oceanography Command, Detachment), *Guide to Standard Weather Summaries and Climatic Services*, NAVAIR 50-1C-534, Asheville, N.C., 1980.

Quegan, S., G.J. Bailey, R.J. Moffett, R.A. Heelis, T.J. Fuller-Rowell, D. Rees, and R.W. Spiro, "A Theoretical Study of the Distribution of Ionization in the High Latitude Ionosphere and the Plasmasphere: First Results on the Midlatitude Trough and the Light-Ion Trough," *J. Atmos. Terr. Phys.*, **44:**619, 1982.

Rees, D., T.J. Fuller-Rowell, and R.W. Smith, "Measurements of High-Latitude Thermospheric Winds by Rocket and Ground-Based Techniques and Their Interpretation using a Three-Dimensional Time-Dependent Dynamical Model," *Planet. Space Sci.*, **28:** 919, 1980.

Rees, D., T.J. Fuller-Rowell, R. Gordon, T.L. Killeen, P.B. Hays, L. Wharton, and N. Spencer, "A Comparison of Wind Observations of the Upper Thermosphere from the Dynamics Explorer Satellite with the Predictions of a global Time-Dependent Model," *Planet. Space Sci.*, **31:** 1299–1314, 1983.

Reiff, P.H., R.W. Spiro, and T.W. Hill, "Dependence of Polar Cap Potential Drop on Interplanetary Parameters," *J. Geophys. Res.* **85:** 7639–7648, 1981.

Rieger, E., "Neutral Air Motions Deduced from Barium Release Experiments – I. Vertical Winds," *J. Atmos. Terr. Phys.*, **36:** 1377, 1974.

Richmond, A.D., and S. Matsushita, "Thermospheric Response to a Magnetic Substorm," *J. Geophys. Res.*, **80:** 2839–2850, 1975.

Rishbeth, H., "Thermospheric Winds and the F-Region: A Review," *J. Atmos. Terr. Phys.*, **34:** 1–48, 1972.

Roble, R.G., "The Thermosphere," in *Physics of the Upper Atmosphere and Magnetosphere*, National Academy of Sciences, Washington, D.C., **68:** 1977.

Roble, R.G., "Dynamics of the Earth's Thermosphere," *Rev. Geophys. Space Phys.*, **21:** 217–233, 1983.

Roble, R.G., R.E. Dickinson, and E.C. Ridley, "Seasonal and Solar Cycle Variations of the Zonal Mean Circulation in the Thermosphere," *J. Geophys. Res.*, **82:**5493–5504, 1977.

Roble, R.G., R.E. Dickinson, and E.C. Ridley, "Global Circulation and Temperature Structure of Thermosphere and High-Latitude Plasma Convection," *J. Geophys. Res.*, **87:** 1599–1614, 1982.

Roble, R.G., R.E. Dickinson, E.C. Ridley, B.A. Emery, P.B. Hays, T.L. Killeen, and N.W. Spencer, "The High Latitude Circulation and Temperature Structure of the Thermosphere Near Solstice," *Planet. Space Sci.*, **31:** 1479–1499, 1983.

Rottger, J., "VHF Radar Observations of a Frontal Passage," *J. Appl. Meteorol.*, **18:** 85–91, 1979.

Scoggins, J.R., and W.W. Vaughan, "Cape Canaveral Wind and Shear Data (1 through 80 km) for Use in Vehicle Design and Performance Studies," *Tech. Note D-1274*, NASA, 1962.

Shellard, H.C., "The Estimation of Design Wind Speeds, Win Effects on Buildings and Structures," *Nat'l Phys. Lab. Symp., Teddington, England*, **16:** 30–51, 1965.

Sherlock, R.H., "Variation of Wind Velocity and Gusts with Height," Paper No. 2553, *Proc. Am. Soc. Civil Eng.*, **1:** 463–508, 1952.

Sherlock, R.H., "Gust Factors for the Design of Buildings," *Int. Assoc. for Bridge and Structural Engineering*, **8:** 207–235, 1947.

Sipler, D.P., and M.A. Biondi, "Midlatitude F-Region Netural Winds and Temperatures during the Geomagnetic Storm of March 26, 1976," *J. Geophys. Res.*, **84:** 37–40, 1979.

Sipler, D.P., B.B. Lukkala, and M.A. Biondi, "Fabry-Perot Determinations of Mid-Latitude F-Region Neutral Winds and Temperatures from 1975 to 1979," *Planet. Space Sci.*, **30:** 1025–1032, 1982.

Sissenwine, N., P. Tattelman, D.D. Grantham, and I.I. Gringorten, "Extreme Wind Speeds, Gustiness, and Variations with Height for MIL-STD-210B," AFCRL-TR-73-0560, AD774044, 1973.

Spencer, N.W., R.F. Theis, L.E. Wharton, and G.R. Carignan, "Local Vertical Motions and Kinetic Temperature from AE-C as Evidence for Aurora-Induced Gravity Waves," *Geophys. Res. Lett.*, **3:** 313, 1976.

Spencer, N.W., L.E. Wharton, H.B. Niemann, A.E. Hedin, G.R. Carignan, and J.C. Maurer, "The Dynamics Explorer Wind and Temperature Spectrometer," *Space Sci. Instr.*, **5:** 417–428, 1981.

Spencer, N.W., L.E. Wharton, G.R. Carignan, and J.C. Maurer, "Thermosphere Zonal Winds, Vertical Motions and Temperature as Measured from Dynamics Explorer," *Geophys. Res. Lett.*, **9:** 953–956, 1982.

Straus, J.M., "Dynamics of the Thermosphere at High Latitudes," *Rev. Geophys. Space Phys.*, **16:** 183–194, 1978.

Tattelman, P., "Surface Gustiness and Wind Speed Range as a Function of Time Interval and Mean Wind Speed," *J. Appl. Meteorol.*, **14:** 1271–1276, 1975.

Volland, H., "Models of Global Electric Fields Within the Magnetosphere," *Ann. Geophys.*, **31:** 154–173, 1975.

Volland, H., "Semiempirical Models of Magnetospheric Electric Fields," in *Quantitative Modeling of Magnetospheric Processes*, Geophys. Monogr. Ser., Vol 21, edited by W.P. Olson, 261–280, American Geophysical Union, Washington, D.C., 1979a.

Volland, H., "Magnetospheric Electric Fields and Currents and Their Influence on Large Scale Thermospheric Circulation and Composition," *J. Atmos. Terr. Phys.*, **41:** 853–866, 1979b.

Wu, S.T., S. Matsushita, and L.L. DeVries, "An Analysis of the Upper Atmospheric Wind Observed by Logacs," *Planet. Space Sci.*, **22:** 1036, 1974

Chapter 18

OPTICAL AND INFRARED PROPERTIES OF THE ATMOSPHERE

R. W. Fenn
S. A. Clough
W. O. Gallery
R. E. Good
F. X. Kneizys
J. D. Mill Lt. Col. USAF
L. S. Rothman
E. P. Shettle
F. E. Volz

This chapter deals with the absorption, scattering, emission and refractive properties of the natural atmosphere from the ultraviolet through the microwave region. Nonlinear propagation processes (such as thermal blooming) are not covered.

The natural atmosphere includes the range of those average conditions which are not directly affected by limited local processes or sources, such as gases or particulates from a fire or industrial plant, a dust cloud from vehicular traffic etc. The natural atmosphere does include the effects of molecules, aerosol (haze) particles, clouds, fogs, rain, and snow. The vertical extent of the atmosphere is from the surface up to 100 km altitude. A brief discussion is included on the earth's surface reflectance properties since they also affect the atmospheric optical properties.

The chapter includes descriptions of the basic physical relationships of optical propagation as well as the current state of experimental knowledge and modeling of the atmospheric propagation medium. An extensive annotated bibliography covering much of the material of this chapter was compiled by Wiscombe [1983] in Reviews of Geophysics and Space Physics.

18.1 ATMOSPHERIC GASES

18.1.1 Atmospheric Molecules, Models of the Atmospheric Composition

The propagation of optical and infrared radiation through the atmosphere depends on the composition and variability of the atmosphere. Systematic variations in the density, pressure, temperature, water vapor, and ozone as a function of latitude and season have been known for many years. The development of model atmospheres goes back to the 1920s. Early Standard Atmosphere Models represented mean conditions of temperature, density, and pressure as a function of altitude. During the 1950s and 1960s extensions and supplemental atmospheres were derived that describe in addition the seasonal and latitudinal variability of atmospheric structure (see Chapters 14 and 15).

For the purpose of modeling the optical properties of the atmosphere, specifically as described in Sections 18.4.1 (FASCODE) and 18.4.2 (LOWTRAN), the 1962 U.S. Standard Atmosphere and supplemental model atmospheres were used as models. Although, updated Standard Atmospheres have been published as recently as 1976, differences are limited to altitudes above 50 km where the adopted radiance models were not well characterized.

Pressure, temperature, water vapor density, and ozone density profiles as a function of altitude are provided to describe these molecular atmospheric models. These profiles, taken from McClatchey et al. [1972], correspond to the *U.S. Standard Atmosphere, 1962* and five supplementary models [Cole et al., 1965 or *U.S. Standard Atmosphere Supplements 1966*]: Tropical (15 N), Midlatitude Summer (45 N, July), Midlatitude Winter (45 N, January), Subarctic Summer (60 N, July), and Subarctic Winter (60 N, January). These profiles are identical to the ones used in the LOWTRAN code [Kneizys et al., 1980]. The water vapor and ozone latitude profiles added to the 1962 U.S. Standard Atmosphere by McClatchey et al. [1972] were obtained from Sissenwine et al. [1968] and Hering and Borden [1964] respectively, and correspond to mean annual values. The water vapor densities for the 1962 U.S. Standard Atmosphere correspond to relative humidities of approximately 50% for altitudes up to 10 km, whereas the relative humidity values for the other supplementary models tend to decrease with altitude from approximately 80% at sea level to approximately 30% at 10 km altitude. Above 12 km, the water vapor density profiles depicted here have been replaced by

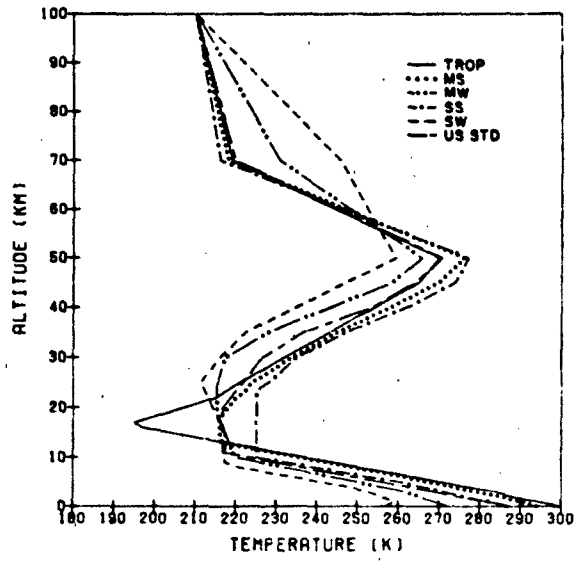


Figure 18-1. Temperature vs altitude for the six model atmospheres: tropical (TROP), midlatitude summer (MS), midlatitude winter (MW), subarctic summer (SS), subarctic winter (SW), and US standard (US STD).

more recent measurements [Remsberg et al., 1984] and are approximately 5 parts per million by volume(ppmv.) For all model atmospheres, the gases CO_2, N_2O, CO, CH_4, N_2, and O_2 are considered uniformly mixed, with mixing ratios of 330, 0.28, 0.075, 1.6, 7.905×10^5, and 2.095×10^5 ppmv, respectively.

The temperature profiles for the six model atmospheres as a function of altitude are shown in Figure 18-1. The pressure profiles are given in Figure 18-2. Figures 18-3 a and b show the water vapor density versus altitude from 0 to 100 km, and an expanded profile from 0 to 30 km. Figures

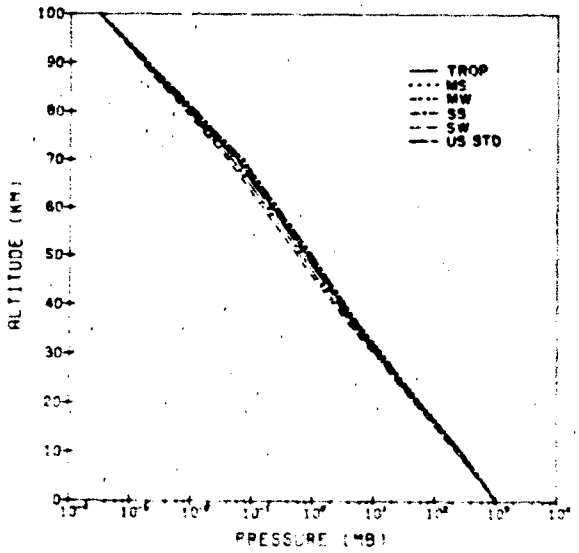
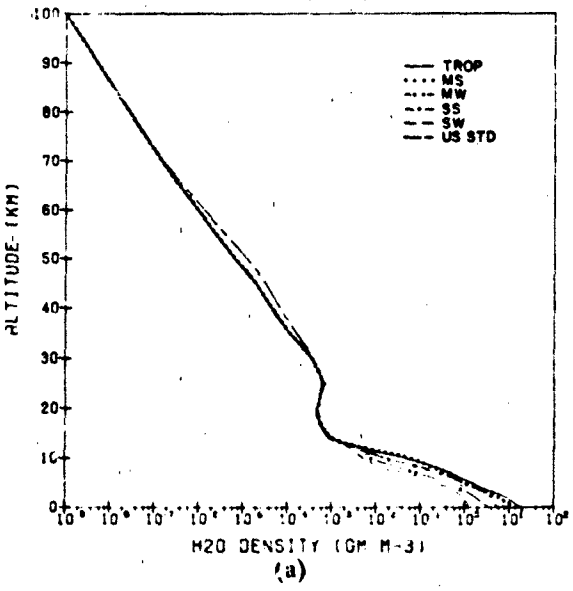


Figure 18-2. Pressure vs altitude for the six model atmospheres.



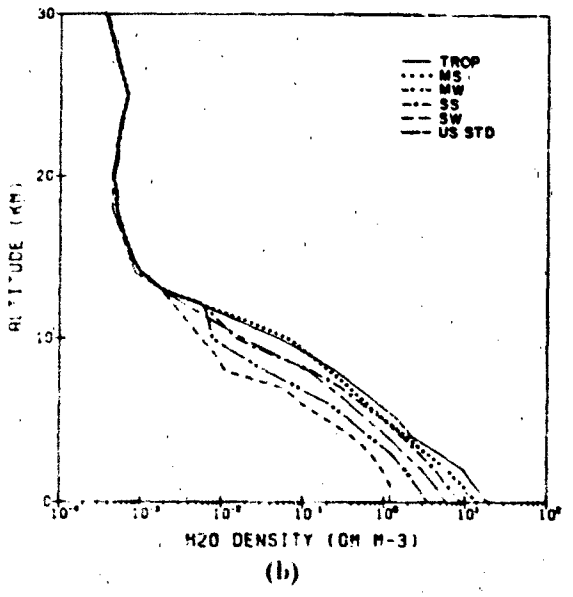


Figure 18-3. (a) Water vapor density profiles vs altitude for the six model atmospheres.
(b) Water vapor density profiles vs altitude for the six model atmospheres with the region from 0 to 30 km expanded.

18-4 a and b and Figures 18-5 a and b show similar profiles for ozone and relative air density.

Measurements made from balloon flights [Murcray et al., 1968] have shown the existence of nitric acid in the earth's atmosphere. Although nitric acid (HNO_3) is of only minor importance in atmospheric transmittance calculations, it has been shown to be a significant source of stratospheric emission, particularly in the atmospheric window region from 10 to 12 μm. The concentration of atmospheric nitric acid varies with altitude and also appears to depend on

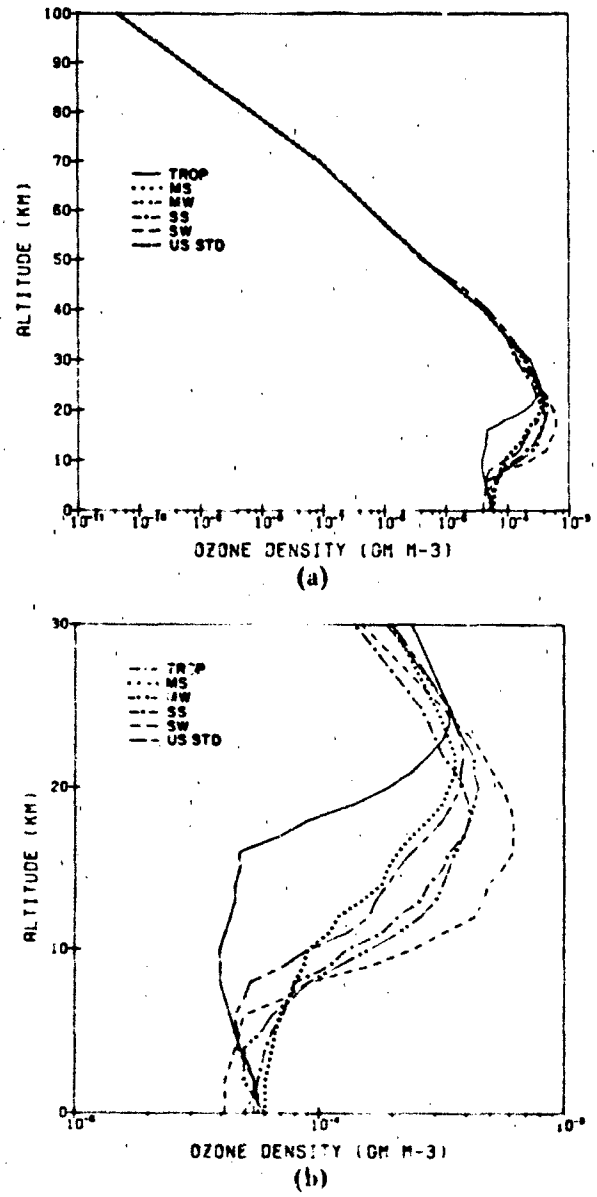


Figure 18-4 (a) Ozone density profiles vs altitude for the six model atmospheres.
(b) Ozone density profiles vs altitude for the six model atmospheres with the region from 0 to 30 km expanded.

latitude and season. Figure 18-6 shows the volume mixing ratio profile of atmospheric nitric acid as a function of altitude from the measurements of Evans et al. [1975]. This profile has been chosen to represent a mean nitric acid profile for the six model atmospheres.

Since all of the above profiles were constructed, knowledge of the state of the atmosphere has increased. This is particularly true regarding the stratosphere and the concentration of minor constituents. For example, stratospheric water vapor concentrations for the six profiles given in Figure 18-3a and b are now known to be too high. The model atmospheres can still be considered representative of their respective conditions up to about 50 km for temperature, 30 km for ozone densities, and the tropopause (approximately from 8 km in the article to 15 km in the tropics) for water vapor. These models are still used for cases dominated

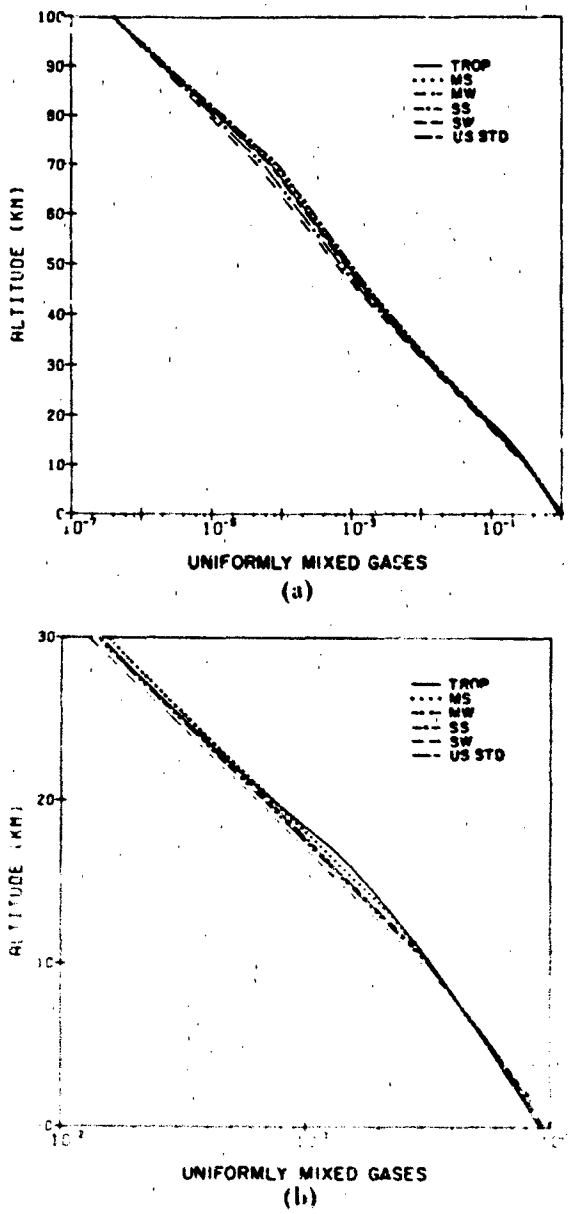


Figure 18-5. (a) Profile of $(P/P_s)(T_s/T)$, the relative air density, vs altitude for the six model atmospheres. The density of the uniformly mixed gases is proportional to this quantity. $P_s = 1013$ mb and $T_s = 273$ K.
(b) Profile of $(P/P_s)(T_s/T)$, the relative air density, vs altitude for the six model atmospheres with the regions 0 to 30 km expanded.

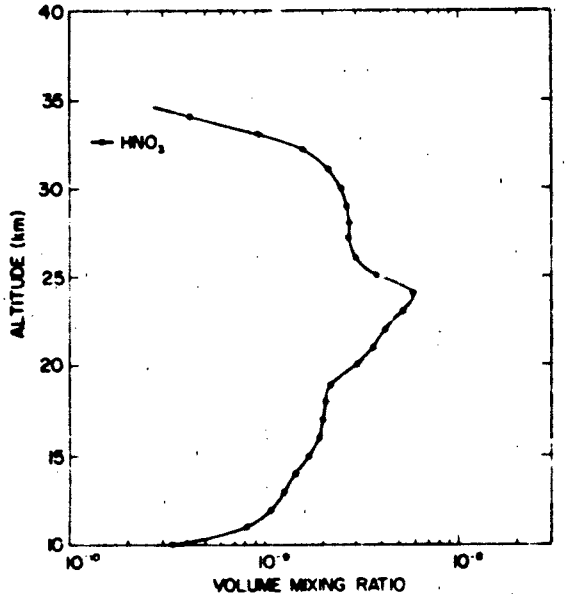


Figure 18.6 Volume mixing ratio profile for nitric acid vs altitude from the measurements of Evans et al. [1975]. This single profile is used with all of the six model atmospheres

by conditions in the troposphere. For cases dominated by stratospheric conditions or where the distribution of minor constituents is significant, modelers should supply their own profiles.

There are several recent sources for profiles of temperature and minor constituent density. The *U.S. Standard Atmosphere 1976* updates the 1962 Standard for temperature above 50 km and provides revised estimates for the surface concentrations of what was termed previously the "uniformly mixed gases". The new values for the volume mixing ratios of CO_2, N_2O, CO, and CH_4 are 330, 0.27, 0.19 and 1.50 ppmv. Actually the concentrations of these gases do show significant variations from these values, particularly with altitude in the stratosphere. Cole and Kantor [1978] provide sets of monthly mean temperature profiles up to 90 km at 15° intervals between the equator and the pole. Along with statistics on the variability of these profiles, they also provide models that portray longitudinal variations in monthly mean values of temperature during winter months and the vertical variation that occurs during stratospheric warming and cooling events in the winter arctic and subarctic. Houghton [1977] provides seasonal profile temperatures at 10°N, 40°N and 70°N up to 105 km plus the original references for the data.

For profiles of the minor constituents including ozone and stratospheric water vapor, WMO [1982] provides an up-to-date and exhaustive source. See also Chapter 21. Much of the profile data from this source plus some more recent measurements have been compiled as annual averages in 2 km steps in Smith [1982].

18.1.2 Molecular Absorption and Spectroscopic Parameters

18.1.2.1 The Role of Molecular Absorption. Molecular absorption, that is, attenuation of electromagnetic radiation by the mechanism of molecules absorbing quanta of energy to alter vibrational and rotational states, is of prime significance in propagation in the atmosphere. Although the terrestrial atmosphere consists primarily of nitrogen, oxygen, and argon, absorption in the infrared region is dominated by species with very low concentrations but very active vibration-rotation bands such as water vapor, carbon dioxide, ozone, and nitrous oxide. Other active species also play an important role depending on the region of the electromagnetic spectrum being investigated or the layer of the atmosphere being probed.

In order to calculate transmittance due to a given spectral line in the atmosphere, it is useful to describe the absorption coefficient as a function of frequency for each line. Assuming superposition of the contribution from individual spectral transitions, the absorption coefficient $k(\nu)$ as a function of wavenumber ν (cm^{-1}) can be considered in general to be a product of the radiation field, the molecular system, and the coupling between them. This can be expressed [Clough et al., 1981] as

$$\bar{k}(\nu) = \nu \tanh(hc\nu/2kT) \times \sum_i p(m,) \bar{S}_i(T) [f(\nu, \nu_i) + f(\nu, -\nu_i)], \quad (18.1)$$

where $p(m,)$(molec/cm^3) is the density for the molecular species m, with transition wavenumber ν_i responsible for the i'th transition. $\bar{S}_i(T)$ (molec/cm^2)$^{-1}$ is the intensity of the transition at temperature T(K) appropriate to the line shape $f(\nu,\nu_i)$ (1/cm^{-1}). For most molecular species, the line intensity is a function of the expectation value of an expansion of the electric dipole moment. For some species, for example nitrogen and oxygen [Rothman and Goldman, 1981], the much weaker electric quadrupole and magnetic dipole moments are the means by which the quantum states change. In terms of intensity definition used in the AFGL line compilations, $S_i(T)$, discussed in the next section, we have the relationship

$$S_i(T) = \nu_i \tanh(hc\nu/2kT)\, \bar{S}(T) = \nu_i \frac{1 - e^{-hc\nu_i/kT}}{1 + e^{-hc\nu_i/kT}} \bar{S}_i(T). \quad (18.2)$$

The line shape function $f(\nu,\nu_i)$ is dependent on molecular species, broadening density and temperature. For collisional broadening in the impact limit, the form factor is given by the Lorentz line shape

$$f(\nu_i) = \frac{1}{\pi} \frac{\alpha_i}{(\nu - \nu_i)^2 + (\alpha_i)^2} \quad (18.3)$$

where α_i (cm^{-1}) is the collision broadened halfwidth at half maximum (HWHM). The monochromatic transmittance through a layer of homogeneous medium of thickness ℓ is obtained by the Lambert-Beers law

$$\tau_\nu = e^{-k(\nu)\ell} \quad (18.4)$$

where the exponent is called the optical thickness.

The line intensity is temperature dependent through the Boltzmann factor and the internal partition function

$$S_i(T) = d_i(1 - e^{-hc\nu_i/kT}) \frac{e^{-hcE_i''/kT}}{Q(T)} \quad (18.5)$$

where Q(T) is the internal partition sum, E_i'' (cm^{-1}) the energy of the lowerstate of the transition, and d_i is the transition strength. The collisional halfwidth depends on the specific molecule and can be expressed in terms of density and temperature as

$$\alpha_i(p,T) = \alpha^o \left(\frac{p}{p_o}\right) \left(\frac{T}{T_o}\right)^{X_T} \quad (18.6)$$

(the exponent X_T is 1/2 with the usual classical theory assumption of temperature-independent collision diameters). The ratio of partition sums does not require the knowledge of any further parameters aside from the fundamental frequencies of the vibrational modes. Thus, at least for local thermodynamic equilibrium conditions, calculations of molecular absorption (transmission and emission) at high resolution basically require the knowledge of four essential parameters (1) the resonant frequency of the transition ν_i, (2) the intensity per absorbing molecule $S_i(T_o)$, (3) the Lorentz line width parameter α_i^o, and (4) the energy of the lower state E_i. The question of line shape will be discussed in more detail in Section 18.4.1. The effects of molecular (Rayleigh) scattering and of aerosol scattering and absorption will be discussed in Sections 18.1.4 and 18.2.1.5.

18.1.2.2 The Atmospheric Absorption Line Parameter Compilation. From the preceding discussion it can be seen that a compilation of spectroscopic data on individual molecular transitions would facilitate the so-called line-by-line calculations of spectra. These high resolution calculations, that is, calculations capable of reproducing spectra obtained by such instruments as Michelson interferometers, tunable diode lasers, and heterodyne spectrometers have been developed along with the advances in computer technology. In the 1960s a program was initiated at the Air Force Cambridge Research Laboratories (now AFGL) to create such a compilation and the associated codes to produce synthetic spectra [McClatchey et al., 1973]. The scope originally aimed at providing a complete set of data for all vibration-rotation lines of naturally occurring molecules of significance in the terrestrial atmosphere. Since that time these data have been expanded in terms of the frequency covered, addition of trace species, and inclusion of weaker transitions [Rothman, 1981; Rothman et al., 1981, 1983a,b]. Historically, these data have been divided into two atlases, a Main Compilation comprising transitions of the most active terrestrial absorbers H_2O, CO_2, O_3, N_2O, CO, CH_4, and O_2, and a Trace Gas Compilation. The species presently covered by these data bases are summarized in Table 18-1. The isotopes are abbreviated by the code 161 ≡ $H_2{}^{16}O$, 162 ≡ $HD^{16}O$, etc. The first seven species (incorporated in the Main Atlas) include all transitions contributing to at least 10% absorption over a maximum atmospheric path. This has been determined from standard atmospheric models of the uniformly mixed gases and the two nonuniformly mixed gases, water vapor and ozone. The standard temperature chosen was 296 k. The gases in the Trace Gas Compilation were made available for a variety of problems including stratospheric probing, pollution monitoring, temperature profile retrieval, and laboratory studies. The Main Atlas covers the frequency range 0 to 20 000 cm^{-1}. The low frequency or millimeter and submillimeter limit represents primarily pure rotational transitions while the high frequency domain represents for the most part excited vibration-rotation transitions of water vapor. The Trace Gas Compilation covers the range 0 to 10 000 cm^{-1}. In addition to the four basic parameters, there are included for each transition the unique quantum identification, the molecule and isotope, and a reference code.

The goal of generating these parameters has been to create a theoretically self-consistent set. Ideally, the performance of quantum mechanical calculations from the development of mathematical models of the molecular system would provide the necessary parameters. This has been satisfactorily accomplished in many cases, the hydrogen halides being a prime example. In cases involving complex resonances or anomalous behavior, theory has lagged behind experiment and it has been necessary to utilize available observed parameters. The latter method does not directly allow for interpolating the unobserved or weaker transitions that might play a role in long path or high temperature observations. Nevertheless, a fair amount of success has been achieved deriving the tabulated parameters ν_i, S_i, α_i, E_i and their unique quantum identifications within the framework of general theories of molecular spectroscopy applied to experimental observations. Details of the basic theory, specific data, and treatment of exceptional cases can be found in references contained in the articles on the compilations [Rothman, 1981; Rothman et al., 1981; 1983a; 1983b]. As of this writing, the compilations, representing the culmination of international efforts in molecular spectroscopy, contain about a third of a million transitions. The maintenance, updating, and improvement of these data is ongoing.

Table 18-1. Molecular species on atmospheric absorption line parameter atlases.

Molecule	Isotope	Relative Natural Abundance	Molecule	Isotope	Relative Natural Abundance
H_2O (1)	161	0.9973	HNO_3 (12)	146	0.9891
	181	0.0020			
	171	0.0004	OH (13)	61	0.9975
	162	0.0003		81	0.0020
				62	0.00015
CO_2 (2)	626	0.9842			
	636	0.0111	HF (14)	19	0.99985
	628	0.0040			
	627	0.0008	HCl (15)	15	0.7576
	638	0.00044		17	0.2423
	637	0.000009			
	828	0.000004	HBr (16)	19	0.5068
	728	0.000002		11	0.4930
O_3 (3)	666	0.9928	HI (17)	17	0.99985
	668	0.0040			
	686	0.0020	ClO (18)	56	0.7559
				76	0.2417
N_2O (4)	446	0.9904			
	456	0.0036	OCS (19)	622	0.937
	546	0.0036		624	0.0416
	448	0.0020		632	0.0105
	447	0.0004		822	0.0019
CO (5)	26	0.9865	H_2CO (20)	126	0.9862
	36	0.011		136	0.0111
	28	0.0020		128	0.0020
	27	0.0004			
			HOCl (21)	165	0.7558
CH_4 (6)	211	0.9883		167	0.2417
	311	0.0111			
	212	0.00059	N_2 (22)	44	0.9928
O_2 (7)	66	0.9952	HCN (23)	124	0.9852
	68	0.0040			
	67	0.0008	CH_3Cl (24)	215	0.7490
				217	0.2395
NO (8)	46	0.9940			
			H_2O_2 (25)	1661	0.9949
SO_2 (9)	626	0.9454			
	646	0.0420	C_2H_2 (26)	1221	0.9776
NO_2 (10)	646	0.9916	C_2H_6 (27)	1221	0.9770
NH_3 (11)	4111	0.9960	PH_3 (28)	1111	0.99955
	5111	0.0036			

18.1.3 Index of Refraction

The theory of molecular optics [Born and Wolf, 1975] derives the index of refraction for gases from the scattering properties of molecules (Lorentz-Lorenz formula). In the atmosphere the index of refraction m is very close to 1 so that it is convenient to define the refractive modulus N as

$$N = (m - 1) \times 10^6. \qquad (18.7)$$

N is a function of both wavelength and density and is different for dry air and water vapor. For optical wavelengths greater than 0.23 μm and for infrared wavelengths, N for atmospheric conditions is given by the following formula from Edlen [1966]:

$$N = \left[a_0 + \frac{a_1}{1 - (\nu/b_1)^2} + \frac{a_2}{1 - (\nu/b_2)^2} \right] \frac{P}{P_0} \frac{(T_0 + 15.0)}{T} - \left[c_0 - (\nu/c_1)^2 \right] \frac{P_w}{P_0} \qquad (18.8)$$

where P is the total air pressure in mb, T is the temperature in K, $P_0 = 1013.25$ mb, $T_0 = 273.15$, P_w is the partial pressure of water vapor in mb, and $\nu = 10^4/\lambda$ is the frequency in cm^{-1} for the wavelength λ in micrometers.

$$a_0 = 83.42$$
$$a_1 = 185.08$$
$$a_2 = 4.11$$
$$b_1 = 1.140 \times 10^5$$
$$b_2 = 6.24 \times 10^4$$
$$c_0 = 43.49$$
$$c_1 = 1.70 \times 10^4$$

The formula is valid from 0.23 μm to the infrared. For millimeter and microwaves, the refractive modulus is much more complicated and includes a strong dependence on water vapor density. For reference, Table 18-2 lists the index of refraction for dry air at 1 atmosphere and 288 K for various wavelengths.

Table 18-2. Refractive modulus of dry air at 1 atmosphere and 288 K.

Wavelength (μm)	Wavenumber (cm^{-1})	$N = (m-1) \times 10^6$
0.23	4.35×10^4	308.0
0.25	4.00×10^4	301.5
0.30	3.33×10^4	291.6
0.40	2.50×10^4	282.7
0.50	2.00×10^4	279.0
0.75	1.33×10^4	275.4
1.0	10000	274.2
2.0	5000	272.9
4.0	2500	272.7
10.0	1000	272.6

18.1.4 Molecular Rayleigh Scattering

When air molecules are subjected to an oscillating electric field such as in a light wave, the molecule temporarily absorbs and immediately re-emits this radiation as a point source. This process is known as scattering. It occurs on particles of any size. If the size of a particle is very small compared to the wavelength of the incident light, certain simplifying conditions exist that also simplify the theoretical description of the scattering process. This type of scattering of light, which causes the blue sky light, was first described by Lord Rayleigh in 1871 [McCartney, 1976].

The total volumetric scattering coefficient $k_s\,\lambda$ for molecules is defined by

$$dI_\lambda = -I_\lambda \cdot k_s(\lambda) \cdot dx\,, \qquad (18.9)$$

with

$$k_s(\lambda) = \frac{24\,\pi^3}{N'\lambda^4} \left(\frac{m^2 - 1}{m^2 + 2} \right)^2 \qquad (18.10)$$

where m is the refractive index of the gas (such as air), N' the concentration of molecules per unit volume, I_λ the intensity at wavelength λ, and dx the path element. This expression is often simplified by taking advantage of the refractive index of air being nearly 1, and writing Equation (18.10) as

$$k_s(\lambda) \approx \frac{8\pi^3}{3N'\lambda^4} (m^2 - 1)^2, \qquad (18.11a)$$

or

$$k_s(\lambda) \approx \frac{32\pi^3}{3N'\lambda^4} (m - 1)^2\,. \qquad (18.11b)$$

For standard temperature and pressure, using Equation (18.11a) introduces an error on the order of 0.04% in the visible and using Equation (18.11b) introduces an error of about 0.025%. Since air molecules are not completely isotropic, a small correction factor for anistoropy [see, for example, Chandrasekhar, 1950; Kerker, 1969] must be ap-

Table 18-3. Depolarization factor Δ of atmospheric gases for incident unpolarized light [Young, 1980].

Gas	Rowell et al. [1971]	Alms et al. [1975]	Baas and van den Hout [1979]	Young [1980]
N_2	0.0214 and 0.0242	0.0210 ± 0.0004	0.0200 ± 0.006	—
O_2	0.0566 and 0.0587	—	0.0580 ± 0.002	—
CO_2	0.073 and 0.075	0.0797 ± 0.0010	0.077 ± 0.002	—
Dry Air	—	—	—	0.0279

plied to the molecular scattering in Equations (18.10) or (18.11). With this correction, Equation (18.10) becomes

$$k_s(\lambda) = \frac{24\pi^3}{N'\lambda^4}\left(\frac{m^2 - 1}{n^2 + 2}\right)^2 \frac{6 + 3\Delta}{6 - 7\Delta}, \quad (18.12)$$

where the depolarization Δ is the ratio of the two polarized intensities i_2/i_1 at a 90° scattering angle (see Section 18.2.1.5). Table 18-3 gives depolarization values for atmospheric gases [Young, 1980]. It is often convenient to use the scattering cross section per molecule

$$\sigma_s(\lambda) = \frac{k_s(\lambda)}{N'} = \frac{24\pi^3}{N'^2\lambda^4}\left(\frac{m^2 - 1}{m^2 + 2}\right)^2 \frac{6 + 3\Delta}{6 - 7\Delta} \quad (18.13)$$

which has the advantage over the scattering coefficient in that it is independent of air density (or temperature and pressure). Values of σ_s [Equation (18.13)] and k_s [Equation (18.12)] are presented in Table 18-4 for wavelengths between 0.25 and 4.0 μm. The largest uncertainty in the tabulated values is due to the uncertainty in the values of the depolarization factor Δ used. Young [1980] indicates an error of a "few percent" in his value of Δ for air (Table 18-3). This corresponds to an uncertainty of ±0.1 to 0.2% in $\sigma_s(\lambda)$ or $k_s(\lambda)$.

The angular distribution of the light scattered by atmospheric molecules is given by the Rayleigh scattering phase function:

$$P(\theta) = \frac{3}{16\pi}\frac{2}{(2 + \Delta)}\left[(1 + \Delta) + (1 - \Delta)\cos^2\theta\right]. \quad (18.14a)$$

Neglecting the correction for depolarization, this simplifies to

Table 18-4. Rayleigh scattering cross section $\sigma_s(\lambda)$ and Rayleigh scattering coefficient $k_s(\lambda)$.

Wavelength (μm)	$\sigma_s(\lambda)$ (cm^2/molec.)	$k_s(\lambda)$ (for t = 273.15 K and p = 1013.25mb)
0.25	1.243E-25	3.339E-1
0.30	5.605E-26	1.506E-1
0.35	2.913E-26	7.829E-2
0.40	1.668E-26	4.482E-2
0.45	1.025E-26	2.754E-2
0.50	6.650E-27	1.787E-2
0.55	4.505E-27	1.211E-2
0.60	3.161E-27	8.496E-3
0.65	2.284E-27	6.139E-3
0.70	1.692E-27	4.547E-3
0.80	9.864E-28	2.651E-3
0.90	6.135E-28	1.649E-3
1.0	4.014E-28	1.079E-3
1.2	1.929E-28	5.184E-4
1.4	1.039E-28	2.793E-4
1.6	6.083E-29	1.635E-4
1.8	3.794E-29	1.020E-4
2.0	2.488E-29	6.695E-5
2.5	1.018E-29	2.736E-5
3.0	4.906E-30	1.318E-5
4.0	1.552E-30	4.169E-6

$$P(\theta) = \frac{3}{16\pi}\left[1 + \cos^2\theta\right], \qquad (18.14b)$$

which is often used for the Rayleigh phase function. The phase function gives the probability distribution for the scattered light, so that $P(\theta)\,d\Omega$ is the fraction of the scattered radiation that enters a solid angle $d\Omega$ about the scattering angle θ.

18.2 AEROSOLS, CLOUD, AND PRECIPITATION PARTICLES

Propagation of electromagnetic radiation through the atmosphere at optical/infrared frequencies is affected by absorption and scattering from air molecules and particulate matter (haze, dust, fog, and cloud droplets) suspended in the air. Scattering and absorption by haze particles or aerosols becomes the dominant factor in the boundary layer near the earth's surface, especially in the visible, and under low visibility conditions at all wavelengths. Atmospheric aerosol particles in the atmosphere vary greatly in their concentration, size, and composition, and consequently in their effects on optical and infrared radiation.

18.2.1 Aerosols, Geographic, and Temporal Variations

18.2.1.1 Composition, Sources and Sinks, Refractive Index, Particle Shapes. Figure 18-7 summarizes the general characteristics of atmospheric aerosols. Aerosols in the boundary layer of 1–2 km have the greatest variability. These aerosols consist of a variety of natural and manmade chemical compounds, inorganic as well as organic. Particles are transported into the atmosphere from their sources at or near the surface, or they may be formed in the atmosphere by chemical reactions from gaseous components, often with the influence of solar radiation through photochemical processes. Since more than two-thirds of the earth's surface is covered by oceans, the maritime aerosol component, which consists largely of sea salt particles from the sea water, forms the most uniquely identifiable aerosol. Over land areas, soil particles and dust are an important component. Organic particles from vegetation sources are also an important aerosol component. Table 18-5 gives a breakdown of the global atmospheric aerosol composition [SMIC, 1971].

In the troposphere above the boundary layer the distribution and composition of aerosol particles becomes less dependent on the geography and on the variability of sources near the surface. Experimental data [Whitby and Cantrell, 1976 and Whitby, 1973] point toward an aerosol which is composed of several modes, each having a different origin and history. The most clearly identifiable two components are submicron size particles formed from gaseous components and are still going through processes of coagulation and agglomeration; this component is called the accumulation mode. The second mode is made up of larger particles approximately 1 μm diameter and larger which are produced by mechanical processes such as resuspension from soil or sea spray; this component is called the coarse particle mode.

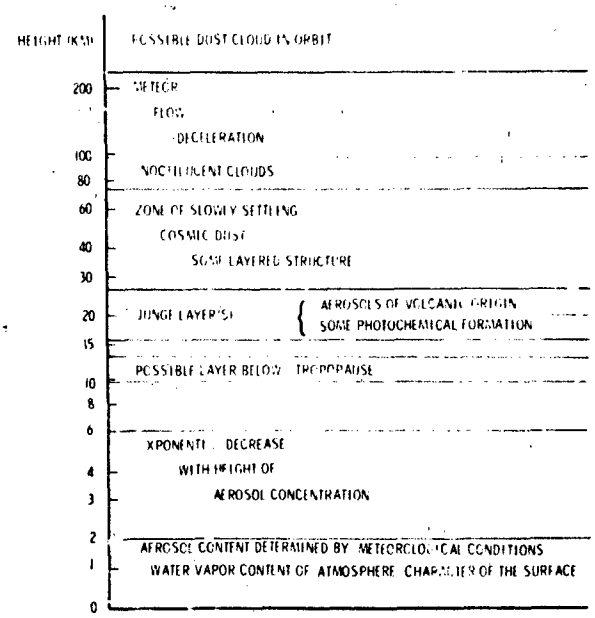


Figure 18-7. Characteristics of atmospheric aerosols.

Table 18-5. Estimates of particles smaller than 20 μm radius emitted into or formed in the atmosphere (10^6 metric tons/year) [SMIC, 1971].

Natural	
Soil and rock debris*	100–500
Forest fires and slash-burning debris*	3–150
Sea salt	300
Volcanic debris	25–150
Particles formed from gaseous emissions:	
Sulfate from H_2S	130–200
Ammonium salts from NH_3	80–270
Nitrate from NO_x	60–430
Hydrocarbons from plant exudations	75–200
Subtotal	773–2200
Man-made	
Particles (direct emissions)	10–90
Particles formed from gaseous emissions:	
Sulfate from SO_2	130–200
Nitrate from NO_x	30–35
Hydrocarbons	15–90
Subtotal	185–415
Total	958–2615

*Includes unknown amounts of indirect manmade contributions.

In the stratospheric region from 10–30 km, measurements have shown a background aerosol that has a rather uniform global distribution. This aerosol is considered to be mostly composed of sulfate particles formed by photochemical reactions. The stratospheric aerosol background can occasionally be increased by factors of 100 or more due to the injection of dust and SO_2 from massive volcanic eruptions. Once these volcanic particles and gases have been injected into the stratosphere they are spread out over large portions of the globe by the stratospheric circulation. While the dust particles may settle out within a few months, the SO_2 is converted to sulfuric acid through photo-chemical reactions. This enhanced layer of approximately 75% sulfuric acid droplets may remain for one to two years. For a complete review of the stratospheric aerosols see Turco et al. [1982].

Only a very small portion of the total aerosol content of the atmosphere exists above 30 km. However, when the effects of the lower atmosphere are small, such as in satellite observations of the limb of the earth's atmosphere or determining the ozone distribution by inverting measurements of backscattered ultraviolet from the atmosphere, the effects of the aerosols in the upper atmosphere can be significant.

Newkirk and Eddy [1964] and later Rosen [1969] concluded that the major component of the normal upper atmospheric aerosols is meteoric dust. Meteoric or cometary dust also form layers occasionally observed in the upper atmosphere. The optical effects of small particles are determined largely by their refractive index.

Aerosol samples from all over the world have been collected and analyzed for their optical and infrared refractive index properties. Based on such measurements, the aerosol in rural, non urbon areas can be assumed to be composed of a mixture of water soluble substances (ammonium and calcium sulfates and organic compounds) and dust-like aerosols [Volz, 1972 a, b]. In urban areas the rural aerosol gets modified by the addition of aerosols from combustion products and industrial sources. There is no experimental evidence for any systematic change in aerosol refractive index as a function of altitude through the troposphere. The refractive index of the volcanic components in the stratospheric aerosols was measured by Volz [1973] on volcanic dust samples. The refractive index for meteoric dust was determined by Shettle and Volz [1976] based on the composition of meteorites falling on the earth.

Although the aerosol particle size (see Section 18.2.1.2) is the most important parameter determining small particle scattering properties, the effect of particle shape cannot be ignored; it becomes significant especially for particles where size is large compared to the wavelength of the incident radiation. However, because of poor understanding of these shape effects, essentially all applications-oriented aerosol models assume spherical particle shapes. For more information on nonspherical particle scattering properties see Schuerman [1980].

18.2.1.2 Size Distributions. Over the past 20 to 30 years, much emphasis has been put on measurements of the aerosol particle size distribution, realizing that natural aerosols were far from being monodisperse. Junge [1963] suggested that most aerosol size distributions over the radius range from 0.1 to 10 μm, and even larger sizes, could be described by an inverse power law $dN/d \log r = cr^{-\gamma}$ with c being a constant and γ being between 2 and 4. This means that the number of aerosol particles decreases rapidly with increasing particle size. It has already been mentioned that more detailed measurements point towards a multimodal distribution. This becomes especially noticeable if the distribution is presented as particle volume distribution dv/dr rather than dN/dr. One can assume that these size distributions are the result of a dynamic equilibrium between the various processes which act on the aerosol population. However the interaction of these various processes is complex and no general theory for aerosol formation processes exists at this time.

18.2.1.3 Vertical Profiles. The change in aerosol properties and distribution with altitude is very closely tied to the vertical structure of the atmosphere, and in the troposphere especially to the "weather" processes. In general the number of aerosol particles decreases much more rapidly with altitude then the molecular air density; however, wide variations in the vertical profile do occur. Within the boundary layer of the atmosphere, aerosol vertical mixing is strongly influenced by the temperature profile.

18.2.1.4 Models of the Aerosol Properties. There are many scientific and technical reasons why it is necessary to develop models for atmospheric aerosol and cloud particles. They are needed to make estimates of the transmittance, angular light scattering distribution, contrast reduction, sky radiance, or other atmospheric optical properties or effects (see Section 18.4).

Models for the optical properties of aerosols have been developed at AFGL and elsewhere. [Elterman, 1964, 1968, 1970; Ivlev, 1967; Deirmendjian, 1964, 1969; McClatchey et al., 1970 and 1972; Shettle and Fenn, 1976; Toon and Pollack, 1976; Hanel and Bullrich, 1978; and Nilsson, 1979]. This chapter describes aerosol models and their optical properties for the lower and upper atmosphere [Shettle and Fenn 1976, 1979]. The models presented below are based on a review of the available experimental data on the nature of aerosols, their sizes, distribution, and variability. However, it must be emphasized that these models represent only a simple, generalized version of typical conditions. It is not practical to include all the details of natural particle distributions nor are existing experimental data sufficient to describe the frequency of occurrence of the different conditions. While these particulate models were developed to be as representative as possible of different atmospheric conditions, the following point should be kept in mind when

Table 18-6. Characteristics of the aerosol models as a function of relative humidity

Aerosol Model	Size Distribution			Type
	N_i	r_i	σ_i	
Rural	0.999875 0.000125	0.03 0.5	0.35 0.4	Mixture of Water Soluble and Dust-Like Aerosols
Urban	0.999875 0.000125	0.03 0.5	0.35 0.4	Rural Aerosol Mixtures with Soot-Like Aerosols
Maritime				
Continental Origin	0.99	0.03	0.35	Rural Aerosol Mixture
Oceanic Origin	0.01	0.3	0.4	Sea Salt Solution in Water
Tropospheric	1.0	0.03	0.35	Rural Aerosol Mixture

using any such model: given the natural variability of the atmosphere, almost any particle model is supported by some measurements and no model (or set of models) will be consistent with all measurements.

The size distributions for the different aerosol models are represented by one or the sum of two log-normal distributions:

$$n(r) = \frac{dN(r)}{dr} = \sum_{i=1}^{2} \left\{ \left[\frac{N_i}{\ln(10) \cdot r \cdot \sigma_i \sqrt{2\pi}} \right] \times \exp\left[- \frac{(\log r - \log r_i)^2}{2\sigma_i^2} \right] \right\}, \quad (18.15)$$

where N(r) is the cumulative number density of particles of radius r, σ is the standard deviation, and r_i, N_i are the mode radius and the number density of r_i. This form of distribution function represents the multimodal nature of the atmospheric aerosols. Following the usual convention, log is the logarithm to the base 10 and ln is the logarithm to the base e.

Four different aerosol models for the atmospheric boundary layer have been developed. They differ in particle size distribution and particle refractive index. Table 18-6 lists the parameters defining the size distributions in accordance with Equation (18.15) for these models. These mode radii correspond to moderate humidities (70% to 80%); values of r_i as a function of humidity are given in Table 18-7.

The choices of N in Table 18-6 are normalized to correspond to 1 particle/cm^3. The actual size distributions can be renormalized to give the correct extinction coefficients for the altitude and the visibility being used. The continental and oceanic components of the maritime model can be used in various proportions depending on the prevailing winds—particularly in coastal regions.

As the relative humidity increases, water vapor con-

Table 18-7. Mode radii for the aerosol models as a function of relative humidity

Relative Humidity	Tropospheric r_1	Rural		Maritime	Urban	
		r_1	r_2		r_1	r_2
0%	0.02700	0.02700	0.4300	0.1600	0.02500	0.4000
50%	0.02748	0.02748	0.4377	0.1711	0.02563	0.4113
70%	0.02846	0.02846	0.4571	0.2041	0.02911	0.4777
80%	0.03274	0.03274	0.5477	0.3180	0.03514	0.5808
90%	0.03884	0.03884	0.6462	0.3803	0.04187	0.7061
95%	0.04238	0.04238	0.7078	0.4606	0.04904	0.8634
98%	0.04751	0.04751	0.9728	0.6024	0.05996	1.1691
99%	0.05215	0.05215	1.1755	0.7505	0.06847	1.4858

denses out of the atmosphere on the particulates suspended in the atmosphere. This condensed water increases the size of the aerosols and changes their composition and effective refractive index. The resulting effect of the aerosol on the absorption and scattering of light will correspondingly be modified. There have been a number of studies on the change of aerosol properties as a function of relative humidity. The most comprehensive of these, especially in terms of the resulting effects on the aerosol optical properties, is the work of Hanel [1976].

The "Rural Model" is intended to represent the aerosol under conditions where it is not directly influenced by urban and/or industrial aerosol sources. The rural aerosols are assumed to be composed of a mixture of 70% water soluble substances (ammonium and calcium sulfate and also organic compounds) and 30% dust-like aerosols. The refractive index for these components based on the measurements of Volz [1972a,b; 1973] is shown in Figure 18-8. The refractive index is in general a complex quantity with a real and imaginary part. These refractive index data weighted by the mixing ratio of the two components are consistent with other direct measurements and with values inferred from *in situ* measurements. For the refractive index of water, the survey of Hale and Querry [1973] was used. While there are some minor differences between the optical constants in Hale and Querry's survey and the more recent measurements, these differences are comparable to the experimental errors and are small compared with the other uncertainties in the model parameters. These refractive index data are shown in Figure 18-9. The resulting number density distributions n(r), are shown in Figure 18-10. To allow for the dependence of the humidity effects on the size of the dry aerosol, the growth of the aerosol was computed separately for the accumulation and coarse particle components. In accounting for the aerosol growth, changes in the width of the size distribution were assumed negligible so only the mode radius r, was

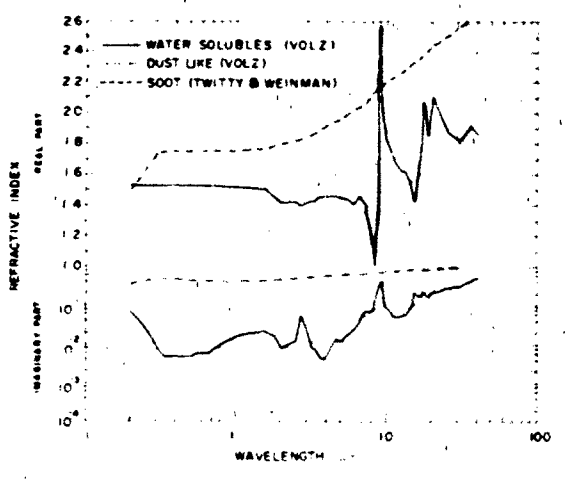


Figure 18-8. Refractive index for the dry rural and urban aerosol components

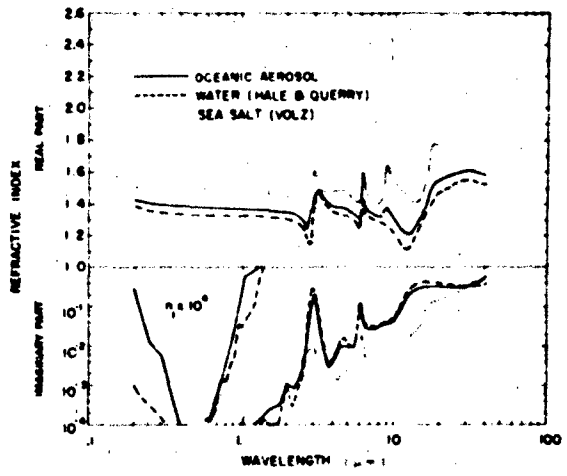


Figure 18-9. Refractive index of oceanic aerosol, water, and sea salt.

modified by humidity changes (Figure 18-10). The effective refractive indices for the two size components were then computed.

In urban areas the air with a rural aerosol background is primarily modified by the addition of aerosols from combustion products and industrial sources. The urban aerosol model therefore was taken to be a mixture of the rural aerosol with carbonaceous aerosols. The proportions of the soot-like carbonaceous aerosols and the rural type of aerosol mixture are assumed to be 20% and 80%, respectively. The soot-like aerosols are assumed to have the same size distribution as both components of the rural model. The refractive index of the soot-like aerosols are based on the soot data in Twitty and Weinman's [1971] survey of the refractive index of carbonaceous materials. As with the rural model, a composite urban aerosol refractive index was determined at each wavelength.

The aerosol compositions and distributions of oceanic origin are significantly different from continental aerosol types. These aerosols are largely sea-salt particles produced by the evaporation of sea-spray droplets that have grown again due to aggregation of water under high relative humidity conditions. However, even over the ocean there is a more or less pronounced continental aerosol background that, mixed with the aerosol of oceanic origin, forms a fairly uniform maritime aerosol. It is representative for the boundary layer in the lower 2–3 km in the atmosphere over the oceans, but also may occur over the continents in a maritime air mass. This maritime model should be distinguished from the fresh sea-spray aerosol that exists in the lower 10–20 m above the ocean surface and is strongly dependent on wind speed.

The maritime aerosol model, therefore, is composed of two components: (1) the sea-salt component and (2) a continental component assumed to be identical to the rural aerosol with the exception that the very large particles are eliminated since they will eventually be lost due to fallout

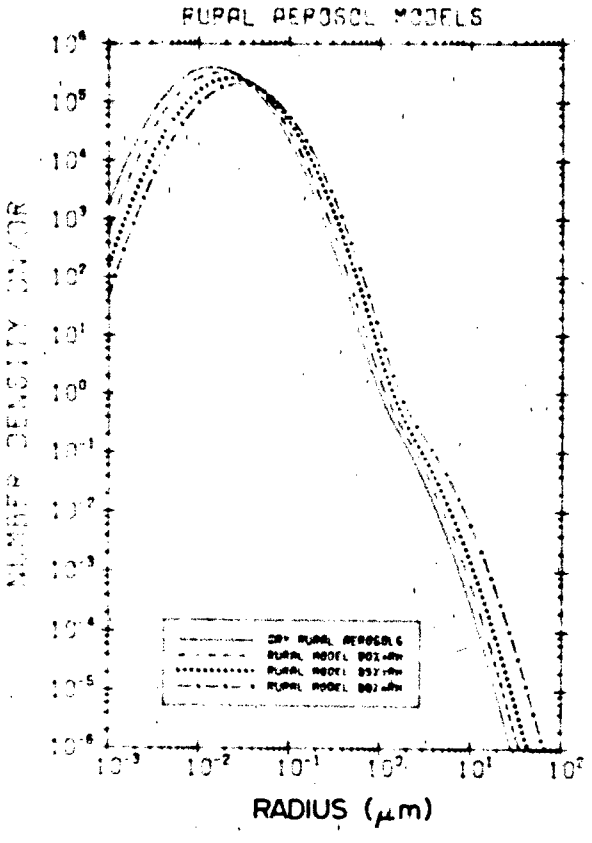


Figure 18-10 Aerosol number distribution (cm $^{-3}$ μm $^{-1}$) for the rural model at different relative humidities with total particle concentrations fixed at 15 000 cm $^{-3}$

as the air masses move across the oceans. For the size distribution of the oceanic component, a log normal distribution is used with r_v = 0.3 μm for moderate relative humidity (=80%) and σ = 0.4. The relative proportions of aerosol of oceanic or continental origins will vary particularly in coastal regions. To account for these variations, the model permits the user to adjust the relative amounts of the oceanic and continental types of aerosol. The number density distribution is shown in Figure 18-11. The refractive index is based on that for a solution of sea salt in water, using a weighted average of the refractive indices of water and sea salt. The refractive index of the sea salt is primarily taken from the measurements of Volz [1972 a, b]. A model for maritime aerosols that accounts for production of oceanic aerosol at the sea surface by white caps and spray as a function of wind speed and the changes in drop size due to varying relative humidity was developed recently by Gathman [1983]. With some modifications this model has been incorporated into LOWTRAN 6 (see Section 18.4.2).

The tropospheric aerosol model represents the aerosols within the troposphere above the boundary layer. These aerosols are assumed to have the same composition as the rural model (70% water soluble and 30% dust-like). The size distribution is modified from the rural model by eliminating the large (or coarse) particle (r_v = 0.5 μm) component of the size distribution because of the longer residence of aerosols above the boundary and the expected differential loss of the larger particles. This leaves the log-normal distribution with the small (or accumulation particle component). This is consistent with the changes in aerosol size distribution with altitude suggested by Whitby and Cantrell [1976]. The dependence of particle size on relative humidity is the same as for the small particle component of the rural model, and is shown in Figure 18-12.

Based on observations of stratospheric aerosol behavior four different vertical distribution models for the stratosphere were developed: a background model and three different volcanic aerosol profiles (moderate, high, and extreme). These models represent either different amounts of volcanic material injected into the stratosphere or the decrease over time from the extreme conditions following a major event such as the Krakatoa explosion. The experimental data also confirm a seasonal trend in the stratospheric aerosol distributions. Figure 18-13 shows these stratospheric models and also for comparison the 1968 Elterman model

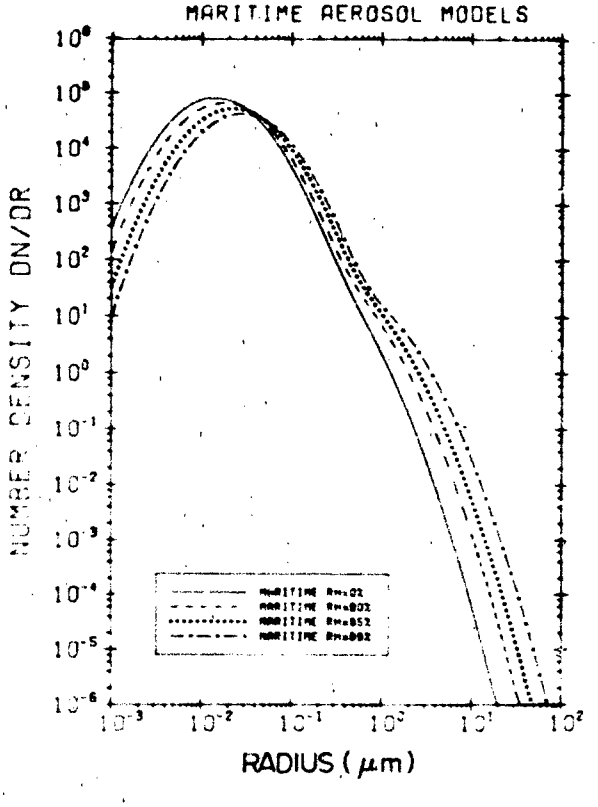


Figure 18-11 Aerosol number distribution (cm $^{-3}$ μm $^{-1}$) for the maritime model at different relative humidities with total particle concentrations fixed at 4000 cm $^{-3}$

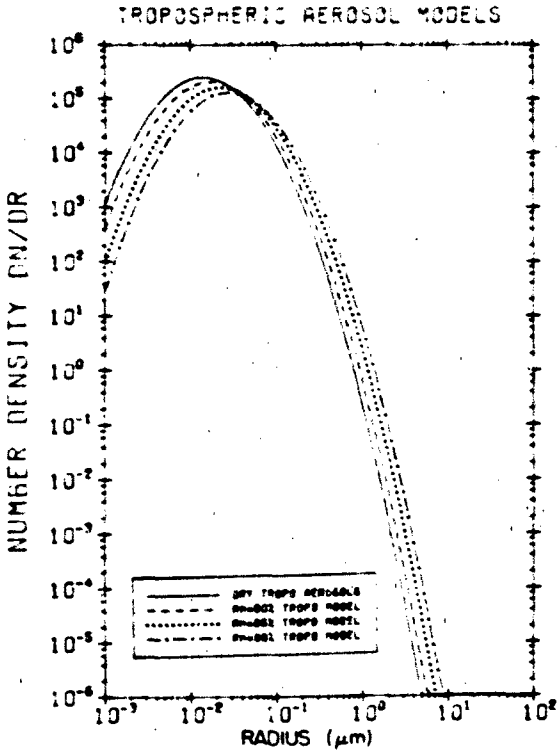


Figure 18-12. Aerosol number distribution (cm^{-3} μm^{-1}) for the tropospheric model at different relative humidities with total particle concentrations fixed at 10 000 cm^{-3}.

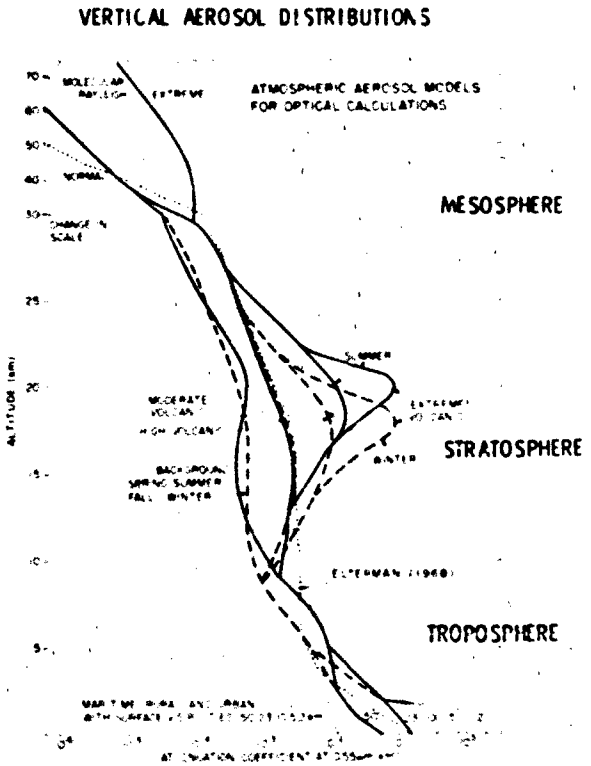


Figure 18-13. The vertical distribution of the aerosol extinction (at 0.55 microns) for the different models. Also shown for comparison are the Rayleigh profile (dotted line) and Elterman's [1968] model. Between 2 and 30 km, where a distinction on a seasonal basis is made, the spring–summer conditions are indicated with a solid line and fall–winter conditions are indicated by a dashed line.

based on data collected during the period several years after the Agung eruption in 1963 and therefore representative for the moderate volcanic conditions.

The stratospheric background aerosols were assumed to be a 75% solution of sulfuric acid in water following the work of Rosen [1971] and Toon and Pollack [1973]. The complex refractive index as a function of wavelength is based on the measurements of Remsberg [1971 and 1973] and Palmer and Williams [1975]. The refractive index for the volcanic models is based on the measurements of Volz [1973] on volcanic dust.

There are two volcanic aerosol size models: a "Fresh Volcanic Model" representing the size distribution of aerosols shortly after a volcanic eruption, and a "Volcanic Model" representing the aerosol about a year after an eruption. Both size distributions were chosen mainly on the basis of Mossop's [1964] measurements following the eruption of Mt. Agung. The size distribution was also made consistent with the observed wavelength dependence of extinction due to volcanic aerosols, and in the case of the "Fresh Volcanic Model" consistent with the observation of optical phenomena such as Bishop's rings, and a blue or green sun sometimes observed following major volcanic eruptions. These size distributions are represented by a modified gamma distribution:

$$\frac{dN}{dr} = n(r) = Ar^{\alpha} \exp(-br^{\gamma}) \qquad (18.16)$$

whose parameters are given in Table 18-8. Either of the two volcanic size distribution models are appropriate to use with the "moderate" or "high" volcanic vertical profiles, depending on whether these profiles represent a major eruption after a couple of years or a recent weaker eruption. It should be noted that the profiles may have a much more layered structure than is shown in the model distributions, particularly shortly after a volcanic injection of dust into the atmosphere.

The major component of the normal upper atmospheric

Table 18-8. Modified gamma distribution parameters.

	A	α	γ	b
Background Stratospheric	324.0	1	1	18
Fresh Volcanic	341.33	1	0.5	8
Aged Volcanic	5 461.33	1	0.5	16

aerosol is considered to be meteoric dust. Meteoric or cometary dust also forms some of the layers occasionally observed in the upper atmosphere. The refractive index of meteoric dust is based on the work of Shettle and Volz [1976] who determined the complex refractive index for a mixture of chondrite dust which represents the major type of meteorite falling on the earth. The size distribution has been represented by a log-normal distribution (equation 18.15 with i = 1) whose parameters are given as

$$N_1 = 1 \text{ cm}^{-3}, r = 0.03 \ \mu\text{m}, \sigma_1 = 0.5 .$$

Figure 18-13 shows in addition to the normal upper atmospheric model an extreme model similar to Ivlev's [1969] model for the upper atmosphere. Attenuation coefficients this large occur in layers with thicknesses no greater than a few kilometers. At different times these layers can be either micrometeoric dust or noctilucent clouds. The wavelength dependent refractive indices for the different aerosol types are listed in Table 18-9.

18.2.1.5 Aerosol Extinction, Scattering, Polarization. Once the size distribution and refractive index of the aerosol models are specified, the optical properties can be determined. The effects of small particles on the propagation of radiation are defined by the coefficients for extinction, scattering and absorption, the angular scattering function, and the polarization of scattered radiation.

The Mie [1908] theory treats the scattering of light by homogeneous particles of arbitrary size and refractive index (for Rayleigh scattering see Section 18.1.4). For detailed discussions on particle scattering see textbooks by van de Hulst [1957], Kerker [1969], and McCartney [1976]. The spatial distribution of scattered light for incident unpolarized light of intensity $I_{o,\lambda}$ is

$$I_\lambda = I_{o,\lambda} \frac{1}{2 \cdot k^2 \cdot R^2} [i_1(\alpha, m, \theta) + i_2(\alpha, m, \theta)], \qquad (18.17)$$

where $k = 2\pi/\lambda$,

R = Distance from the scattering particle center,

α = $2\pi r/\lambda$, the size parameter, r is particle radius,

m = complex refractive index of the particle relative to its surrounding medium,

θ = angle between the incident and the scattered beams ($\theta = 0°$ is defined as forward scattering

i_1 = perpendicular polarized component

i_2 = parallel polarized component.

The dimensionless intensity functions define degree of linear polarization P.

$$P = \frac{i_1 - i_2}{i_1 + i_2} . \qquad (18.18)$$

Integration of the scattered light over all directions (4π steradians) gives

$$\int_0^{4\pi} I_\lambda \, d\Omega = I_{o,\lambda} \cdot C_s . \qquad (18.19)$$

The ratio of C_s (scattering cross section) to the geometrical cross section of the scattering particle, $r^2\pi$, is called scattering efficiency factor Q_s. If light is traveling through a scattering medium containing N particles per unit volume, the loss of light due to scattering per unit path length dx is

$$dI_\lambda = -I_\lambda \cdot C_s \cdot N \cdot dx. \qquad (18.20)$$

Integration of Equation (18.20) gives the Bouguer extinction law:

$$I_{\lambda \text{ transmitted}} = I_{o,\lambda} \exp(-k_s x) \qquad (18.21)$$

with $k_s = N \cdot C_s$ called the scattering coefficient. Similarly to k_s one can define an absorption coefficient k_a for absorption by absorbing aerosol particles.

$k_e = k_s + k_a$ is called the extinction coefficient.

For Mie scattering calculations, the aerosol particles are assumed to be spherical, which in general is not true. While liquid aerosols are approximately spherical, dry particles usually are irregularly shaped. However, it can be argued that many of the measurements of aerosol size distributions directly measure scattered light from the aerosol particles, and the size "assigned" to the particle is the size of a sphere that has similar scattering properties to the measured particle. Furthermore, the irregularly shaped particles are closest to "equivalent" spheres in their scattering properties in the forward direction generally used for size measurements. Therefore, the resulting size distribution is the size distribution of spheres that have similar optical properties to those of the actual aerosol particles. Also, studies by Chylek et al. [1976] and Holland and Gagne [1970] indicate that for particles of equal overall dimension but different shapes, the spherical particle extinction has the highest values.

The computed attenuation coefficients (extinction, scattering, and absorption) for the model aerosols above are shown in Tables 18-10 a thru c and Figures 18-14 thru 18-22.

The attenuation coefficients for the rural aerosol model at 50% relative humidity as a function of wavelength are shown in Figure 18-14. To show the effect of variations of relative humidity on the aerosol extinction, the rural model extinction has been calculated over a range of humidities for a constant total number density. The resulting extinction versus wavelength is shown in Figure 18-15 for several

Table 18-9a Refractive indices for different aerosol types

Wavelength	Sea Salt		Oceanic		Water		Ice		Meteoric	
(μm)	real	imag	real	imag	real	imag	real	imag	real	imag
0.200	1.510	-1.00E-04	1.429	-2.87E-05	1.396	-1.10E-07	1.394	-1.50E-08	1.515	-1.23E-05
0.250	1.510	-5.00E-06	1.404	-1.45E-06	1.362	-3.35E-08	1.351	-8.60E-09	1.515	-2.41E-05
0.300	1.510	-2.00E-06	1.395	-5.83E-07	1.349	-1.60E-08	1.334	-5.50E-09	1.515	-4.18E-05
0.337	1.510	-4.00E-07	1.392	-1.20E-07	1.345	-8.45E-09	1.326	-4.50E-09	1.514	-5.94E-05
0.400	1.500	-3.60E-08	1.385	-9.90E-09	1.339	-1.86E-09	1.320	-2.71E-09	1.514	-9.95E-05
0.488	1.500	-2.00E-08	1.382	-6.41E-09	1.335	-9.69E-10	1.313	-1.75E-09	1.513	-1.81E-04
0.515	1.500	-1.00E-08	1.381	-3.70E-09	1.334	-1.18E-09	1.312	-2.19E-09	1.51.	-2.13E-04
0.550	1.500	-1.00E-08	1.381	-4.26E-09	1.333	-1.96E-09	1.311	-3.11E-09	1.513	-2.61E-04
0.633	1.490	-2.00E-08	1.377	-1.62E-08	1.332	-1.46E-08	1.308	-1.09E-08	1.512	-3.99E-04
0.694	1.490	-1.00E-07	1.376	-5.04E-08	1.331	-3.05E-08	1.306	-2.62E-08	1.511	-5.30E-04
0.860	1.480	-3.00E-06	1.372	-1.09E-06	1.329	-3.29E-07	1.303	-2.15E-07	1.509	-1.02E-03
1.060	1.470	-2.00E-04	1.367	-6.01E-05	1.326	-4.18E-06	1.300	-1.96E-06	1.506	-1.95E-03
1.300	1.470	-4.00E-04	1.365	-1.41E-04	1.323	-3.69E-05	1.295	-1.32E-05	1.501	-3.72E-03
1.536	1.460	-6.00E-04	1.359	-2.43E-04	1.318	-9.97E-05	1.290	-6.10E-04	1.495	-6.34E-03
1.800	1.450	-8.00E-04	1.351	-3.11E-04	1.312	-1.15E-04	1.282	-1.13E-04	1.488	-1.06E-02
2.000	1.450	-1.00E-03	1.347	-1.07E-03	1.306	-1.10E-03	1.273	-1.61E-03	1.482	-1.51E-02
2.250	1.440	-2.00E-03	1.334	-8.50E-04	1.292	-3.90E-04	1.256	-2.13E-04	1.474	-2.24E-02
2.500	1.430	-4.00E-03	1.309	-2.39E-03	1.261	-1.74E-03	1.225	-7.95E-04	1.467	-3.18E-02
2.700	1.400	-7.00E-03	1.249	-1.56E-02	1.188	-1.90E-02	1.163	-2.93E-03	1.462	-4.10E-02
3.000	1.610	-1.00E-02	1.439	-0.197	1.371	-0.272	1.045	-0.429	1.456	-5.73E-02
3.200	1.490	-3.00E-03	1.481	-6.69E-02	1.478	-9.24E-02	1.652	-0.283	1.454	-6.94E-02
3.392	1.480	-2.00E-03	1.439	-1.51E-02	1.422	-2.04E-02	1.510	-4.01E-02	1.454	-8.15E-02
3.500	1.480	-1.60E-03	1.423	-7.17E-03	1.400	-9.40E-03	1.453	-1.61E-02	1.455	-8.82E-02
3.750	1.470	-1.40E-03	1.398	-2.90E-03	1.369	-3.50E-03	1.391	-7.00E-03	1.459	-0.103
4.000	1.480	-1.40E-03	1.388	-3.69E-03	1.351	-4.60E-03	1.361	-1.00E-02	1.466	-0.116
4.500	1.490	-1.40E-03	1.377	-9.97E-03	1.332	-1.34E-02	1.340	-2.87E-02	1.485	-0.131
5.000	1.470	-2.50E-03	1.366	-9.57E-03	1.325	-1.24E-02	1.327	-1.20E-02	1.500	-0.135
5.500	1.420	-3.40E-03	1.333	-9.31E-03	1.298	-1.16E-02	1.299	-2.17E-02	1.508	-0.132
6.000	1.410	-1.10E-02	1.306	-7.96E-02	1.265	-0.107	1.296	-6.47E-02	1.507	-0.126
6.200	1.600	-2.20E-02	1.431	-6.91E-02	1.363	-8.80E-02	1.313	-6.83E-02	1.504	-0.124
6.500	1.460	-5.00E-03	1.374	-2.94E-02	1.339	-3.92E-02	1.320	-5.59E-02	1.497	-0.121
7.200	1.420	-7.00E-03	1.343	-2.49E-02	1.312	-3.21E-02	1.318	-5.44E-02	1.469	-0.119
7.900	1.400	-1.30E-02	1.324	-2.79E-02	1.294	-3.39E-02	1.313	-4.79E-02	1.422	-0.130
8.200	1.420	-2.00E-02	1.324	-3.08E-02	1.286	-3.51E-02	1.306	-3.90E-02	1.395	-0.142
8.500	1.480	-2.60E-02	1.336	-3.36E-02	1.278	-3.67E-02	1.291	-3.91E-02	1.363	-0.162
8.700	1.600	-3.00E-02	1.366	-3.56E-02	1.272	-3.79E-02	1.282	-4.00E-02	1.339	-0.182
9.000	1.650	-2.80E-02	1.373	-3.65E-02	1.262	-3.99E-02	1.269	-4.29E-02	1.302	-0.228
9.200	1.610	-2.60E-02	1.356	-3.71E-02	1.255	-4.15E-02	1.261	-4.46E-02	1.281	-0.273
9.500	1.580	-1.80E-02	1.339	-3.68E-02	1.243	-4.44E-02	1.245	-4.59E-02	1.272	-0.360
9.800	1.560	-1.60E-02	1.324	-3.88E-02	1.229	-4.79E-02	1.219	-4.70E-02	1.310	-0.450
10.000	1.540	-1.50E-02	1.310	-4.06E-02	1.218	-5.08E-02	1.197	-5.10E-02	1.355	-0.488
10.591	1.500	-1.40E-02	1.271	-5.22E-02	1.179	-6.74E-02	1.098	-0.131	1.419	-0.547
11.000	1.480	-1.40E-02	1.246	-7.31E-02	1.153	-9.68E-02	1.093	-0.239	1.509	-0.691
11.500	1.480	-1.40E-02	1.227	-0.105	1.126	-0.142	1.176	-0.360	1.847	-0.634
12.500	1.420	-1.60E-02	1.208	-0.190	1.123	-0.259	1.387	-0.422	1.796	-0.252
13.000	1.410	-1.80E-02	1.221	-0.223	1.146	-0.305	1.472	-0.389	1.711	-0.219
14.000	1.410	-2.30E-02	1.267	-0.271	1.210	-0.370	1.569	-0.283	1.641	-0.217
14.300	1.430	-2.60E-02	1.307	-0.292	1.258	-0.396	1.579	-0.191	1.541	-0.198
15.000	1.450	-3.50E-02	1.321	-0.297	1.270	-0.402	1.572	-0.177	1.510	-0.206
16.400	1.560	-9.00E-02	1.407	-0.331	1.346	-0.427	1.531	-0.125	1.478	-0.467
17.200	1.740	-0.120	1.487	-0.341	1.386	-0.429	1.534	-0.107	1.441	-0.400
18.000	1.780	-0.130	1.525	-0.341	1.423	-0.426	1.522	-8.39E-02	1.354	-0.557
18.500	1.770	-0.135	1.536	-0.339	1.448	-0.421	1.510	-7.60E-02	1.389	-0.705
20.000	1.760	-0.152	1.560	-0.324	1.480	-0.393	1.504	-6.70E-02	1.803	-0.765
21.300	1.760	-0.165	1.568	-0.311	1.491	-0.379	1.481	-3.85E-02	1.797	-0.556
22.500	1.760	-0.180	1.579	-0.316	1.506	-0.370	1.455	-2.91E-02	1.661	-0.592
25.000	1.760	-0.205	1.596	-0.313	1.531	-0.356	1.414	-2.99E-02	1.983	-0.861
27.000	1.770	-0.275	1.612	-0.320	1.549	-0.339	1.358	-4.90E-02	2.023	-0.666
30.000	1.770	-0.300	1.614	-0.320	1.551	-0.328	1.325	-6.50E-02	2.139	-0.665
35.000	1.760	-0.500	1.597	-0.383	1.532	-0.336	1.226	-0.155	2.146	-0.380
40.000	1.740	-1.000	1.582	-0.561	1.519	-0.385	1.202	-0.344	1.979	-0.359

Table 18-9b Refractive indices for different aerosol types

Wavelength (μm)	Water Soluble real	Water Soluble imag	Dust-Like real	Dust-Like imag	Soot real	Soot imag	75% H_2SO_4 real	75% H_2SO_4 imag	Volcanic real	Volcanic imag
0.200	1.530	-7.00E-02	1.530	-7.00E-02	1.500	-0.350	1.498	-1.00E-08	1.500	-7.00E-02
0.250	1.530	-2.00E-02	1.530	-3.00E-02	1.620	-0.450	1.484	-1.00E-08	1.500	-3.00E-02
0.300	1.530	-8.00E-03	1.530	-8.00E-03	1.740	-0.470	1.469	-1.00E-08	1.500	-1.00E-02
0.337	1.530	-5.00E-03	1.530	-8.00E-03	1.750	-0.470	1.459	-1.00E-08	1.500	-8.00E-03
0.400	1.530	-5.00E-03	1.530	-8.00E-03	1.750	-0.460	1.440	-1.00E-08	1.500	-8.00E-03
0.488	1.530	-5.00E-03	1.530	-8.00E-03	1.750	-0.450	1.432	-1.00E-08	1.500	-8.00E-03
0.515	1.530	-5.00E-03	1.530	-8.00E-03	1.750	-0.450	1.431	-1.00E-08	1.500	-8.00E-03
0.550	1.530	-6.00E-03	1.530	-8.00E-03	1.750	-0.440	1.430	-1.00E-08	1.500	-8.00E-03
0.633	1.530	-6.00E-03	1.530	-8.00E-03	1.750	-0.430	1.429	-1.47E-08	1.500	-8.00E-03
0.694	1.530	-7.00E-03	1.530	-8.00E-03	1.750	-0.430	1.428	-1.99E-08	1.500	-8.00E-03
0.860	1.520	-1.20E-02	1.520	-8.00E-03	1.750	-0.430	1.425	-1.72E-07	1.500	-8.00E-03
1.060	1.520	-1.70E-02	1.520	-8.00E-03	1.750	-0.440	1.420	-1.50E-06	1.500	-8.00E-03
1.300	1.510	-2.00E-02	1.460	-8.00E-03	1.760	-0.450	1.410	-1.00E-05	1.500	-8.00E-03
1.536	1.510	-2.30E-02	1.400	-8.00E-03	1.770	-0.460	1.403	-1.37E-04	1.490	-8.00E-03
1.800	1.460	-1.70E-02	1.330	-8.00E-03	1.790	-0.480	1.390	-5.50E-04	1.480	-8.00E-03
2.000	1.420	-8.00E-03	1.260	-8.00E-03	1.800	-0.490	1.384	-1.26E-03	1.460	-8.00E-03
2.250	1.420	-1.00E-02	1.220	-9.00E-03	1.810	-0.500	1.370	-1.80E-03	1.460	-8.00E-03
2.500	1.420	-1.20E-02	1.180	-9.00E-03	1.820	-0.510	1.344	-3.76E-03	1.460	-9.00E-03
2.700	1.400	-5.50E-02	1.180	-1.30E-02	1.830	-0.520	1.303	-5.70E-03	1.460	-1.00E-02
3.000	1.420	-2.20E-02	1.160	-1.20E-02	1.840	-0.540	1.293	-9.55E-02	1.480	-1.30E-02
3.200	1.430	-8.00E-03	1.220	-1.00E-02	1.860	-0.540	1.311	-0.135	1.480	-1.40E-02
3.392	1.430	-7.00E-03	1.260	-1.30E-02	1.870	-0.550	1.352	-0.159	1.490	-1.20E-02
3.500	1.450	-5.00E-03	1.280	-1.10E-02	1.880	-0.560	1.376	-0.158	1.490	-1.10E-02
3.750	1.452	-4.00E-03	1.270	-1.10E-02	1.900	-0.570	1.396	-0.131	1.500	-9.00E-03
4.000	1.455	-5.00E-03	1.260	-1.20E-02	1.920	-0.580	1.398	-0.126	1.503	-7.00E-03
4.500	1.460	-1.30E-02	1.260	-1.40E-02	1.940	-0.590	1.385	-0.120	1.520	-7.50E-03
5.000	1.450	-1.20E-02	1.250	-1.60E-02	1.970	-0.600	1.360	-0.121	1.510	-9.00E-03
5.500	1.440	-1.80E-02	1.220	-2.10E-02	1.990	-0.610	1.337	-0.183	1.510	-1.20E-02
6.000	1.410	-2.30E-02	1.150	-3.70E-02	2.020	-0.620	1.425	-0.195	1.480	-1.50E-02
6.200	1.430	-2.70E-02	1.140	-3.90E-02	2.030	-0.625	1.424	-0.165	1.460	-1.80E-02
6.500	1.460	-3.30E-02	1.130	-4.20E-02	2.040	-0.630	1.370	-0.128	1.450	-2.40E-02
7.200	1.400	-7.00E-02	1.400	-5.50E-02	2.060	-0.650	1.210	-0.176	1.440	-4.50E-02
7.900	1.200	-6.50E-02	1.150	-4.00E-02	2.120	-0.670	1.140	-0.488	1.380	-7.20E-02
8.200	1.010	-0.100	1.130	-7.40E-02	2.130	-0.680	1.200	-0.645	1.340	-9.70E-02
8.500	1.300	-0.215	1.300	-9.60E-02	2.150	-0.690	1.370	-0.755	1.620	-0.121
8.700	2.400	-0.290	1.400	-0.100	2.160	-0.690	1.530	-0.772	1.950	-0.170
9.000	2.560	-0.370	1.700	-0.140	2.170	-0.700	1.650	-0.633	2.200	-0.215
9.200	2.200	-0.420	1.720	-0.150	2.180	-0.700	1.600	-0.586	2.230	-0.240
9.500	1.950	-0.160	1.730	-0.162	2.190	-0.710	1.670	-0.750	2.250	-0.275
9.800	1.870	-9.50E-02	1.740	-0.162	2.200	-0.715	1.910	-0.680	2.280	-0.304
10.000	1.820	-9.00E-02	1.750	-0.162	2.210	-0.720	1.890	-0.455	2.300	-0.320
10.591	1.760	-7.00E-02	1.620	-0.120	2.220	-0.730	1.720	-0.340	2.200	-0.305
11.000	1.720	-5.00E-02	1.620	-0.105	2.230	-0.730	1.670	-0.485	2.150	-0.270
11.500	1.670	-4.70E-02	1.590	-0.100	2.240	-0.740	1.890	-0.374	2.050	-0.240
12.500	1.620	-5.30E-02	1.510	-9.00E-02	2.270	-0.750	1.740	-0.198	1.800	-0.155
13.000	1.620	-5.50E-02	1.470	-0.100	2.280	-0.760	1.690	-0.195	1.760	-0.148
14.000	1.560	-7.30E-02	1.520	-8.50E-02	2.310	-0.775	1.640	-0.195	1.700	-0.145
14.800	1.440	-0.100	1.570	-0.100	2.330	-0.790	1.610	-0.205	1.650	-0.157
15.000	1.420	-0.200	1.570	-0.100	2.330	-0.790	1.590	-0.211	1.650	-0.170
16.400	1.750	-0.160	1.600	-0.100	2.360	-0.810	1.520	-0.414	1.750	-0.200
17.200	2.080	-0.240	1.630	-0.100	2.380	-0.820	1.724	-0.590	1.850	-0.240
18.000	1.980	-0.180	1.640	-0.115	2.400	-0.825	1.950	-0.410	2.000	-0.305
18.500	1.850	-0.170	1.640	-0.120	2.410	-0.830	1.927	-0.302	2.100	-0.325
20.000	2.120	-0.220	1.680	-0.220	2.450	-0.850	1.823	-0.235	2.250	-0.318
21.300	2.060	-0.230	1.770	-0.280	2.460	-0.860	1.780	-0.292	2.400	-0.290
22.500	2.000	-0.240	1.900	-0.280	2.480	-0.870	1.870	-0.315	2.500	-0.350
25.000	1.880	-0.280	1.970	-0.240	2.510	-0.890	1.930	-0.200	2.600	-0.400
27.900	1.840	-0.290	1.890	-0.320	2.540	-0.910	1.920	-0.180	2.500	-0.430
30.000	1.820	-0.300	1.800	-0.420	2.570	-0.930	1.920	-0.180	2.400	-0.450
35.000	1.920	-0.400	1.900	-0.500	2.630	-0.970	1.900	-0.190	2.300	-0.520
40.000	1.860	-0.500	2.100	-0.600	2.680	-1.000	1.890	-0.220	2.250	-0.650

Table 18-10a Vertical distribution of aerosol extinction (km^{-1}) for a wavelength of 0.550 μm with rural model aerosol boundary layer.

Height (km)	50 km Met. Range	23 km Met. Range	10 km Met. Range	5 km Met. Range	2 km Met. Range
0.0	6.62E-02	1.58E-01	3.79E-01	7.70E-02	1.94E+00
1.0	4.15E-02	9.91E-02	3.79E-01	7.70E-01	1.94E+00
1.5	3.26E-02	7.92E-02	3.79E-01	7.70E-01	1.94E+00
2.0	2.60E-02	6.21E-02	6.21E-02	6.21E-02	6.21E-02

	Fall-Winter Profiles		Spring-Summer Profiles	
	50 km Surface Met. Range	2 to 23 km Surface Met. Range	50 km Surface Met. Range	2 to 23 km Surface Met. Range
2.0	2.60E-02	6.21E-02	2.60E-02	6.21E-02
3.0	1.14E-02	2.72E-02	1.46E-02	3.46E-02
4.0	6.43E-03	1.20E-02	1.07E-02	1.85E-02
5.0	4.85E-03	4.85E-03	9.30E-03	9.30E-03
6.0	3.54E-03	3.54E-03	7.71E-03	7.71E-03
7.0	2.30E-03	2.30E-03	6.22E-03	6.22E-03
8.0	1.41E-03	1.41E-03	3.36E-03	3.36E-03
9.0	9.80E-04	9.80E-04	1.81E-03	1.81E-03

	Background Stratosphere	Moderate Volcanic	High Volcanic	Extreme Volcanic	Background Stratospheric	Moderate Volcanic	High Volcanic	Extreme Volcanic
9.0	9.80E-04	9.80E-04	9.80E-04	9.80E-04	1.81E-03	1.81E-03	1.81E-03	1.81E-03
10.0	7.87E-04	1.38E-03	1.70E-03	1.70E-03	1.14E-03	1.85E-03	1.85E-03	1.85E-03
11.0	7.14E-04	1.79E-03	2.31E-03	2.31E-03	8.00E-04	2.11E-03	2.11E-03	2.11E-03
12.0	6.63E-04	2.21E-03	3.25E-03	3.25E-03	6.42E-04	2.45E-03	2.45E-03	2.45E-03
13.0	6.22E-04	2.75E-03	4.52E-03	4.52E-03	5.17E-04	2.80E-03	2.80E-03	2.80E-03
14.0	6.45E-04	2.89E-03	6.40E-03	6.40E-03	4.43E-04	2.89E-03	3.61E-03	3.61E-03
15.0	6.43E-04	2.92E-03	7.80E-03	1.01E-02	3.95E-04	2.92E-03	5.23E-03	5.23E-03
16.0	6.41E-04	2.74E-03	9.42E-03	2.35E-02	3.82E-04	2.74E-03	8.10E-03	8.10E-03
17.0	6.01E-04	2.46E-03	1.07E-02	6.10E-02	4.25E-04	2.46E-03	1.20E-02	1.27E-02
18.0	5.63E-04	2.10E-03	1.10E-02	1.00E-01	5.20E-04	2.10E-03	1.52E-02	2.32E-02
19.0	4.92E-04	1.71E-03	8.60E-03	4.00E-02	5.82E-04	1.71E-03	1.53E-02	4.85E-02
20.0	4.23E-04	1.35E-03	5.10E-03	9.15E-03	5.90E-04	1.35E-03	1.17E-02	1.00E-01
21.0	3.52E-04	1.09E-03	2.70E-03	3.13E-03	5.03E-04	1.09E-03	7.10E-03	5.50E-02
22.0	2.96E-04	8.60E-04	1.46E-03	1.46E-03	4.20E-04	8.60E-04	4.50E-03	6.10E-03
23.0	2.42E-04	6.60E-04	8.90E-04	8.90E-04	3.00E-04	6.60E-04	2.40E-03	2.40E-03
24.0	1.90E-04	5.15E-04	5.80E-04	5.80E-04	1.98E-04	5.15E-04	1.28E-03	1.28E-03
25.0	1.50E-04	4.10E-04	4.10E-04	4.10E-04	1.31E-04	4.10E-04	7.75E-04	7.75E-04
26.0	1.15E-04	3.20E-04	3.20E-04	3.20E-04	9.01E-05	3.20E-04	4.45E-04	4.45E-04
27.0	8.95E-05	2.51E-04	2.51E-04	2.51E-04	6.78E-05	2.51E-04	2.90E-04	2.90E-04
28.0	6.70E-05	2.10E-04	2.10E-04	2.10E-04	5.18E-05	2.10E-04	2.10E-04	2.10E-04
29.0	5.20E-05	1.24E-04	1.24E-04	1.24E-04	4.12E-05	1.24E-04	1.24E-04	1.24E-04
30.0	3.32E-05	7.60E-05	7.60E-05	7.60E-05	3.32E-05	7.60E-05	7.60E-05	7.60E-05

	Normal	Transition from Volcanic to		Extreme
	Upper Atmos	Normal	Extreme	Upper Atmos.
30.0	3.32E-05	7.60E-05	7.60E-05	3.32E-05
35.0	1.65E-05	2.45E-05	7.20E-05	4.25E-05
40.0	8.00E-06	8.00E-06	6.95E-05	5.60E-05
45.0	4.02E-06	4.02E-06	6.60E-05	6.60E-05
50.0	2.10E-06			5.04E-05
55.0	1.09E-06			3.40E-05
60.0	5.78E-07			2.30E-05
65.0	3.05E-07			1.62E-05
70.0	1.63E-07			1.03E-05
75.0	6.95E-08			6.70E-06
80.0	2.90E-08			4.30E-06
85.0	1.20E-08			2.78E-06
90.0	5.10E-09			1.55E-06
95.0	2.15E-09			8.30E-07
100.0	9.30E-10			4.50E-07

Table 18-10b. Vertical distribution of aerosol scattering (km^{-1}) for a wavelength of 0.550 μm with rural model aerosol boundary layer.

Height (km)	50 km Met. Range	23 km Met. Range	10 km Met. Range	5 km Met. Range	2 km Met. Range
0.0	6.35E-02	1.52E-01	3.64E-01	7.39E-01	1.86E+00
1.0	3.98E-02	9.51E-02	3.64E-01	7.39E-01	1.86E+00
1.5	3.13E-02	7.60E-02	3.64E-01	7.39E-01	1.86E+00
2.0	2.49E-02	5.96E-02	5.96E-02	5.96E-02	5.96E-02

	Fall-Winter Profiles		Spring-Summer Profiles	
	50 km Surface Met. Range	2 to 23 km Surface Met. Range	50 km Surface Met. Range	2 to 23 km Surface Met. Range
2.0	2.49E-02	5.96E-02	2.49E-02	5.96E-02
3.0	1.09E-02	2.59E-02	1.39E-02	3.30E-02
4.0	6.13E-03	1.14E-02	9.72E-03	1.76E-02
5.0	4.62E-03	4.62E-02	8.86E-03	8.86E-03
6.0	3.37E-03	3.37E-03	7.35E-03	7.35E-03
7.0	2.19E-03	2.19E-03	5.93E-03	5.93E-03
8.0	1.34E-03	1.34E-03	3.20E-03	3.20E-03
9.0	9.34E-04	9.34E-04	1.72E-03	1.72E-03

	Background Stratosphere	Moderate Volcanic	High Volcanic	Extreme Volcanic	Background Stratospheric	Moderate Volcanic	High Volcanic	Extreme Volcanic
9.0	9.34E-04	9.34E-04	9.34E-04	9.34E-04	1.72E-03	1.72E-03	1.72E-03	1.72E-03
10.0	7.87E-04	1.31E-03	1.45E-03	1.45E-03	1.14E-03	1.75E-03	1.58E-03	1.58E-03
11.0	7.14E-04	1.70E-03	1.98E-03	1.98E-03	8.00E-04	2.00E-03	1.80E-03	1.80E-03
12.0	6.63E-04	2.09E-03	2.78E-03	2.78E-03	6.42E-04	2.32E-03	2.10E-03	2.10E-03
13.0	6.22E-04	2.61E-03	3.87E-03	3.87E-03	5.17E-04	2.65E-03	2.39E-03	2.39E-03
14.0	6.45E-04	2.74E-03	5.47E-03	5.47E-03	4.43E-04	2.74E-03	3.09E-03	3.09E-03
15.0	6.43E-04	2.77E-03	6.67E-03	8.64E-03	3.95E-04	2.77E-03	4.47E-03	4.47E-03
16.0	6.41E-04	2.60E-03	8.06E-03	2.01E-02	3.82E-04	2.60E-03	6.93E-03	6.93E-03
17.0	6.01E-04	2.33E-03	9.15E-03	5.22E-02	4.25E-04	2.33E-03	1.03E-02	1.09E-02
18.0	5.63E-04	1.99E-03	9.41E-03	8.55E-02	5.20E-04	1.99E-03	1.30E-02	1.98E-02
19.0	4.92E-04	1.62E-03	7.36E-03	8.42E-02	5.82E-04	1.62E-03	1.31E-02	4.15E-02
20.0	4.23E-04	1.28E-03	4.36E-03	7.83E-03	5.90E-04	1.28E-03	1.00E-02	8.55E-02
21.0	3.52E-04	1.02E-03	2.31E-03	2.68E-03	5.03E-04	1.03E-03	6.07E-03	4.70E-02
22.0	2.96E-04	8.15E-04	1.25E-03	1.25E-03	4.20E-04	8.15E-04	3.85E-03	5.22E-03
23.0	2.42E-04	6.25E-04	7.61E-04	7.61E-04	3.09E-04	6.25E-04	2.05E-03	2.05E-03
24.0	1.90E-04	4.88E-04	4.96E-04	4.96E-04	1.98E-04	4.88E-04	1.09E-03	1.09E-03
25.0	1.50E-04	3.88E-04	3.51E-04	3.51E-04	1.31E-04	3.88E-04	6.63E-04	6.63E-04
26.0	1.15E-04	3.03E-04	2.74E-04	2.74E-04	9.01E-05	3.03E-04	3.81E-04	3.81E-04
27.0	8.95E-05	2.38E-04	2.15E-04	2.15E-04	6.78E-05	2.38E-04	2.48E-04	2.48E-04
28.0	6.70E-05	1.99E-04	1.80E-04	1.80E-04	5.18E-05	1.99E-04	1.80E-04	1.80E-04
29.0	5.20E-05	1.17E-04	1.06E-04	1.06E-04	4.12E-05	1.17E-04	1.06E-04	1.06E-04
30.0	3.32E-05	7.20E-05	7.20E-05	7.20E-05	3.32E-05	7.20E-05	7.20E-05	7.20E-05

	Normal	Transition from Volcanic to		Extreme
	Upper Atmos.	Normal	Extreme	Upper Atmos.
30.0	3.32E-05	7.20E-05	6.50E-05	2.84E-05
35.0	1.64E-05	2.44E-05	7.16E-05	4.23E-05
40.0	7.96E-06	7.96E-06	6.91E-05	5.57E-05
45.0	4.00E-06	4.00E-06	6.57E-05	6.57E-05
50.0	2.09E-06			5.01E-05
55.0	1.08E-06			3.38E-05
60.0	5.75E-07			2.29E-05
65.0	3.03E-07			1.61E-05
70.0	1.59E-07			1.02E-05
75.0	6.91E-08			6.67E-06
80.0	2.89E-08			4.28E-06
85.0	1.19E-08			2.77E-06
90.0	5.07E-09			1.54E-06
95.0	2.14E-09			8.26E-07
100.0	9.25E-10			4.48E-07

Table 18-10c Vertical distribution of aerosol absorption(km^{-1}) for a wavelength of 0.550 μm with rural model aerosol boundary layer

Height (km)	50 km Met Range	23 km Met Range	10 km Met. Range	5 km Met. Range	2 km Met. Range
0.0	2.70E-03	6.45E-03	1.55E-02	3.14E-02	7.92E-02
1.0	1.69E-03	4.04E-03	1.55E-02	3.14E-02	7.92E-02
1.5	1.33E-03	3.23E-03	1.55E-02	3.14E-02	7.92E-02
2.0	1.06E-03	2.53E-03	2.53E-03	2.53E-03	2.53E-03

	Fall-Winter Profiles		Spring-Summer Profiles	
	50 km Surface Met. Range	2 to 23 km Surface Met. Range	50 km Surface Met. Range	2 to 23 km Surface Met. Range
2.0	1.06E-03	2.53E-03	1.06E-03	2.53E-03
3.0	5.38E-04	1.28E-03	6.89E-04	1.63E-04
4.0	3.03E-04	5.66E-04	4.81E-04	8.73E-04
5.0	2.29E-04	2.29E-04	4.39E-04	4.39E-04
6.0	1.67E-04	1.67E-04	3.64E-04	3.64E-04
7.0	1.08E-04	1.08E-04	2.93E-04	2.93E-04
8.0	6.65E-05	6.65E-05	1.58E-04	1.58E-04
9.0	4.62E-05	4.62E-05	8.54E-05	8.54E-05

	Background Stratosphere	Moderate Volcanic	High Volcanic	Extreme Volcanic	Background Stratospheric	Moderate Volcanic	High Volcanic	Extreme Volcanic
9.0	4.62E-05	4.62E-05	4.62E-05	4.62E-05	8.54E-05	8.54E-05	8.54E-05	8.54E-05
10.0	4.66E-11	7.27E-05	2.46E-04	2.46E-04	6.75E-11	9.75E-05	2.68E-04	2.68E-04
11.0	4.22E-11	9.44E-05	3.34E-04	3.34E-04	4.73E-11	1.11E-04	3.05E-04	3.05E-04
12.0	3.92E-11	1.16E-04	4.70E-04	4.70E-04	3.80E-11	1.29E-04	3.54E-04	3.54E-04
13.0	3.68E-11	1.45E-04	6.54E-04	6.54E-04	3.06E-11	1.48E-04	4.05E-04	4.05E-04
14.0	3.82E-11	1.52E-04	9.26E-04	9.26E-04	2.62E-11	1.52E-04	5.22E-04	5.22E-04
15.0	3.80E-11	1.54E-04	1.13E-03	1.46E-03	2.34E-11	1.54E-04	7.57E-04	7.57E-04
16.0	3.79E-11	1.44E-04	1.36E-03	3.40E-03	2.26E-11	1.44E-04	1.17E-03	1.17E-03
17.0	3.56E-11	1.30E-04	1.55E-03	8.83E-03	2.51E-11	1.30E-04	1.74E-03	1.84E-03
18.0	3.33E-11	1.11E-04	1.59E-03	1.45E-02	3.08E-11	1.11E-04	2.20E-03	3.36E-03
19.0	2.91E-11	9.01E-05	1.24E-03	5.79E-03	3.44E-11	9.01E-05	2.21E-03	7.02E-03
20.0	2.50E-11	7.12E-05	7.38E-04	1.32E-03	3.49E-11	7.12E-05	1.69E-03	1.45E-02
21.0	2.08E-11	5.75E-05	3.91E-04	4.53E-04	2.98E-11	5.75E-05	1.03E-03	7.96E-03
22.0	1.75E-11	4.53E-05	2.11E-04	2.11E-04	2.49E-11	4.53E-05	6.51E-04	8.83E-04
23.0	1.43E-11	3.48E-05	1.29E-04	1.29E-04	1.78E-11	3.48E-05	3.47E-04	3.47E-04
24.0	1.12E-11	2.71E-05	8.39E-05	8.39E-05	1.17E-11	2.71E-05	1.85E-04	1.85E-04
25.0	8.88E-12	2.16E-05	5.93E-05	5.93E-05	7.75E-12	2.16E-05	1.12E-04	1.12E-04
26.0	6.80E-12	1.69E-05	4.63E-05	4.63E-05	5.33E-12	1.69E-05	6.44E-05	6.44E-05
27.0	5.30E-12	1.32E-05	3.63E-05	3.63E-05	4.01E-12	1.32E-05	4.20E-05	4.20E-05
28.0	3.96E-12	1.11E-05	3.04E-05	3.04E-05	3.07E-12	1.11E-05	3.04E-05	3.04E-05
29.0	3.08E-12	6.54E-06	1.79E-05	1.79E-05	2.44E-12	6.54E-06	1.79E-05	1.79E-05
30.0	1.96E-12	4.01E-06	4.01E-06	4.01E-06	1.96E-12	4.01E-06	4.01E-06	4.01E-06

	Normal Upper Atmos.	Transition from Volcanic to Normal	Transition from Volcanic to Extreme	Extreme Upper Atmos.
35.0	8.35E-08	1.24E-07	3.64E-07	2.15E-07
40.0	4.05E-08	4.05E-08	3.52E-07	2.83E-07
45.0	2.03E-08	2.03E-08	3.34E-07	3.34E-07
50.0	1.06E-08			2.55E-07
55.0	5.51E-09			1.72E-07
60.0	2.92E-09			1.16E-07
65.0	1.54E-09			8.20E-08
70.0	8.10E-10			5.21E-08
75.0	3.52E-10			3.39E-08
80.0	1.47E-10			2.18E-08
85.0	6.07E-11			1.41E-08
90.0	2.58E-11			7.84E-09
95.0	1.09E-11			4.20E-09
100.0	4.71E-12			2.28E-09

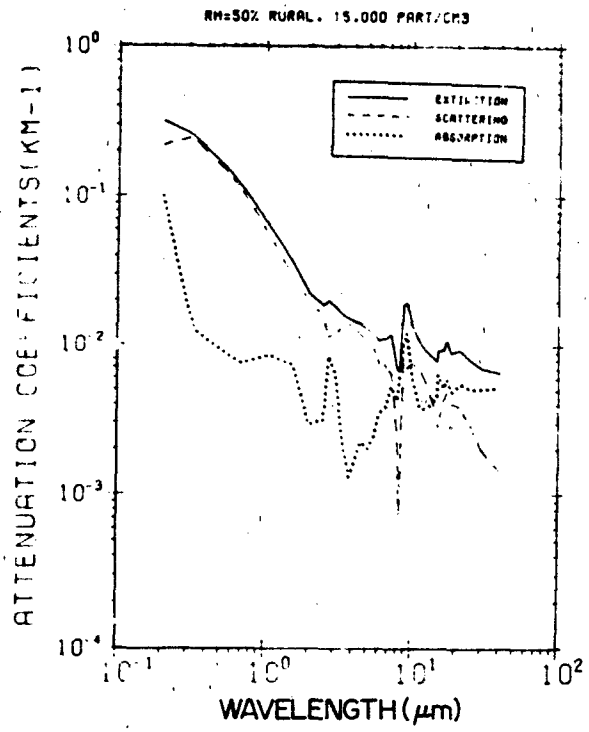


Figure 18-14. Attenuation coefficients vs wavelength for the rural aerosol model at 50% relative humidity.

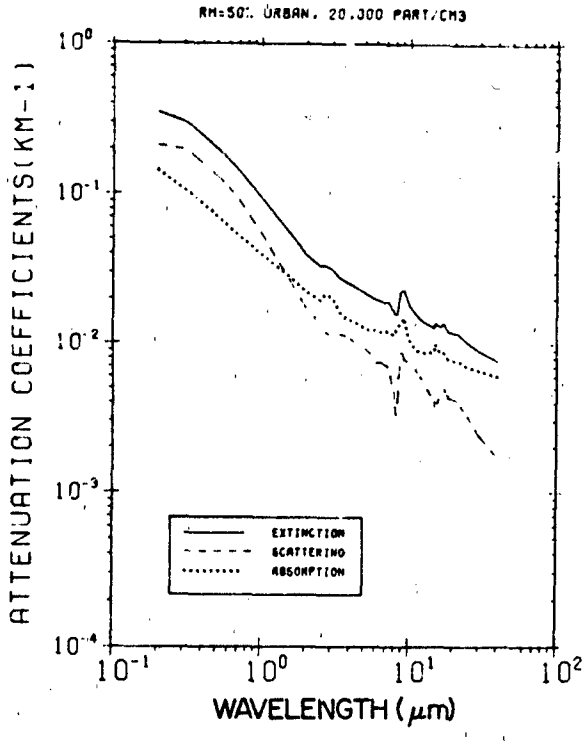


Figure 18-16. Attenuation coefficients vs wavelength for the urban aerosol model at 50% relative humidity.

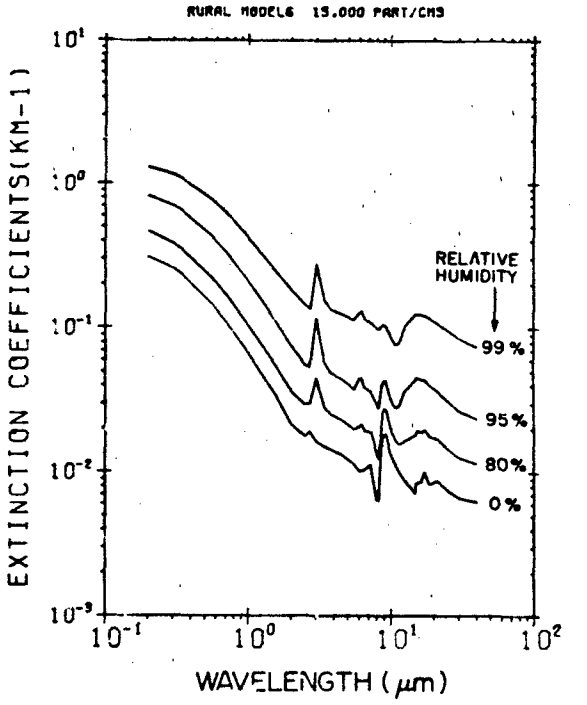


Figure 18-15. Extinction coefficients vs wavelength for the rural aerosol model for different relative humidities and constant number density of particles.

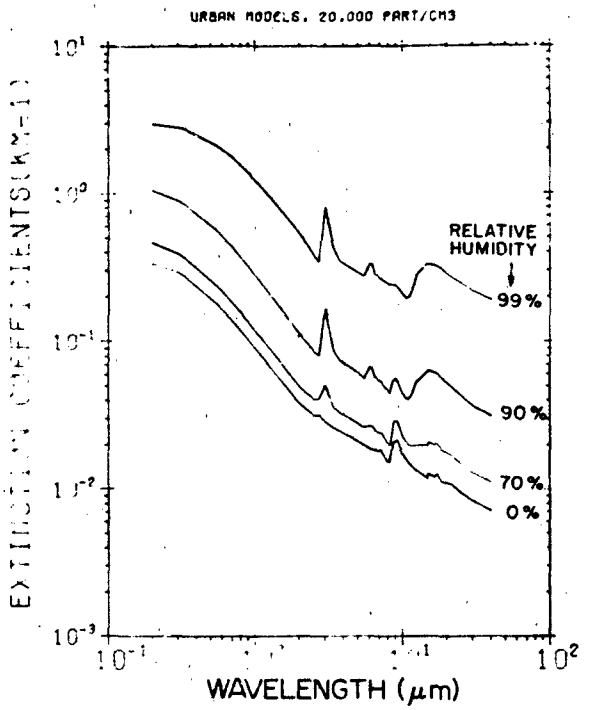


Figure 18-17. Extinction coefficients vs wavelength for the urban aerosol model for different relative humidities and constant number density of particles.

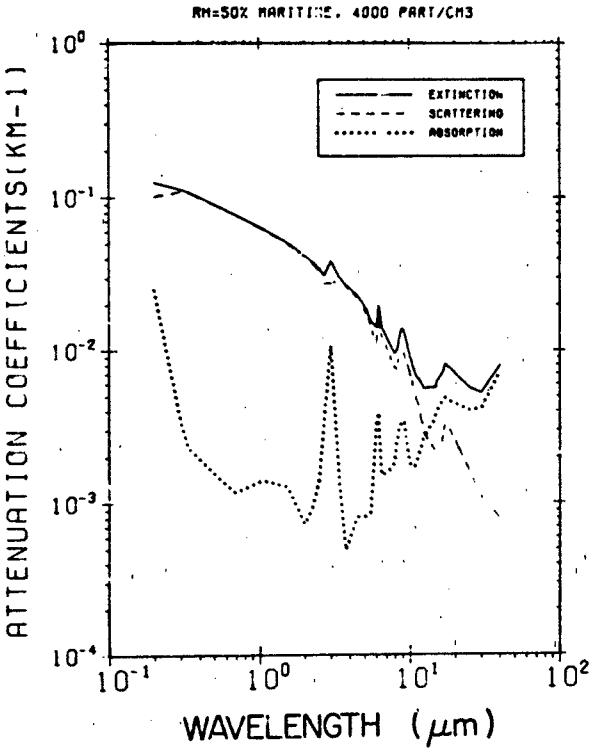


Figure 18-18. Attenuation coefficients vs wavelength for the maritime aerosol model at 50% relative humidity.

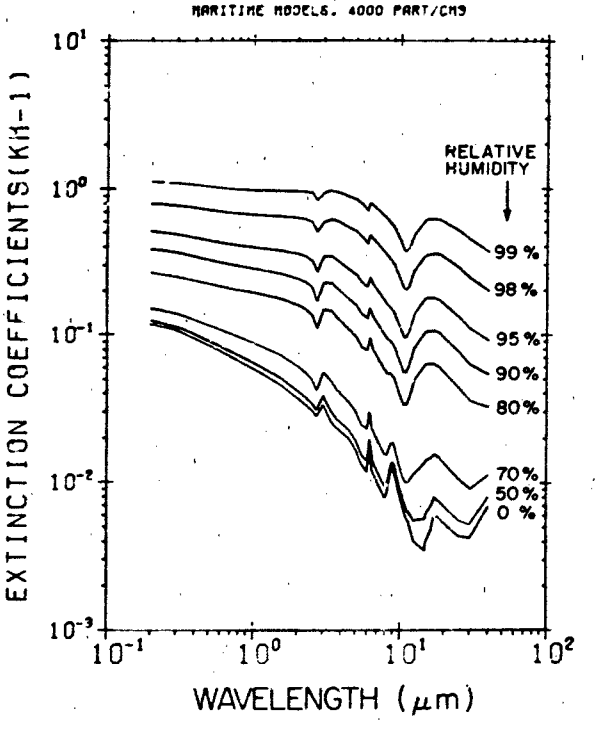


Figure 18-19. Extinction coefficients vs wavelength for the maritime aerosol model for the different relative humidities and constant number density of particles.

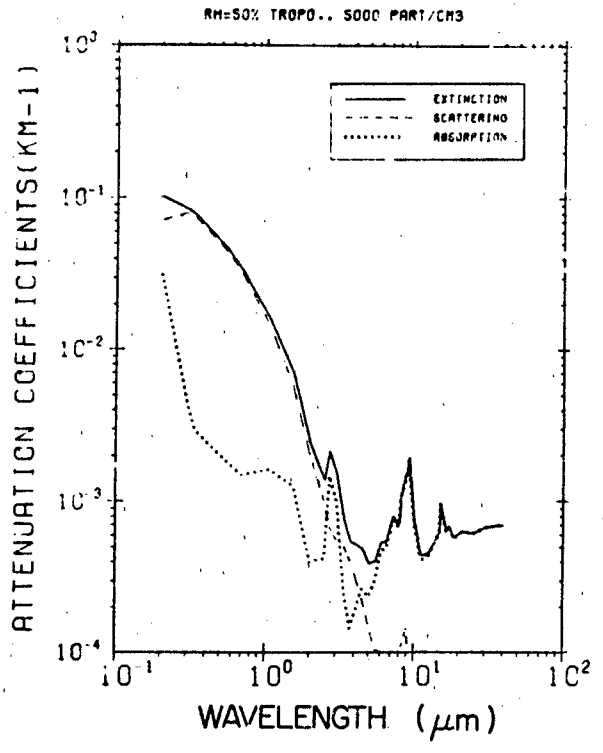


Figure 18-20. Attenuation coefficients vs wavelength for the tropospheric aerosol model at 50% relative humidity.

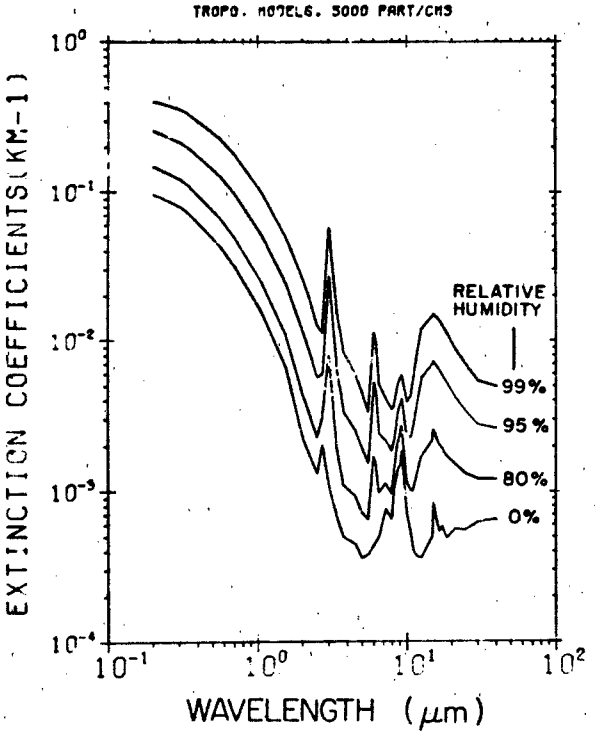


Figure 18-21. Extinction coefficients vs wavelength for the tropospheric aerosol model for different relative humidities and constant number density of particles.

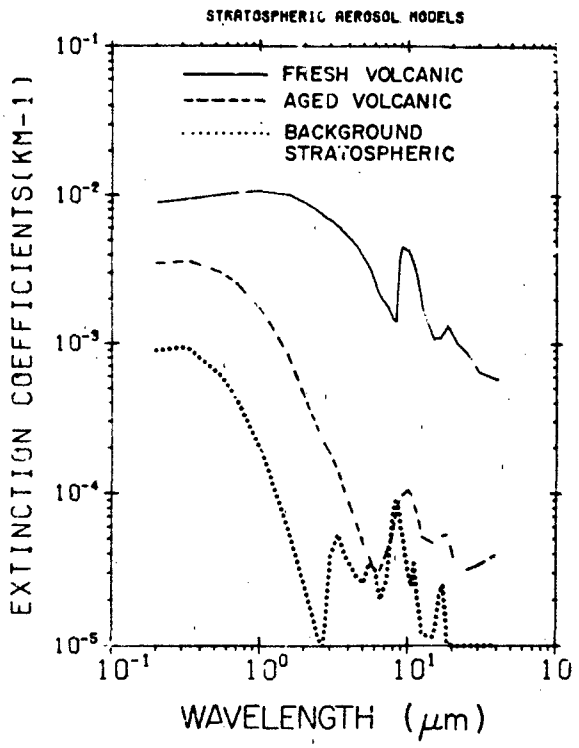


Figure 18-22. Extinction coefficients for the different stratospheric aerosol models (background stratospheric, volcanic, and fresh volcanic). The extinction coefficients have been normalized to values around peak levels for these models (see also Figure 18-13).

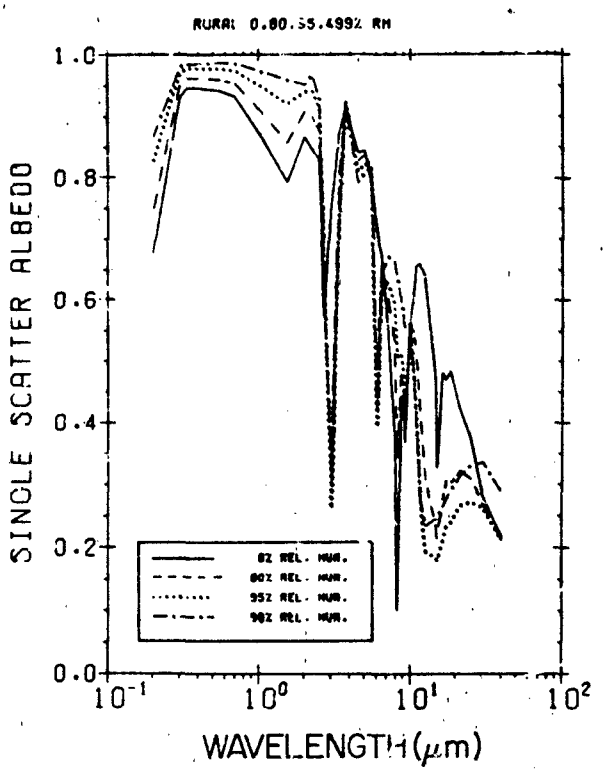


Figure 18-23. Single scattering albedo for the rural aerosol model for different relative humidities.

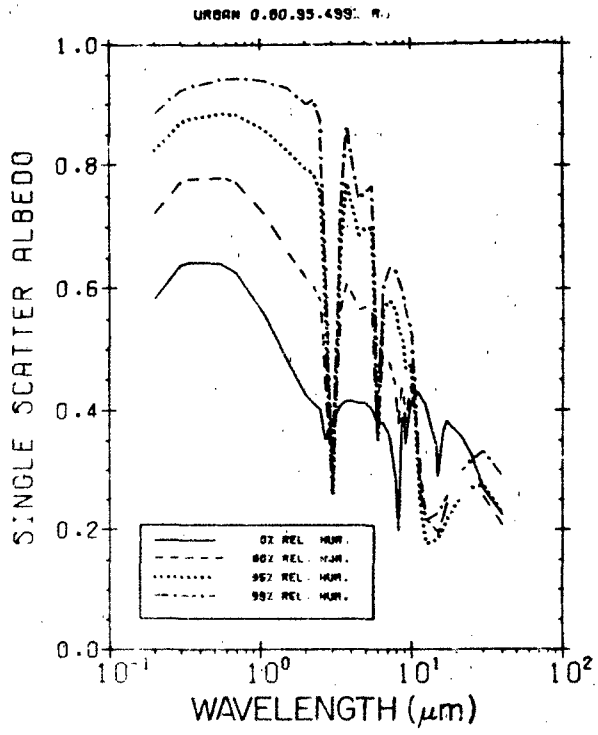


Figure 18-24 Single scattering albedo for the urban aerosol model for different relative humidities.

relative humidities between 0% and 99% where the number density is fixed at 15 000 particles/cm^3, which corresponds to a meteorological range of about 25 km for the dry aerosols and about 5 km at 99% relative humidity.

The attenuation coefficients for the urban aerosol model as a function of wavelength are shown in Figures 18-16 and 18-17 for different humidities. The corresponding results for the maritime and tropospheric aerosol models are shown in Figures 18-18 and 18-19 and Figures 18-20 and 18-21, respectively.

The extinction coefficients for the background stratospheric and volcanic aerosol models are shown in Figure 18-22. Frequently the ratio of the scattering coefficient to total extinction k_s/k_e (single scatter albedo) is used as a measure for the relative contribution of scattering and absorption to the total extinction. Figures 18-23 through 18-27 give the single scatter albedo for the aerosol models discussed above. Notice the rapid change in the relative importance of aerosol scattering and absorption in the middle infrared spectral region. This is due to two factors, an increase in the absorption because of an increase in the imaginary part of the refractive index in the infrared, and a decrease in the contribution from scattering as the wavelength becomes significantly larger than the size of the aerosols.

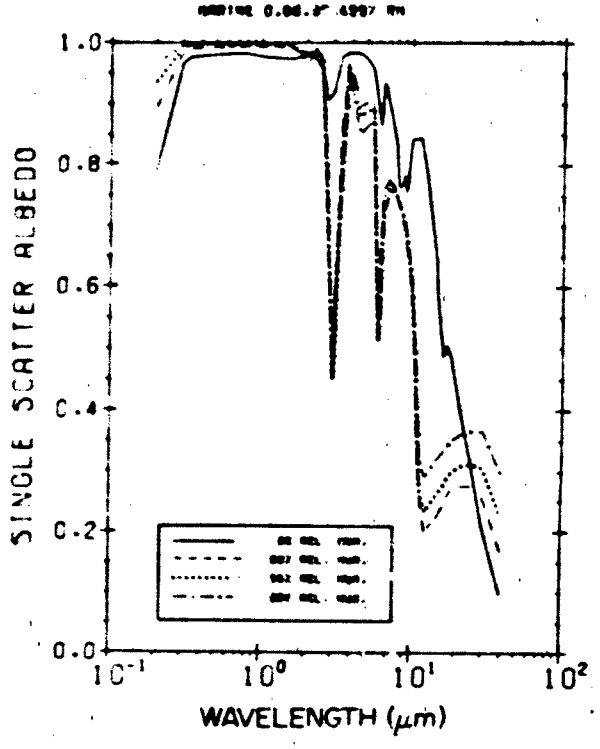


Figure 18-25 Single scattering albedo for the maritime aerosol model for different relative humidities.

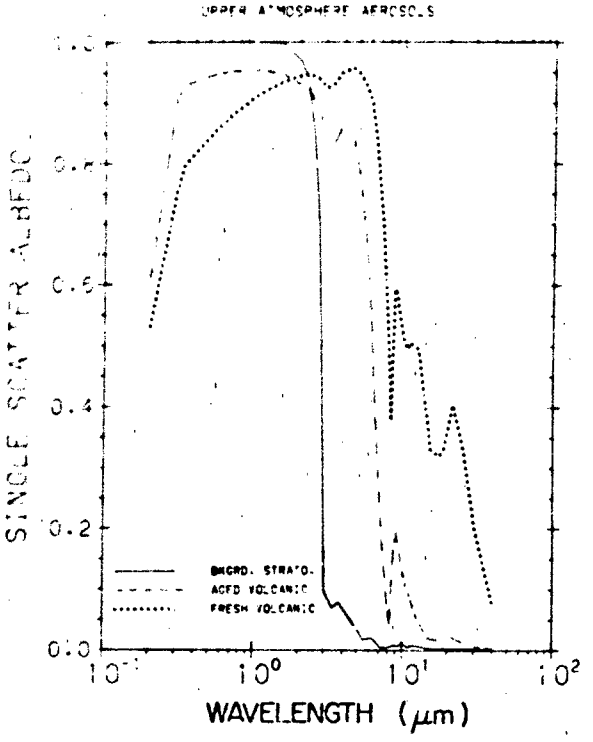


Figure 18-27. Single scattering albedo for the different stratospheric aerosol models versus wavelength.

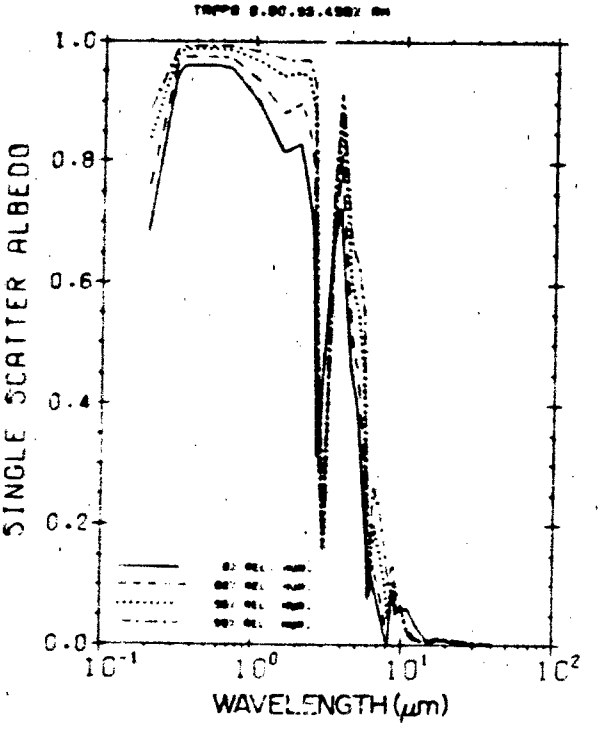


Figure 18-26 Single scattering albedo for the tropospheric aerosol model for different relative humidities.

There are also some noticeable differences between the angular scattering functions of these various aerosol models. Figures 18-28 a, b and 18-29 a, b give examples for two wavelengths 1.06 and 10.6 μm. Both figures give the phase function, defined such that the integral from 0° to 180° becomes 1. For many purposes (see Section 18.5.1) a convenient one-parameter representation for the angular dependence of scattering by aerosols is the asymmetry parameter g. It is defined as the cosine weighted average of the phase function $P(\theta)$:

$$g = \frac{\int_{-1}^{+1} \cos\theta \, P(\theta) \, d(\cos\theta)}{\int_{-1}^{+1} P(\theta) \, d(\cos\theta)} . \qquad (18.22)$$

where $P(\theta)$ is the differential probability of scattering at an angle θ. The asymmetry factor for the various aerosol models is shown in Figures 18-30 to 18-34.

Any quantitative treatment of light scattering must include the polarization of light since in general light becomes polarized in the scattering process, and the angular distribution of scattered light depends on the state of polarization of the light incident on a particle. The degree of polarization P can vary between −1 and +1 (see Equation (18.18)).

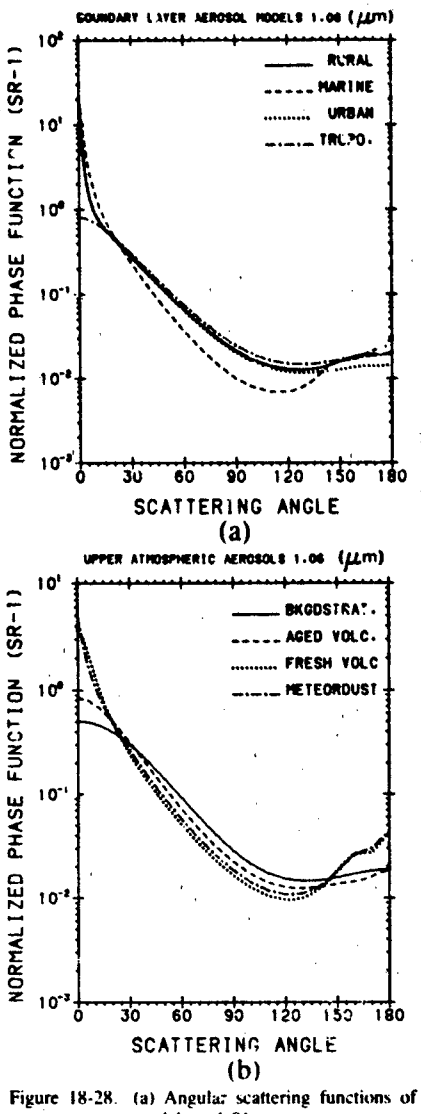


Figure 18-28. (a) Angular scattering functions of low altitude aerosol models at 1.06 μm.
(b) Angular scattering functions of high altitude aerosol models at 1.06 μm.

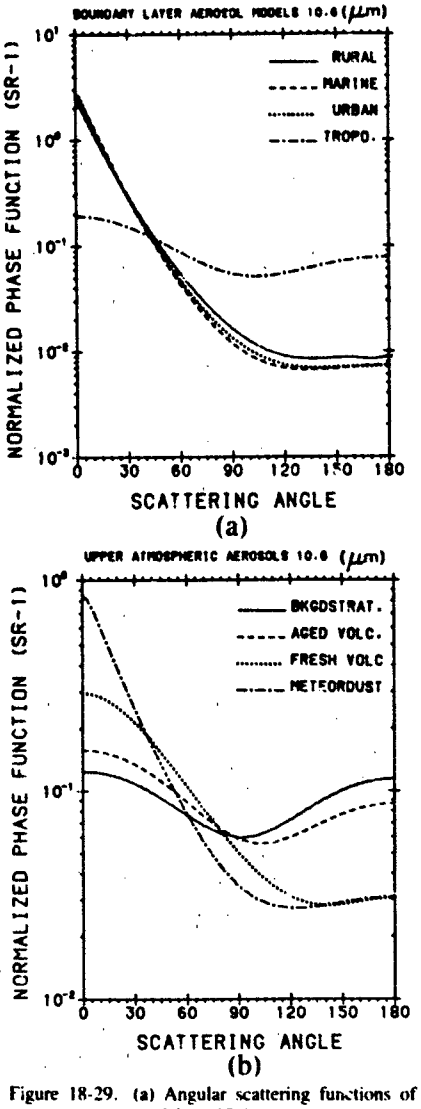


Figure 18-29. (a) Angular scattering functions of low altitude aerosol models at 10.6 μm.
(b) Angular scattering functions of high altitude aerosol models at 10.6 μm.

Figure 18-35 gives examples of the angular variation of i_1 and i_2. The polarization of light scattering is dependent on the particle size and the wavelength of the incident light.

18.2.2 Cloud, Fog, and Precipitation Particles

Cloud, fog, and precipitation particles are discussed separately because of their rather unique and special physical characteristics and life cycle that distinguish them from haze or aerosol particles. First, their composition is uniform, either water or ice, or under some circumstances a combination of both. Second, in liquid phase these particles are drop-shaped, that is, close to spherical. These two characteristics make modeling of these particles much more realistic than for aerosols. Also the formation processes and life cycle for these particles are better understood than those for aerosols. For additional information see Chapter 16. A survey on fog properties was compiled by Stewart and Essenwanger [1982].

18.2.2.1 Particle Types, Water-Ice Refractive Index, Particle Shapes. Liquid, fog, and cloud droplets are the result of condensation of water vapor on condensation nuclei (aerosol particles). This condensation occurs at relative humidities near 100% and is dependent on the physicochemical properties of the condensation nuclei. However, once the condensation has occurred the effect of the condensation nucleus on the total droplet properties becomes negligible.

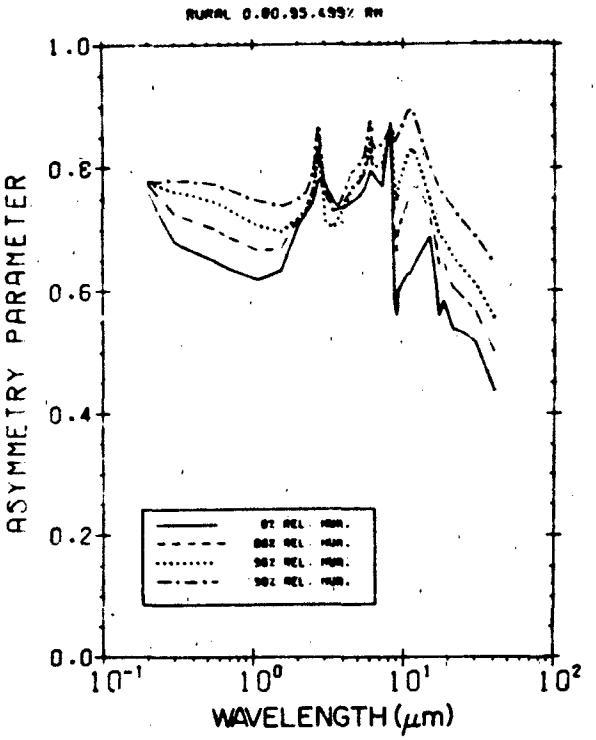


Figure 18-30. Asymmetry parameter of rural aerosol model of different relative humidities.

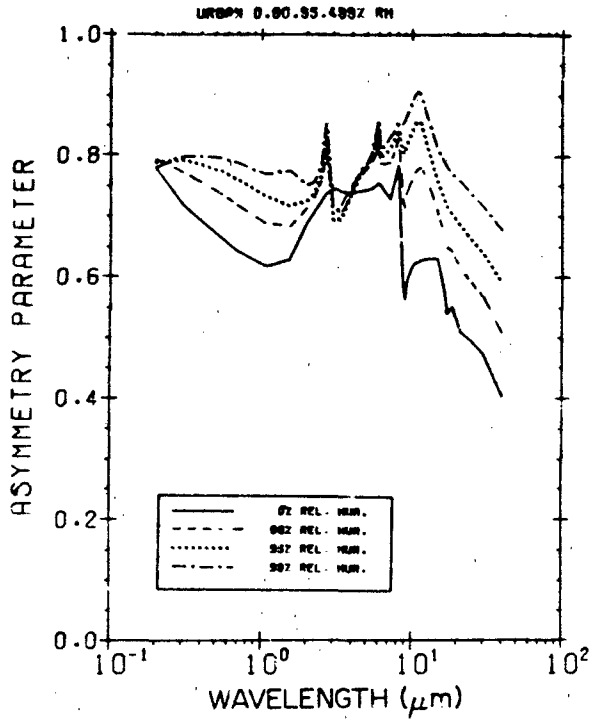


Figure 18-31. Asymmetry parameter of urban aerosol model of different relative humidities.

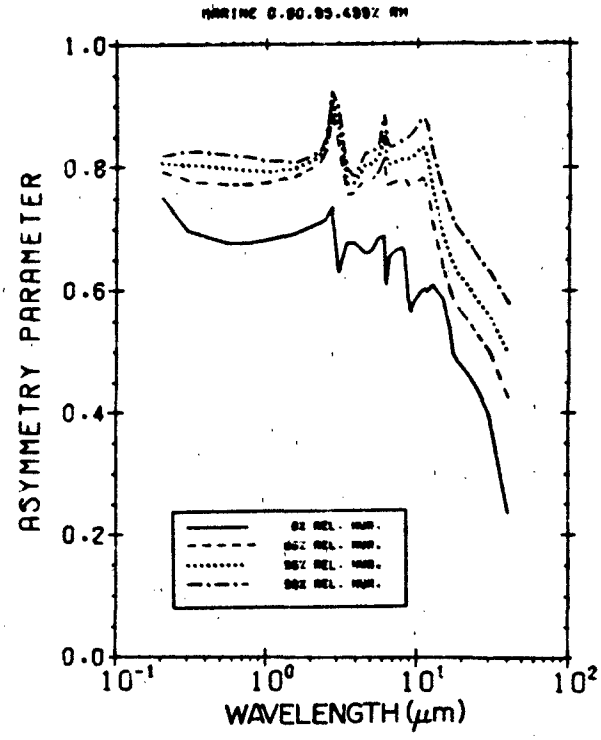


Figure 18-32. Asymmetry parameter of maritime aerosol model of different relative humidities.

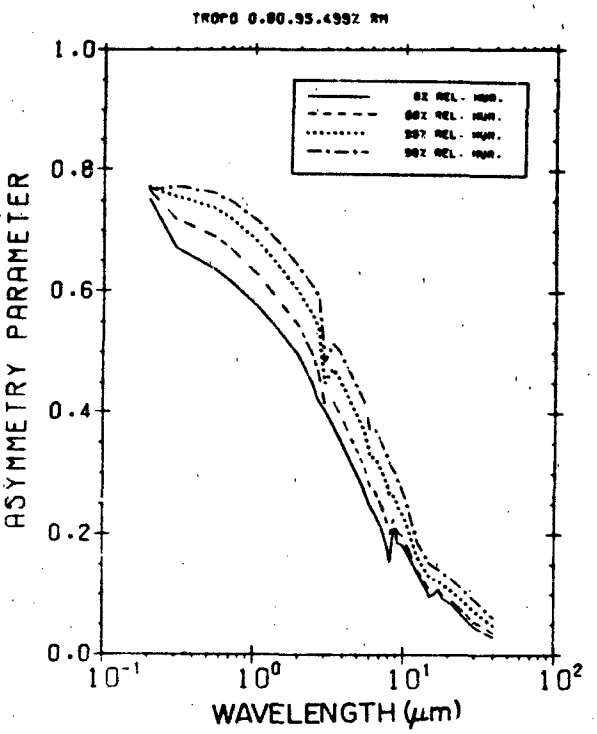


Figure 18-33. Asymmetry parameter of tropospheric aerosol model of different relative humidities.

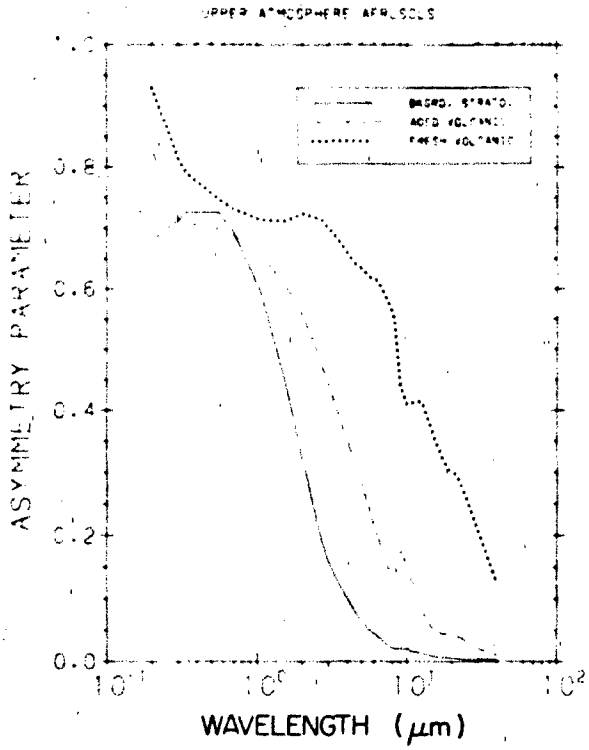


Figure 18-34. Asymmetry parameter of upper atmosphere aerosol models of different relative humidities.

The concentration and size distribution of fog and cloud particles depend in a complex way on the existing meteorological conditions. If fog or cloud droplets become so large that atmospheric buoyancy can no longer balance the gravitational forces, they will fall out as rain.

At temperatures sufficiently below freezing, ice particles will form. Depending on the conditions, they may be single ice crystals such as in ice fog or high altitude cirrus clouds, amorphous ice pellets as in sleet or hail, or complex shaped snowflakes. If ice particles fall from cold air into warmer air, their surface may be coated with a water skin. The refractive index of water at optical and infrared wavelengths is well known [Hale and Querry, 1973; Downing and Williams, 1975]; see Table 18-9.

The complex refractive index of ice in Table 18-9 was derived from several sources [Schaaf and Williams, 1973; Irvine and Pollack, 1967; Grenfell and Perovich, 1981; Hobbs, 1974; and Warren, 1984]. The detailed procedures were as follows:

A. For the real part of the refractive index:
   1. 0.2–0.4 μm: Warren [1984] using the refractive index of water for guidance.
   2. 0.4–0.7 μm: Table 3.1 in Hobbs [1974].
   3. 0.7–2.5 μm: interpolation between Hobbs [1974] data at 0.7 μm and the value of Schaaf and Williams [1973] at 2.5 μm using a two term dispersion equation follows Irvine and Pollack's [1967] data between 0.95 and 1.3 μm but deviates for longer wavelengths.

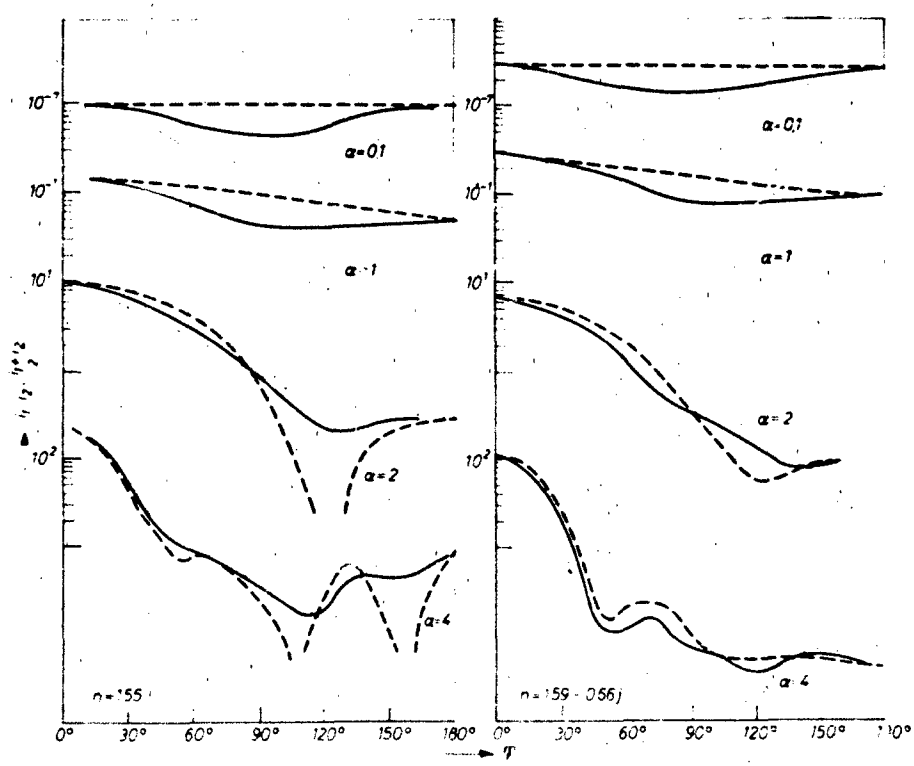


Figure 18-35. Angular dependence of i_1 ---, i_2 ... and $(i_1 + i_2)/2$— for different aerosol particles. n = complex refractive index, α = size parameter.

4. 2.5 to 25 μm: the measurements of Schaaf and Williams [1973], since more accurate than any available at the time of Irvine and Pollacks' [1967] survey.
5. 25–40 μm: Warren [1984].

B. For the imaginary part of the refractive index:
1. 0.2–0.4 μm: Warren [1984]
2. 0.4–0.95 μm: Grenfell and Perovich's [1981] measurements.
3. 0.95–1.40 μm: a weighted average between Grenfell and Perovich [1981] and Irvine and Pollack's [1967] survey.
4. 1.40–2.7 μm: Irvine and Pollack's [1967] values were used since these were based on transmission measurements that are more accurate than the reflectance measurements of Schaaf and Williams for weak absorption.
5. 2.7–25 μm: Schaaf and Williams [1973].
6. 25–40 μm: Warren [1984]

Since the original compilation of Table 18-9, a review of the optical constants of ice has been completed by Warren [1984]. His values have replaced the preliminary ones for wavelengths 0.2–0.4 μm and 25–40 μm. For most of the tabulated data (0.4–25 μm) the differences are the order of experimental uncertainties in the refractive index. For wavelengths outside this range, see Warren [1984].

Cloud and rain drops can be considered more or less perfect spheres. The shape of single ice crystals, although variable depending on conditions, is fairly well known. Under most conditions, simple snow crystals have an approximate hexagonal symmetry; however, size and shape (habit) are influenced by the temperature and humidity conditions extant during their formation and growth. The standard classification scheme is that of Magono and Lee [1966], who describe 80 crystal types and combinations. Further complicating the picture is the essentially infinite variety of agglomerates (what most people refer to as "flakes") that can exist, and the fact that any of the basic crystals can be modified by the attachment of super-cooled drops in a process called riming. For a more detailed discussion of crystal formation and habit, see Chapter 16.

18.2.2.2 Size Distributions. There have been extensive measurements of the the size of fog and cloud droplets.

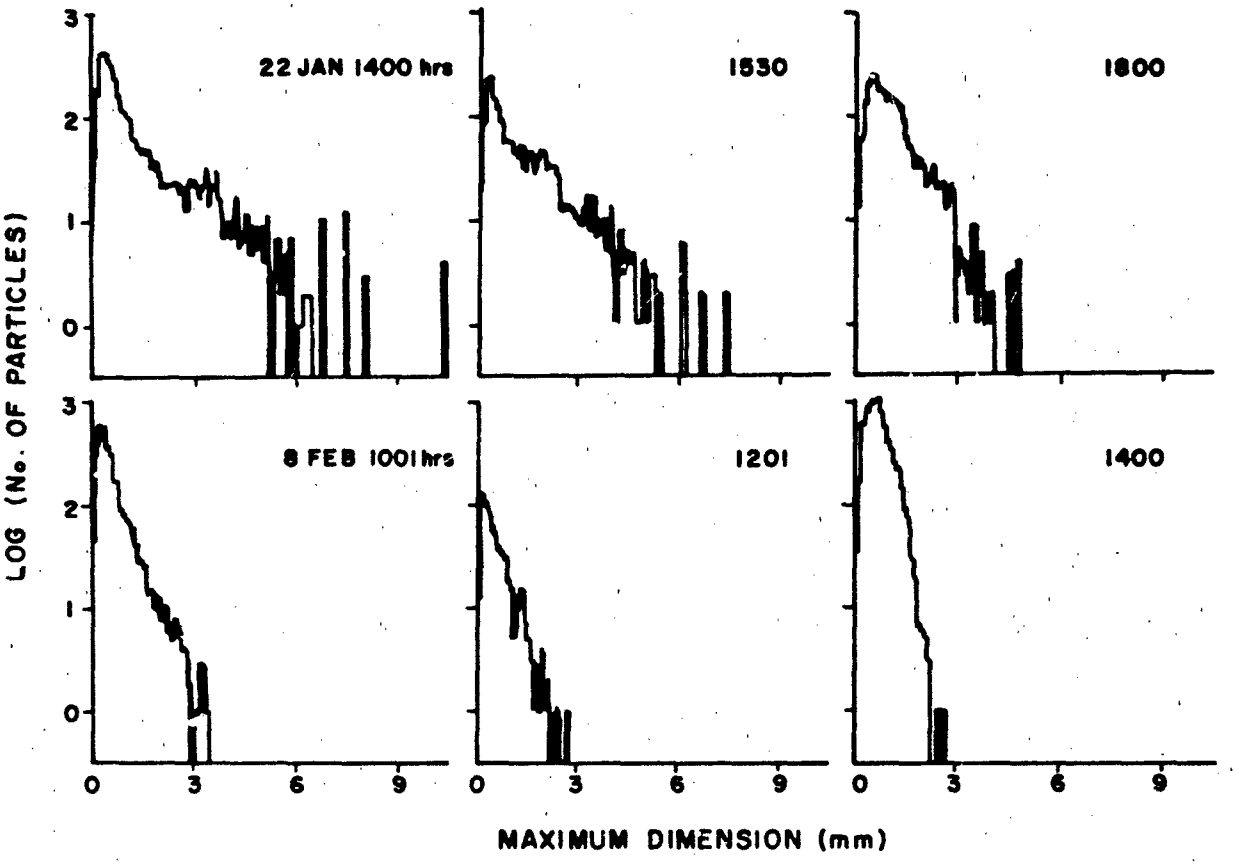


Figure 18-36. Particle size distributions for selected times during snowstorms on 22 January and 8 February 1981 near Burlington, Vermont. Each distribution is for a 5-minute sample period taken with an optical array probe [Berger, 1983].

Sufficient data exist to differentiate between drop size distributions for different types of fogs and clouds. Drop sizes from 1 to 100 μm can occur, the smaller sizes (1–10 μm range) are found in fogs and stratus type clouds, the larger sizes in convective clouds with strong updraft. Ice crystals in clouds (or ice fog) may be as small as a few micrometers although the size of cirrus cloud particles is more typically of the order of 0.5 to 1 μm. Rain drops are distinguished from cloud drops only by the fact that they have grown large (heavy) enough to fall to the surface of the earth (that is, to precipitate out). Rain drops may be as large as 5 mm, larger drops have an increasing tendency to break up. Snow flakes present an interesting problem with regard to optical effects because of their extremely complex shapes and dimensions. Various data (such as Lo and Passarelli [1982]; Berger [1983]; and Gray and Male [1981]) suggest that a "typical" snowfall consists of particles ranging from single crystals 0.1 mm (100 μm) or smaller to agglomerates several millimeters in diameter.

There is little information on the exact nature of the size distributions. Lo and Passarelli [1982] assumed a simple exponential form and found average diameters of 0.5 to 2 mm for several aircraft measurements using optical array probes. Ground-level measurements with aspirated optical array probes by Berger [1983] and Formvar replicating data of Koh and O'Brien [1982] show a wide range of size distributions; typical examples are shown in Figure 18-36. The reasons for this wide variation are not clearly understood, though ongoing modeling work by Martinez-Sanches et al. [1983] is beginning to unravel some of the causes. The majority of the distributions peak between 0.2 and 0.5 mm. The area-weighted average radius, the size measure with the most influence on extinction, is generally in the range 0.1 to 0.3 mm.

18.2.2.3 Cloud–Precipitation Models. Cloud and fog drop size distributions have been modeled by several researchers. Most commonly used are the models given by Deirmendjian [1969] for cumulus clouds and Silverman and Spraglie [1970] for a number of fog types and several types of clouds. Table 18-11 is based on Silverman's compilation. The distribution functions for models 1–12 are shown in Figures 18-37 through 18-39. Advective fog is produced by the transport of moist air over a colder surface resulting in the cooling of the surface layers below their air dew points with condensation taking place in the form of fog. Most sea fogs are advective. Both the size range of particles and the liquid water contents are large (see Models 1 and 2, Table 18-11). Radiative fog is produced when stagnant moist air becomes progressively cooler during the night due to radiative cooling. This type of fog has both a small size range of particles and a small liquid water content.

Cloud droplet size distributions are more difficult to model. However, the optical thickness of most clouds is so large that their transmittance at visible and infrared wavelengths becomes completely diffuse due to multiple scattering (see Section 18.5.1).

The size distribution of rain drops is best expressed in terms of rain rate, since this is the quantity usually measured by the weather services. This formulation goes back to experimental work by Laws and Parsons [1943] and Marshall and Palmer [1948]. According to their work the number of drops of diameter D can be expressed as

$$N_D = N_o \exp(-a \cdot D) \quad (18.23)$$

where $N_o = 0.08\ \text{cm}^{-4}$

and $a = 41\ R^{-0.21}\ \text{cm}^{-1}$

with R being the rain rate in mm/hr.

Table 18-11. Cloud models: drop size distributions and liquid water contents

$n(r) = a\, r^{\alpha} \exp[-br]$						
Cloud Type	Model Number	α	b	N_o (cm^{-4})	a	M (g m^{-3})
Heavy Advection Fog	1	3	0.3	20	0.27	0.37
Moderate Radiation Fog	2	6	3.0	200	667.5	0.02
Cumulus	3	3	0.5	250	2.604	1.00
Altostratus	4	5	1.11	400	6.268	0.41
Stratocumulus	5	5	0.8	200	0.4369	0.55
Nimbostratus	6	1	0.333	100	11.089	0.27
Stratus	7	3	0.667	250	8.247	0.42
Stratus	8	2	0.6	250	27.00	0.29
Stratus-Stratocumulus	9	2	0.75	250	52.734	0.15
Stratocumulus	10	2	0.5	150	9.375	0.30
Nimbostratus	11	2	0.425	200	7.676	0.65
Cumulus-Cumulus Congestus	12	2	0.328	80	1.4115	0.57

M = Liquid water content
N = Total number of particles per unit volume

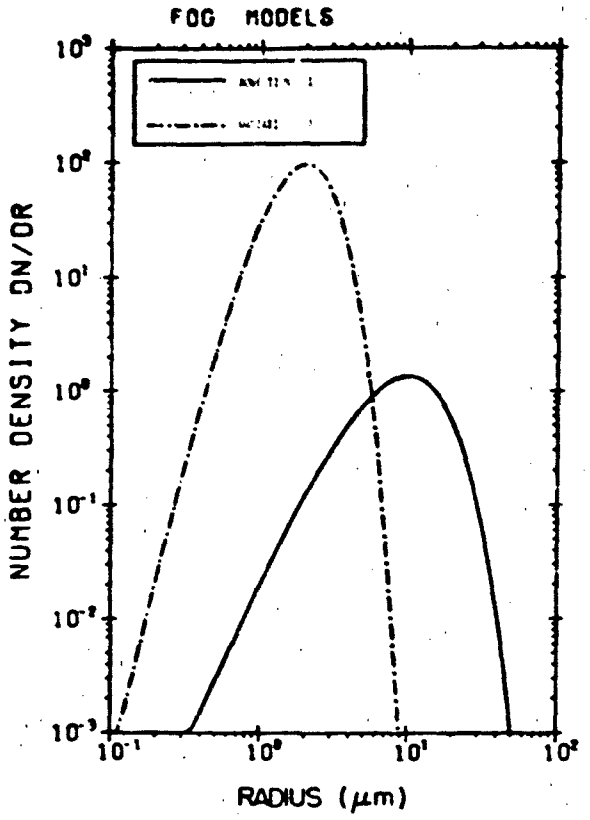


Figure 18-37. The drop size concentration (cm $^{-3}$ μm $^{-1}$) to particle drop radius (μm) for fog models 1–2 in Table 18-11.

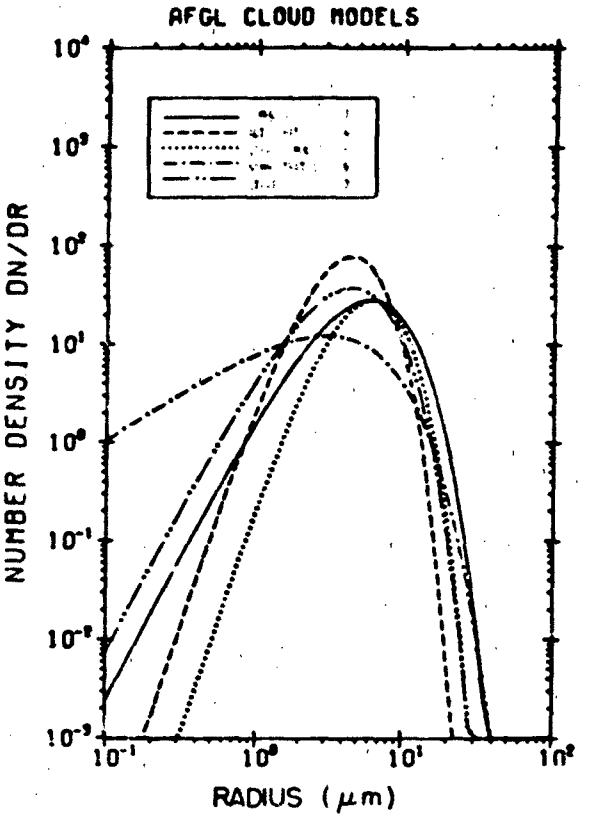


Figure 18-38. The drop size concentration (cm $^{-3}$ μm $^{-1}$) to particle drop radius (μm) for cloud models 3–7 in Table 18-11.

This relationship is strictly empirical and does not explain the physical significance of the distributions.

18.2.2.4 Extinction, Scattering, Polarization. The attenuation coefficients for the fog models given in Section 18.2.2.3 are presented in Figures 18-40 and 18-41. For the droplet concentrations given in Table 18-11 (20 and 200 for Model 1 and 2, respectively), the resulting meteorological ranges (visibility) are 130m and 450m. However, the models can be scaled to visibilities from less than 50m to over 1 km. The extinction coefficients for different cloud models are shown in Figures 18-42 and 18-43. The single scatter albedo and the asymmetry parameters are given in Figures 18-44 through 18-49 to describe further the scattering properties of fogs and clouds. Because of the much larger size of fog and cloud droplets compared to aerosol (haze) particles, the polarization of light scattered by these particles is very small and has little practical importance.

No theory presently exists to calculate exactly the extinction coefficient for complex shapes such as snow crystals. Mie theory in the geometric limit (the minimum Mie size parameter at 10 μm wavelength is about 30) suggests that extinction in snow (and rain) should be nearly independent of wavelength in the visible and infrared. Transmissometer measurements over the past few years by a number of investigators [summarized by Seagraves and Ebersole, 1983], however, have found that the measured infrared extinction coefficient varies from 1.0 to 1.45 times that in the visible, generally increasing with wavelength, and depending somewhat on the type of instrument used.

Mill and Shettle [1983] have shown that most of the apparent wavelength dependence can be explained by instrumental effects. The scattering phase functions of particles that are large compared to the wavelength are characterized by extremely narrow forward lobes due to Fraunhofer diffraction. For snow crystals with average areas equivalent to a sphere of 0.1 to 0.3 mm radius, this peak is comparable to the field-of view (FOV) of typical transmissometer receivers. Therefore, for a given transmissometer FOV, the longer the wavelength the less forward-scattered radiation is sensed resulting in a lower "apparent" transmission. The relationship between the measured transmittance τ^m and the true radiance transmittance τ is given by [Mill and Shettle, 1983]

$$\tau^m = \tau(1 + D' \ln \tau) \qquad (18.24)$$

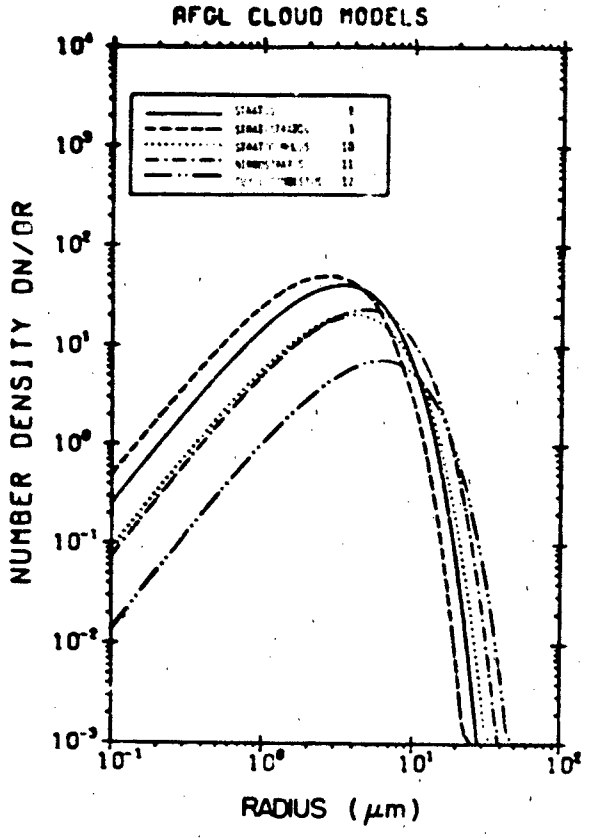


Figure 18-39. The drop size concentration (cm^{-1} μm^{-1}) to particle drop radius (μm) for cloud models 8–12 in Table 18-11.

where

$$D' = x^2 \int_0^{\phi} \int_0^{\pi} \left\{ \frac{J_1[x \sin(\theta' + \phi')]}{x \sin(\theta' + \phi')} \right\}^2 f(\theta')\, d\theta' d\phi' \qquad (18.25)$$

θ = half angle of receiver field of view

ϕ = half angle of source

x = $2\pi r/\lambda$, the particle size parameter

r = particle radius

λ = wavelength

J_1 = Bessel function of the first kind, first order, and

$f(\theta)$ = relative response of the detector over the FOV. ($\equiv 1$ for flat detector response)

This equation assumes that multiple scattering is negligible, θ and ϕ are small (less than 5°), the phase function is adequately represented by the Fraunhofer diffraction, and the size distribution can be represented by a single effective scattering radius. Values of D' have been tabulated [Shettle

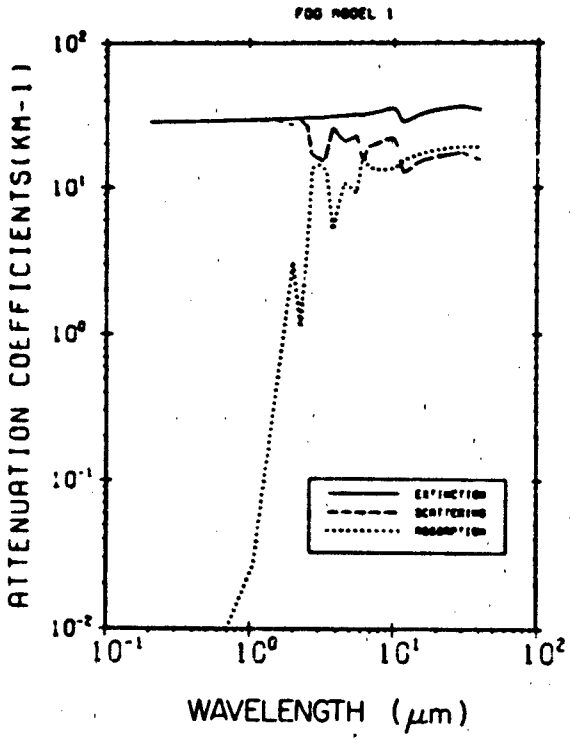


Figure 18-40. Attenuation coefficient versus wavelength: heavy advection fog model 1.

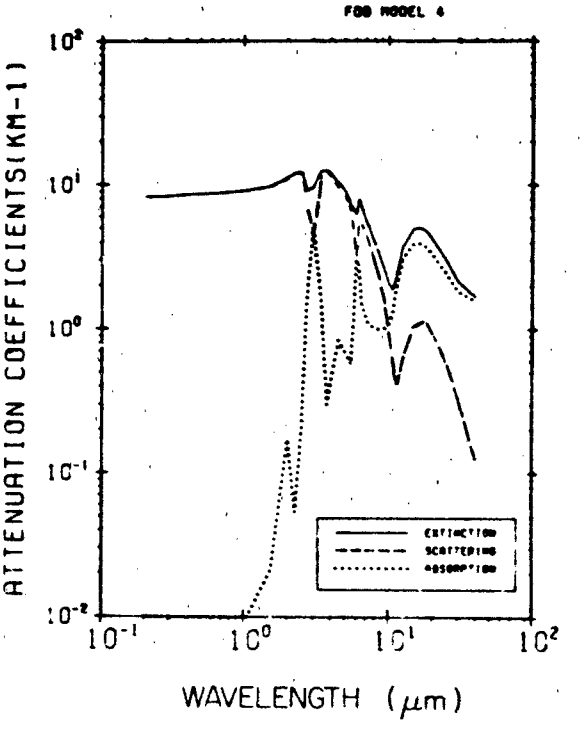


Figure 18-41. Attenuation coefficient vs wavelength, light to moderate radiation fog model 2.

et al., 1985] for various θ, ϕ, and x's and are summarized in Table 18 12. To use the table, calculate the two table parameters as follows:

$$\rho = \text{Max}(\theta, \phi)/\text{Min}(\theta, \phi) \qquad (18.26)$$

$$v = x \cdot \text{Min}(\theta, \phi), \text{ where } \theta \text{ and } \phi \text{ are in radians,}$$

where the expression *Max* means use the largest of the arguments in the parentheses. Similarly, *Min* means use the smallest. Max and Min are, respectively, the larger and smaller of the two arguments. A comparison of uncorrected and diffraction-corrected measured transmittances is shown in Figures 18-50 and 18-51. Seagraves [1983] has compared measured visible and 10.5 μm transmittances to those calculated from geometric optics theory. The calculations assumed a modified gamma size distribution of particle sizes between 84 and 761 μm radius based on replicator data, and used the airborne snow mass concentrations measured by CRREL [see Lacombe, 1983]. These studies strongly suggest that the extinction coefficient in snow is very nearly wavelength independent in the visible and infrared, though particle shape may have a small effect below the resolution of the measurement precision and cannot be ruled out.

Studies by a number of investigators have addressed the question of the relationship between snow fall rate and extinction. Results summarized in Figure 18-52 vary widely

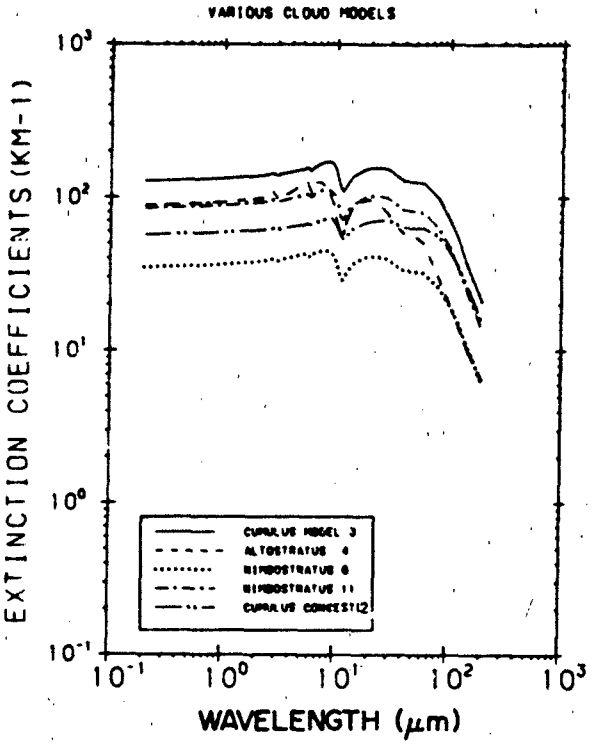


Figure 18-43. Extinction coefficients for cloud models 3, 4, 6, 11, 12 (Table 18-11).

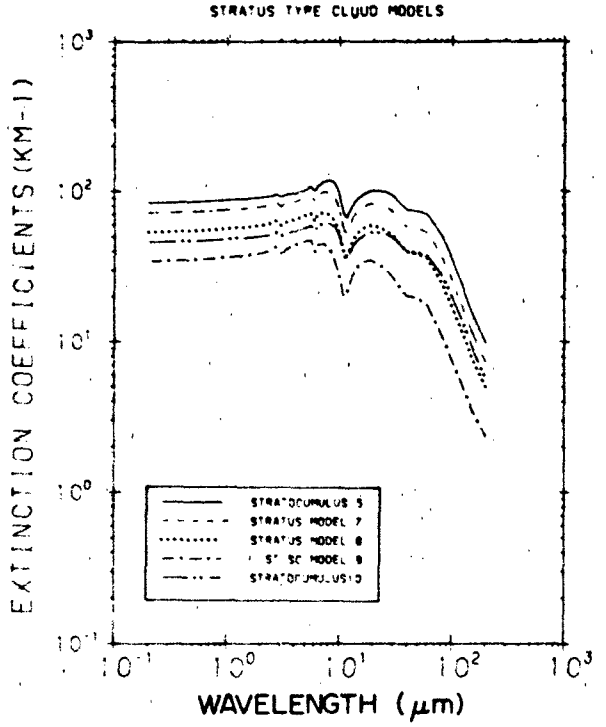


Figure 18-42. Extinction coefficients for stratus cloud models 5, 7, 8, 9, 10 (Table 8-11).

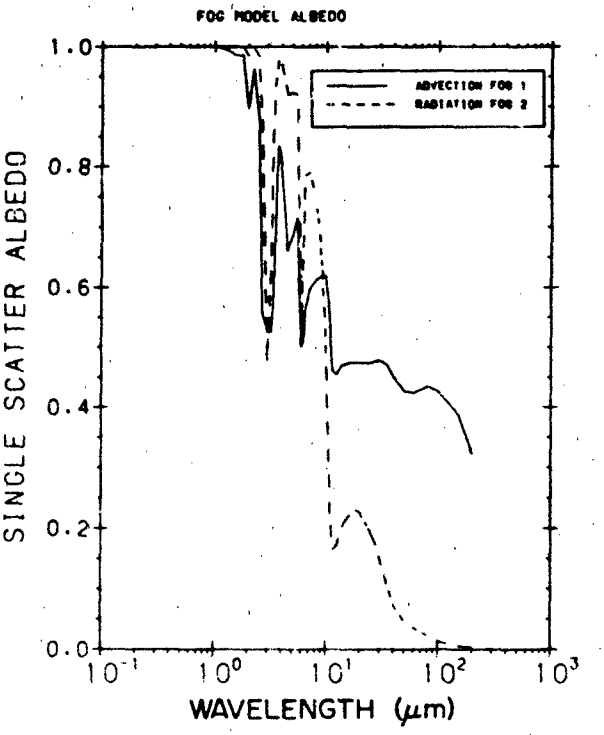


Figure 18-44 Single scatter albedo for fog models 1 and 2 (Table 18-11).

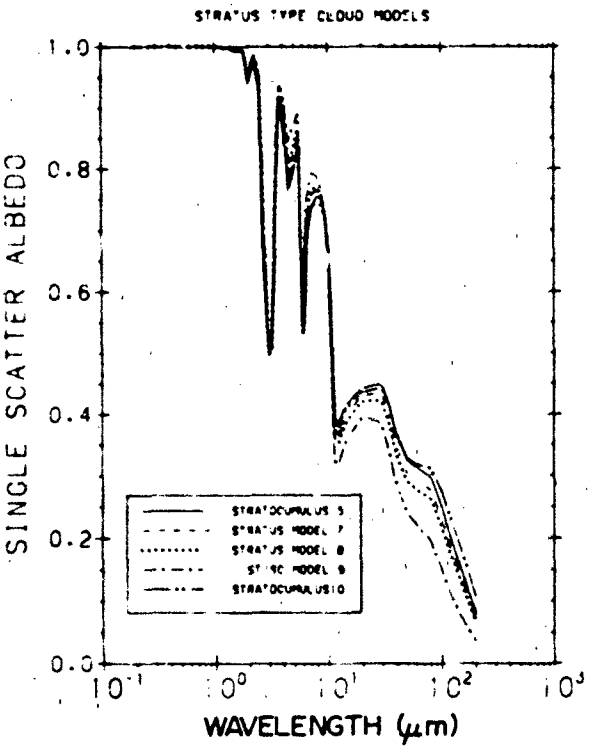


Figure 18-45. Single scatter albedo for stratus cloud models 5, 7, 8, 9, 10 (Table 18-11).

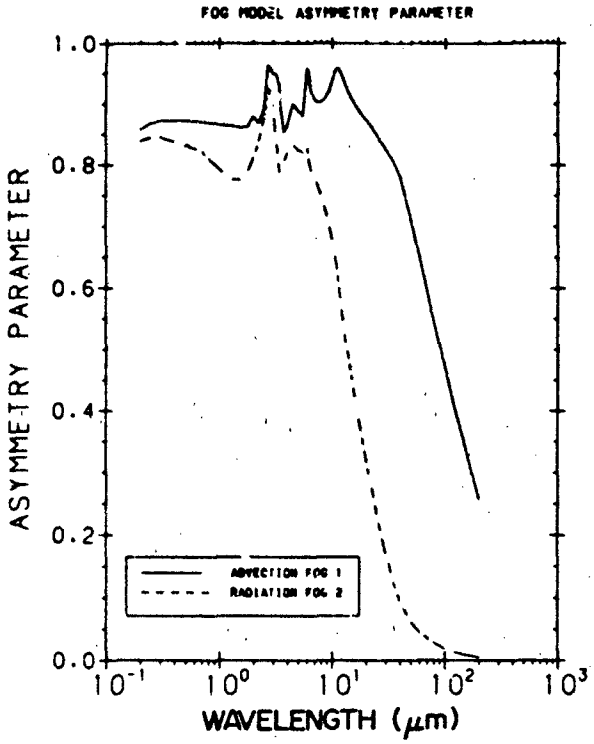


Figure 18-47. Asymmetry parameters for fog models 1 and 2 (Table 18-11).

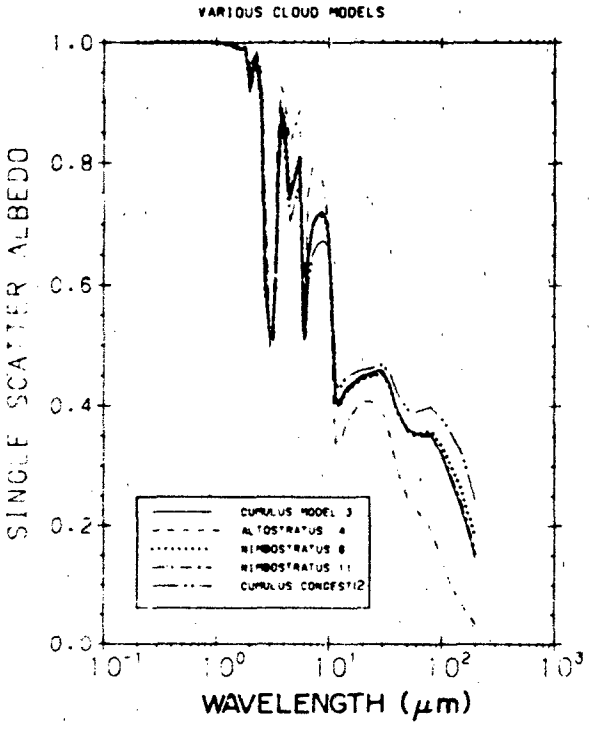


Figure 18-46. Single scatter albedo for cloud models 3, 4, 6, 11, 12 (Table 18-11).

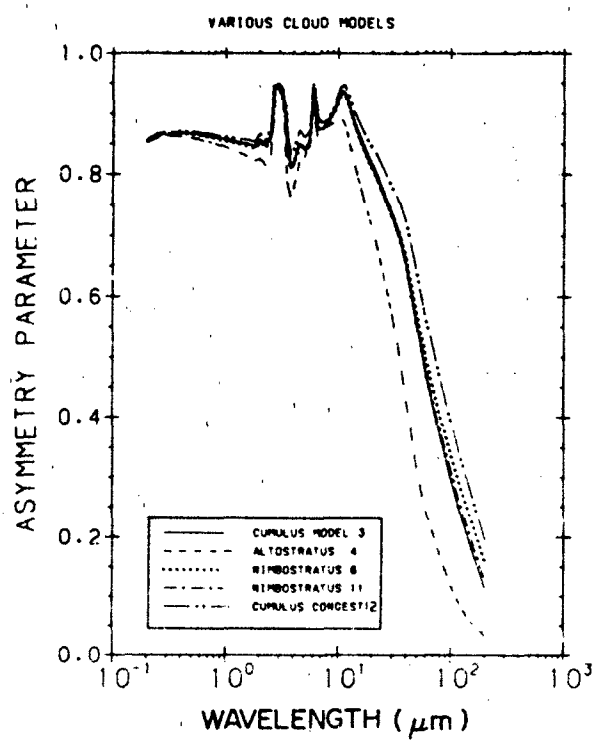


Figure 18-48. Asymmetry parameters for stratus cloud models 5, 7, 8, 9, 10 (Table 18-11).

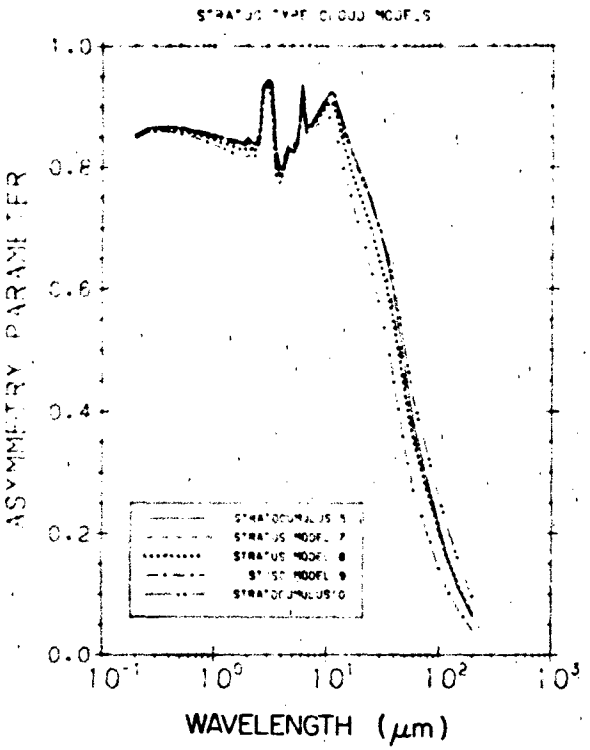


Figure 18-49. Asymmetry parameters for cloud models, 3, 4, 6, 11, 12 (Table 18-11).

Table 18-12. Forward scattering correction factors D' (x, θ, ϕ) for a transmissometer with flat receiver field of view, where ρ = max (θ,ϕ)/min (θ,ϕ) and v = x − min (θ/ϕ), for θ and ϕ in radians.

v \ ρ	100.00	50.00	20.00	10.00	5.00	2.00	1.00
0.04	0.016	0.015	0.008	0.004	0.002	0.001	0.000
0.06	0.025	0.024	0.016	0.009	0.004	0.002	0.001
0.10	0.041	0.040	0.036	0.023	0.012	0.005	0.002
0.15	0.061	0.060	0.059	0.046	0.026	0.011	0.006
0.20	0.080	0.080	0.078	0.070	0.045	0.019	0.010
0.30	0.116	0.116	0.115	0.113	0.089	0.042	0.022
0.40	0.150	0.150	0.149	0.146	0.133	0.072	0.038
0.60	0.210	0.210	0.210	0.208	0.204	0.143	0.081
1.00	0.304	0.304	0.304	0.303	0.299	0.276	0.189
1.50	0.378	0.378	0.378	0.378	0.376	0.364	0.310
2.00	0.418	0.418	0.418	0.418	0.417	0.405	0.381
3.00	0.446	0.446	0.446	0.446	0.446	0.440	0.423
4.00	0.455	0.455	0.455	0.455	0.455	0.452	0.438
6.00	0.468	0.468	0.468	0.468	0.468	0.468	0.461
10.00	0.475	0.475	0.475	0.475	0.475	0.475	0.474
20.00	0.476	0.476	0.476	0.476	0.475	0.476	0.476

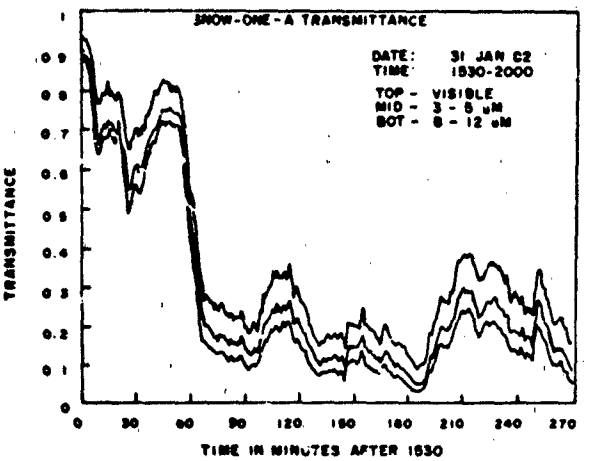


Figure 18-50. Transmittance in the visible, 3–5 μm and 8–12 μm bands during a 31 January 1982 snowstorm near Burlington, Vermont. The apparent increasing extinction with wavelength is typical of measurements in snow [Mill and Shettle, 1983].

due perhaps to both measurement error, particularly in the snow rate data, and to crystal habit or size distribution. Recent results of Lacombe [1983] show that for a given airborne snow mass concentration visible extinction decreases with increasing riming (Figure 18-53) and appears to be greater for plate-like crystals (spacial and plane dendritic, broad branched crystals and plates) than for the more compact types (columns, bullets, side planes, needles and assemblages of plates).

The most extensive set of snow phase function measurements have been made at Michigan Technological University [Winchester et al., 1981]. A typical example is shown in Figure 18-54. These measurements are of high angular resolution but required one to two hours per phase function.

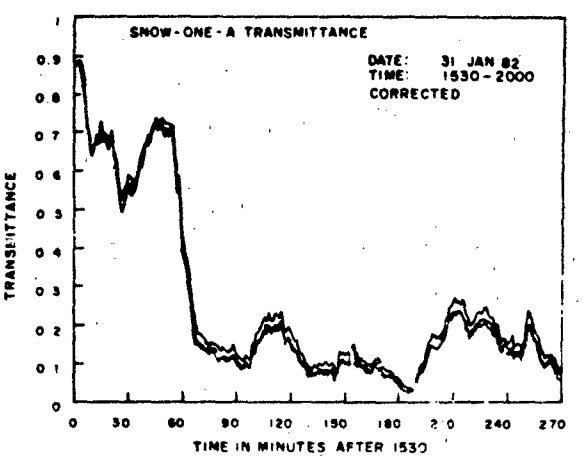


Figure 18-51. Data from Figure 18-44, corrected for forward scatter into the transmissometer field of view, assuming an effective scattering radius of 0.2 mm. Residual wavelength dependence may be due to multiple scattering or shape effects.

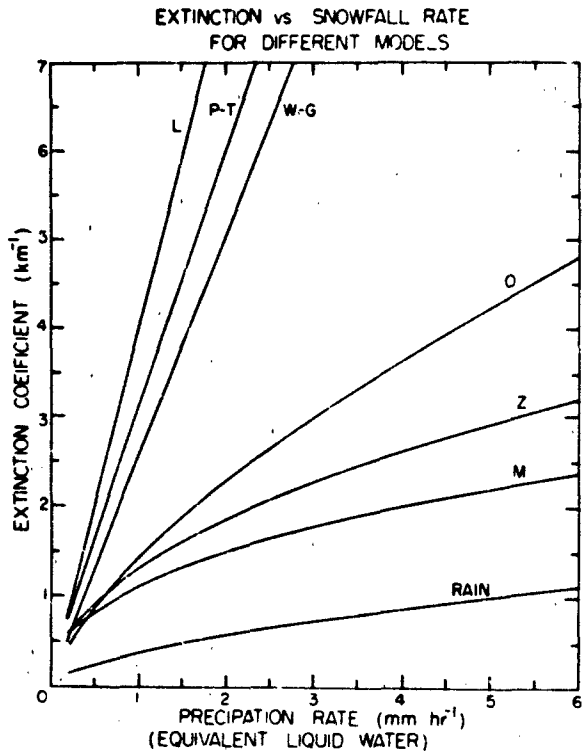


Figure 18-52. Models of visible extinction coefficient versus equivalent snow fall rate. Curve R is for rain while the others are for snow as measured by various investigators [Mason, 1978].

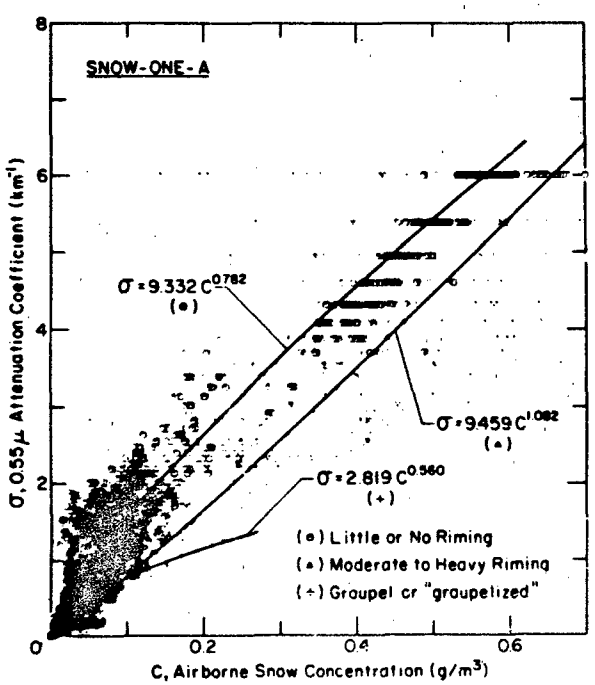


Figure 18-53. Relationship between airborne snow concentration and 0.55 μm extinction for 3 degrees of riming. Data are from SNOW-ONE-A. Curves are power law fits within each category.

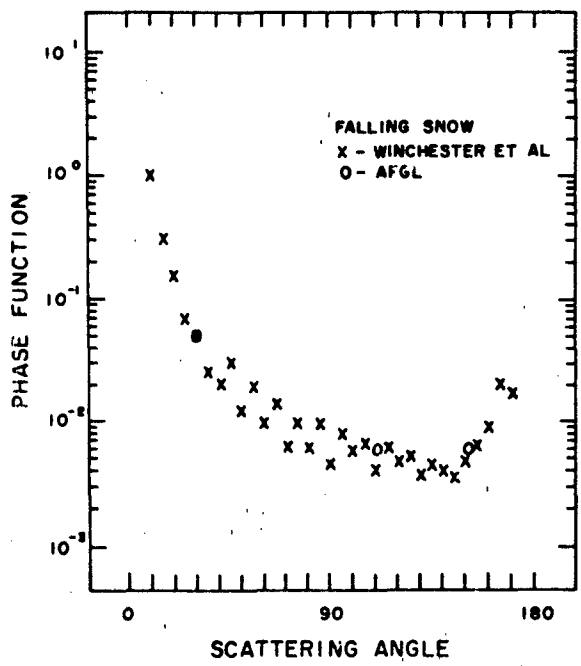


Figure 18-54. Phase function of falling snow measured in Houghton, Mississippi on 26 January 1981 with a polar nephelometer [Winchester et al., 1981]. The dominant crystal was a six-sided dendritic of about 2 mm diameter. Circled points are data from the 31 January 1982 SNOW-ONE-A episode averaged over the 1530–2000 period. Data were taken each minute with a fixed-angle nephelometer at 30°, 110°, and 150° scattering angles.

raising the question whether they represent any one snowfall "type". Measurements made by AFGL are also shown in Figure 18-54. They were made simultaneously at three fixed angles and are shown normalized to the Winchester data at 30°; they show good agreement in relative scattered intensity. An absolute comparison is not possible since Winchester's data are not normalized to total extinction.

18.3 SURFACE REFLECTANCE AND ALBEDO

A fraction of a light beam that is incident on the earth's surface or on clouds will be reflected back into the atmosphere where it may undergo further scattering and absorption. Surface reflectance consequently affects sky radiance seen from the ground, the radiance of the atmosphere as seen from space, and other atmospheric optical quantities (see Section 18.5). The percentage of light reflected from natural terrain or water surfaces varies with angle of incidence and look angle and is also a function of the wavelength of the radiation. Details of these reflection processes, especially laboratory measurements on surfaces, are often treated by the bidirectional reflectance function (BDRF), while integral reflection effects—for example the brightness of a

planet—are usually characterized by "albedo". For mirror surfaces (specular reflection), the incident and reflected beam and the normal to the surface at the point of incidence lie in the same plane; the angle between the normal and the direction of incidence and between the normal and the reflected beam are identical.

A brief derivation of the BDRF following Nicodemus [1967] and Leader [1979] will now be given. The radiant power incident on a surface element dA through a beam of solid angle $d\Omega_i$ (from polar angle θ_i, azimuth angle ϕ_i) is

$$d P_i = N_i \, dA \, d\Omega'_i \, . \qquad (18.27)$$

where N_i is the radiance (W cm^{-2} sr^{-1}) and $d\Omega'_i = \cos\theta_i \, d\Omega_i$ is the projected solid angle. The radiant power reflected into $d\Omega_r$ is

$$d P_r = dN_r \, dA \, d\Omega'_r \, . \qquad (18.28)$$

Thus the ratio of the radiant power is

$$\frac{d P_r}{d P_i} = \frac{d N_r}{N_i \, d\Omega'_i} \, d\Omega'_r = f' \, d\Omega'_r, \qquad (18.29)$$

$f'(\theta_i, \phi_i, \theta_r, \phi_r)$ (sr^{-1}) is called the bidirectional reflectance function (BDRF). This definition is especially useful for narrow (collimated) illuminating beams (sun, laser). If the incident radiance comes from all directions of the hemisphere, then

$$N_r = \int_0^{2\pi} \int_0^{\pi/2} f' \, N_i \sin\theta \cos\theta \, d\theta \, d\phi. \qquad (18.30)$$

A perfectly diffuse reflector is characterized by a constant f' in all directions. If such a surface is diffusely irradiated and the reflected radiance in any direction is measured, the directional reflectance is $f_{dir} = \pi f'$ [dimensionless]. Hence, the BDRF of a reflection standard (barium sulfate) with $f_{dir} = 1$ is $f' = 1/\pi$.

Comparison of incoming and outgoing shortwave radiance of the earth's surface or the earth's atmosphere in the simplest way is made by a "flat" receiver turned up (for downwelling radiation) and down (for upwelling radiation). The ratio of the fluxes, $P\uparrow / P\downarrow$, is often called albedo or diffuse reflectance.

Eaton and Dirmhirn [1979], for example, have made measurements of BDRFs and albedo of several surfaces for shortwave radiation (0.3–3 μm). Those of a salt flat and a field of corn are shown in Figure 18-55 for two solar elevations. Due to the anisotropy, the nadir values $\pi f'$ of the BDRF are to be multiplied by a factor F to obtain the albedo. At low solar elevation angles, F is generally >1, but surfaces with vertical structure may make F < 1 at high sun when partial shadowing occurs).

Spectral reflectance from 0.2 to 4 μm for several surface types is given in Figures 18-56 through 18-58 [Suits, 1978].

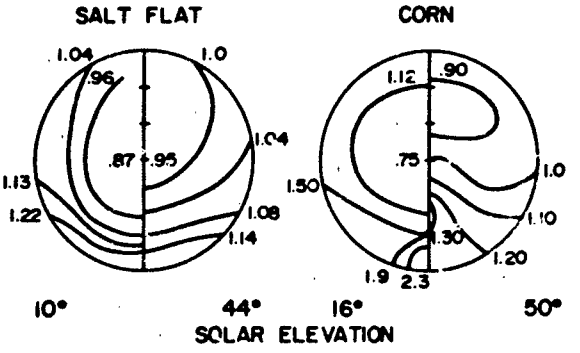


Figure 18-55. Bidirectional reflection function of Bonneville, Utah, Salt Flat (left) and a corn field (right), each for a low and a high solar elevation. $F_r(\pi f')_{Nadir}$ = albedo for a horizontal receiver. Spectral range 0.3 to 3 μm [Eaton and Dirmhirn, 1979].

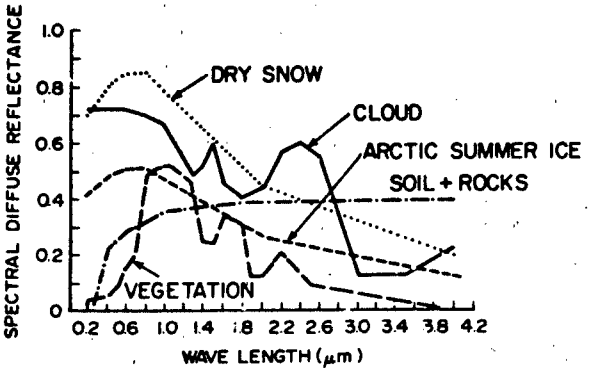


Figure 18-56. Typical spectral diffuse reflectance of snow, ice, soil, and vegetation after Suits [1978, Figure 3-19].

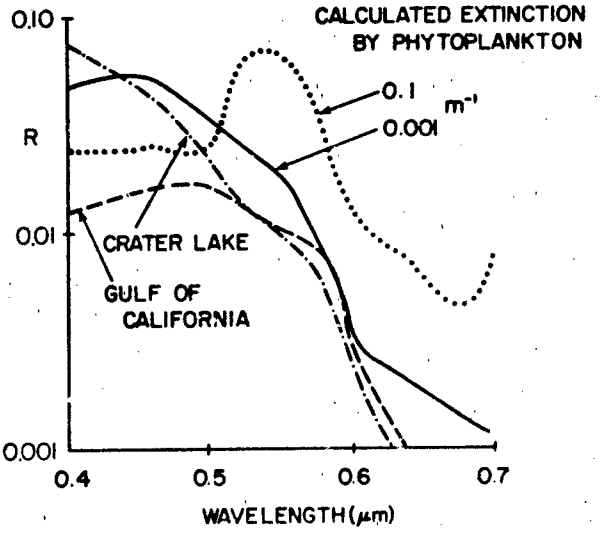


Figure 18-57. Nadir reflectance R 0.4–0.7 μm of some water bodies. Measured value taken from Tyler [1978], calculations from Suits [1978, Figure 3-124].

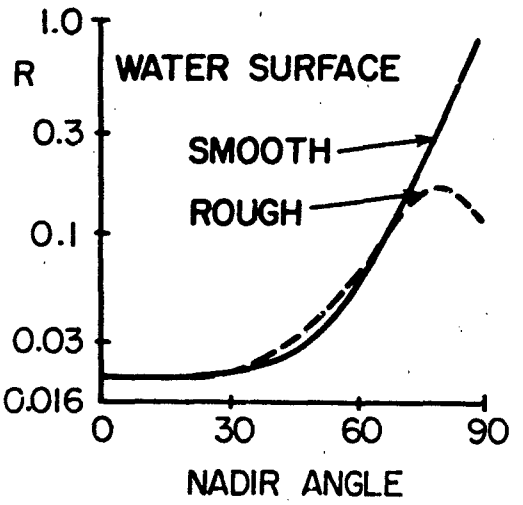


Figure 18-58. Variation with nadir look angle of specular reflectance R of smooth water surface and surface roughened by a beaufort 4 [Suits, 1978, Figure 3-114].

Further examples, citations, and references to data files can be found in Suits [1978]. Fitch [1981] made a theoretical study of the polarized radiance (at 0.4 and 0.6 μm) from the top of the earth's atmosphere for different surfaces and aerosol conditions.

18.4 ATMOSPHERIC TRANSMITTANCE AND RADIANCE

18.4.1 Monochromatic Radiation.

For many purposes, knowledge of the monochromatic optical depth for a specific optical path in the earth's atmosphere is required. The type of calculation necessary to obtain this information is referred to as a line-by-line calculation since it takes into account the contribution of each spectral line to the monochromatic absorption. Consequently, line by line calculations require detailed knowledge of the spectral line parameters associated with each contributing transition. Such computations, performed over an extended wavenumber interval (of the order of 100 cm^{-1}), may require significant computational effort. Several methods for performing such computations have been discussed in the literature. Some are particularly suitable for specialized calculations [Scott and Chedin, 1981; McMillin et al., 1976; Kunde and Maguire, 1973], some make approximations that in certain cases may be acceptable [Karp, 1973; Mankin 1979], and others are accurate and direct but not computationally efficient [Susskind and Searl, 1978]. At AFGL a computational model, FASCODE (Fast Atmospheric Signature Code), has been developed and made available to the scientific community. Line-by-line calculations are discussed in the following section in the general context of atmospheric calculations and the FASCODE model [Clough et al., 1981; Smith et al., 1978]. Specific details of the model are not discussed, but are available in the literature [for example, Clough and Kneizys, 1979; Clough et al., 1977]. A general reference covering the topic of atmospheric molecular absorption that the reader will find particularly useful is the work by Goody [1964].

18.4.1.1 Line-by-Line Computations (FASCODE). The monochromatic optical depth, $\tilde{k}(\nu)$, at wavenumber value $\nu(cm^{-1})$, assuming superposition of the contribution from the individual spectral transitions, is given by

$$\tilde{k}(\nu) = \nu \frac{1 - e^{-hc\nu/kT}}{1 + e^{-hc\nu/kT}} \times \sum_i W(m_i)\, \tilde{S}_i(T)\, [f(\nu,\nu_i) + f(\nu,-\nu_i)] \qquad (18.31)$$

where $W(m_i)$ (mol/cm^2) is the column density for the molecular species, m_i, involved in the i^{th} transition, ν_i is the transition wavenumber and $\tilde{S}_i(T)(1/mol/cm^2)$ is the intensity at temperature T (K) appropriate to the FASCODE line shape, $f(\nu, \nu_i)$ $(1/cm^{-1})$. We note that $W(m_i) = \rho(m_i)l$ for path length l and molecular density $(\rho(m_i))$. This line shape formulation is an extension of that used by Van Vleck and Huber [1977] and discussed by Clough et al. [1980, 1983]. Equation (18.31) may be conveniently rewritten in terms of the radiation field dependent term $\nu \tanh(hc\nu/2kT)$ to obtain

$$\tilde{k}(\nu) = \nu \tanh(hc\nu/2kT) \times \sum W(m_i)\, \tilde{S}_i(T)\, [f(\nu,\nu_i) + f(\nu,-\nu_i)] \qquad (18.32)$$

The intensity, $\tilde{S}_i$, is expressed in terms of the transition strength $|\mu_i^2|$ debye^2, as

$$\tilde{S}_i(T) = \frac{8\pi^3 \times 10^{-36}}{3hc} \times |\mu_i^2| \left(1 + e^{-hc\nu_i/kT}\right) \frac{e^{-hcE_i/kT}}{Q_V(m_i,T)Q_R(m_i,T)} \qquad (18.33)$$

where $E_i(cm^{-1})$ is the lower state energy for the transition and $Q_V(m_i,T)$ and $Q_R(m_i,T)$ are the vibrational and rotational partition functions for the appropriate molecular species, m_i. In terms of the intensity definition, $S_i(T)$, used for the AFGL line compilations [Rothman et al., 1983a,b; McClatchey et al., 1973] we have the relationship

$$\bar{S}_i(T) = -\frac{1 + e^{-hc\nu_i/kT}}{\nu_i\left(1 - e^{-hc\nu_i/kT}\right)} S_i(T). \qquad (18.34)$$

The line shape function must satisfy the normalization condition.

$$\int_{-\infty}^{\infty} d\nu \, [f(\nu,\nu_i) + f(\nu,-\nu_i)] = 1. \qquad (18.35)$$

As a result, the expression for optical depth, Equation (18.32) satisfies two important conditions: the Nyquist condition,

$$\int_{-\infty}^{\infty} d\nu \frac{k(\nu)}{\nu \tanh(hc\nu/2kT)} = \sum_i |\mu_i| \qquad (18.36)$$

that is, the preservation of transition strengths; and the condition that radiation balance be satisfied between emission and absorption for a system in thermal equilibrium.

The line shape function $f(\nu,\nu_i)$ is dependent on molecular species, broadening density, and temperature. For collisional broadening in the impact limit, the form factor is given by

$$f(\nu,\nu_i) = \frac{1}{\pi} \frac{\alpha_i^c}{(\nu - \nu_i)^2 + (\alpha_i^c)^2}. \qquad (18.37)$$

where α_i^c (cm^{-1}) is the collision broadened halfwidth at half maximum (HWHM). It proves convenient for computational purposes to define a dimensionless parameter z such that

$$z = \frac{\nu - \nu_i}{\alpha_i^c}. \qquad (18.38)$$

Thus, $f(\nu,\nu_i)$ becomes $f(z) = (1/\alpha_i^c)\, L(z)$ where

$$L(z) = \frac{1}{\pi} \frac{1}{1 + z^2}. \qquad (18.39)$$

The functional dependence of L(z) is shown as function L in Figure 18-59a.

The dependence of the collision broadened halfwidth on the number density, n, and temperature, T, of the absorbing media is given by the relation,

$$\alpha_i^c(n,T) = \alpha_i^{c'}(n/n_o)(T/T_o)^{X_T} \qquad (18.40)$$

where $\alpha_i^{c'}$ is the collision halfwidth at reference number density n_o defined for 1013 mb and reference temperature T_o (296 K). The quantity X_T, characterizing the temperature dependence of the halfwidth, typically ranges between 0 and 0.5, depending on molecular species and temperature. Classical collision theory gives a value of 0.5 for X_T. Values of X_T appropriate to each molecule type are stored in the program.

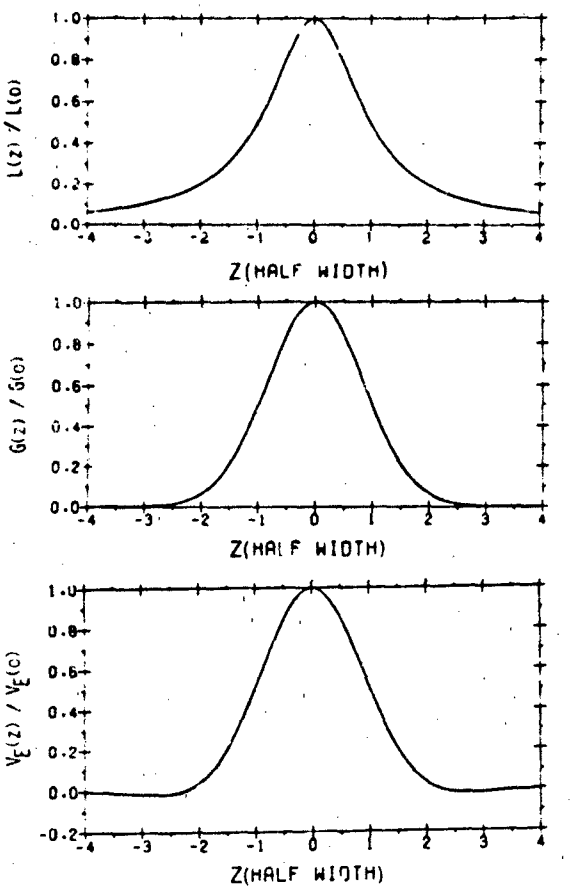


Figure 18-59. (a) The line shape profile appropriate for collisional broadening.
(b) The line shape profile appropriate for velocity broadening.
(c) The function utilized in FASCODE for the construction of the Voigt line profile.

The extensive computational effort required for atmospheric line by line calculations is a result of three principal factors: the large number of spectral lines contributing to the absorption, the small sampling interval necessary to model the attenuation at higher altitudes and the slow convergence to zero of the Lorentz function given in Equation (18.39). The latter consideration has been treated through the algorithms utilized in FASCODE and the details have been made available in the literature. A computational savings of sixteen is achieved over direct methods for a spectral interval of 512 halfwidths.

For the Doppler regime, in which velocity broadening dominates, the line shape is

$$f(\nu,\nu_i) = \frac{1}{\alpha_i^D}\left(\frac{\ln 2}{\pi}\right)^{1/2} \exp\left[-\ln 2\left(\frac{\nu - \nu_i}{\alpha_i^D}\right)^2\right]. \qquad (18.41)$$

Letting

$$z = \frac{\nu - \nu_i}{\alpha_i^D} \qquad (18.42)$$

then $f(z) = (1/\alpha_i^D)\, G(z)$ where

$$G(z) = \left(\frac{\ln 2}{\pi}\right)^{1/2} \exp\left[-(\ln 2)(z^2)\right]. \qquad (18.43)$$

The functional dependence of G(z) is Gaussian and is shown in Figure 18-59b. The Doppler halfwidth (HWHM) is given by

$$\alpha_i^D = \frac{\nu_i}{c}\left[2(\ln 2)\frac{kT}{M/N_0}\right]^{1/2} \qquad (18.44)$$

where M is the gram molecular weight of the molecular species of the transition and N_0 is Avogadro s number.

In the intermediate regime between collisional broadening and velocity broadening, the line shape is obtained from the convolution of the collisional and Doppler line shape giving the Voigt line profile. The computation of the Voigt line shape using a direct approach is too time consuming for general utilization. These direct methods generally give results to significantly higher precision than the uncertainty in the line parameters and the uncertainty in the line shape itself. An effective approximation to the Voigt line shape may be obtained by using an extension of the method proposed by Whiting [1968]. A Voigt parameter, ζ, is defined in terms of the collisional and Doppler widths

$$\zeta = \frac{\alpha^C}{\alpha^C + \alpha^D} \qquad (18.45)$$

for which $\zeta = 0$ in the Doppler limit and $\zeta = 1$ in the collision broadened limit. For $\zeta = 0.5$ the collision and Doppler widths are equal. The Voigt width α^V (HWHM) is obtained through the relation

$$\alpha^V = A_V(\zeta)\,(\alpha^D + \alpha^C). \qquad (18.46)$$

The function $A_V(\zeta)$ shown in Figure 18-60 is obtained from the relations given by Kielkopf [1973] where

$$A_V(\zeta) = (\zeta)\frac{1+\varepsilon}{2} + \left[(\zeta)^2 \left(\frac{1-\varepsilon}{2}\right)^2 + 1 - (1-\zeta)^2\right]^{1/2}. \qquad (18.47)$$

with $\varepsilon = 0.0990\ \ln 2$. An appropriate Voigt line shape, $V(\zeta,z)$, is given as

$$V(\zeta,z) = C_V(\zeta)\,L(z) + C_D(\zeta)\,G(z) + C_I(\zeta)\,V_I(z). \qquad (18.48)$$

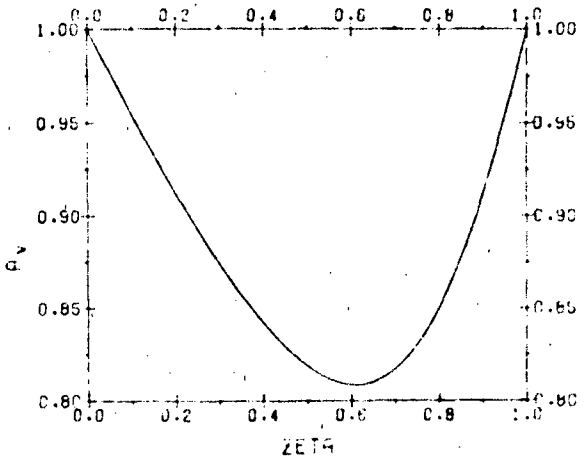


Figure 18-60. The constant relating the Voigt width to the sum of the Doppler and collision widths: collision broadened width over the doman 0.5.

where z is defined as

$$z = \frac{\nu - \nu_i}{\alpha^V} \qquad (18.49)$$

The function $V_I(z)$ is an analytic function having the dependence on z shown in Figure 18-59c. The constants $C_V(\zeta)$, $C_D(\zeta)$, and $C_I(\zeta)$ are determined from least square fits of $V(\zeta,z)$ to the exact Voigt function, The Voigt line profiles for four values of ζ and the FASCODE approximation, $V(\zeta,z)$, are shown in Figure 18-61. The computational savings using this approximate approach is on the order of 100 compared with direct methods. The largest error, ~3%, occurs for $\zeta \approx 0.05$ at a value of z corresponding to 16 halfwidths from line center at which the function is four orders of magnitude less than the peak value.

18.4.1.2 Continuum Absorption. FASCODE includes continuum contribution from self and foreign density dependent water vapor absorption, foreign density dependent carbon dioxide absorption, and the collision induced band of nitrogen. A more extensive treatment of the approach used in FASCODE in which the continua are developed from the wings of strong lines appears in the literature [Clough et al., 1980 and 1983]. For atmospheric applications it is advantageous to express the density dependence in terms of a self and foreign component. The continuum contribution to the optical depth, $\tilde{k}_C$, is given by the expression

$$\tilde{k}_C = W\,\nu\,\tanh(hc\nu/2kT) \times \left[(n_s/n_0)\,\tilde{C}_s(\nu,T) + (n_f/n_0)\,\tilde{C}_f(\nu,T)\right]. \qquad (18.50)$$

where W is the column density of the absorbing molecular species, (n_s/n_0) and (n_f/n_0) the number density ratios for the

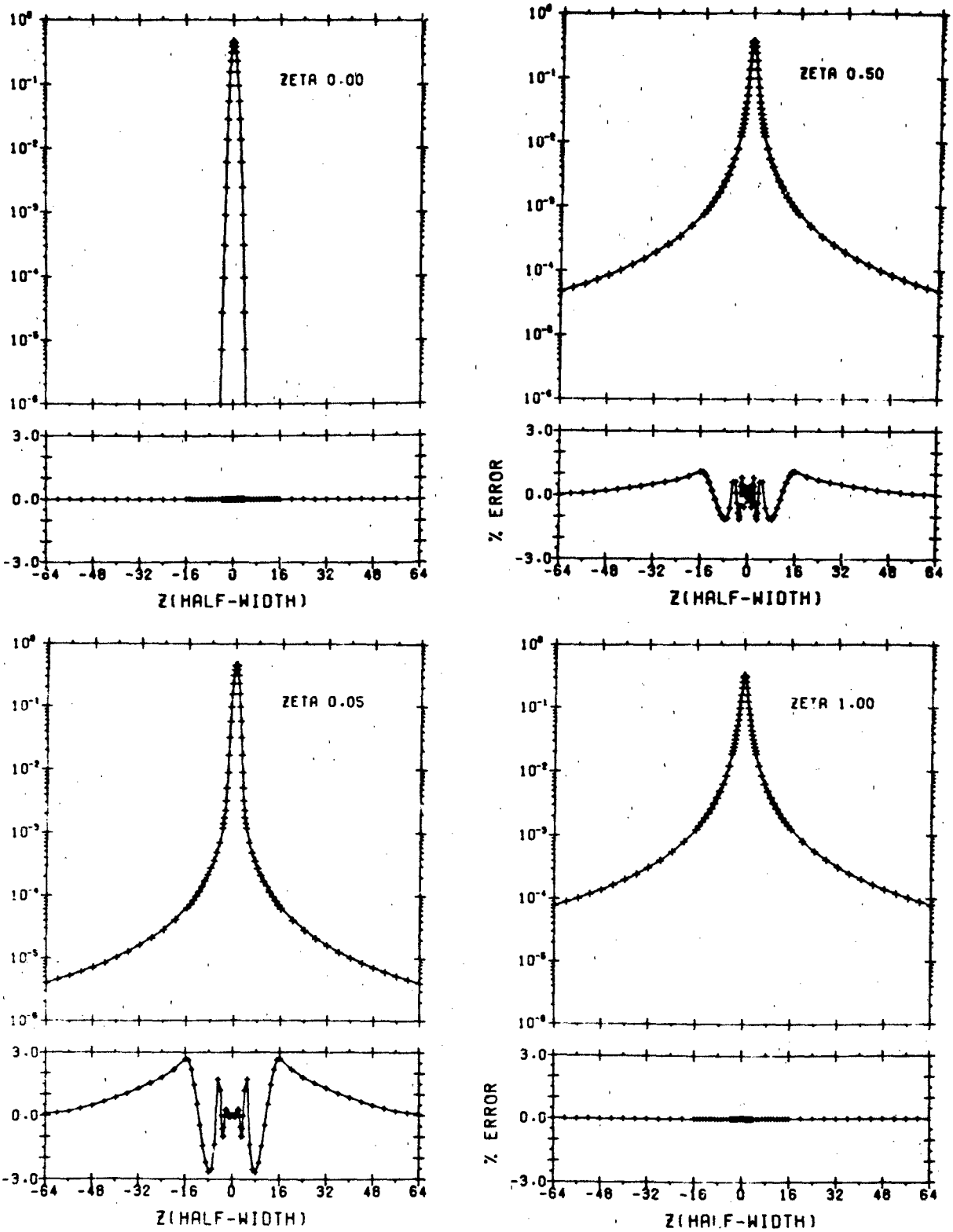


Figure 18-61. Voigt line shape profiles for representative values of zeta. The solid line represents the correct function and the symbol + the FASCOD1 approximation

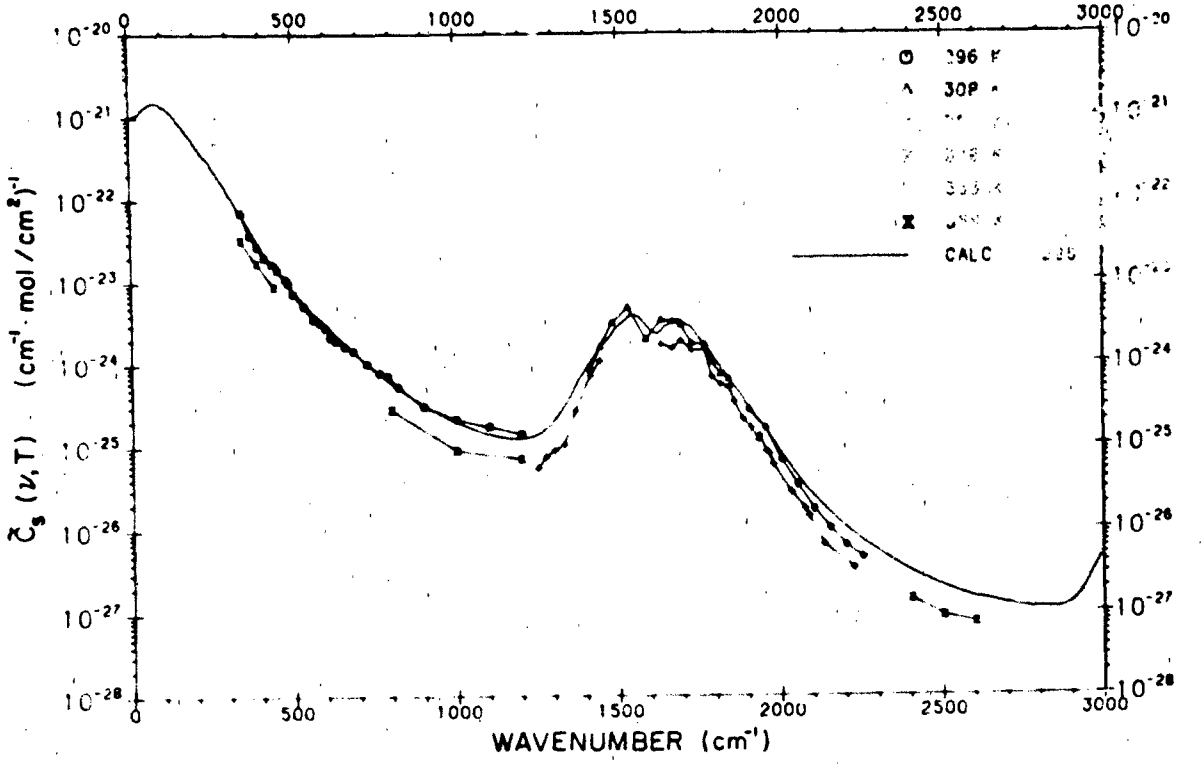


Figure 18-62. The self density dependent continuum values C_s for water vapor as a function of wave number. The experimental values are from Burch [1981].

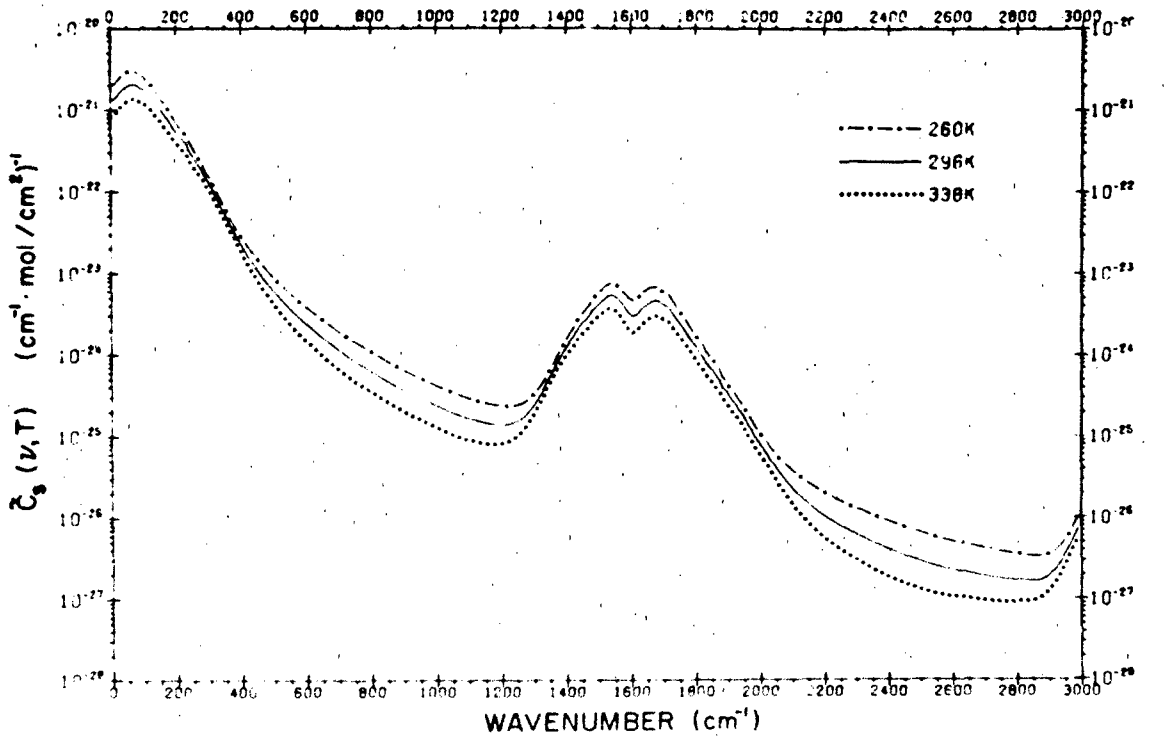


Figure 18-63. The self density dependent continuum values C_s for water vapor as a function of wave number at 260 K, 296 K, and 338 K. The values from 296 K and 338 K are fits to experimental results, the 260 K is extrapolated.

self and foreign continuum, and $\tilde{C}_s$ and $\tilde{C}_f$ [$(cm^{-1}\ mol/cm^2)^{-1}$] the wavenumber dependent continuum absorption parameters for the self and foreign components. The density n_s is the density of the absorbing molecular species and n_f is the density of all other molecular species; consequently, $n_s + n_f$ represents the total density. The quantity, n_{s_0} is the reference number density defined at 1013 mb and 296 K. In some cases the dependence of $\tilde{C}$ on temperature is not known. The present formulation has the advantage that the continuum contribution to the optical depth decreases with increasing temperature through the number density ratio term. The quantities $\tilde{C}_s$ and $\tilde{C}_f$ for water vapor and $\tilde{C}_f$ for carbon dioxide are stored in the program for the spectral range 0 to 20 000 cm^{-1}.

The values for $\tilde{C}_s$ for water vapor at 296 K are shown in Figure 18-62 together with the experimental values obtained by Burch [Burch and Gryvnak, 1978 and 1979; Burch et al., 1971; and Burch, 1970]. The 260 K result was obtained by extrapolating the fits to the 353 K and 296 K data of Burch. The results for the three temperatures are shown in Figure 18-63. The strong temperature dependence of the self density dependent water vapor continuum is treated by using exponential interpolation of the 260 K and 296 K values.

Only values near room temperature are available for the foreign dependence of the water vapor continuum. The continuum values $\tilde{C}_f$ at 296 K are shown in Figure 18-64 and have been obtained by a fit to the data of Burch. There is considerable uncertainty in the values for the spectral window regions at 1000 and 2500 cm^{-1}.

For nitrogen, continuum values at 296 K are included in FASCODE for the collision induced absorption band at 2350 cm^{-1}. For this case $\tilde{C}_s$ is taken to equal $\tilde{C}_f$, so that the effect is dependent on the total density.

18.4.1.3 Radiance and Transmittance for Atmospheric Paths. The layering of the atmosphere is primarily dependent on two considerations: the ratio of the Voigt widths at the layer boundaries and the temperature across the layer. The Voigt width as a function of altitude for the U.S. Standard Atmosphere is shown in Figure 18-65. The sampling interval, DV, for FASCODE is indicated on the figure where nominally DV = α'/4. All calculations are performed commencing at the lowest altitude involved in the problem. Consequently the sampling interval decreases monotonically in the altitude regime for which collisional broadening predominates and becomes essentially constant at altitudes for which the Doppler broadening is dominant. This method of

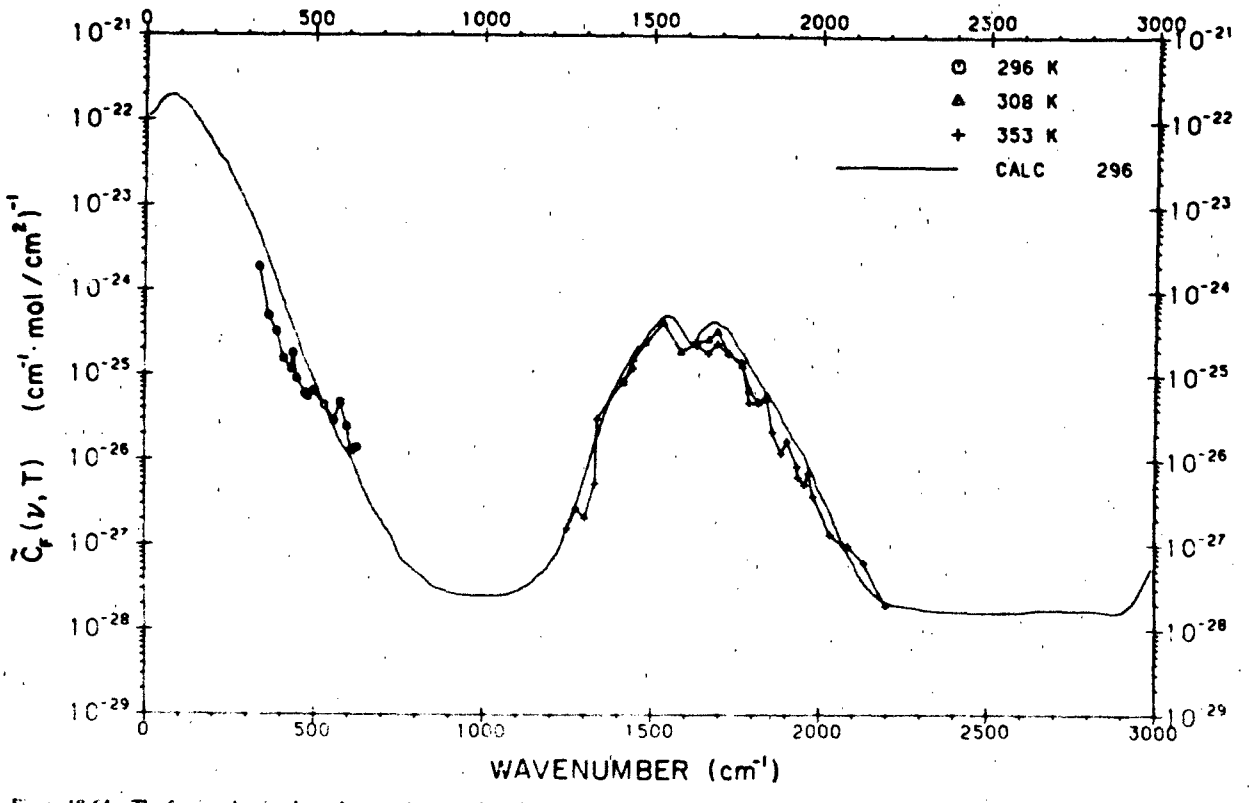


Figure 18-64. The foreign density dependent continuum values C_f for water vapor as a function of wave number. The experimental values are from Burch [1981]

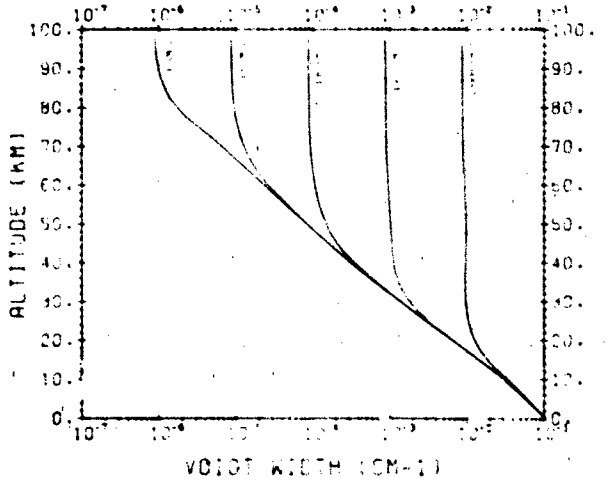


Figure 18-65. The Voigt width (HWHM) as a function of altitude for the U.S. Standard Atmosphere.

performing the calculation ensures that the minimum amount of information is handled at any given stage of the calculation. Figure 18-66 shows the Voigt parameter, ζ, as a function of altitude for the U.S. Standard Atmosphere.

The atmospheric layering is performed by program FSCATM [Gallery et al., 1983]. For specified boundaries and geometric parameters, the program computes the column densities and the density weighted temperature and pressure for each layer. Alternatively, the maximum acceptable ratio of Voigt line widths for the layer boundaries and the maximum temperature differential across the layer may be selected to perform the atmospheric layering. These two parameters are directly related to the accuracy of the radiance and/or the transmittance result.

FASCODE operates in either of two modes: optical depth or radiance/transmittance. In the optical depth mode, the

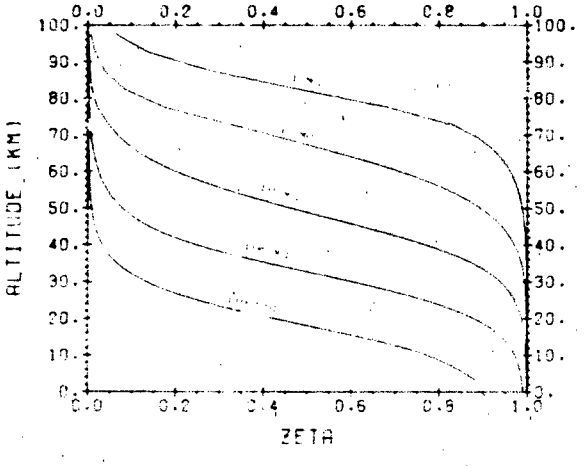


Figure 18-66. The Voigt parameter ε as a function of altitude for the U.S. Standard Atmosphere.

recursion relation for the accumulated optical depth, $\bar{K}_\ell$, after ℓ layers is expressed as

$$\bar{K}_\ell = \bar{K}_{\ell-1} + \bar{k}_\ell \qquad (18.51)$$

where $\bar{k}_\ell$ is the optical depth for layer ℓ and $\bar{K}_0 = 0$. This relationship is implemented layer by layer to obtain the total optical depth for the path.

For the radiance/transmittance mode, the radiant intensity, I_ℓ (W cm^{-2} sr^{-1}/cm^{-1}), originating in layer ℓ, is expressed as

$$I_\ell = (1 - \tau_\ell)\, B(T_\ell), \qquad (18.52)$$

where τ_ℓ is the transmittance and B (W cm^{-2} sr^{-1}/cm^{-1}) is the Planck black body function at temperature T_ℓ for the layer. Since all calculations are performed from lower to higher altitudes, different recursion algorithms are required for the looking down, the tangent path, and the looking up geometric situation. For the looking down geometry, we obtain the recursion relation for the contribution of the current layer ℓ to the upwelling radiation as

$$I_\ell = \tau_\ell\, I_{\ell-1} + (1 - \tau_\ell)\, B(T_\ell) \qquad (18.53)$$

with the condition that I_0, the boundary radiance, is

$$I_0 = \varepsilon_0\, B(T_0), \qquad (18.54)$$

where T_0 is the boundary temperature and ε_0 the boundary emissivity. The transmittance relationship for this case becomes

$$T_\ell = \tau_\ell\, T_{\ell-1} \qquad (18.55)$$

where T_ℓ is the total transmittance after layer ℓ.

For tangent path geometry with symmetry about the tangent height, we obtain the recursion relations

$$I_\ell = \tau_\ell\, I_{\ell-1} + (1 + \tau_\ell\, T_{\ell-1})(1 - \tau_{\ell-1})\, B(T_\ell) \qquad (18.56)$$

for the radiance with $I_0 = 0$ and

$$T_\ell = \tau_\ell\, T_{\ell-1} \qquad (18.57)$$

for the transmittance with $T_0 = 1$. This algorithm for the tangent path case involves only a single line-by-line calculation for each layer and enables this case to be handled consistently with the other cases, that is, the calculation commences with the tangent layer and proceeds to higher altitudes.

For looking up geometries, the recursion relation for the downwelling radiance is

$$I_\ell = I_{\ell-1} + T_{\ell-1}(1 - \tau_\ell)\, B(T_\ell) \qquad (18.58)$$

with $I_{\nu_0} = 0$ and for the transmittance

$$T_t = \tau_t \, T_{t-1} \tag{18.59}$$

with $T_{\nu_0} = 1$. These algorithms are implemented layer by layer starting with the initial layer at the lowest altitude and proceeding to the higher altitude.

In Figure 18-67 we show the comparison of a FASCODE calculation with experimental results obtained by Rice and Ade [Rice and Ade, 1979] in the millimeter spectral region. The data were taken with an interferometer measuring the downwelling radiance at the ground from which atmospheric transmittance was inferred. The calculation was performed with a two layer atmosphere and the water amount given by the measured amount for the Rice and Ade experiment. In Figure 18-68 we show the results for a fifteen layer calculation with a geometry in which the observer is at 100 km looking through a 15 km tangent height. Figure 18-69 shows an expanded view of the Q branch region shown in Figure 18-68. The radiance calculation is plotted in terms of equivalent black body temperature.

18.4.2 Broad-Band Radiation

Band models assume an array of lines having chosen intensities, half-widths, and spacings which can be adjusted to represent the molecular line structure in some part of a real band. For a particular band model, the mean transmittance can be represented by a mathematical expression (transmittance function) expressed in terms of pressure, temperature effective path length (or absorber concentration), and one or more frequency dependent absorption coefficients.

Several band models have been developed [see Goody, 1952 and 1964; Elsasser, 1942, Plass, 1958; King, 1964; and Wyatt et al., 1964]; the Elsasser and Goody models being the most well known. The Elsasser band model consists of an array of equally spaced identical lines of Lorentz shape. This model has been applied to absorption bands that have a regular line structure, for example, to some bands of CO_2, N_2O, CO, CH_4, and O_3. The Goody model, on the other hand, assumes that the band is composed of spectral lines with an exponential intensity distribution and with random spacing between lines. Again the lines are assumed to have a Lorentz shape. This model has generally been applied to bands that have an irregular line structure, for example, some H_2O and O_3 bands.

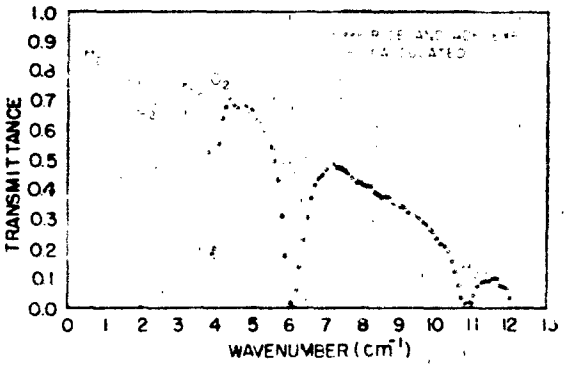


Figure 18-67 Spectral transmittance inferred from a radiance measurement looking up from the ground. The symbols o are the data of Rice and Ade [1979], and the continuous curve is a two-layer FASCOD1B calculation

In practice the wavelength dependent absorption coefficients are determined for each absorbing gas separately using laboratory transmittance values measured under known conditions. The absorption coefficients obtained in this way are then used in the band model transmittance function to determine the average transmittance for each absorber as a function of frequency for other values of path length and pressure. Finally, the total mean transmittance for molecular absorption is given by the product of the mean transmittances of the individual absorbers at each frequency.

Exact analytical expressions have been obtained for most of the band models. However, they are sometimes difficult to use and simpler approximations have been found to apply in two limiting conditions common to all band models. Those simpler expressions are the well known "weak line" and "strong line" approximations [see Goody, 1964, and Plass, 1958] for which the transmittance is a function of the absorber amount, and the product of the pressure and absorber amount, respectively (for a given temperature).

The weak line approximation, which corresponds to the exponential law, is valid when the absorption is small at the line centers (generally for high pressures and low absorber amounts). Unfortunately this case is rarely applicable to conditions existing in the terrestrial atmosphere. The strong line approximation is applicable where the lines are completely absorbing at their centers; the effect of increasing the amount of absorber is then confined to the edges or wings of the lines. The regions of validity of the strong and weak line approximation for the Elsasser and Goody models are discussed by Plass [1958]. For practical purposes most problems fall in either the strong line approximation region or the intermediate region.

In the LOWTRAN band model discussed in the following section, empirical transmittance functions for H_2O, O_3, and the combined contributions of the uniformly mixed gases have been determined from laboratory and calculated transmittance data. Over a wide range of pressures and absorber amounts, the empirical transmittance functions were found to give better agreement with laboratory and calculated transmittance data than the commonly used band models. An excellent review of the LOWTRAN model and other band models is given in La Rocca [1978].

18.4.2.1 Broad-band Transmission (LOWTRAN). The LOWTRAN model was developed to provide a fairly accurate, simple and rapid means of estimating the transmit-

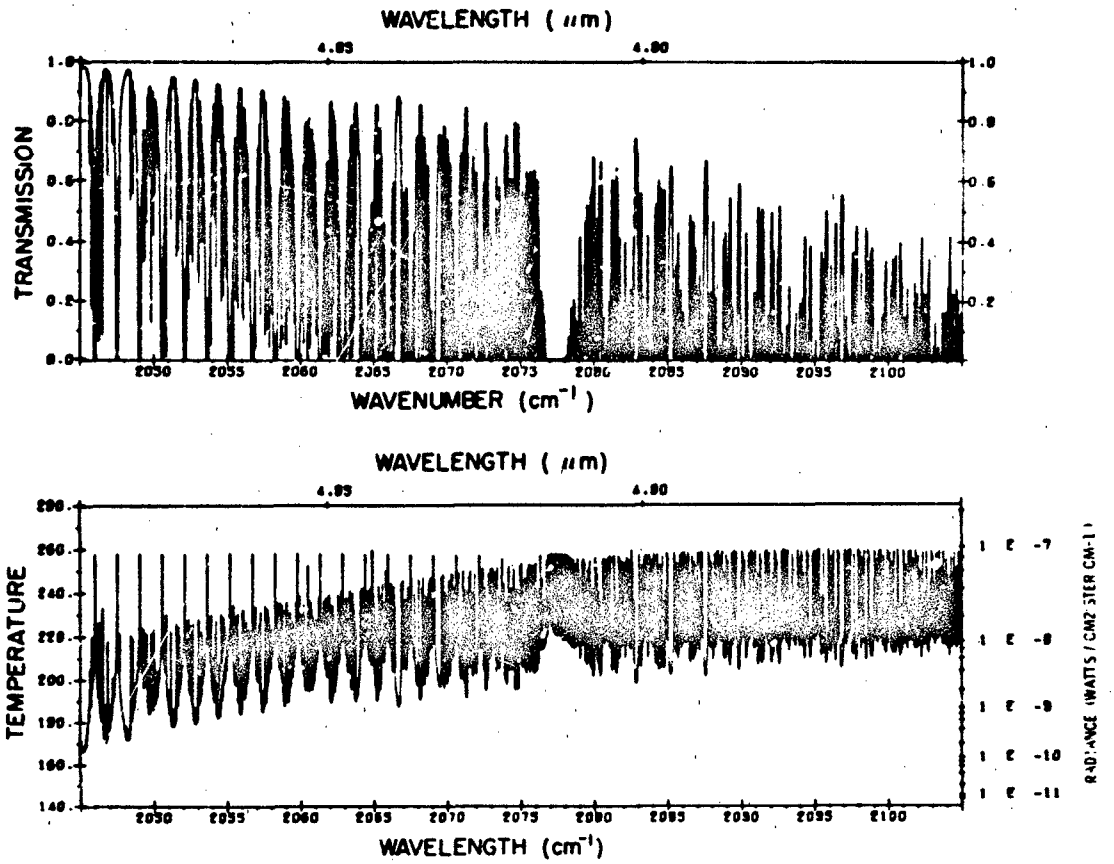


Figure 18-68. The transmittance and equivalent temperature radiance plot at 100 km and a 15 km tangent height. The results are from a 15 layer FASCOD1B calculation.

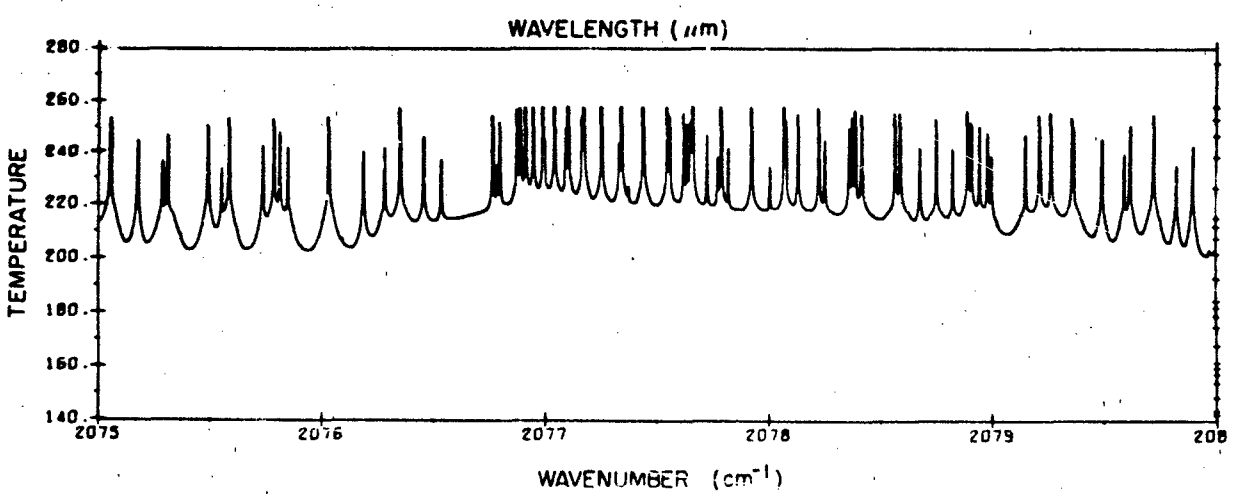


Figure 18-69. An expanded spectral region from the results shown in Figure 18-68.

tance and background radiance of the earth's atmosphere at moderate resolution (20 cm^{-1}) over a broad spectral interval (350 to 40 000 cm^{-1}). The model was derived using a single parameter band model for molecular absorption and includes the effects of continuum absorption, molecular scattering, and aerosol extinction. It is based on an empirical graphic prediction scheme of transmittance using some techniques originally suggested by Altshuler [1961]. The prediction scheme, derived from laboratory transmittance measurements complemented by line-by-line transmittance calculations was reported by McClatchey et al. [1972] and, subsequently, together with atmospheric models developed into the model and computer code LOWTRAN [Selby et al., 1972; 1975; 1976; 1978; and Kneizys et al., 1980; 1983]. The model contains representative (geographical and seasonal) atmospheric models and representative aerosol models with an option to replace them with user derived or measured values and computes atmospheric transmittance or both atmospheric transmittance and radiance for any given slant path geometry.

18.4.2.2 Assumptions of the LOWTRAN Model. The basic assumptions made in the LOWTRAN Model are as follows:

1. The atmosphere can be represented by a 33 layer model between sea level and 100 km. Each model atmosphere contains the variation of pressure, temperature, water vapor, and ozone with altitude. A mean nitric acid profile is incorporated into the model. It is assumed that the mixing ratios of CO_2, N_2O, CH_4, CO, O_2 and N_2 remain constant with altitude. Each layer is assumed to be in thermal equilibrium.

2. Refraction and earth curvature effects are included in the calculation of absorber amounts for slant atmospheric paths.

3. The average transmittance over a 20 cm^{-1} interval (due to molecular absorption) can be represented by a single parameter empirical transmittance function. The argument of the transmittance function is the product of wavenumber (or wavelength) dependent absorption coefficient, and "an equivalent absorber amount" for the atmospheric path.

4. The total transmittance at a given wavenumber averaged over a 20 cm^{-1} interval is given by the product of the average transmittances due to molecular absorption, molecular scattering, aerosol extinction, and continuum absorption. It is further assumed that the molecular absorption is composed of four components, namely the separate transmittances of water vapor, ozone, nitric acid and the uniformly mixed gases.

5. The variation of aerosol optical properties with altitude is modeled by dividing the atmosphere into four height regions each having a different type of aerosol. The aerosol models for the lower atmosphere (RURAL, URBAN, MARITIME and TROPOSPHERIC) are given as a function of relative humidity. Radiation and advection fog models are also included.

6. Atmospheric and earth radiance are calculated by a numerical evaluation of the integral form of the equation of radiative transfer. The emission from aerosols and the treatment of aerosol and molecular scattering are considered only in the zeroth order.

18.4.2.3 LOWTRAN Atmospheric Transmittance. In the LOWTRAN model, the total atmospheric transmittance at a given wavenumber averaged over a 20-cm^{-1} interval is given by the product of the average transmittances due to molecular band absorption, molecular scattering, aerosol extinction, and molecular continuum absorption. The molecular band absorption is composed of four components, namely the separate transmittances of water vapor, ozone, nitric acid and the uniformly mixed gases (CO_2, N_2O, CH_4, CO, O_2 and N_2).

The average transmittance $\bar{\tau}$ over a 20-cm^{-1} interval (due to molecular absorption) is represented by a single parameter model of the form

$$\bar{\tau} = f(C_\nu, \omega^* DS) \qquad (18.60)$$

where C_ν is the LOWTRAN wavenumber-dependent absorption coefficient and ω^* is an "equivalent absorber density" for the atmospheric path DS at altitude z, defined in terms of the pressure P(z), temperature T(z), concentration of absorber ω and an empirical constant n as follows

$$\omega^* = \omega \left\{ \frac{P(z)}{P_o} \sqrt{\frac{T_o}{T(z)}} \right\}^n \qquad (18.61)$$

where P_o and T_o correspond to STP (1 atm, 273 K). If Equation (18.61) is substituted in Equation (18.60) and n is set to zero and unity respectively, Equation (18.61) reverts to the well-known weak-line and strong-line approximations common to most band models.

The form of the function f and parameter n was determined empirically using both laboratory transmittance data and available molecular line constants. In both cases, the transmittance was degraded in resolution to 20 cm^{-1} throughout the entire spectral range covered here. It was found that the functions f for H_2O and the combined contributions of the uniformly mixed gases were essentially identical, although the parameter n differed in the two cases. Mean values of n were determined to be 0.9 for H_2O, 0.75 for the uniformly mixed gases, and 0.4 for ozone.

Figure 18-70 shows the LOWTRAN empirical transmittance functions defined by Equation (18.60) versus the $\log_{10}$ of the effective optical depth ($C_\nu \omega^* DS$). The solid function shown is used for water vapor and the uniformly mixed gases. The dashed function is applicable to ozone.

Absorption coefficients for water vapor, ozone, and the combined effects of the uniformly mixed gases are included as data for LOWTRAN. The absorption coefficients for water vapor, ozone, and the uniformly mixed gases are

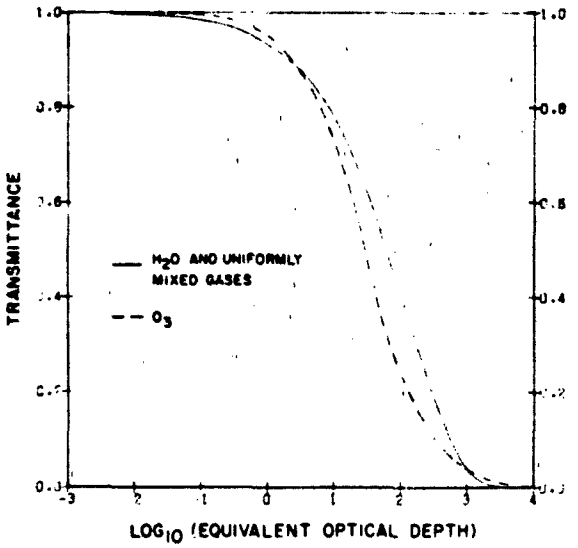


Figure 18-70. LOWTRAN empirical transmittance functions versus Log_{10} of the effective optical depth (C_{ab}*DS).

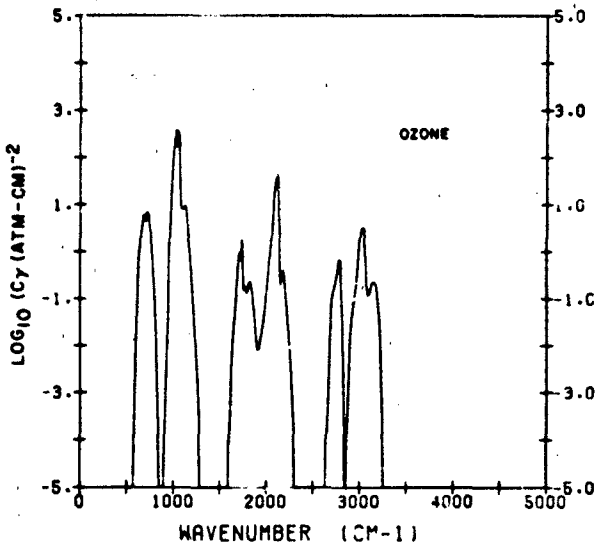


Figure 18-72. Absorption coefficient C_ν for ozone from 350 to 5000 cm^{-1}.

shown in Figures 18-71, 18-72 and 18-73 for the spectral region from 350 to 5000 cm^{-1}.

18.4.2.4 LOWTRAN Atmospheric Radiance. The LOWTRAN model and code also calculate atmospheric and earth radiance. A numerical evaluation of the integral form of the equation of radiative transfer is used. The emission from aerosols and the treatment of aerosol and molecular scattering is considered only in the zeroth order. Additional contributions to atmospheric emission from radiation scattered one or more times are neglected. Local thermodynamic equilibrium is assumed in the atmosphere.

The average atmospheric radiance (over a 20-cm^{-1} interval) at the wavenumber, ν, along a given line-of-sight in terms of the LOWTRAN transmittance parameters is given by

$$I(\nu) = \int_{\bar{\tau}^b}^{1} d\,\bar{\tau}_a\, B(\nu, T)\,\bar{\tau} + B(\nu, T_b)\,\bar{\tau}^b_t, \qquad (18.62)$$

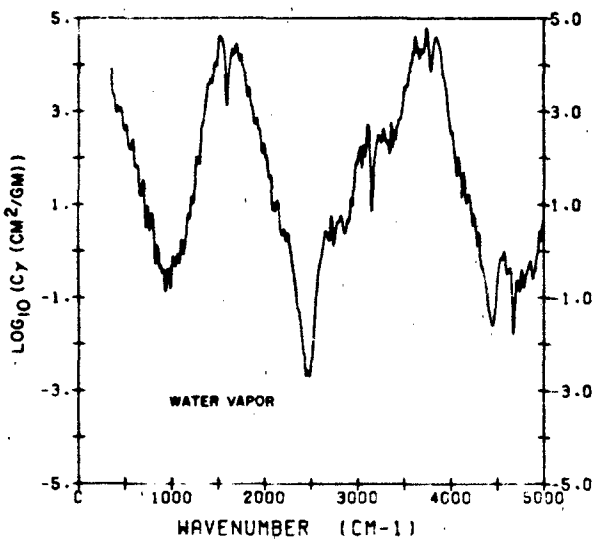


Figure 18-71. Absorption coefficient C_ν for water vapor from 350 to 5000 cm^{-1}.

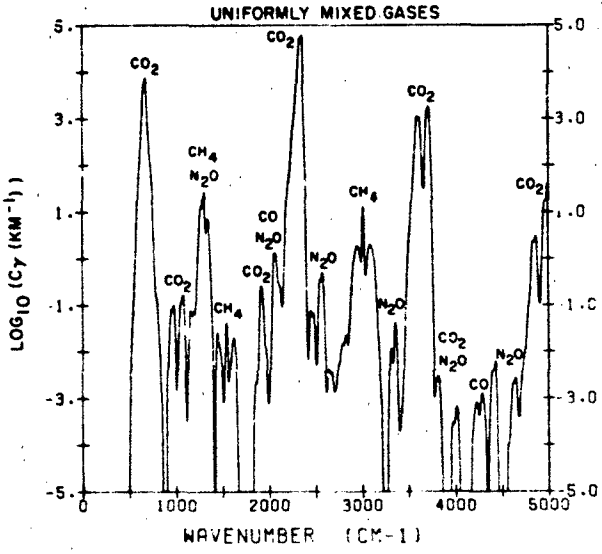


Figure 18-73. Absorption coefficient C_ν for the uniformly mixed gases from 350 to 5000 cm^{-1}.

where the integral represents the atmospheric contribution and the second term is the contribution of the boundary, (for example, the surface of the earth), and

$\bar{\tau}_a$ = average transmittance due to absorption,

$\bar{\tau}_s$ = average transmittance due to scattering,

$\bar{\tau}_t = \bar{\tau}_a \bar{\tau}_s$ = average total transmittance,

$\bar{\tau}_a^b$, $\bar{\tau}_t^b$ = average total transmittances from the observer to boundary,

$B(\nu,T)$ = average Planck (blackbody) function corresponding to the frequency ν and the temperature T of an atmospheric layer,

T_b = temperature of the boundary.

The emissivity of the boundary is assumed to be unity.

The numerical analogue to Equation (18.62) has been incorporated in the LOWTRAN model.

18.4.2.5 Examples of Transmittance and Radiance Spectra. Some examples of transmittance and radiance spectra obtained from the LOWTRAN model are presented in Figures 18-74 through 18-77.

Figure 18-74 shows the transmittance from ground to space from 0.25 to 4μm. This calculation used the U.S. Standard model atmosphere and the rural aerosol model with a 23-km met. range.

Figure 18-75 shows the variation in transmittance in the spectral region between 400 and 4000 cm^{-1} for the rural, maritime, urban, and tropospheric aerosol models. The calculation is for a 10-km horizontal sea-level path using the U.S. Standard model atmosphere and a 23-km meteorological range.

Figure 18-76 shows the transmittance of the two fog models in LOWTRAN for a 0.2-km horizontal sea-level path and a 1-km meteorological range in the spectral region from 400 to 4000 cm^{-1}.

Figures 18-77 a and b show the transmittance and radiance spectra for a vertical path at 0° zenith angle looking to space as a function of observer altitude (0, 20 and 40

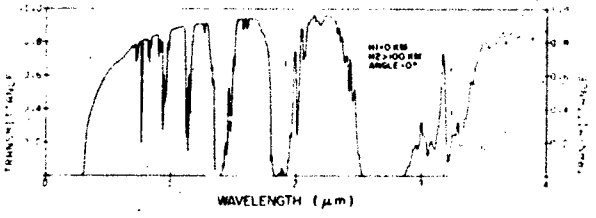


Figure 18-74. Transmittance spectra for a vertical path from ground to space from 0.25 to 4 μ, using the rural aerosol model, 23-km VIS and the U.S. Standard Model Atmosphere.

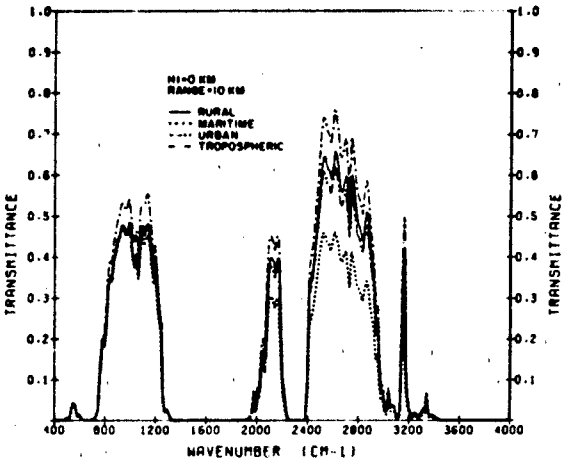


Figure 18-75. Transmittance spectra for a 10-km horizontal path at sea level for the rural, maritime, urban, and tropospheric aerosol models using the U.S. Standard Model Atmosphere and a VIS of 23 km.

km). This calculation for the spectral region from 400 to 4000 cm^{-1} used the rural aerosol model (23 km met. range) and the U.S. Standard model atmosphere.

18.4.2.6 Comparisons of LOWTRAN with Measurements. Comparisons of LOWTRAN with measurements are presented in Figures 18-78 through 18-87. Figures 18-78 and 79 show transmittance comparisons of LOWTRAN with laboratory measurements of Burch et al. [1962] for some important water vapor and carbon dioxide absorption bands.

Figure 18-80 shows a transmittance comparison with a sea-level measurement by Ashley et al. [1973] (General Dynamics). The measurement, made with an interferometer

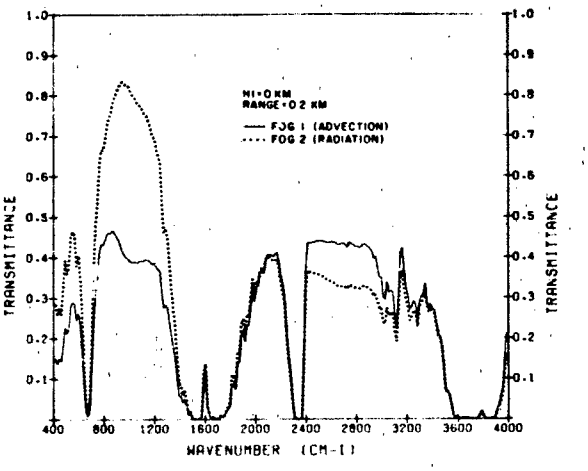


Figure 18-76. Transmittance spectra for the advection fog (Fog 1) and the radiation fog (Fog 2) models, for a 0.2-km horizontal path at sea level, with the U.S. Standard Model Atmosphere and a 1-km VIS, from 400 to 4000 cm^{-1}.

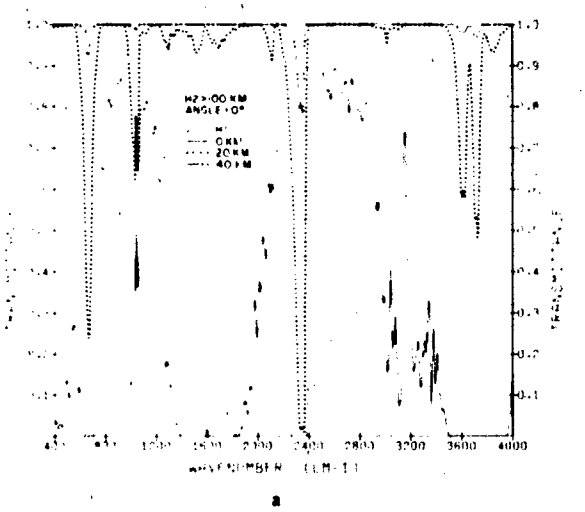


Figure 18-77 Transmittance and radiance spectra for a vertical path looking to space from H1 (H1 = 0, 20 km, 40 km, H2 ≥ 100 km, ANGLE = 0°) the rural aerosol model (IHAZE = 1, VIS = 23 km) and the U.S. Standard Atmosphere (MODEL = 6), from 400 to 4000 cm^{-1}: (a) transmittance, (b) radiance.

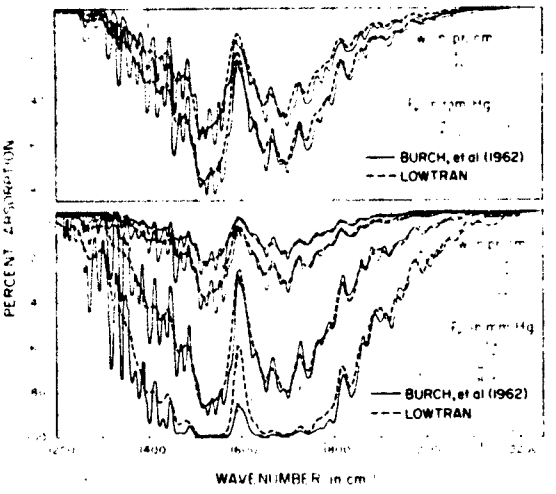


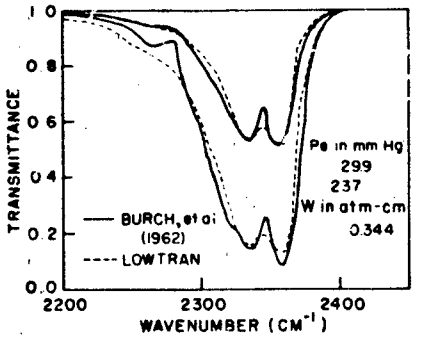
Figure 18-78. Representative absorption curves for the 6.3 μm H_2O Band.

of ~ 4 cm^{-1} resolution from 1.8 to 5.4 μm, is for a 1.3 km sea-level horizontal path.

Figure 18-81 shows a comparison with Chaney [1969] of the calculated upward atmospheric radiance and an interferometer measurement from a balloon flight over northern Nebraska. The measurement was taken at a float altitude of 34 km. The calculated radiance used the midlatitude winter model with a 23 km met. range and a ground temperature of 280 K.

Figure 18-82 shows a comparison of an interferometer measurement made from the Nimbus 3 satellite [Conrath et al., 1970] looking down over the Gulf of Mexico with the calculated atmospheric radiance. The resolution of the interferometer was 5 cm^{-1} as compared to the 20 cm^{-1} resolution of LOWTRAN. As shown in Figure 18-82, two theoretical models (tropical and midlatitude summer) were used for comparison and are displaced two divisions above and below the measured radiance for clarity. Both models assumed a 23 km met. range and used the temperature at 0 km in the model atmosphere as the boundary temperature.

Figure 18-83 shows the comparison of atmospheric ra-



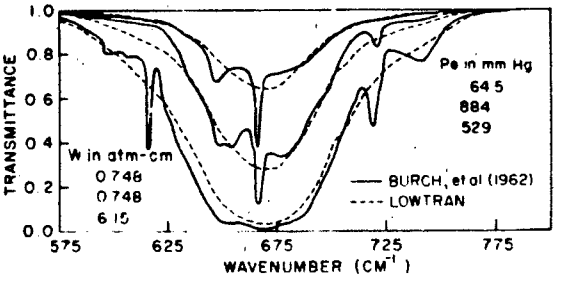


Figure 18-79. Comparison LOWTRAN calculations and Burch et al [1962] calculations for CO_2 bands at 4.3 μm and 15μm.

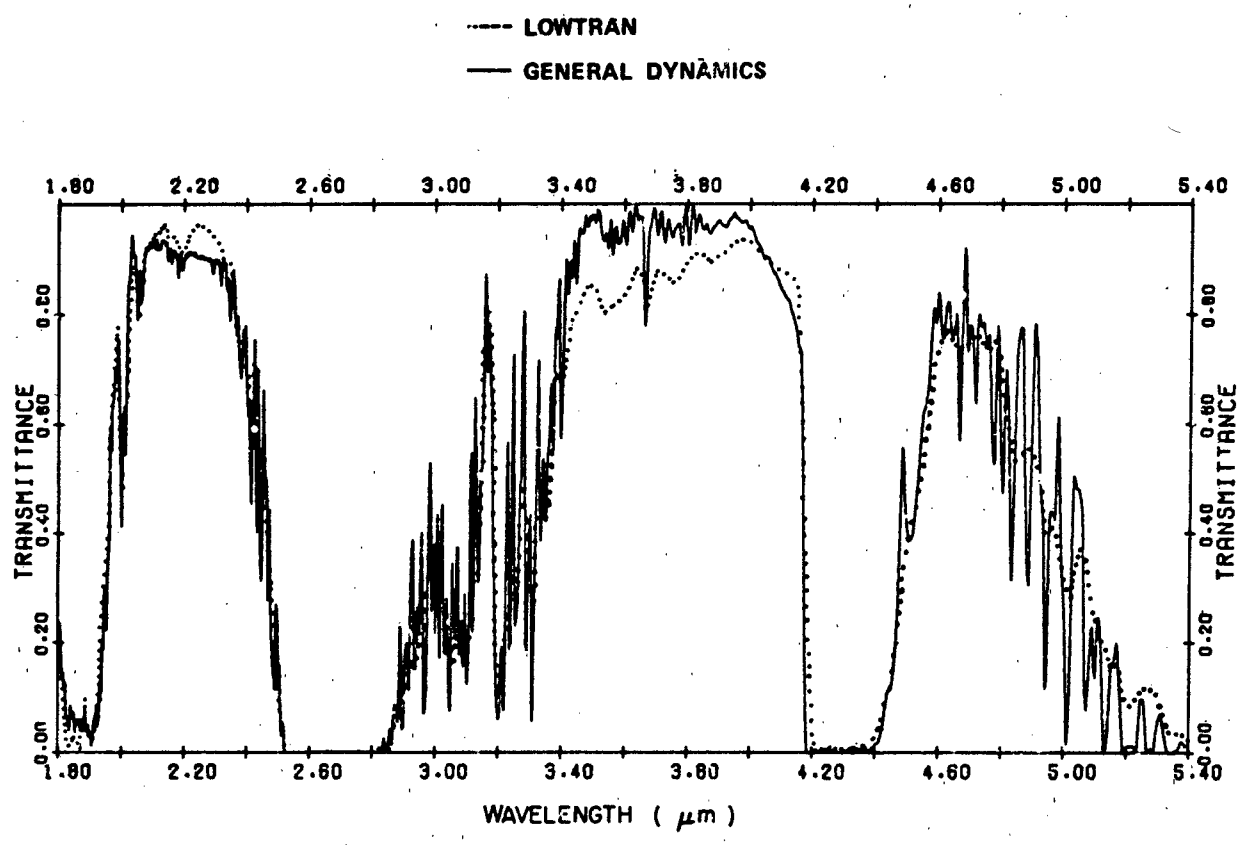


Figure 18-80. Comparison between LOWTRAN and General Dynamics measurements; range = 1.3 km at sea level.

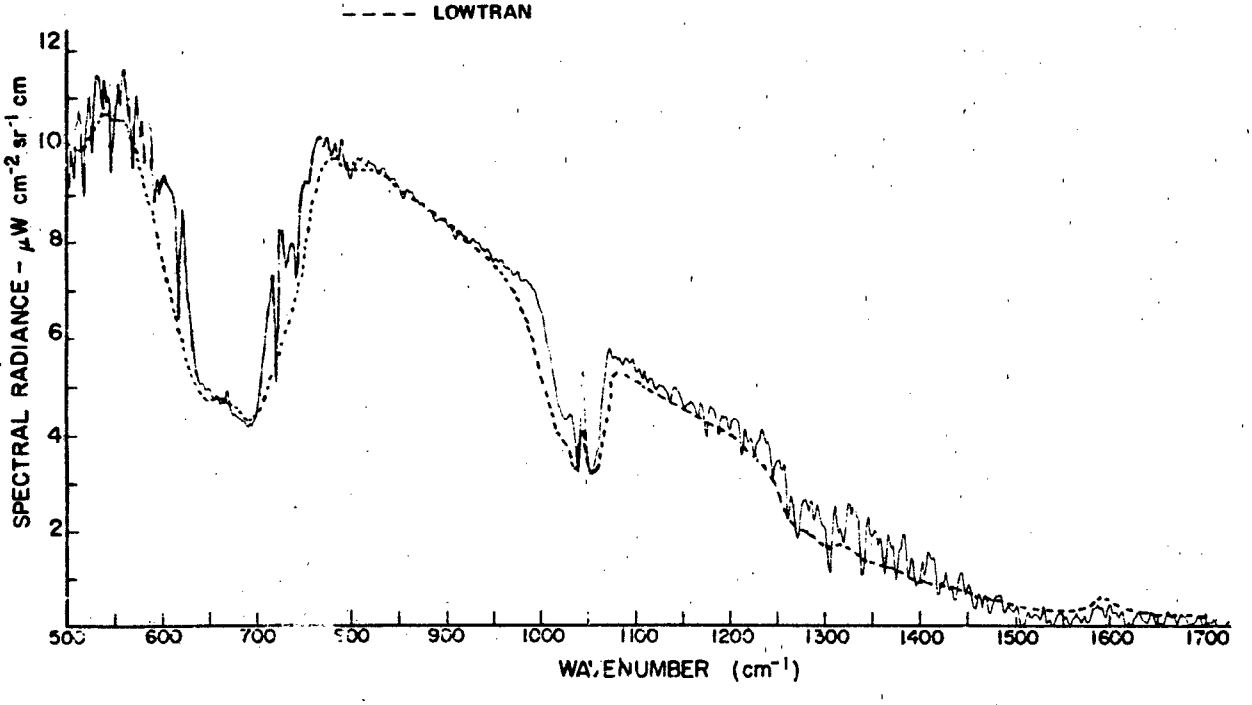


Figure 18-81. Comparison between LOWTRAN prediction and University of Michigan balloon measurement of atmospheric radiance over northern Nebraska.

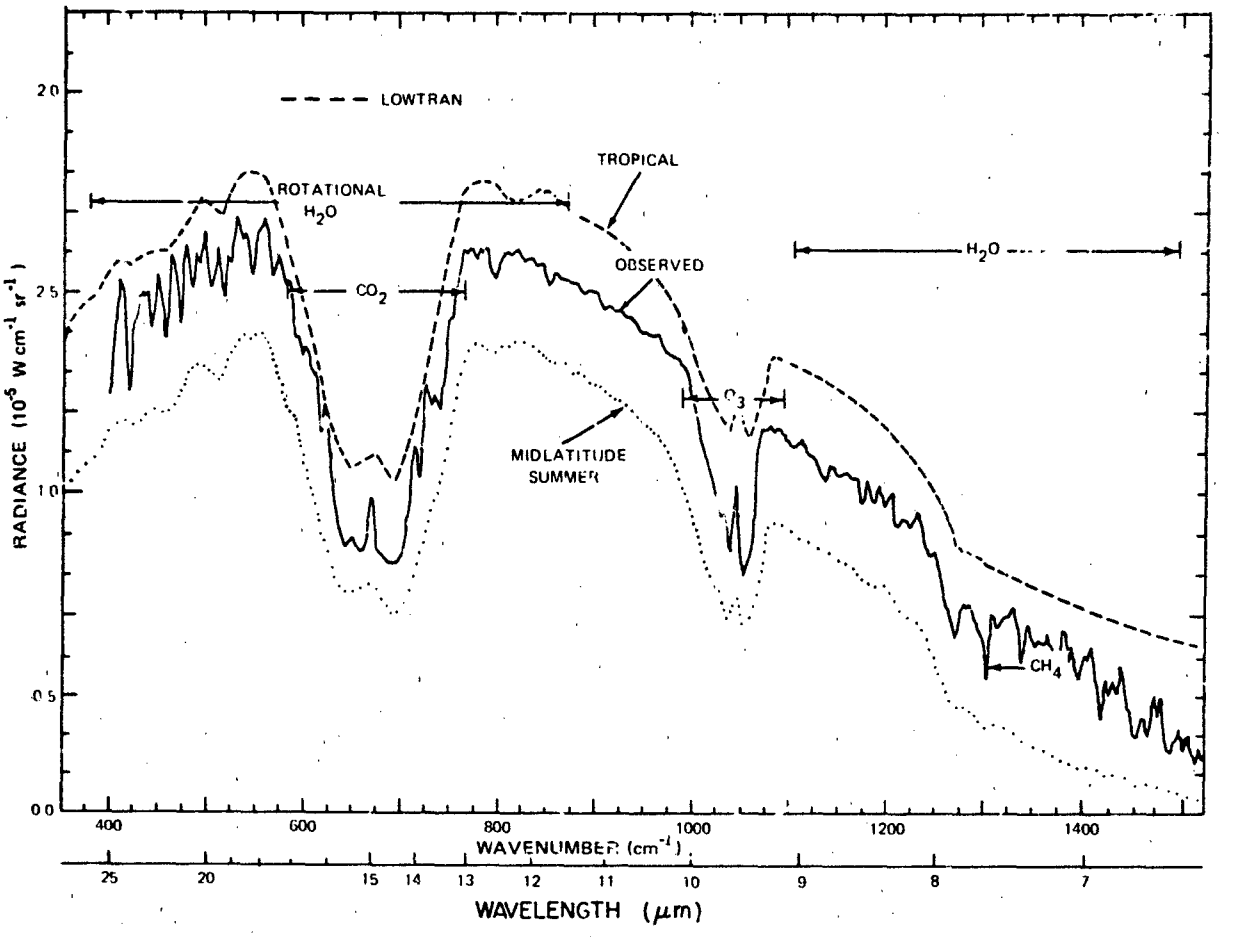


Figure 18-82. Comparison between LOWTRAN prediction and NIMBUS 3 satellite measurement of atmospheric radiance over the Gulf of Mexico.

diance as seen from space between the LOWTRAN calculation and measurements from the Nimbus 4 satellite [Hanel and Conrath, 1970] for three different geographic locations. The spectra, obtained with a Michelson interferometer of resolution 2.8 cm^{-1}, were measured over the Sahara Desert, the Mediterranean, and the Antarctic. The calculated LOWTRAN radiances used the midlatitude winter model and a ground temperature of 320 K for the Sahara, the midlatitude winter model and a ground temperature of 285 K for the Mediterranean; and an arctic winter cold model taken from the *Handbook of Geophysics and Space Environments* [Cole et al., 1965] and a ground temperature of 190 K for the Antarctic comparison. All three calculations assumed a 23 km met. range for aerosols.

Figures 18-84 through 87 show comparisons of calculated and observed atmospheric spectral radiance versus wavelength in the 8 to 14 μm spectral region. The measurements were made on a balloon flight launched from Holloman AFB, New Mexico by Murcray et al. [1977] University of Denver. The instrument used for these observations was a grating spectrometer, operated in the first and second order of the grating. The resolution was 0.03 μm in the 8 to 14 μm region. The data in these figures are presented as a function of altitude and as a function of zenith angle. The LOWTRAN radiance calculation used the pressure, temperature, ozone, and nitric acid profiles from the Murcray report, and the midlatitude winter water vapor profile contained in LOWTRAN.

18.4.2.7 Limitations of LOWTRAN. The overall accuracy in transmittance from LOWTRAN is better than 10%. The largest errors may occur in the distant wings of strongly absorbing bands in regions where such bands overlap appreciably. The reason for this error is twofold. First, the LOWTRAN spectral coefficients are based on a single absorber parameter and cannot be defined for a wide range of atmospheric paths without some loss in accuracy.

Secondly, the transmittance in the window regions between strong bands generally lies in the weak line approximation region where the transmittance is a function of the quantity of absorber present and not of the product of absorber amount and pressure. The one-dimensional prediction

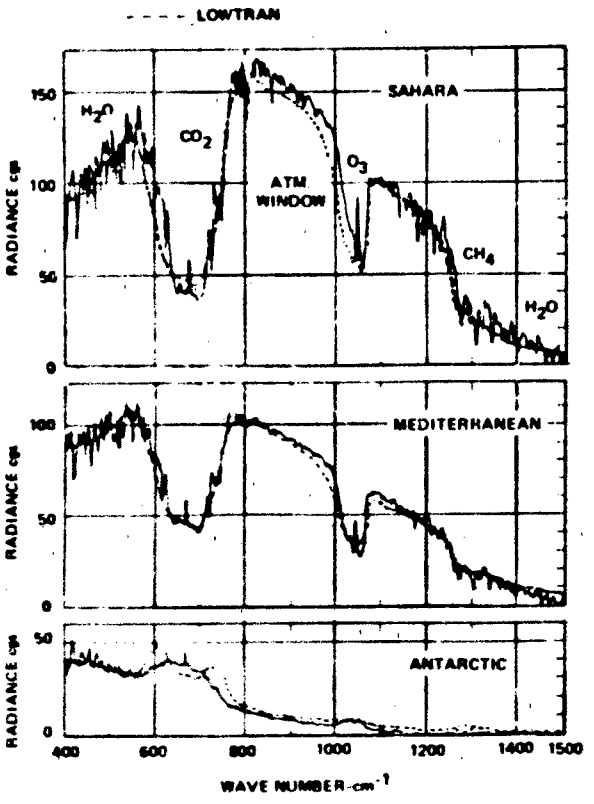


Figure 18-83. Comparison between LOWTRAN predictions and NIMBUS 4 satellite measurements of atmospheric radiance over the Sahara Desert, the Mediterranean, and the Antarctic.

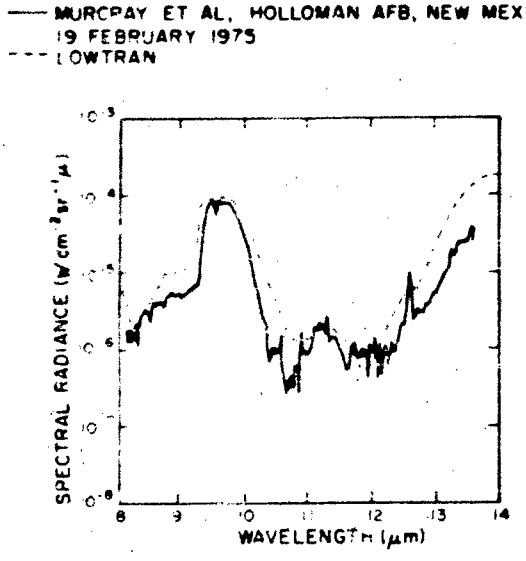


Figure 18-84. Sample spectrum of short wavelength region observed at an altitude of 9.5 km and a zenith angle of 63° on 19 February 1975 and LOWTRAN comparison.

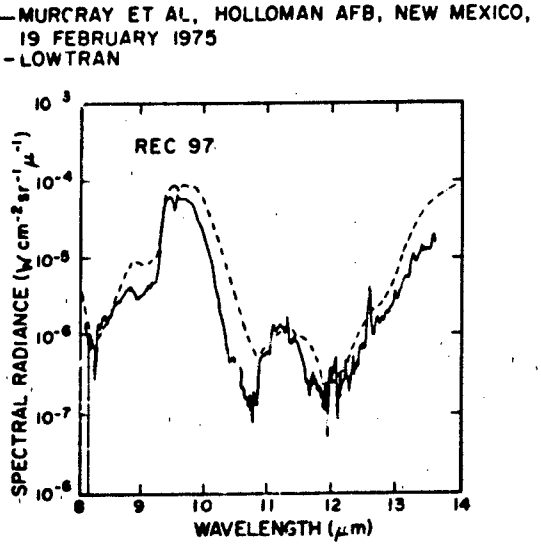


Figure 18-85. Sample spectrum of short wavelength region observed at an altitude of 13.5 km and a zenith angle of 63° on 19 February 1975 and LOWTRAN comparison.

scheme in LOWTRAN is less accurate for such conditions. The digitized spectral data used in LOWTRAN were obtained for conditions representative of moderate atmospheric paths and will tend to overestimate the transmittance for very long paths and underestimate the transmittance for very short paths in the spectral regions described above. As the transmittance approaches 1.0, the percentage error in trans-

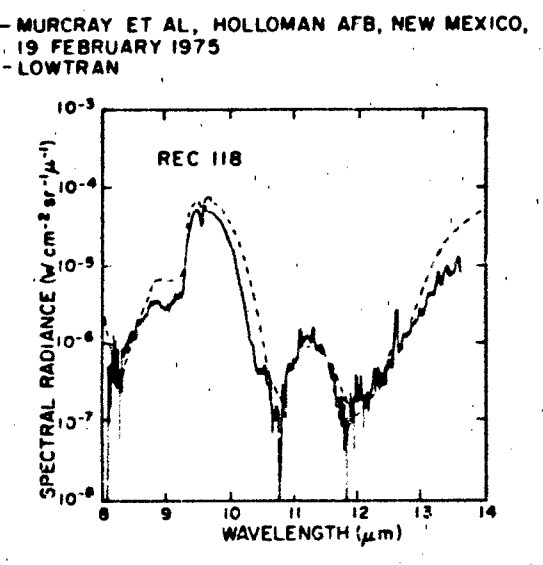


Figure 18-86. Sample spectrum of short wavelength region observed at an altitude of 18.0 km and a zenith angle of 63° on 19 February 1975 and LOWTRAN comparison.

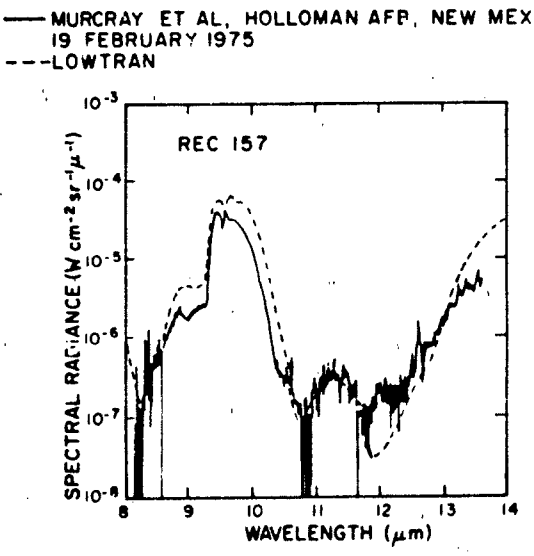


Figure 18-87. Sample spectrum of short wavelength region observed at an altitude of 24.0 km and a zenith angle of 63° on 19 February 1975 and LOWTRAN comparison.

mittance decreases toward zero but the uncertainty in the absorptance (or radiance) increases.

Additional constraints on both the validity of the model as well as the range of applicability are introduced for atmospheric radiance calculations. As mentioned above, the atmospheric radiance becomes less accurate for very short paths. In addition, the radiance calculations assume local thermodynamic equilibrium exists in each layer of the model atmospheres. This assumption will break down for radiance calculations in the upper atmosphere.

For the shorter wavelengths (<5 μm), scattered solar radiation becomes an important source of background radiation. This contribution for radiance calculations at the shorter wavelengths with a sunlit atmosphere would add to the atmospheric radiance calculation using the LOWTRAN model. Single-scattered solar and lunar radiation has been added to the latest LOWTRAN model [Kneizys et al., 1983].

18.5 RADIATION TRANSFER BY ATMOSPHERIC SCATTERING

18.5.1 Multiple Scattering

Scattering of visible and infrared radiation occurs in the atmosphere by air molecules, haze particles, and water droplets in the air (see Sections 18.1.4 and 18.2). Scattering effects from molecules and aerosols are additive. Multiple scattering occurs when scattered light undergoes subsequent scattering on other particles or molecules. The probability for multiple scattering events increases with increasing optical thickness $t = k_e \cdot x$ [see Equation (18.21)]. Van de Hulst [1957] states that if the optical thickness $t < 0.1$ single scattering prevails, for $0.1 < t < 0.3$ a correction for double scattering may be necessary. For still larger values of the optical depth the full complexities of multiple scattering have to be considered.

Monographs by Chandrasekhar [1950], Sobolev [1975], and van de Hulst [1980] are classical references on theory and approaches to radiation transfer. Recent reviews of multiple scattering have been presented by Hansen and Travis [1974], Herman et al. [1978], and by the Radiation Commission of the International Association of Meteorology and Atmospheric Physics (IAMAP/RC) [Lenoble, 1977 and Fouquart et al., 1980].

The basic radiative transfer equation in a scattering atmosphere defining the diffuse radiance I_ν can be written as

$$\mu \frac{dI_\nu(t_\nu, \mu, \phi)}{dt_\nu} = -I_\nu(t_\nu, \mu, \phi) + J_\nu(t_\nu, \mu, \phi) \tag{18.63}$$

where the source term, J_ν, is given by

$$J_\nu = \frac{\omega_\nu(t_\nu)}{4\pi} \int_0^{2\pi} \int_{-1}^{+1} P_\nu(\mu, \phi; \mu', \phi') I_\nu(t_\nu, \mu', \phi') d\mu' d\phi'$$
$$+ \frac{1}{4} \omega_\nu(t_\nu) P_\nu(\mu, \phi; \mu_0, \phi_0) \exp(-t_\nu/\mu_0) F_{\nu_0}$$
$$+ [1 - \omega_\nu(t_\nu)] B_\nu(t_\nu, T) \tag{18.64}$$

and where the following notation is used:

$t_\nu = \int_z^\infty k_\nu(z')dz'$ optical depth

$k_\nu(z)$ = extinction coefficient

$\omega_\nu(t_\nu)$ = albedo for single scattering

z = height

$\mu = \cos\theta$

θ = angle between a given direction and the direction of increasing optical depth, t (that is, towards the ground)

ϕ = azimuth angle

$P_\nu(\mu, \phi, \mu', \phi')$ = phase function defining the function of the light incident at μ', ϕ' which is scattered in the direction μ, ϕ

$B_\nu(t, T)$ = Planck blackbody function at temperature T

πF_{ν_0} = solar irradiance perpendicular to the direction of incidence, (μ_0, ϕ_0)

The first term on the right side of Equation (18.63) is the energy lost from the beam of radiation traveling in the direction (μ, ϕ). This loss is due to either absorption or scattering out of the beam. The second term (known as the source term) represents the diffuse radiation contributed to the direction (μ, ϕ) by scattering processes or self emission by the atmosphere. This source term is expanded in Equation (18.64) into three components (in order): the diffuse radiation scattered into the beam from other directions, the direct sunlight scattered into the beam direction (μ, ϕ), and the thermal radiation emitted by the atmosphere at frequency υ. For simplicity, the frequency υ will be omitted in subsequent equations. However it should be kept in mind that these equations only apply to spectral intervals sufficiently narrow that the atmospheric absorption and scattering processes do not change significantly with wavelength.

The full radiative transfer Equation (18.63) with the source term as in Equation (18.64) is too cumbersome to solve in general. However for most scattering problems in the earth's atmosphere, there is a natural separation of the parts of the source term. In the temperature region of interest for the earth's atmosphere, $B_\upsilon(t,T)$ contributes only in the longwave region (infrared and beyond). For wavelengths less than 3 to 4 μm, generally the thermal radiation $B_\upsilon(t,T)$ can be neglected compared to the solar irradiation $F_{\upsilon\infty}$ and for longer wavelengths $F_{\upsilon\infty}$ can be neglected compared to $B_\upsilon(t,T)$.

Also, for the longer wavelengths (greater than 3 to 4 μm) scattering is negligible except in the presence of clouds or fog. These cases where the source term is limited to the thermal radiation $B_\upsilon(t,T)$ are discussed in Section 18.4. Sections 18.5.2–18.5.4 will focus on the solution to Equation (18.63) when the thermal radiation can be neglected although many of the methods of solving the multiple scattering problem can easily be generalized to include atmospheric emission. The reduced expression for the source term is given by

$$J(t,\mu,\phi) = \frac{\omega(t)}{4\pi}\int_0^{2\pi}\int_{-1}^{+1} P(\mu,\phi;\mu',\phi')I(t,\mu',\phi')d\mu'd\phi' + \frac{\omega(t)}{4} P(\mu,\phi;\mu_\infty,\phi_\infty)\exp(-t/\mu_\infty)F_{\upsilon\infty} \qquad (18.65)$$

The phase function P used here is a weighted average of the phase functions for molecular scattering and aerosol scattering discussed in Sections 18.1.4 and 18.2.1.5 respectively. The weights are proportional to the scattering coefficients for scattering by air molecules and by aerosol particles.

18.5.1.1 Formal Solution to the Multiple Scattering Problem. The formal solution to the radiative transfer equation is given by

$$I(t,\mu,\phi) = \begin{cases} I^-(t^*,\mu,\phi)\exp\left[-\dfrac{(t^*-t)}{|\mu|}\right] + \dfrac{1}{\mu}\displaystyle\int_t^{t^*} J(t',\mu,\phi)\exp\left[-\dfrac{(t'-t)}{|\mu|}\right]dt', & \mu<0 \quad (18.66a) \\ J(t,\mu,\phi), & \mu=0 \quad (18.66b) \\ I^+(0,\mu,\phi)\exp\left(-\dfrac{t}{\mu}\right) + \dfrac{1}{\mu}\displaystyle\int_0^{t} J(t',\mu,\phi)\exp\left[-\dfrac{(t-t')}{\mu}\right]dt', & \mu>0 \quad (18.66c) \end{cases}$$

where t^* is the total optical depth of the atmosphere and where $I^-(t^*,\mu,\phi)$ for $\mu < 0$ and $I^+(0,\mu,\phi)$ for $\mu > 0$ (denoting the upward radiance at the ground and the downward radiance at the top of the atmosphere respectively) are specified by the boundary conditions: At the ground, the upwelling radiance is related to the downward radiance by an appropriate surface reflection law (see Section 18.3), and at the "top" of the atmosphere the downward diffuse radiance $I^+(0,\mu,\phi)$ for $\mu > 0$ is usually taken to be zero (that is, the downward radiance at the top of the atmosphere is due only to the incident sunlight). The formal solution [Equation (18.66)] does not completely solve the radiative transfer equation since J is defined in terms of I [Equation (18.65)]. However it is used in the development of several of the methods discussed below.

18.5.1.2 Method of Successive Orders of Scattering. The method of successive orders of scattering expands the radiance as [see e.g. Dave, 1964, 1965, or Irvine, 1965]

$$I(t,\mu,\phi) = \sum_{n=1}^{\infty} I^{(n)}(t,\mu,\phi) \qquad (18.67)$$

where $I^{(n)}$ is the part of the radiance scattered exactly n times. $I^{(n)}(t,\mu,\phi)$ is found by evaluating the formal solution [Equation (18.66)] using $J = J^{(n)}(t,\mu,\phi)$ which is given by Equation (18.65) with $I = I^{(n-1)}$ and starts with $I^{(0)} = 0$. This method has the advantage of easily handling vertically inhomogeneous atmospheres. The main disadvantage is that the computation time becomes very long as ω approaches 1.0 (absorption becomes weak) or as the optical depth becomes large.

18.5.1.3 Gauss-Seidel Iterative Method. The Gauss-Seidel iterative method originally developed by Herman and

Browning [1965] and others [Dave and Gazdag, 1970] also starts with Equations (18.65) and (18.66). Here they initially assume I = 0 and numerically integrate Equation (18.66) for small steps Δt for downward radiances ($\mu > 0$) until they reach the ground. They then integrate stepwise from the ground back to the top of the atmosphere using the downward radiances just calculated for the integral in Equation (18.65). The process is then repeated until the solution converges. Recently Herman et al. [1980] improved the computation speed by approximating $I(t,\mu,\phi)$ by polynomials in t which allowed the use of larger step sizes Δt for optically thick atmospheres.

18.5.1.4 The Matrix Operator or Layer Adding Methods. The radiative transfer problem has been solved by using different formations of the matrix operator or layer adding techniques reviewed by Plass, et al. [1973]. One of the first applications to light scattering in the atmosphere was the doubling method discussed by van de Hulst in an unpublished report in 1963 and later used by him and others [e.g. van de Hulst and Grossman, 1968; van de Hulst, 1971]. Hansen [1969a and b] has developed a modified version of this method and used it extensively [Hansen, 1971a and b; Hansen and Pollack, 1970].

An equivalent procedure was independently developed by Twomey et al. [1966] and has been used by them to study multiple scattering of light [Twomey et al., 1967]. A similar approach has been formulated by Grant and Hunt [1968, 1969a, b] based on invariance principles [Chandrasekhar, 1950] and the discrete space theory [Preisendorfer, 1965]. They have applied this matrix operator method to a variety of problems.

If we consider a plane-parallel medium of optical depth t and an incident radiation field $I_{in}(0, \mu', \phi')$, $(0 \leq \mu' \leq 1)$, it is convenient to introduce a scattering function $S(t; \mu, \phi, \mu', \phi')$ and a transmission function $T(t; \mu, \phi, \mu', \phi')$ [see Chandrasekhar, 1950]. Then the angular distribution of the reflected and transmitted light will be given by

$$I_{ref}(0,\mu,\phi) = \frac{1}{4\pi|\mu|}\int_0^1\int_0^{2\pi} S(t;\mu,\phi,\mu',\phi')\, I_{in}(0,\mu',\phi')\, d\phi' d\mu', \qquad 0 \geq \mu \geq -1 \qquad (18.68a)$$

$$I_{trans}(t,\mu,\phi) = \frac{1}{4\pi|\mu|}\int_0^1\int_0^{2\pi} T(t;\mu,\phi,\mu',\phi')\, I_{in}(0,\mu',\phi')\, d\phi' d\mu', \qquad 1 \geq \mu \geq 0 \qquad (18.68b)$$

The basic principle of the layer-adding methods is that given the scattering functions $S(t_1)$ and $S(t_2)$ and transmission function $T(t_1)$ and $T(t_2)$ of two layers, it is possible to find $S(t_1 + t_2)$ and $T(t_1 + t_2)$ for a combination of the layers in terms of the scattering and transmission functions for the individual layers. For a homogenous atmosphere, it is most efficient to start with $t_1 = t_2 = t_0$, where t_0 is small, and build up to $t = 2^p t_0$ in p steps, by a doubling procedure.

The Neumann series method (iteration in orders of scattering [see Irvine, 1965] has been used to find $S(t_0)$ and $T(t_0)$ for $t_0 \simeq 0.125$ by van de Hulst and Grossman [1968] and Irvine [1968b]. Hansen started with an initial optical thickness of $t_0 \simeq 2^{-25}$ where he could approximate $S(t_0)$ and $T(t_0)$ by the appropriate expressions for single scattering.

The formalism developed by van de Hulst and Hansen has only been applied to radiation emerging from an atmosphere and in its present form cannot be used to calculate the radiation within an atmosphere. The approach developed by Grant and Hunt overcomes this drawback by incorporating an algorithm for determining the radiances between the layers used in the calculation. However, this can be quite time consuming for a number of different depths. The advantage in computation time is also weakened if the approach is extended to vertically inhomogeneous atmospheres.

18.5.1.5 Monte Carlo Technique. The Monte Carlo method basically utilizes a random walk approach, following a large number of photons in turn as they scatter through the atmosphere until they are absorbed or emerge from the atmosphere. This procedure was first extensively applied to the study of light in the atmosphere by Collins, Wells, and their associates [Collins and Wells, 1965; Thompson and Wells, 1971; Collins et al., 1972; Blättner et al., 1974]. Subsequently, Plass and Kattawar [1968a, 1968b & 1972] have independently developed a Monte Carlo type of program. The Russian work in this area is covered in the monograph by Marchuk et al. [1980].

The major advantage of the Monte Carlo approach is its great flexibility. It can allow for all of the complications of a realistic atmosphere usually neglected by other methods: these include the spherical nature of the atmosphere, molecular band model type absorption, and horizontal inhomogeneities. The major drawback is that the errors in the calculations are proportional to $N^{-1/2}$, where N is the number of photons counted. This means that to increase the accuracy by an order of magnitude, the number of photons followed (and the computer time) must be increased by two orders of magnitude.

18.5.1.6 Discrete Ordinates. One of the best known solutions of the transfer equation is Chandrasekhar's [1950] method of discrete coordinates. Here, the phase function is expanded as a series of spherical harmonics:

$$P(\mu, \phi, \mu', \phi') = \sum_{m=0}^{N} (2 - \delta_{0,m}) \times \left[\sum_{l=m}^{N} \omega_l^m P_l^m(\mu) P_l^m(\mu)\right] \cos m(\phi' - \phi), \qquad (18.69)$$

where

$$\delta_{0,m} = \begin{cases} 0, & m \neq 0 \\ 1, & m = 0 \end{cases}$$

and the P_l^m are associated Legendre polynomials. Then the radiance is expanded in a cosine series, leading to the following system of equations:

$$\mu \frac{dI^m(t,\mu)}{dt} = -I^m(t,\mu) + \frac{1}{2}\sum_{m}^{n} \omega_l^m P_l^m(\mu) \int_{-1}^{+1} I^m(t,\mu') P_l^m(\mu') d\mu' + \frac{1}{4}(2-\delta_{0,m}) \sum_{m}^{n} \omega_l^m P_l^m(\mu_0)$$

$$m = 0, 1, \ldots, N. \qquad (18.70)$$

where

$$I(t,\mu,\phi) = \sum_{m=0}^{n} I^m(t,\mu)\cos m(\phi_0 - \phi). \qquad (18.71)$$

After suitable manipulations and evaluation of the integral in Equation (18.70) by Gaussian quadrature of order $n(4n^2 + 2N - 1)$, Equation (18.70) is replaced by a system of linear differential equations in terms of $I^m(t,\mu_i)$, where μ_i ($i = \pm 1, \pm 2, \ldots, \pm N$) are the quadrature points. This system of equations can be evaluated quite readily. In principle, as N (and n) $\rightarrow \infty$, the solution becomes exact. Various forms of this approach have been used by Samuelson [1967 and 1969], Liou [1973], Shettle and Green [1974], Stammes and Swanson [1981], and Stammes and Dale [1981].

18.5.1.7 Dodecaton Approach to Radiative Transfer (DART). A different discrete stream approach has been developed by Whitney [1972, 1974], and Whitney and Malchow [1978]. This approach appears to have considerable potential especially for problems requiring consideration of the curved geometry of the earth's atmosphere such as observations of the earth's limb or twilights. This method has several important conceptual differences from the discrete ordinate or two-stream and multichannel methods discussed elsewhere in this section. The DART method achieves discretion by angular integration of the radiance I in a given direction, $\hat{r}$, with a cosine to the n^{th} power weighting function:

$$S\hat{r} = \int_{\hat{p}\cdot\hat{r}>0} (\hat{p}\cdot\hat{r})^n I(\hat{r})\, d\Omega \qquad (18.72)$$

where $\hat{p} \cdot \hat{r}$ is the dot product of the unit vectors $\hat{p}$ and $\hat{r}$ and $S\hat{r}$ is a radiation stream whose nominal propagation direction is $\hat{p}$. Application of this integral operator to the radiative transfer equation leads to a system of linear differential equations in terms of the streams, $S\hat{r}$, in the $(n+1)(n+2)$ different directions $\hat{p}$.

For $n = 2$, there are 12 streams that can be arranged to be centered on the faces of a regular dodecahedron. This arrangement minimizes the number of different scattering angles that must be considered. Whitney has used her dodecaton approach in combination with exact single-scattering calculations and has used the shape of the single-scattering results to estimate the multiple-scattered radiances in directions other than for the 12 streams.

Whitney has compared the computation speed of her DART method with the Monte Carlo technique, which is the only other method that has been applied to radiative transfer calculations in realistic spherical atmospheres. She found the DART method to be considerably faster, although it is hard to judge how much faster because of the differences in computers and programming languages and the question of whether more angles and photon histories than necessary were used for the Monte Carlo calculation. In a comparison with other methods such as spherical harmonic or matrix operator for a homogeneous plane-parallel atmosphere [Lenoble, 1977], the DART method appeared to be slower and less accurate.

The primary advantage of the DART method is its applicability to a spherical geometry with reasonable accuracy and computer time. It is not as advantageous for plane parallel geometries or irradiance calculations.

18.5.1.8 Spherical Harmonic Method. The development of the spherical harmonic method goes back half a century [Jeans, 1917 and Eddington, 1926], however, it has been used more widely for neutron transport problems. In recent years, it has begun to be used more for atmospheric light scattering problems (for example, Devaux and Herman, 1971; Canosa and Penafiel, 1973; Dave, 1975a,b; and Karp et al., 1981). Like the discrete ordinate method, the spherical harmonic method expands the phase function as a sum of associated Legendre polynomial [Equation (18.69)]. However, it differs from the discrete ordinate method in that the radiance is also expanded in a series of spherical harmonics. Equivalently, the Fourier components I^m are expressed as an expansion of Legendre functions:

$$I^m(t,\mu) = \sum_{l=0}^{n} C_l^m(t) P_l^m(\mu). \qquad (18.73)$$

Substituting Equation (18.73) allows the integrals in Equation (18.70) to be evaluated analytically using the orthogonality of the Legendre functions. After additional manipulations of the equations, a linear system of first order differential equations for the $C_l^m(t)$ is obtained.

The major advantage of the spherical harmonic method is that it can give the radiances at all depths in the atmosphere and not only those emerging from the atmosphere. Those

timing comparisons made with other methods (the matrix operator method, successive orders of scattering, and Gauss-Seidel iteration) tend to favor the spherical harmonic method [see Lenoble, 1977], especially for cases where the sun is directly overhead and the azimuthal symmetry allows the system of Equation (18.70) to be limited to m = 0. The discrete ordinate method should be comparable to the spherical harmonic method in terms of computational efficiencies. The primary disadvantage of these methods is that the computational load increases with the number of terms kept in the expansion of I and P (or the asymmetry of the phase function used); also, the radiance values calculated from the basic solution often have appreciable oscillations about the correct solution unless some smoothing process is used [Dave and Armstrong, 1974 or Karp, 1981].

18.5.1.9 Small-Angle Approximations. When light is scattered by a particle whose dimensions are larger than the wavelength of the incident radiation, it is diffracted in a narrow forward cone. It is this sharp diffraction peak that greatly increases the difficulty of solving the multiple scattering by increasing the number of angles required to evaluate the integral in Equation (18.65). Romanova [1962 and 1963] has developed a procedure in which she separates the radiance for the small-angle scattering $\hat{I}$ from the rest of the radiance:

$$I(\tau, \mu, \phi) = \hat{I}(\tau, \mu, \phi) + \tilde{I}(\tau, \mu, \phi) \qquad (18.74)$$

and solves for $\hat{I}$ exactly. The rest of the radiation field $\tilde{I}$ can then be found from a modified form of the transfer equation. Irvine [1968a and b] has made a comparison of Romanova's approach with Neumann series and a doubling procedure and concluded that her method gives radiance accurate to at least 5%.

Weinman [1968] developed an approach in which the diffraction peak of the phase function was represented by a Gaussian distribution and the rest of the phase function by a Legendre expansion. He then solved for the small-angle scattering $\hat{I}$, using a Neumann solution. It was possible to do the integration analytically because of the use of the Gaussian distributions. The large-angle scattering $\tilde{I}$ was then found by applying the discrete coordinate solution.

Potter [1970] approximated the forward peak as a Dirac delta distribution. He then considered this forward-scattered radiation as not being scattered at all but as part of the transmitted radiation and used a truncated phase function. This necessitated a rescaling of the scattering coefficient and correspondingly, the optical depth and the albedo for single scatter, the rescaling depending on the fraction of radiation included in the forward peak that was eliminated. His approach gave quite good results except in the immediate vicinity of the diffraction peak in contrast with Weinman's approach which is most accurate for the diffraction peak.

Wiscombe [1977] discusses the delta-M Method which provides a systematic procedure for representing the phase function as the sum of a Dirac delta-function and a Legendre expansion (with 2M terms). This is a generalization of the approach for the delta-Eddington Approximation [Joseph et al., 1976]. Potter [1970] had basically just truncated the phase function by eye. McKellar and Box [1981] have provided a discussion showing the relationship between these delta-function approximations and the necessary rescaling of the transfer equation, with other scaling laws and similarity relations used in radiative transfer theory.

18.5.1.10 Approximations of Irradiance. Several different methods of calculating radiances within, or emerging from, a scattering atmosphere have been discussed. When interest is mainly in the irradiances F, less accurate solutions of the transfer equation can be used because the errors may cancel out in averaging the radiances over a hemisphere

$$F(\tau) = \int\int \mu I(\tau, \mu, \phi) d\phi d\mu \qquad (18.75)$$

where $\mu > 0$ corresponds to $F\downarrow$ and $\mu < 0$ to $F\uparrow$. Often it is adequate to have the irradiances computed accurately to a few percent, such as calculating heating rates or the irradiance incident at the ground. Several approximations for irradiances often used are basically simplifications of the closely related discrete ordinate and spherical harmonic methods discussed above. When considering the irradiance, only the first term (m = 0) in the cosine expansion of the radiance in Equation (18.71) must be included, since the higher order terms drop out in doing the azimuthal or ϕ integration in Equation (18.75). The most common are the Eddington approximation [Irvine, 1968b; Shettle and Weinman 1970], various formulations of the "two stream" approximation [Chu and Churchill, 1955, Sagan and Pollack, 1967, Coakley and Chylek, 1975], the delta-Eddington [Joseph et al., 1976], and the delta-two-stream [Schaller, 1979] approximations. These have been discussed and unified in the studies by Meador & Weaver [1980] and by Zdunkowski et al. [1980], [also see Wiscombe & Grams, 1976 and Welch & Zdunkowski, 1982].

They showed that the different Eddington and two-stream approximations all reduce the radiative transfer equation to a pair of coupled differential equations for the upwelling and downwelling irradiances.

$$\frac{dF\downarrow}{d\tau} = -\gamma_1 F\downarrow + \gamma_2 F\uparrow + \frac{\omega_0}{4}\pi\gamma_3 \exp(-\tau/\mu_0) F_0 \qquad (18.76a)$$

$$\frac{dF\uparrow}{d\tau} = +\gamma_2 F\downarrow - \gamma_1 F\uparrow - \frac{\omega_0}{4}\pi\gamma_4 \exp(-\tau/\mu_0) F_0 \qquad (18.76b)$$

where the values of the γ_i depend on the choice of approximation used.

18.5.1.11 Approximations for the Radiance. The different approaches to solving the radiative transfer equation described above generally must be solved numerically on a computer and permit tradeoffs between the accuracy of the solution and the computer memory and time. These methods in principle can be made arbitrarily accurate by increasing the number of terms kept in the expansions (such as Equation (18.69) or Equation (18.71)), by decreasing the step size used in numerically evaluating any necessary integrals, or by increasing the number of photon histories following the Monte Carlo Method.

There are also several techniques [Kaufman, 1979; Davies, 1980; and Hering, 1981] developed for solving the transfer equation which are aimed at retaining reasonable accuracy (10% to 20%) for the radiance field while minimizing the computer time requirements. These methods have a commonality in that they all use one of the two-stream/Eddington type solutions for the radiances to allow the integral in Equation (18.65) to be evaluated analytically. They differ in the details of the solution beyond that point; for example, Davies [1980] and Hering [1981] both explicitly include the exact single scattered radiances in their solutions whereas Kaufman [1979] does not.

18.5.2 Background (Sky) Radiance

Atmospheric scattering of solar radiation is responsible for the skylight. For a pure molecular atmosphere, analytic solutions were given by Coulson et al. [1960]. They provide tables of sky radiance and polarization including the Stokes parameters. Since then, much effort has been put into developing capabilities to calculate and model sky radiance for real aerosol containing atmospheres. These attempts were based on the different calculation methods for radiative transfer discussed in the previous section. The intensity, wavelength, and polarization of skylight have been studied over the decades to also derive information on atmospheric aerosols.

For some applications, simplifying assumptions can lead to useful results. As long as only single scattering is important and a homogeneous atmosphere is assumed (aerosol/air mixing ratio constant with height), the angular distribution of skylight at a specific wavelength can be obtained simply from

$$B(M, M', \Psi) = I_o \left(\frac{M}{M - M'}\right)\left(e^{-tM} - e^{-tM'}\right)\frac{F(\Psi)}{t_s} \quad (18.77)$$

where

I_o = extraterrestrial solar irradiance (W cm^{-2})

B = sky radiance (W cm^{-2} sr^{-1})

M = secant of the solar zenith angle θ_0

M' = secant of the line of sight angle θ

t = total optical thickness per unit air mass (scattering and absorption by molecules and aerosols

t_s = scattering optical thickness

$F(\Psi)$ = angular scattering intensity (molecular + aerosol) per unit air mass

The scattering angle Ψ is defined by

$$\cos\Psi = \cos\theta_0 \cos\theta - \sin\theta_0 \sin\theta \cos(\phi - \phi_0)$$

where $\phi - \phi_0$ is the angular azimuth difference between the sun direction and the line of sight. However, if tM or tM' becomes larger than 0.10, higher order scattering can no longer be neglected; this means that illumination of the scattering volume from the sky and the earth's reflecting surface becomes increasingly important. Higher order scattering becomes dominant for tM or tM' > 0.5, particularly for high ground albedo. This is particularly true for the radiation leaving the top of the atmosphere (radiation to space). The assumption of a homogeneous (constant mixing-ratio) atmosphere is usually not valid and leads to errors.

The most recent revision of the LOWTRAN transmittance radiance codes, LOWTRAN 6, [Kneizys et al., 1983] contains a subroutine which performs solar (or lunar) single scattering sky radiance calculation for either the atmospheric models built into LOWTRAN or any user provided atmospheric data with nonhomogeneous vertical profiles. Since LOWTRAN also allows a calculation of thermal emitted radiation from atmospheric molecules, the continuity from visible scattered to infrared thermal emitted sky radiance is provided in LOWTRAN for any lookangle, up- or downward. One must always keep in mind that single scattering calculations underestimate sky radiances; for a scattering optical depth of less than 0.7, the ratio of multiply to single radiation is in general less than 1.5. Single scattering is a good approximation for lookangles near the sun.

Procedures for multiple scattering calculations are generally time consuming and complex, especially for cases of strongly anisotropic aerosol scattering. Examples of Monte Carlo calculations for the distribution of the radiance of the hazy atmosphere, seen from both the ground and from space are shown in Figure 18-88 [McClatchey et al., 1972]. Dotted curves of the radiance (per unit solid angle and unit incident solar flux) are for the sun in the zenith and solid lines are for the sun at $\theta = 86.3°$. In Figure 18-88a, the downward or "transmitted" radiance generally increases with decreasing wavelength when the sun is in the zenith. Forward scattering causes the radiance to peak near the sun. Near the horizon, the radiance increases again, except at short wavelengths, and the albedo influence is large. At low solar elevations, downward radiance is generally much smaller especially near the zenith at long wavelengths, and the al-

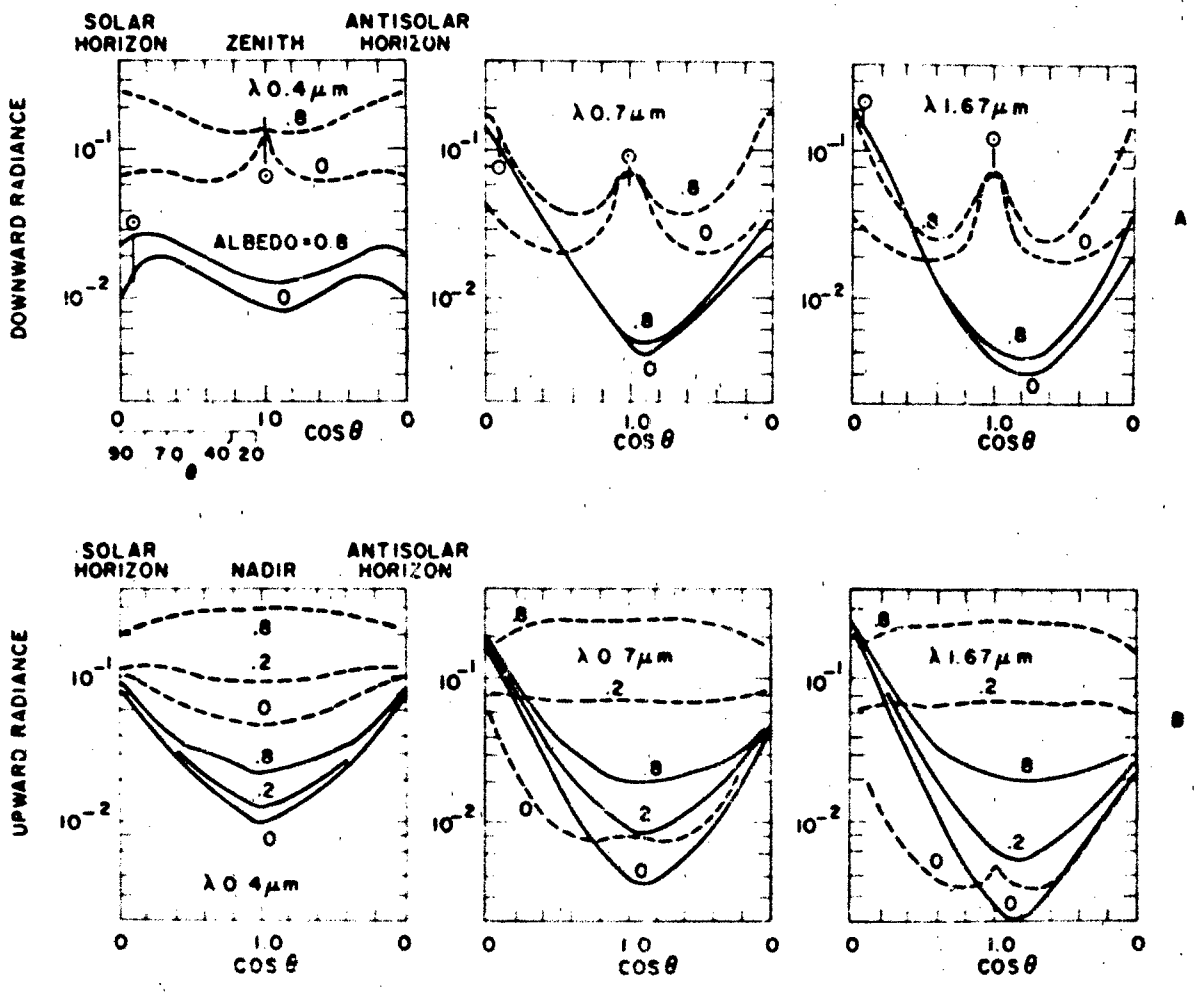


Figure 18-88. Downward (a) and upward (b) radiance computed by Monte Carlo techniques. Solid curves are for the sun at $\theta = 86.3°$. The indicated parameter is the surface albedo and λ is the wavelength. Radiances are per unit solid angle and unit incident solar flux.

bedo influence is significant only at short wavelengths. Since the sky radiance values have been averaged over intervals of 0.1 in cos θ (that is, over angular intervals ranging from 5° to 25°), the radiance near the sun actually should be much higher than shown.

In Figure 18-88b upward or reflected radiance, as seen from outside of the atmosphere, is strongly dependent on the albedo of the ground. The contribution of reflected ground radiance becomes dominant, especially in the near infrared and in the nadir direction where the atmospheric backscattered flux becomes small. This effect will be less pronounced in more hazy atmospheres. The distribution of sky radiance in the ultraviolet is dominated by multiple scattering and below 0.35 μm by ozone absorption. At high solar elevations, downward as well as upward radiances at 0.30 μm are only about 1% of their respective values at 0.40 μm.

In the case of the radiance from clouds, see Figure 18-89 for a rather dense nimbostratus [Kattawar and Plass,

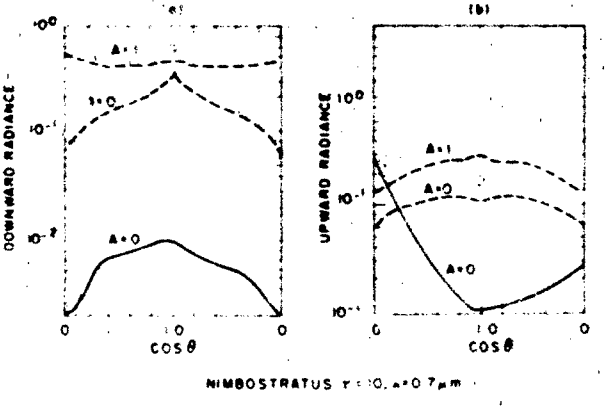


Figure 18-89. Downward (a) and upward (b) radiance computed for a dense Nimbostratus cloud. Dashed curves are for the sun in the zenith. Solid curves are for the sun at $\theta = 86.3°$.

1968]. Aerosol and molecular scattering have been neglected in this model because they would have little effect. The strong downward radiance (Figure 18-89a) for high sun and low albedo is largest near the sun, but with high albedo is nearly independent of the angle of observation. At low sun, the much smaller cloud radiance is largest in the zenith. The radiance of the cloud top (Figure 18-89b) again is very large at high sun. At low sun there is little discrimination between a cloudy and a clear sky (see Figure 18-88b, $\lambda = 0.7$ μm).

Polarization of the cloudless sky is symmetrical to the sun meridian and varies with sun elevation. The degree of polarization [see Equation (18.18)] varies from 0 to as much as 50 to 60% with a maximum around 90° from the sun. Sky light polarization decreases with increasing aerosol content in the atmosphere.

18.5.3 Illuminance-Irradiance

The irradiance at the earth's surface due to direct sunlight, diffuse skylight, moonlight, starlight and airglow varies over approximately nine orders of magnitude (Figure 18-90).

A still frequently referenced publication on natural illumination conditions which provides charts for predicting day and night time illumination levels is a report by Brown [1952]. This report is also included as an appendix in a more easily obtained document by Biberman et al. [1966] on the same subject. Figure 18-91 gives the solar spectral irradiance curves at sea level for various optical air masses. The total solar plus sky irradiance at the surface varies on a clear day between about 0.1 to 1.15 kW/m².

A considerable amount of data on total irradiance over the whole solar spectrum, dependence on time of day, season, and geographic distribution has been accumulated in the literature [for example, Robinson, 1966, Schulze, 1970; SOLMET, 1979].

In addition to astronomical parameters such as sun elevation or moon phase the irradiance at the surface depends largely on the atmospheric scattering properties.

The ratio of diffuse skylight to direct sunlight changes from about 0.6 (at a sun elevation of 10°) to 0.08 (with the sun near the zenith). For a very hazy atmosphere the corresponding ratios would be between 1 and 0.13. Clouds may reduce the total radiance by as much as 90% or more.

During recent years computer programs have been developed at several places to calculate illuminance levels as a function of time (day and night) and location. Unfortunately we are not aware of any that have been published. Figures 18-92 and 18-93 are extracted from the report by Brown [1952] and allow an estimate of the natural illuminance levels of the surface during day or night. Figure 18-92 shows the illumination (footcandles) as a function of altitude (or elevation angle) of the sun or moon and Figure 18-93 the relative intensity of moonlight as a function of the phase angle.

It should be pointed out that except for extremely remote regions of the continents night time illumination levels are always more or less affected by illumination from artificial lights, especially with reflection from clouds under overcast conditions.

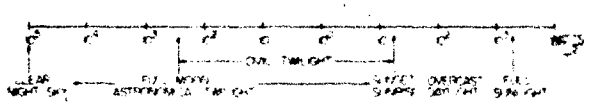


Figure 18-90 Range of natural irradiance levels.

18.5.4 Contrast Transmittance–Visibility

The "quality" of any image is largely a function of the contrast between different elements in the image scene. Several definitions for contrast exist. One definition frequently used is modulation contrast

$$C_{\text{modulation}} = \frac{\text{Max. Brightness} - \text{Min. Brightness}}{\text{Max. Brightness} + \text{Min. Brightness}} \tag{18.78}$$

The contrast between two elements in the object plane (the inherent contrast) is reduced along the atmospheric path due to scattering and turbulence in the visible and due to absorption and thermal atmospheric emission in the infrared. In a transitional spectral region all processes may be of importance.

The apparent contrast in the image plane is proportional to the inherent object contrast $C_a = C_o \cdot \tau_c$ where τ_c is called contrast transmittance. If the spatial resolution is defined as the minimum separation between two object scene elements that the receiver optical system can resolve, it becomes apparent that spatial resolution is a function of atmospheric contrast reduction.

The contrast transmittance for a target with sinusoidally varying brightness is called the modulation transfer function (MTF). In a turbulent atmosphere (see also Section 18.7), the MTF decreases with increasing spatial frequency of the object scene. For near zero spatial frequencies, turbulence phenomena vanish and contrast loss is due only to atmospheric scattering processes. In this case the apparent radiance of a distant object is the sum of two radiation flux components: residual image forming radiation from the object after propagating through the atmosphere and radiance caused by scattering of ambient light incident on the optical path into the direction of the receiver. This second quantity is called path radiance; it obviously does not contain any information about the object scene. Duntley et al. [1957] has shown that if the contrast is defined as

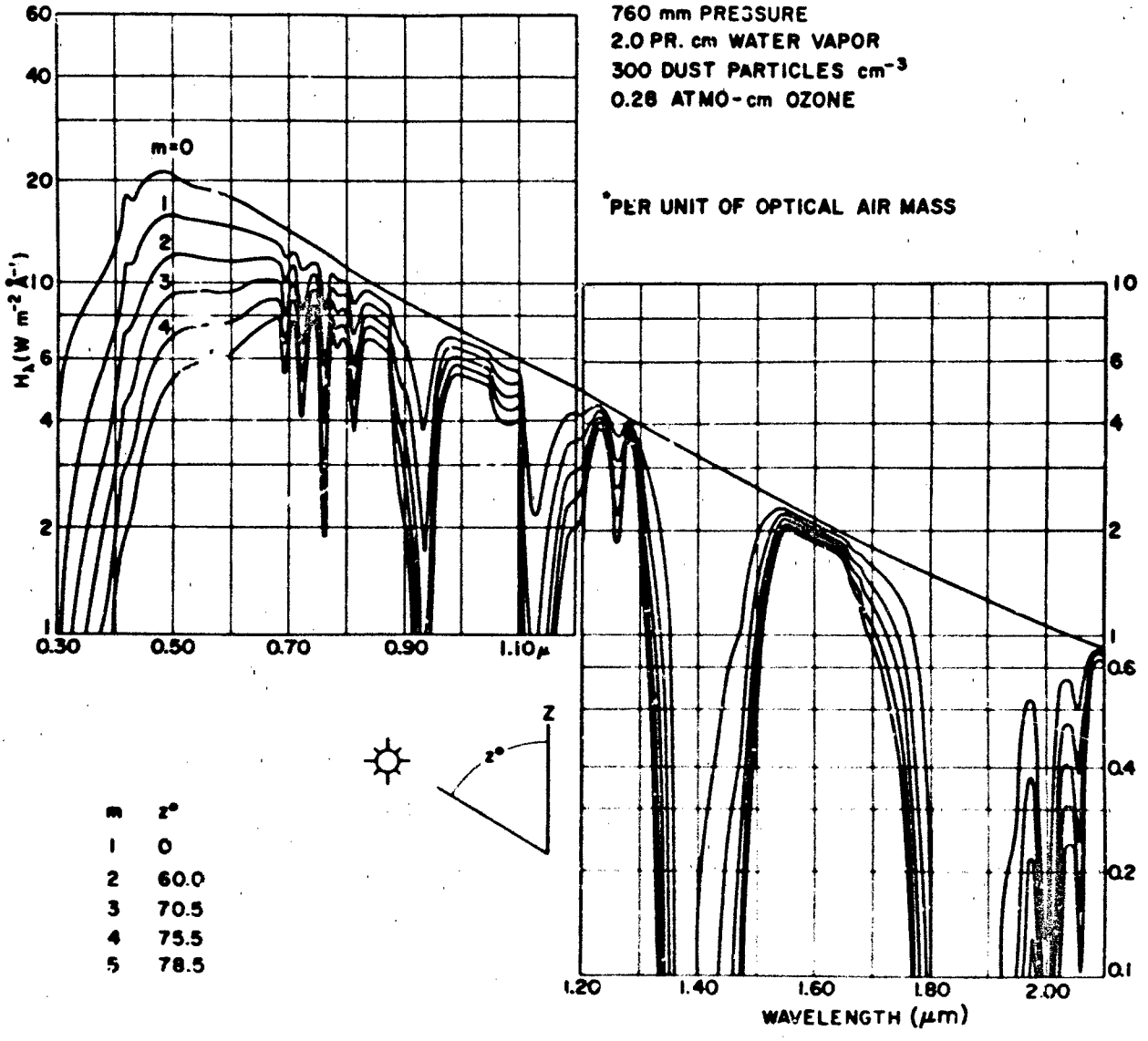


Figure 18-91 Solar spectral irradiance curves at sea level for various optical air masses. The value of the solar constant in this calculation was 1322 W m⁻² [Moon 1940].

$$C_{universal} = \frac{\text{object radiance} - \text{background radiance}}{\text{background radiance}} \qquad (18.79)$$

a universal contrast transmittance can be derived as

$$\tau_{atc,\ universal} = \frac{1}{1 + P/(N_{bb} \cdot \tau)},$$

where P is the path radiance, N_{bb} the inherent background radiance, and τ the beam transmittance. The universal contrast transmittance is only a function of atmospheric and background properties, not the target.

Whereas the path radiance in the visible spectrum is due to scattered radiation, the physical process for path radiance in the infrared is thermal emitted radiation by each path element. This path emission is a function of the atmospheric composition and its emittance and also of the temperature of the path element. The path emission is related to the absorption by Kirchhoff's law, and it is those gases (and aerosols) that absorb infrared radiation that also emit radiation.

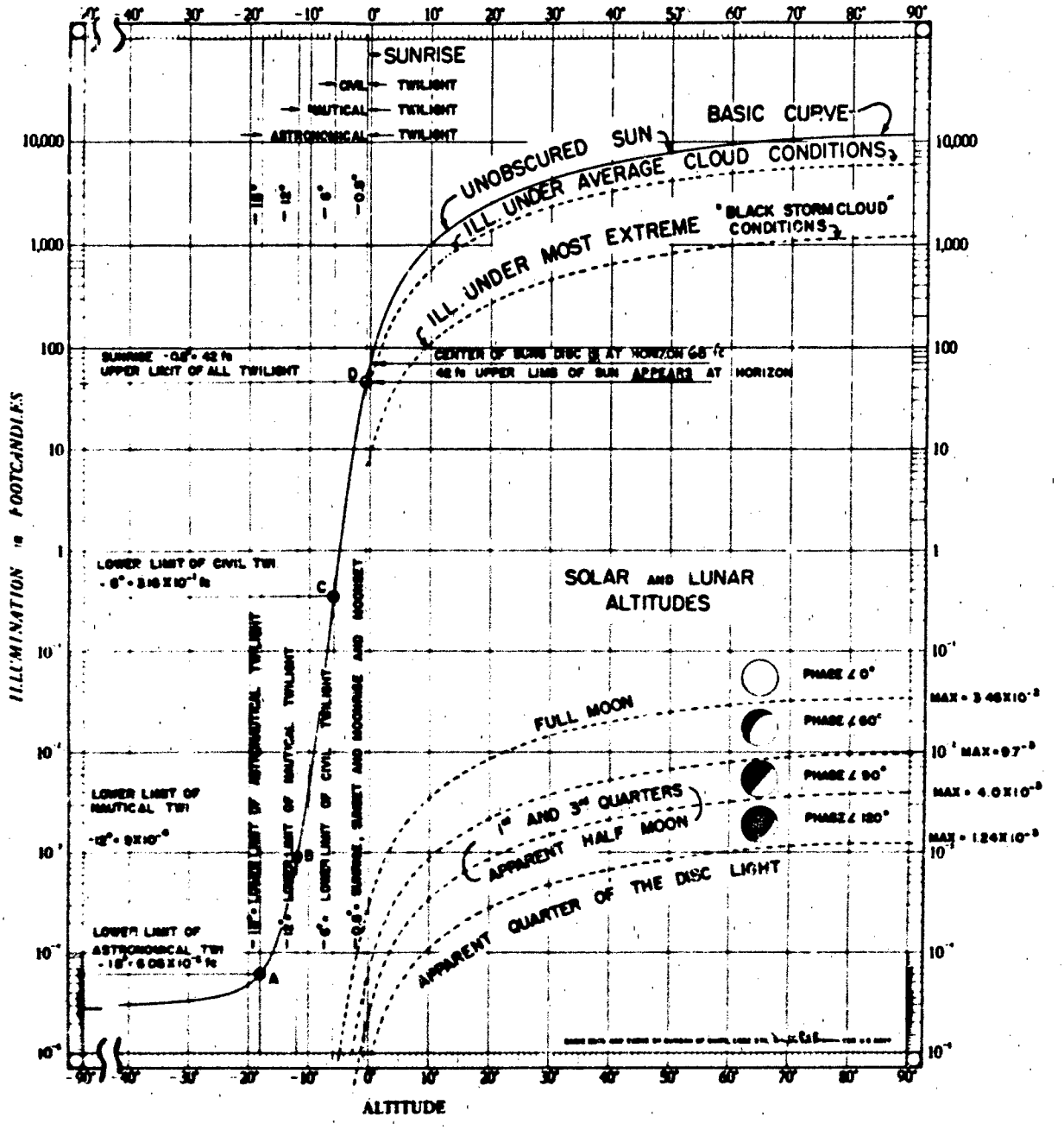


Figure 18-92. Total range of natural illumination levels

Reduction of scene contrast due to atmospheric scattering varies over several orders of magnitude depending on wavelength, atmospheric turbidity, source-slant path geometry, and ground albedo.

Figure 18-94 gives the universal contrast transmission:

$$\tau_{atc} = \frac{C'_{apparent}}{C'_{inherent}} \qquad (18.80)$$

for a receiver located on top of the atmosphere looking straight down. The ground albedo, A in Figure 18-94 is 10, 30, and 60% (ground albedo is the ratio of total incident to total reflected radiance from the ground, see Section 18.3). The contrast transmittance is shown for two different atmospheric models, one corresponding to clear conditions with 23 km surface visibility, and one for a hazy atmosphere with 3 km visibility on the ground. The figure shows the

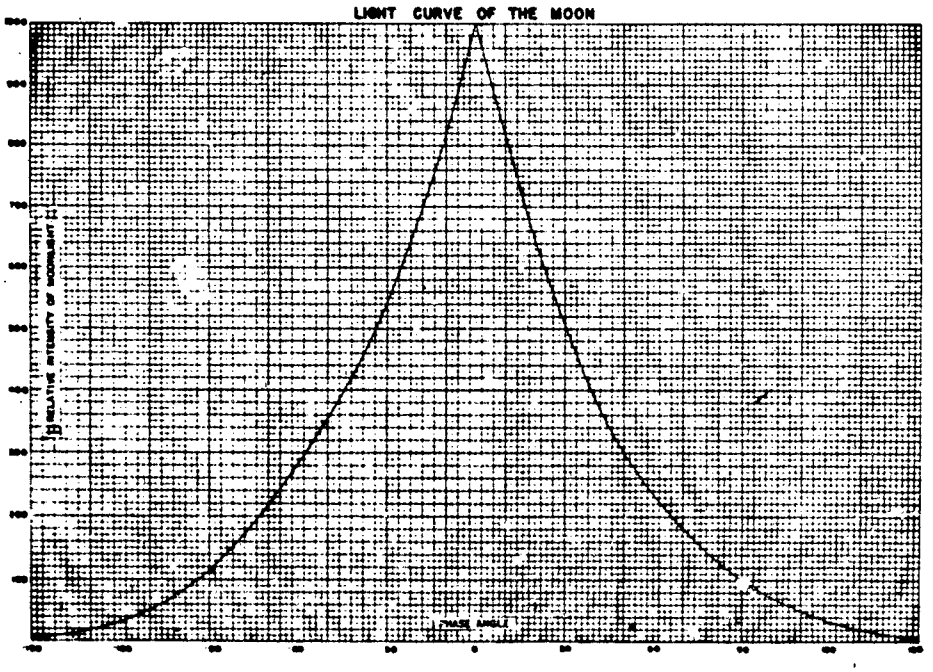


Figure 18-93. Relative intensity of the moon's illumination as a function of phase angle.

dependence of τ'_c on wavelength, sun zenith angle, and albedo. The results of these model calculations illustrate the range of contrast transmittance values that, for instance, will have to be considered in the interpretation of imagery from a satellite platform.

Figure 18-95 shows the effect of different types of aerosols on contrast transmission. In this figure the ratio of contrast transmittance for an urban versus a rural aerosol (see Section 18.2.1.4) is shown under identical visibility of

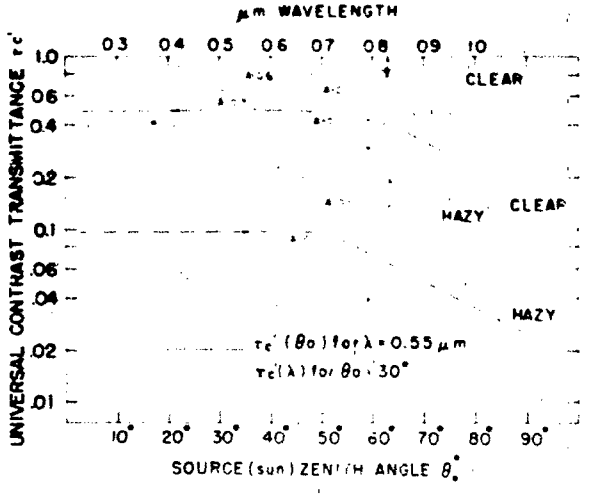


Figure 18-94. Contrast transmittance for a downward looking vertical path through the whole atmosphere.

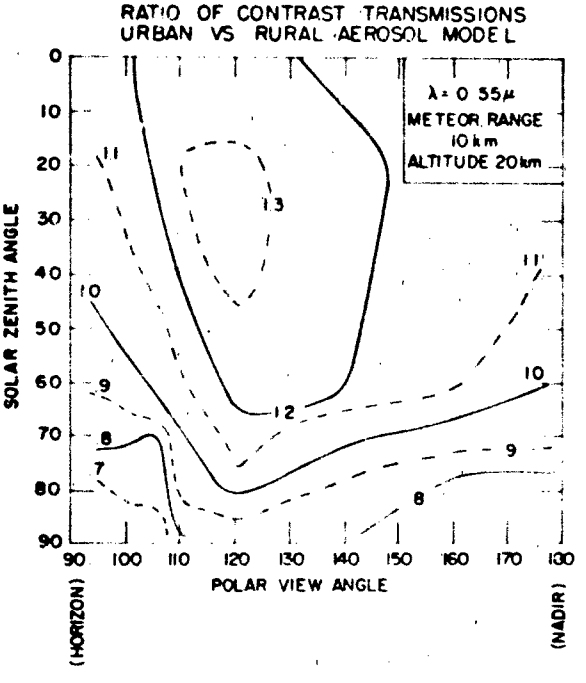


Figure 18-95. Ratio of contrast transmittance τ_c, urban/τ_c, rural for an urban vs rural aerosol model over downward looking paths. The wavelength is 0.55 μm, the surface meteorological range is 10 km, sensor altitude is 20 km.

10 km for various sun and look angles. The difference in contrast transmittance is due to the different scattering properties of the two aerosol models.

The theoretical determination of contrast transmittance involves, in the term of the path radiance [Equation (18.79)], the treatment of multiple scattering. Several of the radiation transfer computational methods described in Section 18.5.1 have been used for path radiance and contrast transmittance modeling. The examples given in Figures 18-94 and 18-95 above have been obtained from Monte Carlo calculations.

Figures 18-96 and 18-97 show an example of contrast transmittance and derived target acquisition ranges that were obtained by applying the delta-Eddington method [Hering, 1981] mentioned earlier.

A special case of viewing an object through the atmosphere is the viewing of objects along a horizontal path against the horizon sky. This case is described by the often quoted (but frequently misunderstood) Koschmieder theory [Middleton, 1952]. It states that the brightness of a black object B_b at distance r is given by

$$B_b = B_h (1 - e^{-\sigma r}) \qquad (18.81)$$

where B_h is the horizon sky brightness and σ = extinction coefficient. If one assumes that (based on tests) the contrast threshold necessary to see the object against the background $(B_h - B_b)/B_h$ is 0.02, the limiting visible distance, called meteorological visible range becomes

$$V = \frac{3.912}{\sigma} \qquad (18.82)$$

It must be pointed out that this relationship is based on the assumption of a black (dark) object of an angular extent larger than a few minutes of arc. Table 18-13 gives [Hulburt,

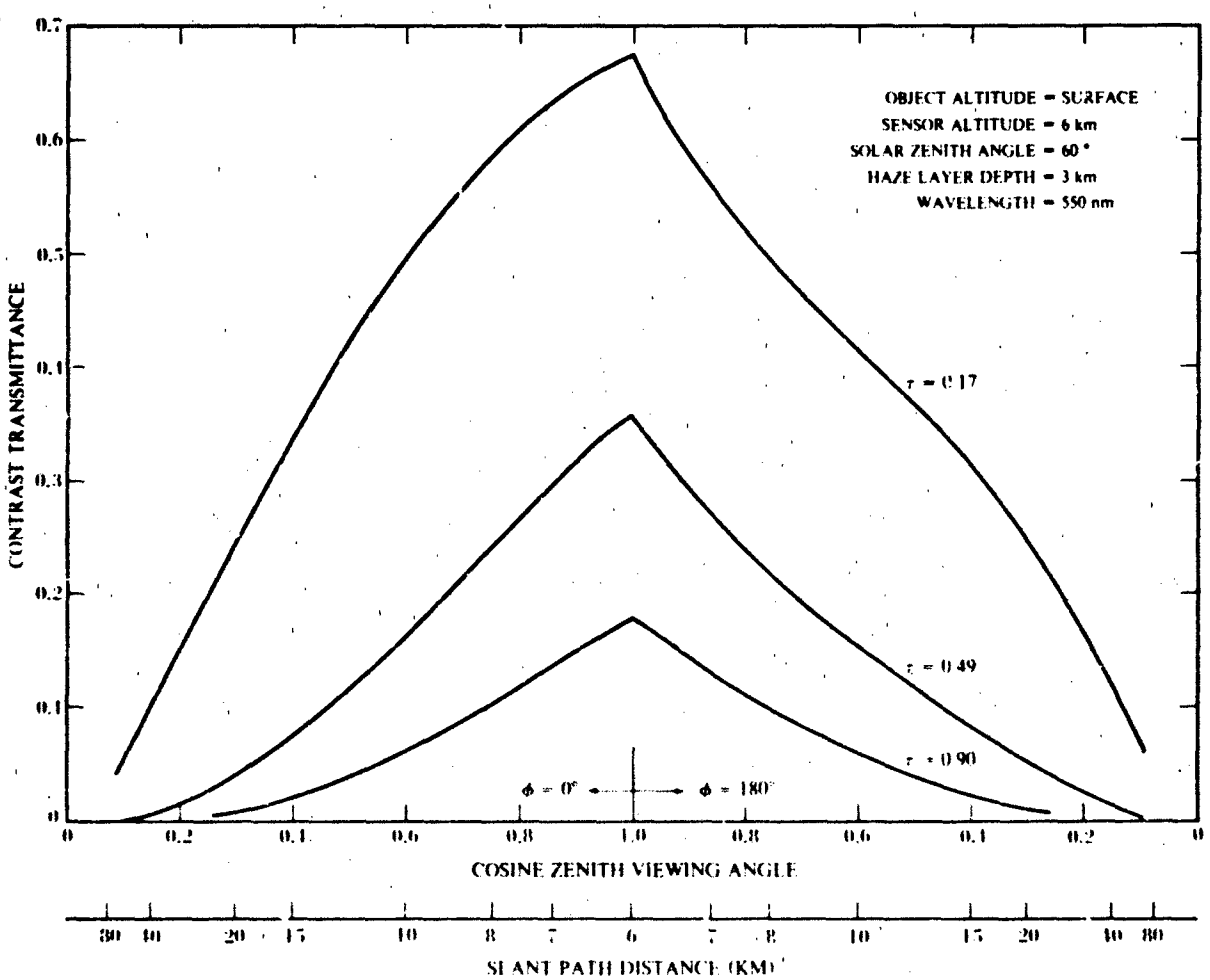


Figure 18-96. Calculations of contrast transmittance as a function of zenith viewing angle for a target viewed against a terrain background. Values of τ are total atmosphere optical depth

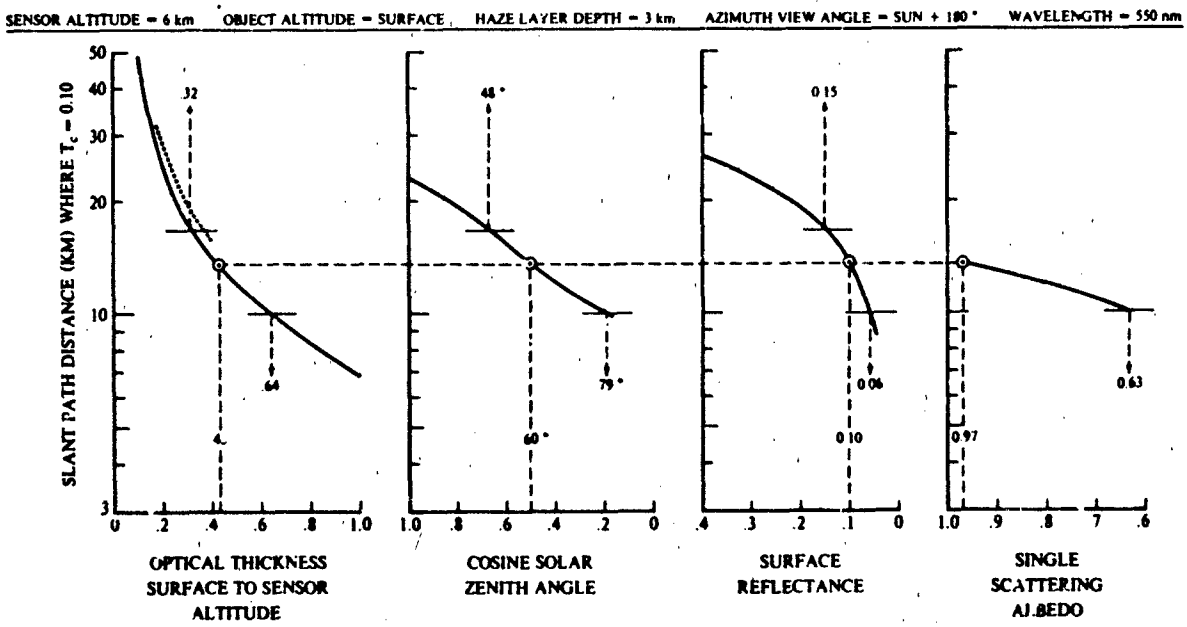


Figure 18-97. Variations in target acquisition range (contrast transmittance τ_c = 0.10) associated with departures of selected parameters from the assumed reference values. The horizontal dashed line denotes parameter values for this reference atmosphere. The horizontal bars identify a change of ± 25% from the slant path distance where T_c = 0.10 for the reference atmosphere, for example, 15.5 ± 3.4 km.

1941 or McCartney, 1976] the international visibility code and corresponding weather condition, meteorological range, and scattering coefficients. Since visibility marks are generally not black and often their angular extent is small, a threshold contrast of 0.05 gives usually a better definition for the meteorological range. Table 18-13 contains both the 0.02 and 0.05 contrast threshold scattering coefficients. In going from a measured scattering to an expected observed visibility, the assumption of an 0.02 contrast threshold gives a 30% higher visibility than 0.05. The values for pure air are based on molecular scattering only, for air at standard conditions, and for a spectrally weighted average wavelength for daylight.

18.6 ATMOSPHERIC REFRACTION

Changes in molecular density and the resulting gradient in the index of refraction cause the bending of a light ray traveling through the atmosphere. Over long paths at sunset, refraction causes such phenomena as the flattening of the solar disc and the "green flash." Strong temperature gradients over short paths are responsible for such effects as mirages and looming. Microscale and time dependent variations of the index of refraction cause optical turbulence and scintillation—the twinkling of stars. This section will consider only the large scale effects of refraction.

18.6.1 Refractive Bending

In a spherically symmetric atmosphere, a light ray follows the path defined by the equation

$$m(r) \times r \times \sin\theta = C \qquad (18.83)$$

where m is the atmospheric refractive index (Section 18.1.3), r is the radius vector from the center of the earth, θ is the angle formed by the ray and the local zenith, and C is a constant. The radius of curvature R of a light ray is given by

$$\frac{1}{R} = -\frac{1}{m}\frac{dm}{dr}\sin 2\theta. \qquad (18.84)$$

The refractive modulus N (see Equation 18.7, Section 18.1.3) is very nearly proportional to the total air density so that it can be approximated, at least over a limited vertical range, by an exponential of the form

$$N(r) = N_o e^{-(r-r_o)/H} \qquad (18.85)$$

where N_o is the value of N at some reference radius r_o, and H is the scale height for N. In this case

$$\frac{1}{R} = \sin\theta\left[H\left(\frac{1}{N_o}e^{-(r-r_o)/H} + 1\right)\right]^{-1} \qquad (18.86)$$

Table 18-13. International visibility code, meterological range, and scattering coefficient.

		Meterological Range, R_m		Scattering Coefficient σ (km^{-1})	
Code No.	Weather condition	Metric	English	for 0.05 contrast	for 0.02 contrast
0	Dense fog	< 50 m	<55 yd	> 59.9	> 78.2
1	Thick fog	50–200 m	55–219 yd	59.9–15.0	78.2–19.6
2	Moderate fog	200–500 m	219–547 yd	15.0–5.99	19.6–7.82
3	Light fog	500–1000 m	547–1095 yd	5.99–3.00	7.82–3.91
4	Thin fog	1–2 km	1095 yd–1.1 naut. mile	3.00–1.50	3.91–1.96
5	Haze	2–4 km	1.1–2.2 naut. mile	1.50–0.749	1.96–0.954
6	Light Haze	4–10 km	2.2–5.4 naut. mile	0.749–0.300	0.954–0.391
7	Clear	10–20 km	5.4–11 naut. mile	3.00–0.150	3.91–0.196
8	Very Clear	20–50 km	11–27 naut. mile	0.150–0.060	0.196–0.078
9	Exceptionally clear	> 50 km	727 naut. mile	< 0.060	< 0.078
—	Pure air	277 km	149 naut. mile	—	0.0141

In a typical case from the *U.S. Standard Atmosphere 1976* for a horizontal path at the ground ($\theta = 90°$), the ratio of the radius of the earth r_E to R is 0.15 at 5 μm.

Since the total air density normally decreases with altitude, light rays are normally bent toward the earth. However, density inversions can occur in the atmosphere, typically as a thin layer above a strongly heated surface such as the desert or a road surface. In such cases, light is bent upwards.

The bending ψ of a ray is shown in Figure 18-98, where θ and ϕ are the zenith angles of the ray at r_1 and r_2 and β is the earth centered angle, and

$$\psi = \pi + \beta - \theta - \phi \qquad (18.87)$$

The bending along a path can be calculated by integrating one of the following equations:

$$\frac{d\psi}{dr} = \frac{l}{R} \tan\theta$$

$$\frac{d\psi}{d\beta} = \frac{r}{R} \qquad (18.88)$$

$$\frac{d\psi}{dS} = \frac{\sin\theta}{R}$$

where S is the distance along the curved path from some point.

The total bending for various paths through the atmosphere is shown in Figures 18-99a,b and 18-100. Figure

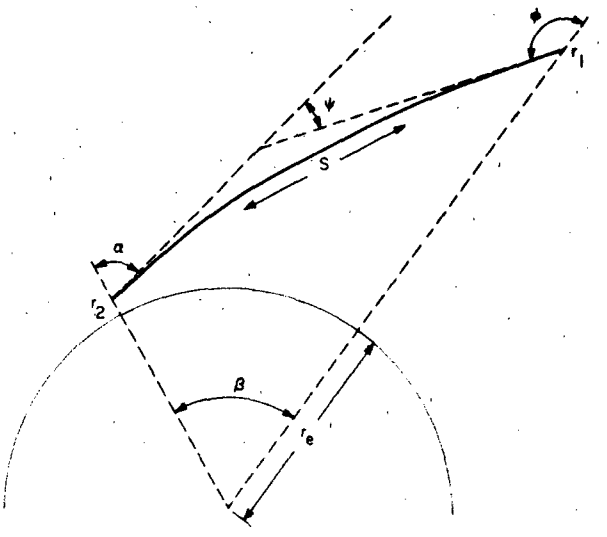
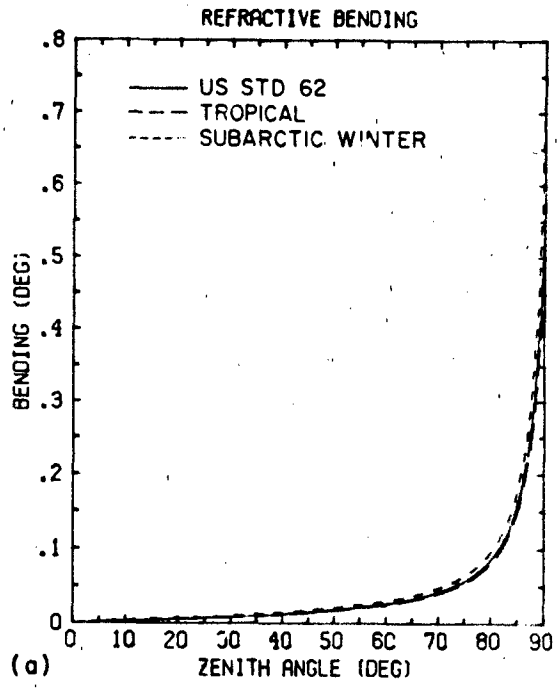


Figure 18-98. Geometry of the refracted path through a single layer.



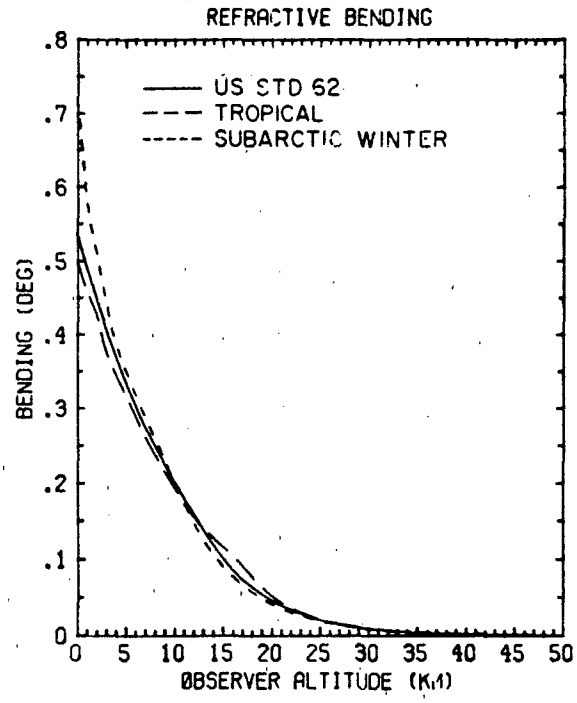


Figure 18-100. Refractive bending vs observed altitude for three atmospheric profiles for path from space to observed altitude tangent height.

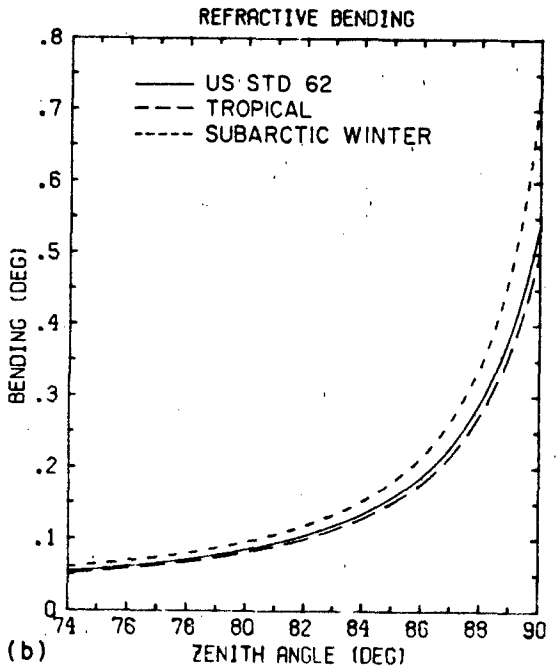


Figure 18-99. Refractive bending vs zenith angle for three atmospheric profiles for a path from 0 to 100 km altitude for zenith angles. (a) 0°–90°, (b) 74°–90°

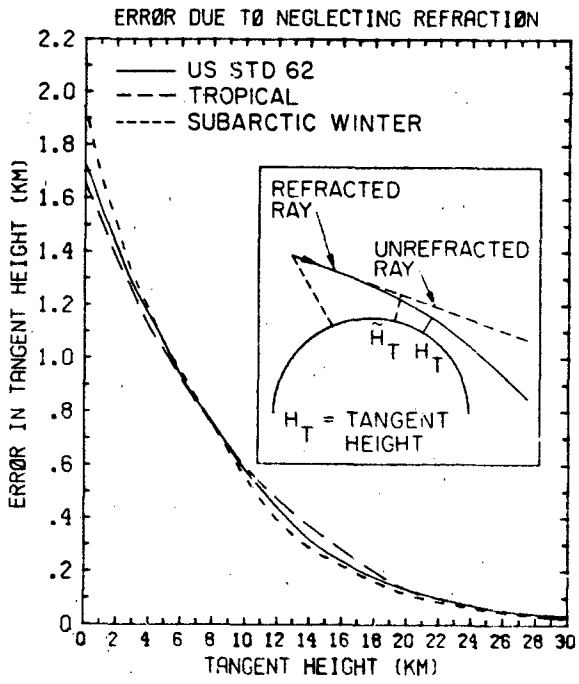


Figure 18-101. Error in predicted tangent height defined as tangent height minus refracted tangent height vs. refracted tangent height for three atmospheric profiles.

18-99a,b shows the bending or zenith angle at the ground for a path from ground to space, for three different atmospheric profiles. The three profiles are the U.S. Standard 1976, the Tropical (15°N), and Subarctic Winter (60°N Jan) profiles [McClatchey et al., 1972]. Figure 18-100 shows the bending for a path looking out horizontally to space from the observer altitude.

Figure 18-101 shows the difference between the refracted tangent height and the unrefracted tangent height versus the refracted tangent height for a ray coming in from infinity. This path corresponds to, for instance, a satellite instrument scanning the earth limb.

18.6.2 Optical Air Mass

Calculations of the absorption and emission of electromagnetic radiation in the atmosphere require knowledge of the amount of the absorbing gases along the path. Examples of such situations are the absorption of solar radiation along a path to the ground and the emission of infrared radiation along a tangent path through the atmosphere.

The quantity of interest is the integrated absorber amount μ along the path, given by

$$\mu = \int \rho ds \qquad (18.89)$$

where ρ is the density of the absorbing gas and ds is the element of length along the path. μ is also known as the "column density" and has units of molecules cm^{-2} if ρ is the number density. The value of μ for air for a vertical path from ground (1013.25 mb) to space is μ_v [molecules cm^{-2}]. The ratio of μ for air for a given path to μ_v is known as the "air mass;" for example, the air mass value for a path from ground to space where the zenith angle at the ground is 90° is 38.1.

For a spherically symmetric atmosphere and for paths where the zenith angle is less than about 70°, the absorber amount can be calculated to better than 1% using the plane parallel approximation, that is:

$$\mu = \int \rho ds = \sec\theta \int \rho dz \qquad (18.90)$$

where θ is the zenith angle along the path and z is the vertical coordinate. For a gas whose density follows an exponential profile

$$\rho(z) = \rho(z_0) e^{-(z - z_0)/H} \qquad (18.91)$$

the vertical integral of the density has a particularly simple form

$$\int_{z_0}^{\infty} \rho dz = H[\rho(z_0) - \rho(\infty)] \qquad (18.92)$$

where the scale height H is

$$H = -\frac{z_2 - z_1}{\ln[\rho(z_2)/\rho(z_1)]}. \qquad (18.93)$$

The total air density and the density of uniformly mixed gases are well represented by profiles of the form of Equation (18.91) especially over altitude ranges of less than a scale height (the density scale height under normal atmospheric conditions varies between 6 and 9 km).

For zenith angles greater than 70°, the effects of the earth's curvature and of refraction become significant. In this case, plane parallel approximation is no longer valid and a more elaborate integration of Equation (18.92) is necessary. In a case where the density profile is exponential but refraction can be neglected, the absorber amount can be calculated using the "Chapman functions." Non-exponential profiles and cases where refraction is important generally require a numerical integration of Equation (18.92). A computer program for calculating absorber amounts for any profile and an arbitrary path is available.

There are a number of tables of air mass for various geometries and profiles [Chapman, 1931; Smith and Smith, 1972; Gallery et al., 1983; Bemporad, 1968; Kasten, 1966 and 1967; Sneider and Goldman, 1974; Sneider, 1975]. Table 18-14 gives the air mass for selected zenith angles for a path from the ground (1013.25 mb) to space, including the effects of refraction. The profile used is the *U.S. Standard Atmosphere 1976*, and the values are accurate to within 1% for wavelengths from 0.5 micrometers to the far infrared. Table 18-15 gives the air mass for various tangent heights for a path from space to space through the tangent height.

Air mass values for paths with large zenith angles cannot in general be applied to non-uniformly mixed gases, such as water vapor and ozone. This fact is shown in Figures

Table 18-14. Air mass versus zenith angle for a path from ground (1013.25 mb) to space, *U.S. Standard Atmosphere, 1976*, 0.23 μm to the far infrared.

Zenith Angle (Deg)	Air Mass	Zenith Angle (Deg)	Air Mass
70.0	2.90	88.0	19.4
72.0	3.21	88.2	20.5
74.0	3.58	88.4	21.8
76.0	4.07	88.6	23.1
78.0	4.71	88.8	24.6
80.0	5.58	89.0	26.2
82.0	6.87	89.2	28.1
84.0	8.85	89.4	30.2
85.0	10.3	89.6	32.5
86.0	12.3	89.8	35.1
87.0	15.2	90.0	38.1
87.5	17.1		

(For zenith angles less than 70°, the secant approximation is good to better than 1%.)

Table 18-15. Air mass versus tangent height (H_T) for a path from space to space through the tangent height, *U.S. Standard Atmosphere, 1976*, 0.23 μm to the far infrared.

H_T	Air Mass	H_T	Air Mass	H_T	Air Mass
0.0	76.2	10.0	21.4	30.0	0.920
1.0	67.9	12.0	17.7	35.0	0.43
2.0	60.4	14.0	11.4	40.0	0.208
3.0	53.6	16.0	8.28	45.0	0.107
4.0	47.5	18.0	6.01	50.0	0.0563
5.0	41.9	20.0	4.38		
6.0	36.9	22.0	3.18		
7.0	32.4	24.0	2.32		
8.0	56.6	26.0	1.70		
9.0	49.4	28.0	1.25		

18-102a,b which show the "air mass" value for air, water vapor, and ozone versus zenith angle for a path from ground to space. The "air mass" values for water vapor and ozone are the ratio of the absorber amount for the given zenith angle to the absorber amount for a vertical path for the same profile. The temperature profile used is the *U.S. Standard Atmosphere 1976* while the profiles of water vapor and ozone used are shown in Figure 18-103. At 90°, the "air mass" value for water vapor is 72.2 and for ozone 14.4 compared to 38.1 for air. For the profiles of water vapor and ozone, the absorber amounts can be computed to better than 1% using the secant approximation up to 80° for water vapor and 60° for ozone (compared to 72° for air or a uniformly mixed gas). These numbers depend upon the profile used and should be taken as guidelines only.

Figure 18-104 shows the air mass value versus tangent height for the three components for a path from the tangent height to space. These curves are similar in shape to the density profiles of the components themselves, since the bulk of the integrated amount is within a few kilometers (vertically) of the tangent height.

Figure 18-105 shows the air mass versus zenith angle for a typical stratospheric balloon-borne spectroscopic measurement: the observer is at 30 km scanning between a zenith angle of 85° to 95.5° at which point the path intersects the earth. (The zenith angle is the apparent or measured zenith angle, not the astronomical zenith angle.) Also shown are the tangent height versus zenith angle and the angular diameter of the sun. If the sun is used as the source for an observation, the absorber amounts to different points on the face of the sun may differ by a factor of 2 for air and an order of magnitude for water vapor. The variation in air mass due to this effect can be the major source of uncertainty in a measurement and must be considered carefully.

18.7 ATMOSPHERIC OPTICAL TURBULENCE

Astronomers have long known that atmospheric turbulence can cause optical propagation effects. The famous "twinkling" of stars is due to such turbulence. When turbulence is at a minimum, astronomers can photograph planets.

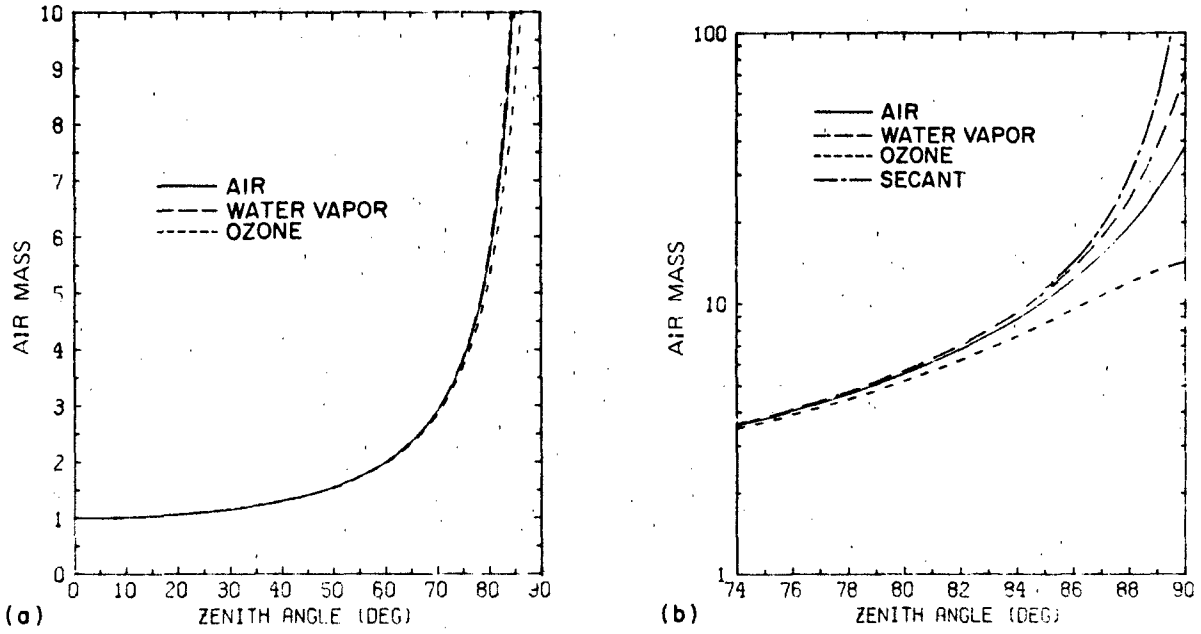


Figure 18-102. Air mass vs zenith angle for a path from 0 to 100 km altitude, *U.S. Standard Atmosphere, 1962*, including refraction (a) 0°-90°, (b) 74°-90°.

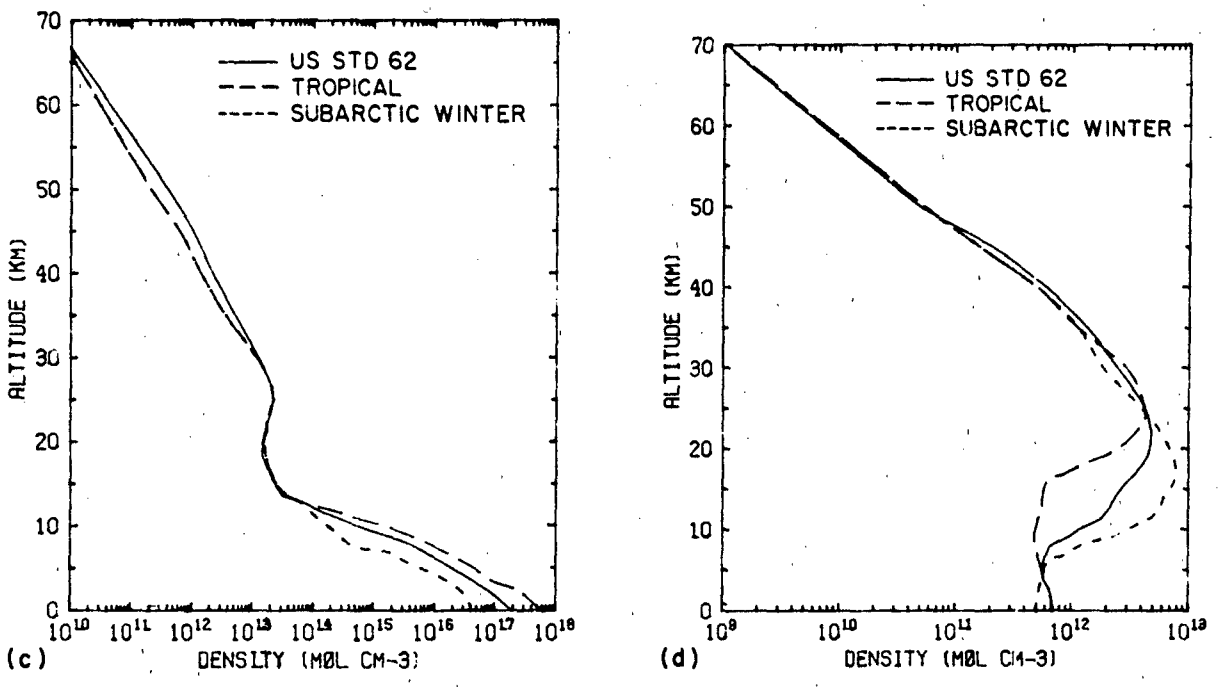


Figure 18-103. a) Water vapor profile of the U.S. Standard model, 1962. b) Ozone profile of the U.S. Standard model, 1962.

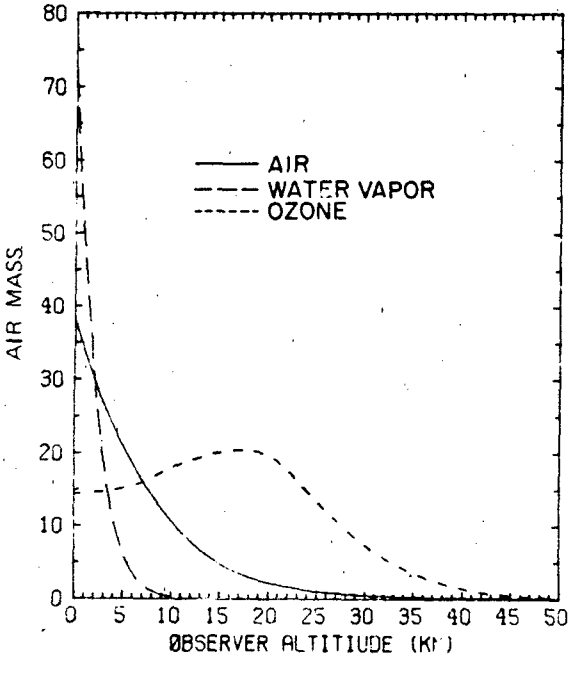


Figure 18-104. Air mass vs observer altitude for a path from 100 km to the observer, zenith angle 90°, *U.S. Standard Atmosphere, 1962*, including refraction.

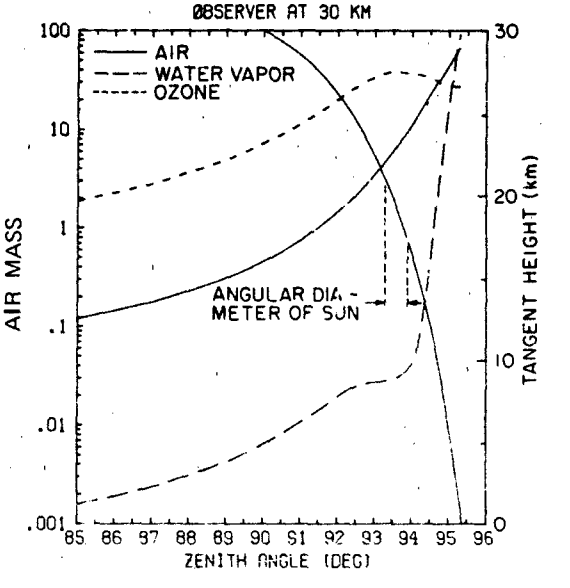


Figure 18-105. Air mass vs zenith angle for a path from 30 to 100 km altitude, *U.S. Standard Atmosphere, 1962*, including refraction. Also shown are the tangent height versus zenith angle and the angular diameter of the sun.

etc., with a minimum of blur because with less turbulence there is less so-called "image dancing." Brightness fluctuations, or scintillations, also diminish when there is low turbulence. In general, such conditions are called "good seeing conditions." The basic cause of the optical effects is the presence of fluctuations in the index of refraction along the optical path.

A number of excellent reviews exist describing the turbulence effects on optical propagation [Pratt, 1969; Lawrence and Strohbehn, 1970; Hufnagel, 1978; Dewan, 1980]. The following is a brief description of commonly known turbulence effects.

Beam steering: The laser beam can be deviated from the line of sight so that part or all of it will miss the receiving aperture. This is most likely for the case where the turbulence occurs close to the source, as when the receiver is in space and the transmitter is on the ground (partly because a deviated beam will drift farther from the target in proportion to its subsequent distance of travel). Beam steering effects arise when the size of the atmospheric inhomogeneities are larger than the width of the beam.

Image dancing: The atmosphere can cause a modulation of the angle of arrival of the beam's wavefront. This will cause the image of a source in the receiver to be focused at different places in the focal plane. In a photographic image this would cause blurring. Image dancing effects can be reduced by increasing the receivers aperture.

Beam spreading: Numerous inhomogeneities distributed across the beam cross section cause many small angle scattering events to occur. This has the effect of spreading the beam energy over a wider cross section reducing the signal intensity at the receiver and, in consequence, introducing a decrease in the signal-to-noise ratio. Beam spreading occurs if the beam width is significantly larger than the eddies; beam spreading occurs, therefore, for the downlink of an earth-space laser system (receiver on the ground and transmitter in space or high altitude aircraft).

Spatial coherence degradation: Inhomogeneities in the beam's path also cause losses of phase coherence across the wavefront of the beam. This rapid change of phase with respect to the radial position within the beam cross section reduces the beam's coherence properties.

Scintillation: Within the beam cross-section interference effects can cause destructive and constructive interaction. This causes the power to vary widely from point to point within the cross section, spatially and temporally. Scintillation or amplitude variation effects can be explained by regarding the eddies, or index of refraction variations, as a random distribution of weak lenses distributed within a spatial volume. This volume is to be envisioned as being convected across the beam and thus causing effects that vary with time.

Additional turbulence effects are in transmission, anisoplantism, thermal blooming, and information band width [Greenwood, 1977], to name a few.

18.7.1 Amplitude Fluctuations (Scintillation)

Derivations of the equations given below will be found in the references cited. Tatarski [1961] has shown that the mean square fluctuations of the logarithm of the plane wave amplitude A are related to a quantity C_n^2.

$$\left\langle \left[\ln \left(\frac{A}{A_o} \right) \right]^2 \right\rangle = 0.56\, k^{7/6} \int_o^L C_n^2 (r)\, x^{5/6}\, dx \qquad (18.94)$$

C_n^2 is a measure of the strength of the atmospheric turbulence [see for example Hufnagel, 1978]. A_o in Equation (18.94) is the mean amplitude, k is the wave number of the radiation and x is the distance along the path of the beam. L is the total length of the beam path and r is the position vector. It is assumed that $\ell_i \ll\ll \sqrt{\lambda L} \ll\ll \ell_o$, that is, that the quantity $\sqrt{\lambda L}$ which is known as the size of the first "Fresnel zone," falls into the size range of inertial range eddies (ℓ_i and ℓ_o are the "inner" and "outer" lengths respectively).

It is important to notice that these scintillation effects depend on where they occur along the optical path. This is indicated by the term $x^{5/6}$ in Equation (18.94) under the integral sign where x is the distance from the receiver.

For short path lengths where $\sqrt{\lambda L} \ll \ell_i$ (that is, the Fresnel zone is much smaller than the inner scale), Tatarski [1961] has shown that

$$\left\langle \left[\ln \left(\frac{A}{A_o} \right) \right]^2 \right\rangle = 7.37\, \ell_i^{-7/3} \int_o^L C_n^2 (r)\, x^2\, dx \qquad (18.95)$$

In the simple case where C_n^2 is constant along the path, as it might be for the case of surface to surface propagation, Equation (18.94) becomes

$$\left\langle \left[\ln \left(\frac{A}{A_o} \right) \right]^2 \right\rangle = 0.31\, C_n^2\, k^{7/6}\, L^{11/6} \qquad (18.96)$$

One of the important parameters in the design of optical receivers is the parameter r_o [Pratt, 1969 and Fried, 1967] the transverse coherence length. Physically r_o is the distance such that, if a receiver diameter is increased beyond r_o, there is significantly less improvement in turbulence degradation. As the turbulence degrades the phase coherence of the radiation, r_o is reduced. As C_n^2 becomes larger, r_o becomes smaller. The transverse coherence length is [Fried, 1966]

$$r_o = 2.1 \left[1.46\, k^2 \int_o^L C_n^2 (x)\, dx \right]^{-3/5}. \qquad (18.97)$$

The isoplanatic angle, that is, the angle over which the optical transfer function is constant, represents a radial coherence scale [Fried, 1979]

$$\theta_o = \left[2.91\ k^2 \int_o C_n^2\ x^{5/3}\ dx\right]^{-3/5}. \quad (18.98)$$

The angle θ_o can be considered as the outer limit of an isoplanatic patch of turbulence.

Measurements of C_n^2 can be inferred from path measurements of amplitude fluctuations or transverse coherence length. These approaches require assumption of a uniform distribution of C_n^2 along the path. Point measurements of C_n^2 are possible using an alternative parameter, C_T^2, which is the mean-square statistical average of the difference in temperature between two points separated by a distance r.

$$C_T^2 = \langle (T_1 - T_2)^2 \rangle r^{2/3} \quad (18.99)$$

where the angle brackets denote an ensemble average. C_n^2 is related to C_T^2 through the partial derivative of the air density with respect to the temperature.

$$C_n^2 = (79 \times 10^{-6} P/T^2)^2 C_T^2. \quad (18.100)$$
$$P\ (mb),\ T(K)$$

Expressions defining C_T^2 in terms of atmospheric parameters have been developed by Tatarski [1961] and Hufnagel [1974, Hufnagel and Stanley, 1964].

Experimental data on C_n^2 is limited to few locations (mostly astronomical observatories). Near ground level C_n^2 is found to be dependent upon many local variables. Figure 18-106 represents the variation of C_n at a height of 2 m above the ground on top of a flat mesa in Boulder, Colorado [Lawrence et al., 1970]. Daytime C_n^2 values near ground level can range from 10^{-16} $m^{-2/3}$ to 10^{-12} $m^{-2/3}$. Order-of-magnitude changes can occur in minutes.

Models of C_T^2 in the atmospheric boundary layer have been developed for stable and unstable conditions. Wyngaard et al. [1971] predicts a power law dependence of C_T^2 with height as

Height dependence	Conditions
$C_T^2 \propto Z^{-4/3}$	local free convection
$C_T^2 \propto Z^{-2/3}$	neutral stability
constant	strongly stable

Actual measurements show some departure from these ideal conditions. Measurements in the New Mexico desert

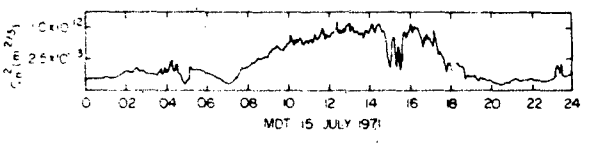


Figure 18-106 Typical diurnal variation of the strength of refractive-index turbulence near the ground. The daytime peak results from heating of the ground; the dips at 15 hours are caused by clouds passing in front of the sun

and mountains indicate minimum C_n^2 values occur shortly after sunrise and before sunset when there is no temperature difference between the ground and atmosphere. For noon conditions, dominated by a convective boundary layer [Walters and Kunkel, 1981]

$$C_n^2 = (7.7 \pm 1.2) \times 10^{-13} \left(\frac{Z}{Z_o}\right)^{-(1.16 \pm .03)} \text{ for } Z < 1000 \text{ m}, \quad (18.101)$$

(where Z_o = 1m). At midnight,

$$C_n^2 = (3.2 \pm 1.1) \times 10^{-16} e^{-(0.001 \pm 0.00008)Z}. \quad (18.102)$$

The C_n^2 data generally have a log normal probability distribution as illustrated by the scintillometer observations at 9.4 km shown in Figure 18-106.

Altitude profiles of C_n^2 obtained by performing log averages are shown in Figure 18-107. The Loos and Hogge [1979] data are from White Sands Missile Range, New Mexico; Miller and Zieski [1978] data are obtained at Mt. Haleakala, Maui, Hawaii, and the Barletti et al. [1977] data are obtained from three locations in Italy. The Brown and Good [1984a] data are from Mt. Haleakala, Hawaii, the Brown and Good [1984b] data are from Westford, Mass. The upper altitude profiles (Z > 5 km) appears to be more constant in time except for possible tropopause effects associated with jet streams diurnal effects. Several models have been presented to describe the C_n^2 altitude profile. Hufnagle [1966] introduced the model

$$C_n^2(z) = 2.7 \times 10^{-16} \left[e^{-z/1500} + 3W^2 \left(\frac{Z}{10\,000}\right)^{10} e^{-z/1000}\right] \quad (18.103)$$

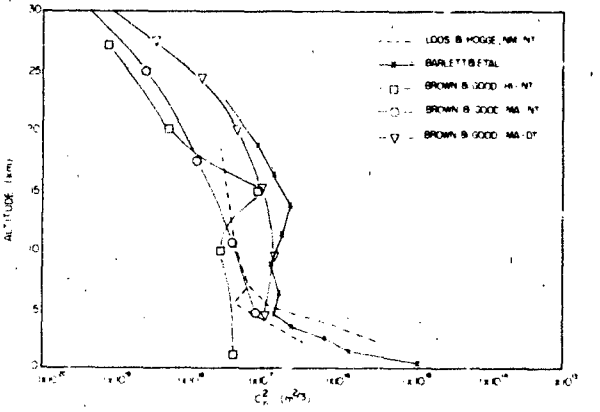


Figure 18-107. Altitude profiles of C_n^2.

where $W = \left[1/15 \int_{5\ km}^{20\ km} V^2(Z)dZ \right]^{1/2}$ is a root mean square of the horizontal wind, V(Z) over the altitude range of the model. Hufnagel [1974] also has a model with a delta function additive term, 2×10^{-16} (h − 12 km), to represent the tropopause contribution. VanZandt et al. [1981] had developed a model based on statistical behavior of the atmospheric stability and wind shears. This model uses observed rawinsonde measurements of temperature and winds at 300 m altitude resolution to infer optical turbulence at higher altitude resolution.

CHAPTER 18

REFERENCES

Alms, G.R., A.K. Burnham, and W.H. Flygare, "Measurement of the Dispersion in Polarity Anisotropies," *J. Chem. Phys.*, **63**:3321–3326, 1975.

Altshuler, T.L., *Infrared Transmission and Background Radiation by Clear Atmospheres*, GE Report 61 SD 199, AD401923, 1961.

Ashley, G.W., L. Gastineau, and B. Blay, (private communication) 1973.

Baas, F. and K.D. van den Hout, "Measurements of Depolarization Ratios and Polarizability Anisotropies of Gaseous Molecules," *Physica*, **95A**:597–601, 1979.

Barletti, R., G. Ceppatelli, L. Paterno, A. Righini, and N. Speroni, "Astronomical Site Testing with Balloon-Borne Radiosondes," *Astron. Astrophys.*, **54**:649–659, 1977.

Bemporad, reprinted in List, R.J., *Smithsonian Meteorological Tables 6th Rev. Ed.*, Smithsonian Institution Press, Washington, D.C., 1968.

Berger, R.H., "Characterization of Snow for Evaluation of its Effects on Electromagnetic Wave Propagation," *Optical Engineering for Cold Environments, Proc. Soc. Photo-Optical Inst. Eng.*, **414**:35, 1983.

Biberman, M., L. Dunkelman, M.L. Fickett, R.G. Fink, "Levels of Nocturnal Illumination," Institution for Defense Analyses Research Paper, **SD-50,** January 1966.

Blättner, W.G., H.G. Horak, D.G. Collins, and M.B. Wells, "Monte Carlo Studies of the Sky Radiation at Twilight," *Appl. Opt.*, **13**:534–547, 1974.

Born, M. and E. Wolf, *Principles of Optics*, Pergamon Press, New York, 1975.

Brown, D., "Natural Illumination Charts," Dept. of the Navy R&D Project NS 714–100, September 1952.

Brown, J.H. and R.E. Good, "Thermosonde C_n^2 Measurements in Hawaii—August 1982, AFGL TR-84-0110, ADA148016, 1984a.

Brown, J.H. and R.E. Good, "Thermosonde and UHF C_n^2 Measurements at Westford, Massachusetts–July 1981," AFGL TR-84-0109, ADA145398, 1984b.

Burch, D.E., *Semi-Annual Technical Report*, Aeroneutronic Report No. U-4784, January 1970.

Burch, D.E., "Continuum Absorption by H_2O," AFGL TR-81-0300, ADA112264, 1981.

Burch, D.E. and D.A. Cryvnak, *Infrared Absorption by CO_2 and H_2O*, AFGL TR-78-0154, ADA060079; Aeronutronic Report No. U-6417, May 1978.

Burch, D.E. and D.A. Gryvnak, *Method of Calculating H_2O Transmission Between 333 and 633 cm^{-1}*, AFGL TR-79-0054, ADA072850; Aeronutronic Report No. U-6503, April 1979.

Burch, D.E., D. Gryvnak, E.B. Singleton, W.L. France, and D. Williams, *Infrared Absorption by Carbon Dioxide, Water Vapor, and Minor Atmospheric Constituents*, AFCRL62-698, AD287406, 1962.

Burch, D.E., D.A. Gryvnak, and J.D. Pembrook, *Investigation Absorption by Atmospheric Gases*, AFCRL-71-0124, AD 882876; Aeronutronic Report No. U-4897, January 1971.

Canosa, J. and H.R. Penafiel, "A Direct Solution of the Radiative Transfer: Application to Rayleigh and Mie Atmospheres," *J. Quant. Spectrosc. Radiat. Transfer*, **13**:21–39, 1973.

Chandrasekhar, S., *Radiative Transfer*, Oxford University Press, London, 393, 1950.

Chaney, L.W., *An Experimental Fourier Transform Asymmetrical Interferometer for Atmospheric Radiance Measurements*, University of Michigan Technical Report 05863-18-T, 1969.

Chapman, S., "The Abosrption and Dissociative or Ionizing Effect of Monochromatic Radiation in an Atmosphere on a Rotating Earth II Grazing Incidence," *Proc. Phys. Sci.* (London) **43**:483–501, 1931.

Chu, C.M. and S.W. Churchill, "Numerical Solution of Problems in Multiple Scattering of Electromagnetic Mediation," *J. Phys. Chem.*, **59**:855–863, 1955.

Chýlek, P., G.W. Grams, and R.G. Pinnick, "Light Scattering by Irregularly Randomly Oriented Particles," *Science*, **193**:480–482, 1976.

Clough, S.A., and F.X. Kneizys, "Convolution Algorithm for the Lorentz Function," *Appl. Opt.*, **18**:2329, 1979.

Clough, S.A., F.X. Kneizys, and J.H. Chetwynd, Jr., "Algorithm for the Calculation of Absorption Coefficient-Pressure Broadened Molecular Transitions," AFGL TR-77-0164, ADA047515, 1977.

Clough, S.A., F.X. Kneizys, R. Davies, R. Gamache, and R.H. Tipping, "Theoretical Line Shape for H_2O Vapor; Application to the Continuum," in *Atmospheric Water Vapor*, edited by A. Deepak, T.D. Wilkerson and L.H. Ruhnke, Academic Press, New York, 1980.

Clough, S.A., F.X. Kneizys, L.S. Rothman, and W.O. Gallery, "Atmospheric Spectral Transmittance and Radiance: FASCODE," *Proc. Soc. Photo-Optical Inst. Eng.*, **277**:152, 1981.

Clough, S.A., R.W. Davies, and R.H. Tipping, "The Line Shape for Collisionally Broadened Molecular Transitions: A Quantum Theory Satisfying the Fluctuation Dissipation Theorem," in *Spectral Line Shapes*, edited by K. Burnett, Walter de Gruyter, New York, 1983.

Coakley, J. and P. Chýlek, "The Two-Stream Approximation in Radiative Transfer: Including the Angle of the Incident Radiation," *J. Atmos. Sci.*, **32**:409–418, 1975.

Cole, A.E. and A.J. Kantor, "Air Force Reference Atmospheres," AFGL TR-78-0051, ADA058505, 1978.

Cole, A.E., A. Court, and A.J. Kantor, "Model Atmospheres," in *Handbook of Geophysics and Space Environments*, edited by S. Valley, McGraw-Hill, New York, NY 1965, (also available from NTIS as ADA056800) 1965.

Collins, D.G. and M.B. Wells, "Monte Carlo Codes for the Study of Light Transport in the Atmosphere," ECOM-00240-F, Radiation Research Associates, Inc. (Rep. RRA-174) Fort Worth, Texas (Avail. from NTIS, AD625225), 1965.

Collins, D.G., W.G. Blattner, M.B. Wells, and H.G. Horak, "Backward Monte Carlo Calculations of the Polarization Characteristics of the Radiation Emerging from Spherical-Shell Atmospheres," *Appl. Opt.*, **11**:2684–2696, 1972.

Conrath, B.J., R.A. Hanel, V.G. Kunde, and C. Prabhakara, *The Infrared Interferometer Experiment on Nimbus 3*, Goddard Space Flight Center, Maryland, Report X-620-70-213, 1970.

Coulson, K.L., J.V. Dave, and Z. Sekera, *Tables Related to Radiation Emerging from a Planetary Atmosphere with Rayleigh Scattering*, University of California Press, Berkeley Calif., 1960.

Dave, J.V., "Meaning of Successive Iteration of the Auxiliary Equation in the Theory of Radiative Transfer," *Astrophys. J.*, **140**:1292–1303, 1964.

Dave, J.V., "Multiple Scattering in a Non-Homogenous Rayleigh Atmosphere," *J. Atmos. Sci.*, **22**:273–279, 1965.

Dave, J., "A Direct Solution of the Spherical Harmonics Approximation to the Radiative Transfer Equation for an Arbitrary Solar Elevation," *J. Atmos. Sci.*, **32**:790–800, 1975a.

Dave, J., "A Direct Solution of the Spherical Harmonics Approximation to the Radiative Transfer Equation for an Arbitrary Solar Evalation. II: Results," *J. Atmos. Sci.*, **32**:1463–1474, 1975b.

Dave, J.V. and B.H. Armstrong, "Smoothing of the Intensity Curve Obtained from a Solution of the Spherical Harmonics Approximation to the Transfer Equation," *J. Atmos. Sci.*, **31**:1934–1937, 1974.

Dave, J.V. and J. Gazdag, "A Modifier Fourier Transform Method for Multiple Scattering Calculations in a Plane-Parallel Mie Atmosphere," *Appl. Opt.*, **9**:1457–1466, 1970.

Davies, R. "Fast Azimuthally Dependent Model of the Reflection of Solar Radiation by Plane-Parallel Clouds," *Appl. Opt.*, **19**:250–255, 1980.

Deirmendjian, D., *Electromagnetic Scattering on Spherical Polydispersions*, American Elsevier, New York, 129–139 and 290, 1969.

Diermendjian, D., "Scattering and Polarization Properties of Water, Clouds and Hazes in the Visible and Infrared," *Appl. Opt.*, **3**:187–196, 1964.

Devaux, C. and M. Herman, "Remarques sur l'Utilization de la Methode des Harmoniques Spheriques dans les Calculs de Transfert Radiatif," *Comptes Rendus, Acad. Sci. Paris*, **B273**:849–852, 1971.

Dewan, E.M., "Optical Turbulence Forecasting: A Tutorial," AFGL TR-80-0030, ADA086863, 1980.

Downing, H. and D. Williams, "Optical Constants of Water in the Infrared," *J. Geophys. Res.* **80**:1656–1661, 1975.

Duntley, S.Q., A.R. Boileau, and R.W. Preisendorfer, "Image Transmission by the Troposphere," *J. Opt. Soc. Am.*, **47**:499–506, 1957.

Earing, D.G. and J.A. Smith, *Target Signatures Analysis Center: Data Compilation*, AFAL, AD489968, 1966.

Eaton, F.D. and I. Dirmhirn, "Reflected Irradiance Indicatrices of Natural Surfaces and Their Effect on Albedo," *Appl. Opt.*, **18**(7), Apr 1979.

Eddington, A.S., *The Internal Constitution of the Stars*, 407, Cambridge University Press, Cambridge, Eng., 1926.

Edlen, B., "The Refractive Index of Air," *Meteorol.*, **2**:71–80, 1966.

Elsasser, W.M., "Mean Absorption and Equivalent Absorption Coefficient of a Band Spectrum," *Phys. Rev.*, **54**:126, 1942. Also as Harvard Meteorological Studies, No. 6, Blue Hill Met. Obs., Harvard Univ., Milton, Mass., 1942.

Elterman, L., "Atmospheric Attenuation Model, 1964, in the Ultraviolet, Visible, and Infrared Regions for Altitudes to 50 km," AFCRL-64-740, AD607859, 1964.

Elterman, L., "UV, Visible, and IR Attenuation for Altitudes to 50 km," AFCRL-68-0153, ADA671933, 1968.

Elterman, L., "Vertical Attenuation Model with Eight Surface Meteorological Ranges 2 to 13 kilometers," AFCRL-70-0200, AD 707488, 1970.

Evans, W.F., J.B. Kerr, and D.I. Wardle, *The AES Stratospheric Balloon Measurements Project: Preliminary Results*, Report No. APRB 30X4, Atmospheric Environment Service, Downsview, Ontario, Canada, 1975.

Falcone, V.J., L.W. Abreu, and E.P. Shettle, *Atmospheric Attenuation of Millimeter Waves, Models and Computer Code*, AFGL TR-79-0253, ADA084485, 1979.

Fitch, B.W., "Effects of Reflection by Natural Surfaces on the Radiation Emerging from the top of the Earth's Atmosphere," *J. Atmos. Sci.*, **38**:2717–2719, 1981.

Fouquart, Y., W.M. Irvine, and J. Lenoble, eds., *Standard Procedures to Compute Atmospheric Radiative Transfer in a Scattering Atmosphere, Vol. II*, Part A. *Review of Methods for Horizontally Inhomogeneous Atmospheres, and Spherical Atmospheres*, Part B. *Problems of Scattering with Gaseous Absorption*, National Center for Atmospheric Research, Boulder, Colorado, for the International Association of Meteorology and Atmospheric Physics, 1980.

Fried, D.L., "Optical Resolution Looking Down Through a Randomly Homogeneous Medium for Very Long and Very Short Exposures," *J. Opt. Soc. Am.*, **56**:1372–1379, 1966.

Fried, D.L., "Optical Heterodyne Detection of an Atmospheric Distorted Signal Wave Frong," *Proc. IEEE*, **55**:57, 1967.

Fried, D.L., "Angular Dependence of the Atmospheric Turbulence Effect in Speckle Interferometry," *Opt. Acta*, **26**:597–613, 1979.

Gallery, W., F.X. Kneizys, and S.A. Clough, *Air Mass Computer Program for Atmospheric Transmittance/Radiance Calculation: FASCATM*, AFGL TR-83-0065, ADA132108, 1983.

Gathman, S.G., "Optical Properties of the Marine Aerosol as Predicted by the Navy Aerosol Model," *Opt. Eng.*, **22**:57–62, 1983.

Goody, R.M. "A Statistical Model for Water Vapour Absorption," *Quart. J. Roy. Meteorol. Soc.*, **78**:165, 1952.

Goody, R.M., *Atmospheric Radiation, Vol. I Theoretical Basis*, Oxford University Press, London, 1964.

Grant, I.G. and G.E. Hunt, "Solution of Radiative Transfer Problems Using the Invariant S_N Method," *Mon. Notices Roy. Astron. Soc.*, **141**:27–41, 1968.

Grant, I.G. and G.E. Hunt, "Discrete Space Theory of Radiative Transfer, 1, Fundamentals," *Proc. Roy. Soc. London*, **A313**:183–197, 1969a.

Grant, I.G. and G.E. Hunt, "Discrete Space Theory of Radiative Transfer, 2, Stability and Non-Negativity," *Proc. Roy. Soc. London*, **A313**:199–216, 1969b.

Gray, D.M. and D.H. Male, (eds.) *Handbook of Snow*, Pergamon Press, New York, 1981.

Greenwood, D.P., "Bandwidth Specification for Adaptive Optics System," *J. Opt. Soc. Am.*, **67**:390–392, 1977.

Grenfell, T.C. and D.K. Perovich, "Radiation Absorption Coefficients of Polycrystalline Ice from 400–1400 nm," *J. Geophys. Res.*, **86**:7447–7450, 1981.

Hale, G.M. and M.R. Querry, "Optical Constants of Water in 200 nm to 200 μm Wavelength Region," *Appl. Opt.*, **12**:555–563, 1973.

Hanel, G., "The Properties of Atmospheric Aerosols Particles as Functions of the Relative Humidity at Thermodynamic Equilibrium with the Surrounding Moist Air," in *Advances in Geophysics*, **19**:73–188, edited by H.E. Landsberg and J. van Miefhem, 1976.

Hanel, R.A. and B.J. Conrath, *Thermal Emission Spectra of the Earth and Atmosphere Obtained from the Nimbus 4 Michelson Interferometer Experiment*, Goddard Space Flight Center, Greenbelt, Maryland, Report X-620-70-244, 1970.

Hänel G. and K. Bullrich, "Physico-Chemical Property Model of Tropospheric Aerosol Particles," *Beitr. Phys. Atm.*, **51**:129–130, 1978.

Hansen, J.E., "Radiative Transfer by Doubling Very Thin Layers," *Astrophys. J.*, **155**:565–573, 1969a.

Hansen, J.E., "Exact and Approximate Solutions for Multiple Scattering by Cloudy and Hazy Planetary Atmospheres," *J. Atmos. Sci.*, **26**:478–487, 1969b.

Hansen, J.E., "Multiple Scattering of Polarized Light in Planetary Atmospheres, Part 1. The Doubling Method," *J. Atmos. Sci.*, **28**:120–125, 1971a.

Hansen, J.E., "Multiple Scattering of Polarized Light in Planetary Atmospheres, Part 2. Sunlight Reflected by Terrestrial Clouds," *J. Atmos. Sci.*, **28**:1400–1426, 1971b.

Hansen, J.E. and J.B. Pollack, "Near-Infrared Light Scattering by Terrestrial Clouds," *J. Atmos. Sci.*, **27**:265–281, 1970.

Hansen, J.E. and L.D. Travis, "Light Scattering in Planetary Atmospheres," *Space Sci. Rev.*, **16**:527–610, 1974.

Hering, W.S., "An Operational Technique for Estimating Visible Spectrum Contrast Transmittance," AFGL TR-81-0198, ADA111823, 1981.

Hering, W.S. and T.R. Borden, "Ozonesonde Observation Over North America," AFCRL-64-30, vol. 2, AD604880, 1964.

Herman, B.M., "Multiple Scatter Effects on the Radar Return from Large Hail," *J. Geophys. Res.*, **70**:1215–1225, 1965.

Herman, B.M. and S.R. Browning, "A Numerical Solution to the Equation of Radiative Transfer," *J. Atmos. Sci.*, **22**:559–566, 1965.

Herman, B.M., S.R. Browning, and R.J. Curran, "The Effect of Atmospheric Aerosols on Scattered Sunlight," *J. Atmos. Sci.*, **28**:419–428, 1971.

Herman, B.M., A.J. LaRocca and R.E. Turner, "Atmospheric Scattering" Chapter 4 in *The Infrared Handbook*, edited by W.L. Wolfe and G.J. Zissis, Office of Naval Research, Washington, D.C., 1978.

Herman, B.M., W. Asous, and S. Browning, "A Semi-Analytic Technique to Integrate the Radiative Transfer Equation Over Optical Depth," *J. Atmos. Sci.*, **37**:1828–1838, 1980.

Hobbs, P.V., *Ice Physics*, p. 202, Clarendon Press, Oxford, U.K., 1974.

Holland, A.C. and G. Gagne, "The Scattering of Polarized Light of Polydisperse Systems of Irregular Particles," *Appl. Opt.*, **9**:1113–1121, 1970.

Houghton, J.T., *The Physics of Atmospheres*, Cambridge University Press, Cambridge, Eng., 1977.

Hufnagel, R.E., *Restoration of Atmospherically Degraded Images* Vol. 2, Woods Hole Summer Study, July 1966, National Academy of Sciences—National Research Council, Washington, D.C., 1966.

Hufnagel, R.E., "Variations of Atmospheric Turbulence," *Proc. Topical Meeting on Optical Propagation through Turbulence*, Optical Society of America, Washington, D.C. 9–11 July, 1974.

Hufnagel, R.E., "Propagation Through Atmospheric Turbulence," in *The Infrared Handbook*, edited by W.L. Wolfe and G.J. Zissis, Office of Naval Research, Washington, D.C., 1978.

Hufnagel, R.E., and N.R. Stanley, "Modulation Transfer Function Associated with Image Transmission through Turbulent Media," *J. Opt. Soc. Am.*, **54**:52–61, 1964.

Hulburt, E.O., "Optics of Atmospheric Haze," *J. Opt. Soc. Am.*, **31**:467–476, 1941.

Irvine, W.M., "Multiple Scattering by Large Particles," *Astrophys. J.*, **142**:1563–1575, 1965.

Irvine, W.M., "Diffuse Reflection and Transmission by Cloud and Dust Layers," *J. Quant. Spectrosc. Radiat. Transfer*, **8**:471–485, 1968a.

Irvine, W.M., "Multiple Scattering by Large Particles, 2, Optical Thick Layers," *Astrophys. J.*, **152**:823–834, 1968b.

Irvine, W.M. and J.B. Pollack, "Infrared Optical Properties of Water and Ice Spheres," *Icarus*, **8**:324–360, 1967.

Ivlev, L.S., "Aerosol Model of the Atmosphere," *Prob. Fiz. Atmos.* No. 7: 125–160, Leningrad, translated by Foreign Science and Technology Ctr, Dept. of the Army, NTIS No. AD760393, 1967.

Ivlev, L.S., "Atmospheric Aerosol," in *Radiation Characteristics of the Atmosphere and the Earth's Surface*, edited by K. YaKondratev, Amerind, New York, NTIS No. TT-71-58003, 1969.

Jeans, J.M., "The Equations of Radiative Transfer," *Mon. Notices Roy. Astron. Soc.*, **78**:28–36, 1917.

Johnson, J.C. *Physical Meterology*, New York Technical Press, MIT, and Wiley, 393, 1954.

Joseph, J.H., W.J. Wiscombe, and J.A. Weinman, "The Delta Eddington Approximation for Radiation Flux Tranfer," *J. Atmos. Sci.*, **33**:2452–2459, 1976.

Junge, Ch.E., *Air Chemistry and Radioactivity*, p. 382, Academic Press, New York, 1963.

Karp, A., "Efficient Computation of Spectral Line Shapes," *J. Quant. Spectrosc. Radiat. Transfer*, **14**:803–817, 1973.

Karp, A., "Computing the Angular Dependence of the Radiation of a Planetary Atmosphere," *J. Quant. Spectrosc. Radiat. Transfer*, **25**:403–412, 1981.

Karp, A., J. Greenstadt, and J. Fillmore, "Radiative Transfer Through an Arbitrarily Thick Scattering Atmo-

sphere," *J. Quant. Spectrosc. Radiat. Trans*, **24**:391–406, 1981.

Kasten, F., "A New Table and Approximation Formula for the Relative Optical Air Mass," *Arch. Meteor. Geophys. Bioklim*, **B14**:206–223, 1966.

Kasten, F., "Effect of Variation of the Vertical Air Density Profile in the Relative Optical Air Mass," *Arch. Meteor. Geophys. Bioklim*, **B15**:62–66, 1967.

Kattawar, G.W. and G.N. Plass, "Influence of Particle Size Distribution on Reflected and Transmitted Light from Clouds," *Appl. Opt.*, **7**:869–878, 1968.

Kaufman, Y.J., "Effect of the Earth's Atmosphere on Contrast for Zenith Observation," *J. Geophys. Res.*, **84**:3165–3172, 1979.

Kerker, M., *The Scattering of Light and Other Electromagnetic Radiation*, Academic Press, New York, 1969.

Kielkopf, J.F., "New Approximation to the Voigt Function with Applications to Spectral Line Profile Analysis," *J. Opt. Soc. Am.*, **63**,987, 1973.

King, J.I.F., "Band Absorption Model for Arbitrary Line Variance," *J. Quant. Spectrosc. Radiat. Transfer*, **4**:705–711, 1964.

Kneizys, F.X., E.P. Shettle, W.O. Gallery, J.H. Chetwynd, Jr., L.W. Abreu, J.E.A. Selby, R.W. Fenn, and R.A. McClatchey, *Atmospheric Transmittance/Radiance: Computer Code LOWTRAN 5*, AFGL TR-80-0067, AD088215, 1980.

Kneizys, F.X., E.P. Shettle, W.O. Gallery, J.H. Chetwynd, Jr., L.W. Abreu, J.E.A. Selby, S.A. Clough, and R.W. Fenn, *Atmospheric Transmittance/Radiance: Computer Code LOWTRAN 6*, AFGL TR-83-184, ADA140012, August 1983.

Koh, G. and H. O'Brien, "Snow Crystal Habit," SNOW-ONE-A Data Report, edited by Aitken, CRREL Special Report 82-8, U.S. Army Corps. of Engineers Cold Regions Research and Engineering Lab., Hanover, N.H., 1982.

Krinov, E.L., "Spectral Reflectance Properties of Natural Formations" (in Russian 1942), Nat. Res. Council of Canada (Ottawa) Tech. Transl. TT-439, 1953.

Kunde, V.G. and W.C. Maguire, "Direct Integration Transmittance Model," *J. Quant. Spectrosc. Radiat. Transfer*, **14**:303–817, 1973.

Lacomb, J., "Techniques for Measuring the Mass Concentration of Falling Snow," *Optical Engineering for Cold Environments, Proc. Soc. Photo-Optical inst. Eng.*, **414**:17, 1983.

LaRocca, A.J., "Atmospheric Absorption," Chapter 5 in *Infrared Handbook*, edited by W.L. Wolfe, G.J. Zissis, Office of Naval Research, Washington, D.C., 1978.

Lawrence, R.S., G.R. Ochs, and S.F. Clifford, "Measurements of Atmospheric Turbulence Relevant to Optical Propagation," *J. Opt. Soc. Am.*, **60**:826, 1970.

Lawrence, R.S. and J.W. Strohbehn, "A Survey of Clear Air Propagation Effects Relevant to Optical Communications," *Proc. IEEE*, **58**:1523–1545, 1970.

Laws, J.O. and Parsons, "The Relation of Rain Drop Size to Intensity," *Trans. Am. Geophys. Union*, **24**:452–460, 1943.

Leader, J.L., "Analysis and Prediction of Loss Scattering from Rough Surface Material," *J. Opt. Soc. Am.*, **69** (4), April 1979.

Lenoble, J. (ed.) *Standard Procedures to Compute Atmospheric Radiative Transfer in a Scattering Atmosphere, Vol. I*, International Association of Meteorology and Atmospheric Physics, Radiation Commission, printed at National Center for Atmospheric Research, Boulder, Colorado, 1977.

Liou, K.N., "A Numerical Experiment on Chandrasekhar's Discrete-Ordinate Method for Radiative Transfer: Applications to Cloudy and Hazy Atmospheres," *J. Atmos. Sci.*, **30**:1303–1326, 1973.

Lo, K.K. and R.E. Passarelli, Jr., "The Growth of Snow in Winter Storms: An Airborne Observational Study," *J. Atmos. Sci.*, **39**:697–706, 1982.

Loos, G.C. and C.B. Hogge, "Turbulence of the Upper Atmosphere and Isoplanatism," *Appl. Opt.*, **13**:2654–2661, 1979.

Magono, C. and C. Lee, "Meteorological Classification of Natural Snow Crystals," *J. Fac. Sci.*, Hokkaido University, Series VII, **2**:321–335, 1966.

Mankin, W.G., "Fourier Transform Method for Calculating the Transmittance of Inhomogeneous Atmospheres," *Appl. Opt.*, **18**:3426–3433, 1979.

Marchuk, G.I., G.A. Mikhailov, M.A. Nazaraliev, R.A. Darbinjan, B.A. Kargin, and B.S. Elepov, *The Monte Carlo Methods in Atmospheric Optics*, Springer-Verlag, New York, 1980.

Marshall, J.S. and W.M.K. Palmer, "The Distribution of Raindrops with Size," *J. Meteor.*, **5**:165–166, 1948.

Martinez-Sanchez, M., V. Yousefian, D. Dvore, and R. Vaglio-Laurin, "Numerical Modeling of Optically Significant Characteristics of Falling Snow," *Opt. Eng.*, **22**:78–85, 1983.

Mason, J.B., "Light Attenuation in Falling Snow," ASL-TR-0018, U.S. Army Atmospheric Sciences Laboratory, White Sands Missile Range, New Mexico, 1978.

McCartney, E.J., *Optics of the Atmosphere*, Wiley, New York, 1976.

McClatchey, R.A., R.W. Fenn, J.E.A. Selby, F.E. Volz, and J.S. Garing, "Optical Properties of the Atmosphere," AFCRL TR70-0527, AD715270, 1970

McClatchey, R.A., W.S. Benedict, S.A. Clough, D.E. Burch, R.F. Calfee, K. Fox, L.S. Rothman, and J.S. Garing, "AFCRL Atmospheric Absorption Line Parameters Compilation," AFCRL TR-0096, AD 762904, 1973.

McClatchey, R.A., R.W. Fenn, J.E.A. Selby, F.E. Volz, and J.S. Garing, "Optical Properties of the Atmosphere," (Third Edition), AFCRL 72-0497, AD 753075, 1972.

McKellar, B. and M. Box, "The Scaling Group of the Radiative Transfer Equation," *J. Atmos. Sci.*, **38**:1063–1068, 1981.

McMillin, L.M., H.E. Fleming, and M.L. Hill, "Atmospheric Transmittance of an Absorbing Gas: A Computationally Fast and Accurate Transmittance Model for Absorbing Gases with Variable Mixing Ratios," *Appl. Opt.*, **18**:1600–1606k 1976.

scattering of so-called "spongy" hail (a mixture of ice and

is useful to express Z as a function of either the pre- ipitation

CHAPTER 18

Meador, W. and W. Weaver, "Two-Stream Approximation to Radiative Transfer in Planetary Atmospheres: A Unified Description of Existing Methods and a New Improvement," *J. Atmos. Sci.*, **37**:630–643, 1980.

Middleton, W.E.K., *Vision Through the Atmosphere*, U. of Toronto Press, 1952.

Mie, G., "Bertrage z. Phys. Truber Medien, spezeziell kolloidaler Metallosungen," *Am. Physik*, **25**:377–445, 1908.

Mill, J.D. and E.P. Shettle, "A Preliminary LOWTRAN Snow Model," AFGL TR-83-0148, AD A129826, 1983.

Miller, M.G. and P.L. Zeiske, "Experimental Turbulence Profile and Parameters," *SPIE Symp. on Optical Properties of the Atmosphere, Proc. Soc. Photo-Optical Inst. Eng.*, **142**:98–103, 1978.

Moon, P. "Proposed Standard Solar Radiation Curves for Engineering Use," *J. Franklin Inst.*, **230**(5): 583, 1940.

Mossop, "Volcanic Dust Collected at an Altitude of 20 km," *Nature*, **103**:824–827, 1964.

Murcray, D.G., T.G. Kyle, F.H. Murcray, and W.G. Williams, "Nitric Acid and Nitric Oxide in the Lower Stratosphere," *Nature*, **218**:78, 1968.

Murcray, D.G., J.N. Brooks, A. Goldman, J.J. Kosters, and W.J. Williams, *Water Vapor Nitric Acid and Ozone Mixing Ratio Height Derived from Spectral Radiometric Measurements*, University of Denver, Contract Report No. 322, 1977.

National Climatic Center, NOAA: Environmental Data Service, Federal Building, Asheville, North Carolina 28801.

Newkirk, G., Jr., and J.A. Eddy, "Light Scattering by Particles in the Upper Atmosphere," *J. Atmos. Sci.*, **21**:35-60, 1964.

Nicodemus, F.E., "Radiometry" in *Applied Optics and Optical Engineering*, edited by R. Kingslake, Academic Press, New York, 1967.

Nilsson, B., "Meteorological Influence on Aerosol Extinction in the 0.2–40 μm Wavelength Range," *Appl. Opt.*, **18**:3457–3473, 1979.

Palmer, K.F. and D. Williams, "Optical Constants of Sulfuric Acid: Application to the Clouds of Venus," *Appl. Opt.*, **14**:208–219, 1975.

Penndorf, R., "Luminous and Spectral Reflectance as well as Colors of Natural Objects," AFCRLTR-56-203, AD98766,1956.

Plass, G.N., "Models for Spectral Band Absorption," *J. Opt. Soc. Am.*, **48**:690–703, 1958.

Plass, G.N. and G.W. Kattawar, "Influence of Single Scattering Albedo on Reflected and Transmitted Light from Clouds," *Appl. Opt.*, **7**:361–367, 1968a.

Plass, G.N. and G.W. Kattawar, "Monte Carlo Calculations of Light Scattering from Clouds," *Appl. Opt.*, **7**:415–419, 1968b.

Plass, G.N. and G.W. Kattawar, "Radiative Transfer in the Earth's Atmosphere-Ocean System," *J. Phys. Ocean.*, **2**:139–156, 1972.

Plass, G.N., G.W. Kattawar and F.E. Catchings, "Matrix Operator Theory of Radiative Transfer, 1. Rayleigh Scattering," *Appl. Opt.*, **12**:314–329, 1973.

Potter, J.F., "The Delta Function Approximation in Radiative Transfer Theory," *J. Atmos. Sci.*, **27**:945–951, 1970.

Pratt, W.K., *Laser Communication Systems*, Wiley, New York, 1969.

Preisendorfer, R., *Radiative Transfer on Discrete Spaces*, Pergamon Press, New York, 1965.

Remsberg, E.E., "Radiative Properties of Several Probable Constituents of Atmospheric Aerosols," Ph.D. Thesis Dept. Meteorology, University of Wisconsin, Madison, 1971.

Remsberg, E.E., "Stratospheric Aerosols Properties and Their Effect on Infrared Radiation," *J. Geophys. Res.*, **78**:1401–1408, 1973.

Remsberg, E.E., J.M. Russel III, L.L. Gordley, J.C. Gille, and P.L. Bailey, "Implications of the Stratospheric Water Vapor Distribution as Determined from the Nimbus 7 LIMS Experiment," *J. Atmos. Sci.* **41**:2934–2945, 1984.

Rice, D.P. and P.A.R. Ade, "Absolute Measurements of the Atmospheric Transparency at Short Millimeter Wavelengths," *Infrared Phys.*, **19**:575, 1979.

Robinson, N., ed., *Solar Radiation*, Elsevier, Amsterdam, 1966.

Romanova, L.M., "The Solution of the Radiative Transfer Equation for the Case When the Indicatrix of Scattering Greatly Differs from the Spherical One," *Opt. Spektrosk.*, **13**:429–435, 1962.

Romanova, L.M., "Radiative Field in Plane Layers of a Turbid Medium with Highly Anisotropic Scattering," *Opt., Spektrosk.*, **14**:262–269, 1963.

Rosen, J.M., "The Boiling Point of Stratospheric Aerosols," *J. Appl. Meteorol.*, **10**:1044–1046, 1971.

Rosen, J.M., "Stratospheric Dust and Its Relationship to the Meteoric Influx," *Space Sci. Rev.*, **9**:58–89, 1969.

Rothman, L.S., "AFGL Atmospheric Absorption Line Parameters Compilation: 1980 Version," *Appl. Opt.*, **20**:291, 1981.

Rothman, L.S. and A. Goldman, "Infrared Electric Quadrupole Transitions of Atmospheric Oxygen," *Appl. Opt.*, **20**:2182, 1981.

Rothman, L.S., A. Goldman, J.R. Gillis, R.H. Tipping, L.R. Brown, J.S. Margolis, A.G. Maki, and L.D.G. Young, "AFGL Trace Gas Compilation: 1980 Version," *Appl. Opt.*, **20**:1323, 1981.

Rothman, L.S., A. Goldman, J.R. Gillis, R.R. Gamache, H.M. Pickett, R.L. Poynter, N. Husson, and A. Chedin, "AFGL Trace Gas Compilation: 1982 Version," *Appl. Opt.*, **22**:1616–1627, 1983a.

Rothman, L.S., R.R. Gamache, A. Barbe, A. Goldman, J.R. Gillis, L.R. Brown, R.A. Toth, J.-M. Flaud, and C. Camy-Peyret, "AFGL Atmospheric Absorption Line Parameters Compilation: 1982 Edition," *Appl. Opt.*, **22**:2247–2256, 1983b.

Rowell, R.L., G.M. Aval, and J.J. Barrett, "Rayleigh-Raman Depolarization of Laser Light Scattered by Gases," *J. Chem. Phys.*, **54**:1960–1964, 1971.

Sagan, C. and J.B. Pollack, "Anisotropic Nonconservative Scattering and the Clouds of Venus," *J. Geophys. Res.*, **72**:469–477, 1967.

Samuelson, R.E., "The Transfer of Thermal Infrared Radiation in Cloudy Planetary Atmospheres," Ph.D. Thesis, Georgetown University, p. 309, 1967.

Samuelson, R.E., "The Thermal Radiation Field Emitted

by Anisotropically Scattering Cloudy Planetary Atmospheres," *Icarus*, **10**:258–273, 1969.

Schaaf, J.B. and D. Williams, "Optical Constants of Ice in the Infrared," *J. Opt. Soc. Am.*, **63**:726–732, 1973.

Schaller, E., "A Delta-Two-Stream Approximation in Radiative Flux Calculations," *Beitr. Phys. Atmos.*, **52**:17–26, 1979.

Schuerman, D.W. (ed.) *Light Scattering by Irregularly Shaped Particles*, Plenum Press, New York, 1980.

Schulze, R., *Strahlenklima der Erde*, D. Steinkopff Verlag, Darmstadt, Germany, 1970.

Scott, N.A. and A. Chedin, "A Fast Line-by-Line Method for Atmospheric Absorption Computations: The Automatized Absorption Atlas," *J. Appl. Meteor.*, **20**:802–812, 1981.

Seagraves, M.A., "Visible and Infrared Extinction and Precipitation Rate in Falling Snow," *Optical Engineering for Cold Environments, Proc. Soc. Photo-Optical Inst. Eng.*, **414**:78, 198?

Seagraves, M.A. and S.F. Ebersole, "Visible and Infrared Transmission through Snow," *Opt. Eng.*, **22**:90–93, 1983.

Selby, J.E.A. and R.A. McClatchey, "Atmospheric Transmittance from 0.25 to 28.5 μm: Computer Code LOWTRAN 2," AFCRL 72-0745, AD 763721, 1972.

Selby, J.E.A. and R.A. McClatchey, "Atmospheric Transmittance from 0.25 to 28.5 μm: Computer code LOWTRAN 3," AFCRL TR75-0255, AD 017734, 1975.

Selby, J.E.A., E.P. Shettle, and R.A. McClatchey, "Atmospheric Transmittance from 0.25 to 28.5 μm: Supplement LOWTRAN 3B," AFCRL TR-76-0258, AD 040701, 1976.

Selby, J.E.A., F.X. Kneizys, J.H. Chetwynd Jr., and R.A. McClatchey, *Atmospheric Transmittance/Radiance: Computer Code LOWTRAN 4*, AFGL TR-078-0053, ADA 058643, 1978.

Shettle, E.P. and R.W. Fenn, "Models of the Atmospheric Aerosols and Their Optical Properties," in AGARD Conference Proceedings No. 183, *Optical Propagation in the Atmosphere*. Presented at the Electromagnetic Wave Propagation Panel Symposium, Lyngby, Denmark, 27–31 October 1975, AGARD-CP-183, ADA028-615, 1976.

Shettle, E.P. and R.W. Fenn, "Models of the Aerosols of the Lower Atmosphere and the Effects of Humdiity Variations on Their Optical Properties," AFGL TR-79-0214, ADA 085951, 1979.

Shettle, E.P. and A.E.S Green, "Multiple Scattering Calculations of the Middle Ultraviolet Reaching the Ground," *Appl. Opt.*, **13**:1567–1581, 1974.

Shettle, E.P. and F.E. Volz, "Optical Constants for Meteoric Dust Aerosol Models" in *Atmospheric Aerosols: Their Optical Properties and Effects*, a Topical Meeting on Atmospheric Aerosols sponsored by Opt. Soc. Am. and NASA Langley Res. Cen., Williamsburg, Virginia, NASA Conference Publication CP 2004, 13–15 Dec 1976.

Shettle, E.P. and J.A. Weinman, "The Transfer of Solar Irradiance Through Inhomogenous Turbid Atmospheres Evaluated by Eddington's Approximation," *J. Atmos. Sci.*, **27**:1048–1055, 1970.

Shettle, E.P., R.W. Fenn, and J.D. Mill, "The Optical and Infrared Properties of Atmospheric Particulates," AFGL to be published, 1985.

Sidran, M., "Broadband Reflectance and Emissivity of Specular and Rough Water Surfaces," *Appl. Opt.*, **20**:3126, 1981.

Silverman, B.A. and E.D. Sprague, "Airborne Measurements of In Cloud Visibility," Preprints, National Conference on Weather Modification of the Am. Meteorol. Soc., Santa Barbara, California, 6–9 April 1970.

Sissenwine, N., D.D. Grantham, and H.A. Salmela, "Humidity Up to the Mesopause," AFCRLTR-68-0550, AD 679996, 1968.

SMIC, *Inadvertent Climate Modification: Report of the Study of Man's Impact on the Climate*, MIT Press, Cambridge, Mass., 1971.

Smith, F.L. and C. Smith, "Numerical Evaluation of Chapman's Gracing Incidence Integral $ch(_{\chi}/_{\psi})$," *J. Soc. Res.*, **77**:3502–3597, 1972.

Smith, H.J.P., D.B. Dube, M.E. Gardner, S.A. Clough, F.X. Kneizys, and L.S. Rothman, "FASCODE—Fast Atmospheric Signature Code (Spectral Transmittance and Radiance)," AFGLTR-78-0081, ADA057359, 1978.

Smith, M.A.H., *Compilation of Atmospheric Gas Concentration Profiles from 0 to 50 km*, NASA TM-83289, 1982.

Sneider, D.E., "Refractive Effects in Remote Sensing of the Atmosphere with Infrared Transmission Spectroscopy," *J. Atmos. Sci.*, **32**:2178–2184, 1975.

Sneider, D.E., and A. Goldman, *Refractive Effects in Remote Sensing of the Atmosphere with Infrared Transmission Spectroscopy*, Ballistic Res. Labs Report 1790, Aberdeen, Maryland ADA011253, 1974.

Sobolev, V.V., *Light Scattering in Planetary Atmospheres*, Translated by W.M. Irvine, Pergamon Press, New York, p. 256, 1975.

SOLMET, *Hourly Solar Radiation—Surface Meteorological Observations*, Vol 2, TD 9724 NOAA Environmental Data Service, Asheville, N.C., Feb. 1979.

Stamnes, K. and H. Dale, "A New Look at the Discrete Ordinate Method for Radiative Transfer Calculations in Anistropically Scattering Atmospheres II. Intensity Computation," *J. Atmos. Sci.*, **38**:2696–2706, 1981.

Stamnes, K. and R. Swanson, "A New Look at the Discrete Ordinate Method for Radiative Transfer Calculations in Anisotropically Scattering Atmospheres," *J. Atmos. Sci.*, **38**:387–399, 1981.

Stewart, D.A. and O.M. Essenwanger, "A Survey of Fog and Related Optical Propagation Characteristics," *Rev. Geophys. Space Phys.*, **20**: 481–195, 1982.

Suits, G.M., *The Infrared Handbook*, Chapter 3, "Natural Sources," edited by W.L. Wolfe and G.L. Zissis, Office of Naval Research, Washington D.C., 1978.

Susskind, J. and J.E. Searl, "Synthetic Atmospheric Transmittance Spectra Near 15 and 4.3 μm," *J. Quant. Spectrosc. Radiat. Transfer*, **19**:195–215, 1978.

Tatarski, V.I., *Wave Propagation in a Turbulent Medium*, McGraw-Hill, New York, 1961.

Thompson, L. (ed.), "Advanced Scanners and Imaging Systems for Earth Observations," NASA SP-335, 1972.

CHAPTER 18

Thompson, B.C. and M.B. Wells, "Scattered and Reflected Light Intensities Above the Atmosphere," *Appl. Opt.*, **10**:1539–1549, 1971.

Toon, O.B. and J.B. Pollack, "A Global Average Model of Atmospheric Aerosols for Radiative Transfer Calculations," *J. Appl. Meteorol.*, **15**:225–246, 1976.

Toon, O.B. and J.B. Pollack, "Physical Properties of the Stratospheric Aerosol," *J. Geophys. Res.* **78**:7057–7056, 1973.

Turco, R.P., R.C. Whitten, and O.B. Toon, "Stratospheric Aerosols: Observation and Theory," *Rev. Geophys. Space Res.*, **20**:233–279, 1982.

Twitty, J.T. and J.A. Weinman, "Radiative Properties of Carbonaceous Aerosols," *J. Appl. Meteorol.*, **10**:725–731, 1971.

Twomey, S., H. Jacobowitz, and H.B. Howell, "Matrix Method for Multiple Scattering Problems," *J. Atmos. Sci.*, **23**:289–296, 1966.

Twomey, S., H. Jacobowitz, and H.B. Howell, "Light Scattering by Cloud Layers," *J. Atmos. Sci.*, **24**:70–79, 1967.

Tyler, J.E., "Optical Properties of Water," Section 15 in *Handbook of Optics*, edited by W.G. Dirscoll and W. Vaughan, McGraw-Hill, New York, 1978.

U.S. Standard Atmosphere 1962, U.S. Government Printing Office, Wash., D.C., 1962.

U.S. Standard Atmosphere Supplements 1966, U.S. Government Printing Office, Wash., D.C., 1966.

U.S. Standard Atmosphere 1976, NOAA S/T 76-1562, U.S. Government Printing Office, Wash. D.C., 1976.

van de Hulst, H.C., *Light Scattering by Small Particles*, Wiley, New York, 1957.

van de Hulst, H.C., "Multiple Scattering in Planetary Atmospheres," *J. Quant. Spectrosc. Radiat. Transfer*, **11**:785–795, 1971.

van de Hulst, H., *Multiple Light Scattering Tables, Formulas and Applications*, Vols. 1 and 2, Academic Press, New York, 1980.

van de Hulst, H.C. and K. Grossman, "Multiple Scattering in Planetary Atmospheres" in *The Atmosphere of Venus and Mars*, edited by J.C. Brandt and M.B. McElroy, 35–55, Gordon and Breach, New York, 1968.

Van Vleck, J.H. and D.L. Huber, "Absorption, Emission and Linebreadths: A Semi-historical Perspective," *Rev. Mod. Phys.*, **49**:939, 1977.

VanZandt, T.E., K.S. Gage, and J.M. Warnock, "An Improved Model for the Calculations of Profiles of C_n^2 and in the Free Atmosphere from Background Profiles of Wind, Temperature, and Humidity," in *Preprints*, Radar Meteorology, Conf. 20th, 1981, Boston, Mass., Amer. Meterol. Soc., Boston, Mass., 1981.

Volz, F.E., "Infrared Absorption by Atmospheric Aerosol Substance," *J. Geophys. Res.*, **77**:1017–1031, 1972a.

Volz, F.E., "Infrared Refractive Index of Atmospheric Aerosol Substances," *Appl. Opt.*, **11**:755–759, 1972b.

Volz, F.E., "Infrared Optical Constants of Ammonium Sulfate, Sahara Dust, Volcanic Pumice and Flyash," *Appl. Opt.*, **12**:564–568, 1973.

Walters, D.L. and K.E. Kunkel, "Atmospheric Modulation Transfer Function for Desert and Mountain Locations: The Atmospheric Effects on r_o," *J. Opt. Soc. Am.*, **71**:397–405, 1981.

Warren, S.G. "Optical Constants of ice from the Ultraviolet to the Microwave," *Appl. Opt.*, **23**: 1206–1225, 1984.

Weinman, J.A., "Axially Symmetric Transfer of Light Through a Cloud of Anisotropically Scattering Particles," *Icarus*, **9**:67–73, 1968.

Welch, R. and W. Zdunkowski, "Backscattering Approximations and Their Influence on Eddington-Type Solar Flux Calculations," *Contrib. Atmos. Phys.*, **55**.28–42, 1982.

Whitby, K.T., "On the Multimodal Nature of Atmospheric Aerosol Size Distributions," Presented at VII Intern. Conf. on Nucleation, Leningrad, USSR, September 1973.

Whitby, K.T. and B. Cantrell, "Atmospheric Aerosols—Characteristics and Measurements," Proc. Intern. Conf. on Environ. Sensing and Assessment, 3rd., Las Vegas, Nev., 14–19 September, IEEE, New York, 1976.

Whitby, K.T., R.B. Hugar, and B.Y. Liu, "The Aerosol Size Distribution of Los Angeles Smog," *J. Colloid and Interface Sci.*, **30**:177–204, 1972.

Whiting, E.E., "An Empirical Approximation to the Voigt Profile," *J. Quant. Spectrosc. Radiat. Trans.*, **8**:1379, 1968.

Whitney, C.K., "Implications of a Quadratic Stream Definition in Radiative Transfer Theory," *J. Atmos. Sci.*, **29**:1520–1530, 1972.

Whitney, C.K., "Efficient Stream Distributions in Radiative Transfer Theory," *J. Quant. Spectrosc. Radiat. Transfer*, **14**:591–611, 1974.

Whitney, C.K. and H.L. Malchow, "Study of Radiative Transfer in Scattering Atmospheres," AFGL-TR-78-0101, ADA057359, June 1978.

Winchester, L.W., Jr., G.G. Ginstead, and R.B. Wetzel, "Measurements of the Phase Function of Natural Particles," in *Atmospheric Effects on Systems Performance, Proc. Soc. Photo-Optical. Int. Eng.*, **304**:106–118, 1981.

Wiscombe, W., "The Delta-M Method: Rapid Yet Accurate Radiative Flux Calculations for Strongly Asymmetric Phase Functions," *J. Atmos. Sci.*, **34**:1408–1422, 1977.

Wiscombe, W., "Atmospheric Radiation 1975–1983," *Rev. Geophys. Space Phys.*, **21**:997–1021, 1983.

Wiscombe, W. and G. Grams, "The Backscattered Fraction in Two-Stream Approximations," *J. Atmos. Sci.*, **33**:2440–2451, 1976.

WMO, "The Stratosphere 1981 Theory and Measurements," WMO Global Ozone Research and Monitoring Project Report No. 11, World Meteorological Organization, Geneva, 1982.

Wyatt, P.J., V.R. Stull, and G.N. Plass, "Infrared Transmittance of Water Vapor," *Appl. Opt.*, **3**:229, 1964.

Wyngaard, J.C., Y. Izumi, and S.A. Collins, "Behavior of the Refractive-Index Structure Parameter Near the Ground," *J. Opt. Soc. Am.*, **61**:1646–1650, 1971.

Young, A.T., "Revised Depolarization Corrections for Atmospheric Extinction," *Appl. Opt.*, **19**:3427–3428, 1980.

Zdunkowski, W., R. Welch, and G. Korb, "An Investigation of the Structure of Typical Two-Stream Methods for the Calculations of Solar Fluxes and Heating Rate in Clouds," *Contrib. Atmos. Phys.*, **53**:147–165, 1980.

Chapter 19

ELECTROMAGNETIC WAVE PROPAGATION IN THE LOWER ATMOSPHERE

Section 19.1 V. J. Falcone, Jr.
Section 19.2 R. Dyer

19.1 REFRACTION IN THE LOWER TROPOSPHERE

The speed of propagation of an electromagnetic wave in free space is a constant, c, which is equal to 3×10^8 m/s. In a material medium such as the atmosphere, the speed of propagation varies. Even small variations in speed produce marked changes in the direction of propagation, that is, refraction.

In the atmosphere, the speed of propagation varies with changes in composition, temperature, and pressure. At radio wavelengths, speed does not vary significantly with the wavelength, but in the optical region the speed depends strongly on the wavelength. In the lower 15 km of the atmosphere, water vapor is the most highly variable of the atmospheric gases, and at radio wavelengths the speed of propagation is strongly affected by water vapor. Temperature and pressure variations are principally functions of altitude, although for propagation at small elevation angles significant variations may occur along horizontal distances. From the standpoint of effect on the speed of propagation, temperature variations at any given altitude are more significant than pressure variations.

In its most general form, the refractive index is a complex function. The real term of the complex function is called the *phase refractive index*, n;

$$n = \frac{c}{v} \tag{19.1}$$

where c is the speed of light in a vacuum and v, the *phase velocity*, is the speed of propagation in a particular medium. In the troposphere where n is nearly equal to one, it is convenient to define the quantity

$$N = (n - 1) \times 10^6. \tag{19.2}$$

N is called the *refractive modulus*; units of $(n - 1) \times 10^6$ are called N-units.

19.1.1 Optical Wavelengths

An approximate relation between the optical refractive modulus and atmospheric pressure and temperature is

$$N_x = 77.6 \frac{P}{T} \tag{19.3}$$

where N_x is the refractive modulus for wavelengths ≥ 20 μm, P is atmospheric pressure in millibars, and T is atmospheric temperature in degrees kelvin.

The dispersion formula of Edlen [1953], which has been adopted by the Joint Commission for Spectroscopy, is

$$N_s = 64.328 + \frac{29498.10}{146 - 1/\lambda^2} + \frac{255.40}{41 - 1/\lambda^2}, \tag{19.4}$$

where N_s is the refractive modulus at a wavelength λ for a temperature of 288 K and a pressure of 1013.25 mb, and λ is the wavelength in micrometers. A somewhat less precise but more convenient dispersion formula is

$$N = N_x \left[1 + \frac{7.52 \times 10^{-3}}{\lambda^2}\right]. \tag{19.5}$$

Equations (19.3) and (19.5) can be combined to give the refractive modulus as a function of pressure, temperature, and wavelength;

$$N = \frac{77.6\ P}{T} + \frac{0.584\ P}{T\ \lambda^2}. \tag{19.6}$$

Refractive moduli calculated by using Equation (19.6) will be in error no more than one N-unit over the temperature range 243 to 303 K for wavelengths from 0.2 to 20 μm. Thus Equation (19.6) covers the spectrum from the far ultraviolet through the near infrared. A more accurate relationship is given in Chapter 18.

CHAPTER 19

19.1.2 Radio Wavelengths

At radio wavelengths the relationship of refractive modulus to pressure, temperature, and water-vapor pressure is given by

$$N = \frac{77.6P}{T} + \frac{373000P_{wv}}{T^2}, \qquad (19.7)$$

where P_{wv} is the partial pressure of water vapor in millibars, P is pressure in millibars, and T is temperature in degrees kelvin. Equation (19.7) is accurate to 0.1 N-unit from the longest radio wavelengths in use down to about 6 mm (50 GHz); 5 N-units from 6 to 4 mm (50 to 75 GHz); and 1 N-unit from 4 to 2.6 mm (75 to 115 GHz). A more accurate description of refraction and its effects in the 30 to 1000 GHz region (EHF range) has been investigated by Liebe [1980].

Absorption by atmospheric constituents begins to rise to significant proportions with decreasing wavelength beginning near 1.5 cm. Water vapor content is by far the leading factor in causing changes in N, followed in order of importance by temperature and pressure. For example, for a temperature of 288 K, pressure of 1013 mb near ground level, and a relative humidity of 60% (P_{wv} = 10 mb), the fluctuation, ΔN, is

$$\Delta N = 4.5\ \Delta P_{wv} - 1.26\ \Delta T + 0.27\ \Delta P. \qquad (19.8)$$

As Equation (19.8) shows, a fluctuation in water vapor pressure has 16 times the effect on the refractive index as the same amount of fluctuation in total pressure and 3.5 times the effect as the same fluctuation in temperature. Equation (19.7) may also be used in the windows of relative transparency for submillimeter waves ($\lambda > 100$ μm, $f < 3 \times 10^6$ GHz) with an error of 10 to 20 N-units.

Equation (19.7) is a function of temperature, pressure, and vapor pressure all of which are height dependent, that is, elevation (h) above the surface; thus N(h) is the *refractivity structure*. In reality, surfaces of constant refractivity are not planes, but are concentric spheres about the earth's center. In characterizing the atmospheric layers that affect radio wave propagation a *modified refractivity structure* M(h) is defined.

$$M = \left(n + \frac{h}{a} - 1\right) \times 10^6 \qquad (19.9)$$

$$M(h) = N(h) + 157h \text{ in units of M} \qquad (19.10)$$

where h is in kilometers, a is the radius of the earth (6370 km) and $1/a = 157 \times 10^{-6}$ km^{-1}.

When the lapse rate of N is less than −157 N-units per kilometer [Equation (19.10)], the slope of M becomes negative indicating a ducting condition. Figure 19-1 illustrates ducts at 1000 and 700 mb [Morrissey, personal communication, 1982].

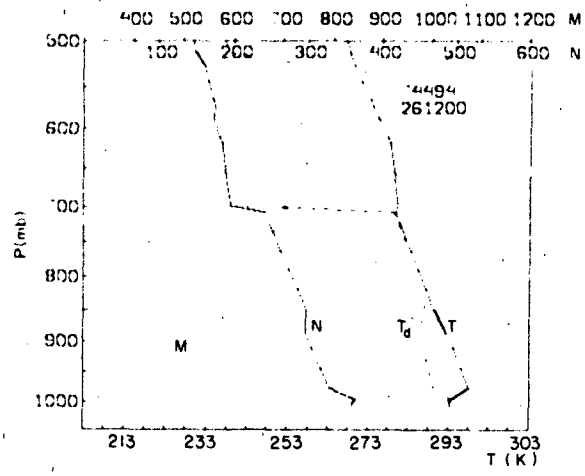


Figure 19-1. Data from Chatham, Mass. radiosonde release of 26 July 1982.

19.1.3 Standard Profiles of Refractive Modulus

The vertical distribution of the refractive modulus can be calculated from Equation (19.3) using vertical distributions of vapor pressure and temperature as a function of pressure. Under normal conditions, N tends to decrease exponentially with height. An exponential decrease is usually an accurate description for heights greater than 3 km; below 3 km, N may depart considerably from exponential behavior. The median value for the gradient dN is typically −0.0394 N/m for the first few thousand meters above ground level.

For many purposes it is desirable to have standard refractive-modulus profiles for the atmosphere. By using the equations of the model atmospnere, an exact analytical expression for the standard optical refractive modulus can be derived. A simplified approximation to this is

$$N_s = 273 \exp\left(-\frac{Z}{9.82}\right), (Z \leq 7.62) \qquad (19.11)$$

Z is the altitude in thousands of km.

Equation (19.11) can be differentiated to obtain the standard gradient of optical refractive modulus;

$$\frac{dN_s}{dZ} = 27.8 \exp\left(-\frac{Z}{9.82}\right), (Z \leq 7.62). \qquad (19.12)$$

Equations (19.11) and (19.12) may be corrected for dispersion through use of Equation (19.5).

For the radio wavelengths, it is necessary to assume a distribution of water vapor in order to obtain an expression for the refractive modulus. Assuming $P_{wv} = 10.2$ $(1 - 0.064Z)$, for $Z \leq 7.62$, a simplified approximation is

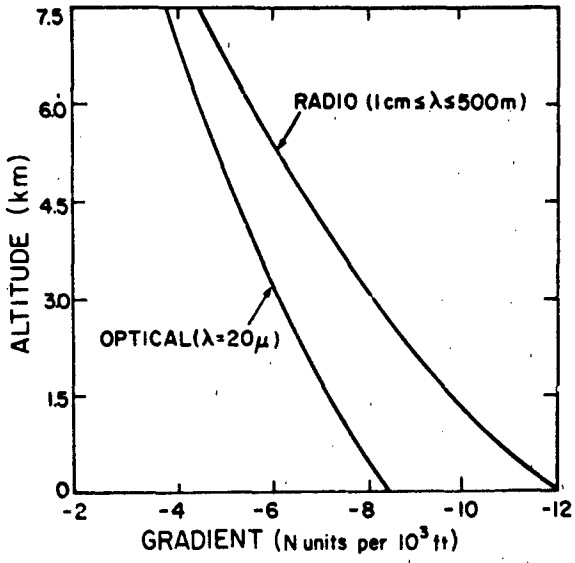


Figure 19-2. Variation of standard gradient of refractive modulus with altitude.

$$N = 316 \exp\left(-\frac{Z}{8.08}\right), (Z \leq 7.62). \quad (19.13)$$

The standard gradient of radio-wave refractive modulus is then:

$$\frac{dN}{dZ} = -39.1 \exp\left(-\frac{Z}{8.08}\right), (Z \leq 7.62). \quad (19.14)$$

Figures 19-2 and 19-3 are graphs of standard profiles calculated from Equations (19.12) through (19.14).

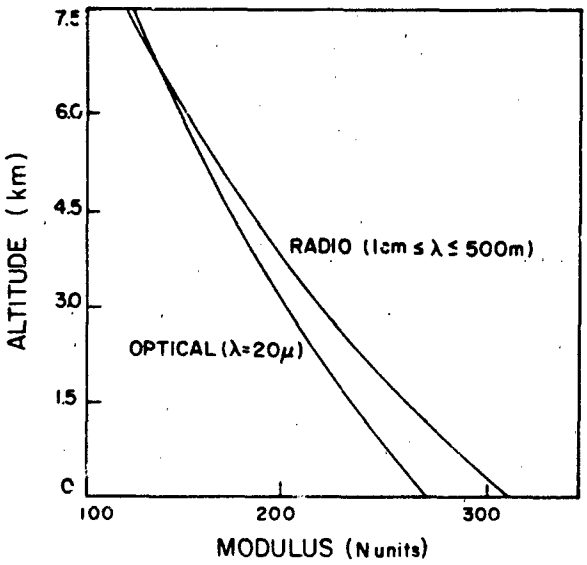


Figure 19-3. Variation of standard refractive modulus with altitude

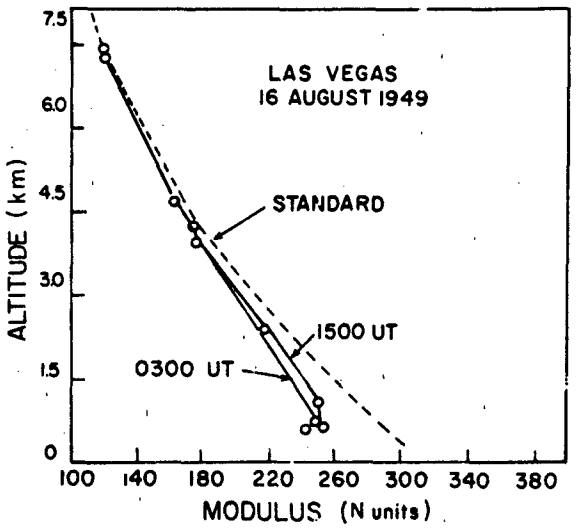


Figure 19-4. Microwave refractive modulus profile in continental tropical air mass.

19.1.4 Variations of Refractive Moduli

Actual profiles may differ markedly from the standard profiles. Figures 19-4 through 19-7 show some profiles of refractive modulus at microwave frequencies calculated from radiosonde measurements. These are considered typical for the air masses indicated. Average deviations from a model atmosphere refractive index have been studied extensively; for example, see Bean and Dutton [1968].

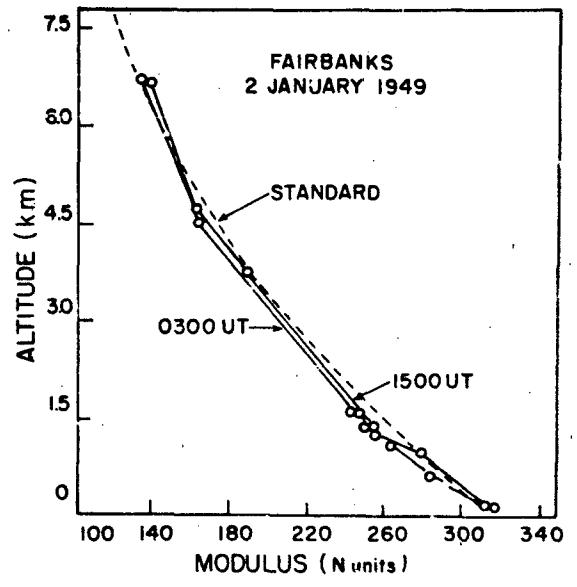


Figure 19-5. Microwave refractive modulus profile in continental polar air mass.

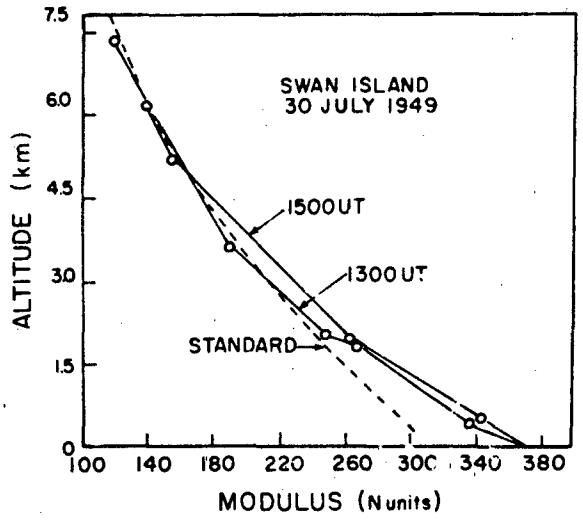


Figure 19-6. Microwave refractive modulus profile in continental polar air mass.

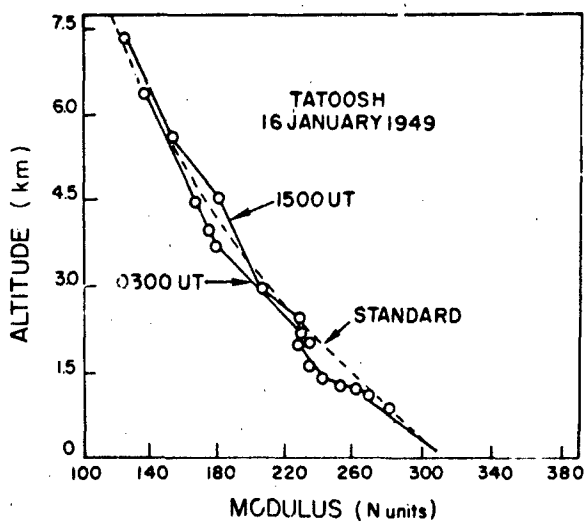


Figure 19-7. Microwave refractive modulus profile in maritime polar air mass.

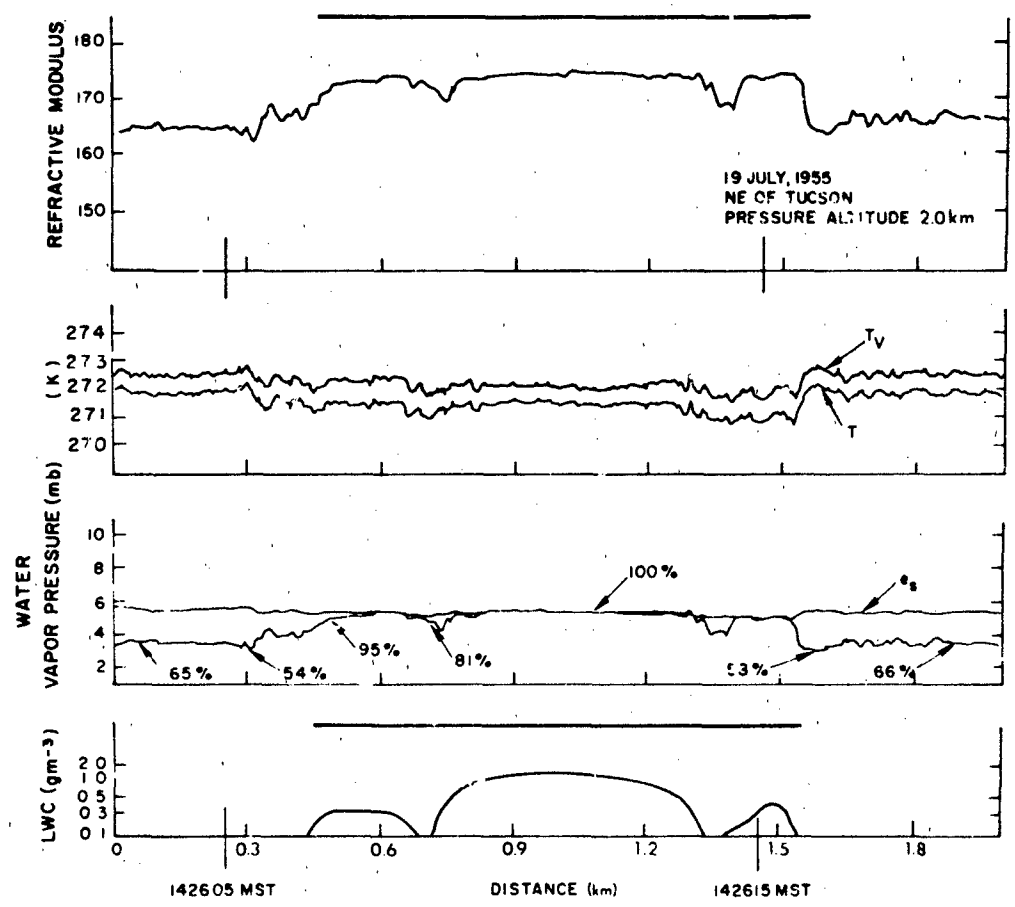


Figure 19-8. Aircraft measurements through a cumulus cloud. The heavy lines show the time during which an observer in the aircraft indicated that the plane was within the visible cloud. The calculated virtual temperature, T_v, was corrected for liquid water content; T is the measured temperature. Relative humidity in percent is shown on the ambient water-vapor pressure curve; the curve labeled e_s is the calculated saturation water-vapor pressure at the temperature encountered. The bottom curve shows the liquid water content (LWC).

Cumulus clouds are evidence of the existence of a very inhomogeneous field of water vapor within the atmosphere. Figure 19-8 shows some measurements of refractive modulus and associated parameters within a fair-weather cumulus cloud. The time response of the instruments from which refractive modulus, temperature, and water-vapor pressure were obtained was such that changes occurring in distances as small as 1.5 m could be measured. However, the instrument for measuring liquid water content had a much slower response. Figure 19-9 shows a composite cloud which summarizes data from 30 cloud passes.

Figures 19-10a and b show the average ΔN between cloud and clear air to be expected in various parts of the United States at the midseason months. The chances of having cumulus clouds at 1500 h local time for these months is also shown. Additional climatological data on ΔN and cumulus clouds is given by Cunningham [1962].

The deviations in refractive modulus are principally in the vertical direction. Regions of more or less constant gradient of the refractive modulus are called *stratified layers*. The horizontal extent of these layers may vary from a few kilometers to hundreds of kilometers depending on the meteorological processes by which they are produced. When N decreases with height inside a specific layer much faster than it does above or below the layer, the layer is said to be *super-refracting* for propagation. One cause of this is a temperature inversion. Layers of negative height-gradient of N in association with regions in which the temperature gradient is positive (or less negative) than the gradients of the layers just above and below are known as *subsidence inversion layers*. These layers generally have a large horizontal extent. Layers in close proximity to the earth's surface are strongly influenced by the local conditions of the earth's surface and for this reason show more variability than the layers described above.

In the inversion layer, the temperature may change by a few degrees in intervals of from fifty to a few hundred meters in altitude. This temperature difference accounts for a change of only a few N-units. However, an inversion usually indicates the presence of a humid air mass under a dry one. The transition from humid to dry air causes a marked change of N in the super-refracting layer, typically of 20 to 50 N-units; changes as pronounced as 80 N-units have been measured. Super-refracting layers may be clear-weather phenomena, or can be accompanied by haze (aerosols) in the lower air mass. Invariably they signify stable weather situations, such as occur when a high pressure center stagnates in an area.

The horizontal and temporal extent of super-refractive layers varies widely. In New England it may be only a few tens of kilometers. In mideastern states, the layers extend farther and may last from a half hour up to a week. In the trade-wind zones of the world, the climatic regime (manifested by steady wind directions and speeds throughout most of the year) sustains super-refracting layers, which extend a few thousand miles both east to west and north to south.

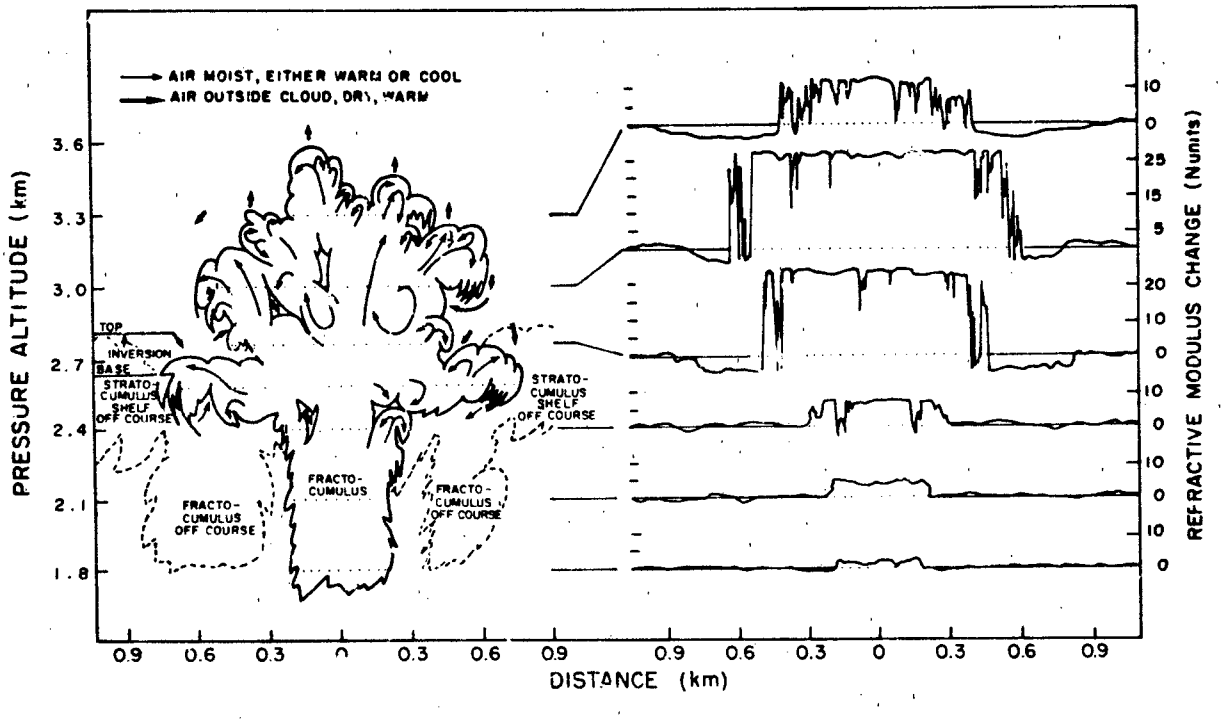


Figure 19-9 Average cloud shape cross section 90° to wind shear and average refractive modulus changes on 30 June 1955 NW of Boston, Mass

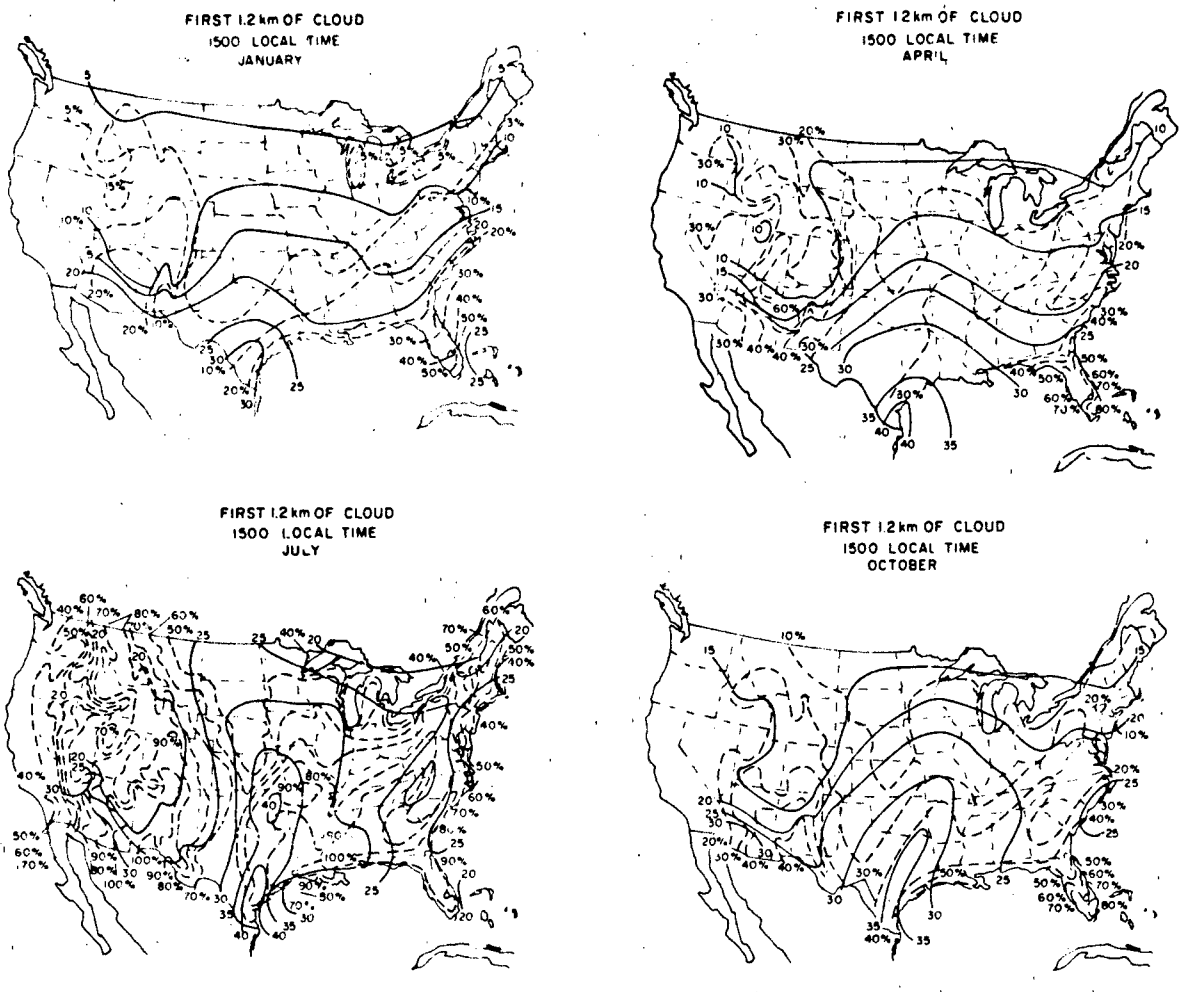


Figure 19-10. Percent frequency of cumulus and cumulonimbus cloud and ΔN, the average change in refractive modulus between the cloud and clear air for January, April, July, and October.

19.1.5 Turbulence

It has long been recognized that the effects produced by the atmosphere upon electromagnetic waves propagating through it are a measure of the nature of the atmosphere. If the atmosphere is considered as a medium with electromagnetic properties that are functions of space and time, then atmospheric properties can be investigated by measurements of wave propagation; these observations are called *remote probing*. The atmospheric (or meteorological) parameters are inferred from their influence on Equations (19.6) and (19.7); it is assumed that scattering is due to random fluctuations in the dielectric constant of the atmosphere. When fluctuations in the refractivity are of interest, they are studied via the correlation function and its Fourier transform, the spectral density. This approach is described by Tatarski [1961 and 1971], Staras and Wheelon [1959], Hufnagle and Stanley [1964], and Strohbehn [1968]. A tutorial review has been written by Dewan [1980].

A locally-homogeneous isotropic turbulent model of the atmosphere is assumed, that is, a model of well mixed random fluctuations. This model is restrictive and requires justification for each new application. One assumes the *spectral density* $\Phi_n(\kappa)$ (the three dimensional spectrum of refractivity) is given by

$$\Phi_n(\kappa) = 0.033\, C_n^2\, \kappa^{-11/3} \exp\left(\frac{-\kappa^2}{\kappa_m^2}\right) [\text{cm}^3] \quad (19.15)$$

where $\kappa_o < \kappa < \kappa_m$. The spatial wave number $\kappa(\text{cm}^{-1})$ is $2\pi/\ell$, ℓ is the size of the turbulent eddy or blob, κ_o is $2\pi/L_o$, L_o is the outer scale of turbulence (typically the order of 10^3 to 10^4 cm, depending on layer height), κ_m is $5.92/\ell_o(\text{cm}^{-1})$, ℓ_o is the inner scale or turbulence, and C_n is specified by Equation (19.16), that is, C_n^2 is a measure of magnitude of the mean squared fine scale gradients. Figure 19-11 shows the typical spectrum of irregularities, $\Phi_n(\kappa)$

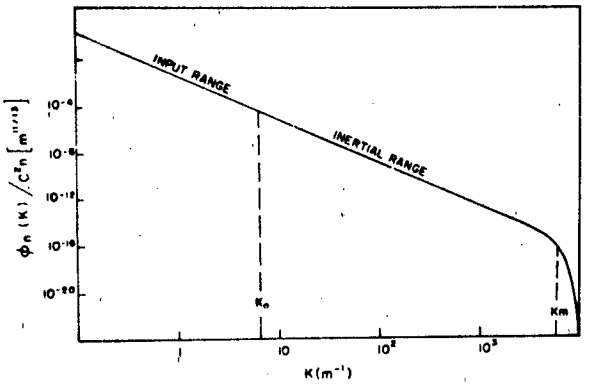


Figure 19-11. Three-dimensional spectrum of refractive index fluctuations.

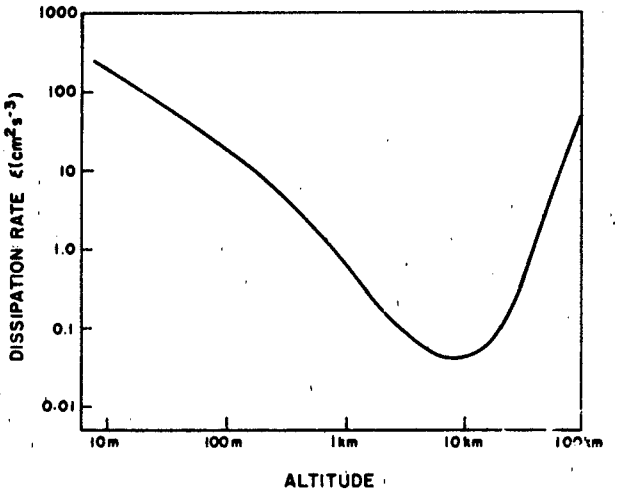


Figure 19-12. Dissipation rate ε vs altitude.

and the ranges of energy input, redistribution, and dissipation. Physically, the energy is put into the turbulence from the largest scale sizes (smallest value of κ) by wind shear and convective heating; the energy-producing eddies are assumed to have a spatial wave number less then κ_o. The region between κ_o and κ_m is the redistribution (inertial) range, where energy is transferred from large edies (small k) to smaller eddies (larger κ) until viscous effects becomes important at $\kappa_m = 5.92/\ell_o$ and the energy is dissipated. Near the ground, ℓ_o is of the order of 0.1 to 1 cm.

Once the value of C_r^2 is determined, the spectral density is known. C_n^2 may be found directly from the dimensionless structure function $D_n(r)$,

$$D_n(\mathbf{r}) = \overline{[N(\mathbf{r} + \mathbf{r}_1) - N(\mathbf{r}_1)]^2} = C_n^2 r^{2/3} \quad (19.16)$$

where $N(\mathbf{r})$ is the normalized fluctuating part of the index of refraction, the bar indicates the average of the squared quantity, and $r(r \gg \ell_o)$ is the size of the inhomogenities determined from the differences in the values of N at two points $\mathbf{r}$ and $\mathbf{r}_1$.

At optical wavelengths, C_n is determined from temperature measurements alone;

$$C_n = 10^{-6}(\rho/\rho_o)C_T \ [cm^{-1/3}] \quad (19.17)$$

where ρ/ρ_o is the ratio of the average atmospheric density at a given altitude to the density at sea level. The structure constant, C_T, is

$$C_T = 2.4\,\varepsilon^{1/3}(\gamma/\beta) \ [K\ cm^{-1/3}] \quad (19.18)$$

where ε ($cm^2 s^{-3}$) is the rate of energy per unit mass dissipated by viscous friction, γ ($K\ cm^{-1}$) is the average vertical gradient of the potential temperature, and β (s^{-1}) is the average shear or vertical gradient of the horizontal wind [Hufnagle and Stanley, 1964 or Chapter 18].

Figure 19-12 shows the average dissipation rate estimated from observed values as a function of altitude. Figures 19-13 and 19-14 are plots of observed geometric-mean values of γ and β respectively as functions of altitude. The values of γ are taken from the 1966 Supplementary Atmosphere; the values of β are computed from reported wind profiles. Figure 19-15 shows C_r as a function of altitude; the values of ρ/ρ_o used to compute C_n are taken from the 1966 Supplementary Atmosphere. At radio wavelengths, Equation (19.17) cannot be used because water vapor must be considered as well as temperature [Crane, 1968, 1980].

C_n for a model atmosphere can be obtained directly from Figure 19-15. When values of ρ/ρ_o are known from radiosonde observations, C_n for the given ρ/ρ_o can be calculated by using values obtained from Figures 19-12 through 19-14 and Equations (19.17) and (19.18).

Hufnagle [1974] synthesized a model for C_n based on

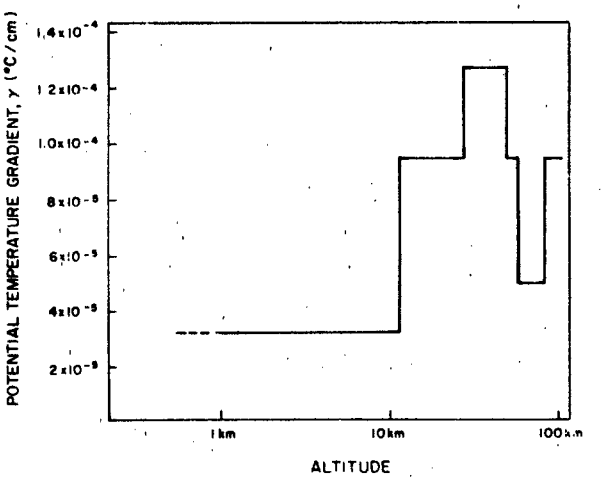


Figure 19-13. Average potential temperature gradient (γ) vs altitude.

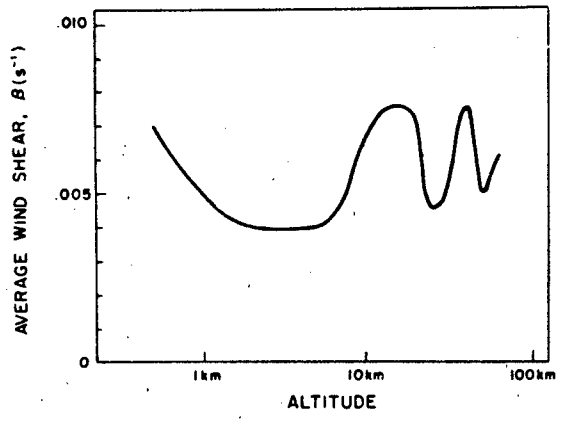


Figure 19-14. Average wind shear (β) vs altitude.

the empirical observation that the best correlation factor to correlate the scintillation spectrum and the meteorological parameter of wind speed is

$$W = \left[\frac{1}{15}\int_{5\ km}^{20\ km} v^2(h)dh\right]^{1/2} \quad \text{(units of m/s)} \tag{19.19}$$

for h in km

$$C_n^2 = \left\{\left[\left(2.2 \times 10^{-53}\right)h^{10}\left(\frac{W}{27}\right)^2\right]\exp\left[\frac{-h}{1000}\right] + (10^{-16})\exp\left(\frac{-h}{1500}\right)\right\}\exp\,[r(h,t)] \tag{19.20}$$

in units of $m^{-2/3}$

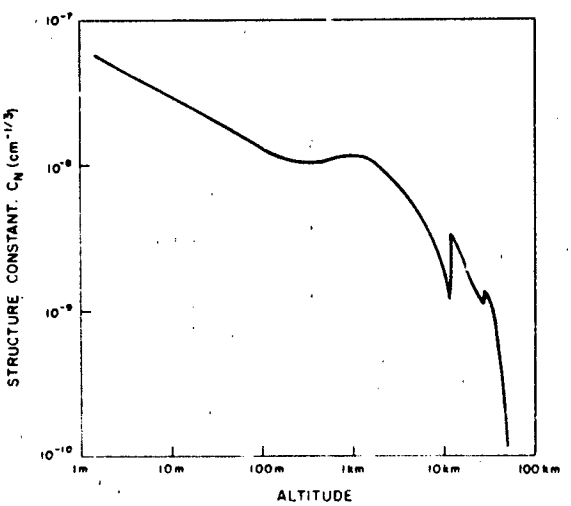


FIgure 19-15 Index of refraction structure (C_n) vs altitude.

where h is height in meters above sea level, r is a zero mean homogeneous Gaussian random variable with a covariance function given by

$$\langle r(h + h', t + \tau)\, r(h,t)\rangle = A(h'/100)e^{-\tau/5} + A(h'/2000)e^{-\tau/80} \tag{19.21}$$

The model for $C_n(h, t)$ is valid for heights between 3 and 24 km. The time interval τ is measured in minutes and $A(h'/L) = 1 - |h/L|$ for $(h') < 1$, and zero otherwise; $\langle r^2\rangle = 2$ and $\langle \exp r\rangle = e = 2.7$ for cases in which fine structure is not of interest. VanZandt et al. [1981] discuss another C_n^2 model. Brown et al. [1982] and Good et al. [1982] present the correlating data. C_n^2 may be determined from radar backscatter measurements in the atmosphere; see for example, Staras and Wheelon [1959], Hardy and Katz [1969], Ottersten [1969], and Gage et al. [1978].

At radio frequencies, the mechanism responsible for backscatter and forward scatter beyond the horizon is the refractive index variation due to fluctuations in properties of the atmosphere. The ratio of the received-to-transmitted power (P_r/P_t) depends upon the integral of the scattering cross section per unit volume over the common volume defined by the antenna pattern or patterns for backscatter and forward scatter respectivity.

$$\frac{P_r}{P_t} = \frac{\lambda^2}{4\pi^2}\int_{Volume}\left(\frac{G_t G_r}{R_t^2 R_r^2}\right)^2 \mu\, d^3r \tag{19.22}$$

where λ is the wavelength employed, G_t and G_r are the gains of the transmitting and receiving antennas respectively, R_t^2 and R_r^2 are the distances from the scattering volume (d^3r) to the respective antennas, and μ (a reciprocal length) is the scattering cross section per unit volume. Figure 19-16 illustrates the path geometry for Equation (19.18).

The scattering cross section per unit volume is directly related to the spectral density:

$$\mu = (\pi^2/\lambda^4)\,\phi_n(\kappa) \tag{19.23}$$

where $|\kappa|$ is $(4\pi/\lambda)\sin(\theta/2)$.

Scattering is not the only effect of atmospheric turbulence on the propagation of electromagnetic waves. As the waves propagate through the atmosphere, fluctuations in amplitude and phase occur. The amplitude and phase fluctuations may be described using the equations of geometrical optics or the smooth perturbation solution of the full wave equation. A summary of the solution, based on work by Tatarski [1961], is provided below.

In order to apply the geometrical optics approximation, the conditions that have to be satisfied are

$$\lambda \ll \ell_o \text{ and } (\lambda L)^{1/2} \ll \ell_o,$$

where λ is the wavelength of the propagating wave, ℓ_o is the inner scale of turbulence, and L is the path length over

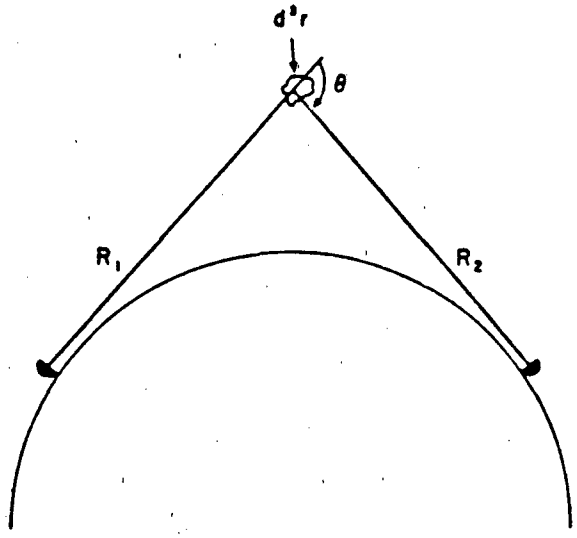


Figure 19-16. Geometry for turbulent scattering.

which the wave propagates. The *phase fluctuation* at a point r along the path is

$$S(\mathbf{r}) = k \int_o^r n(s)\, ds \qquad (19.24)$$

where S(r) is the total phase change, k is $2\pi/\lambda$, n(s) is the index of refraction in the direction of s, and ds is the element of path length. The *amplitude fluctuation* χ at point r along the path is

$$\chi = \ln\left[\frac{A(\mathbf{r})}{A_o(\mathbf{r})}\right] \qquad (19.25)$$

$$\chi = \frac{1}{2k} \int_o^s \nabla_T^2 S_1(\xi,y,z)\, d\xi$$

where A(r) is total amplitude at point r, ∇_T^2 is the transverse Laplacian, and ξ is the direction of propagation. $S_1(\xi,y,z)$ is the phase fluctuation about its mean value.

$$S_1(\mathbf{r}) = k \int_o^r \delta n(s)\, ds \qquad (19.26)$$

where $\delta n(s)$ is the deviation of the index of refraction from its mean value. If n(s) depends only on altitude, then Equation (19.24) may be used to determine the average phase change by substituting the average index of refraction in the integrand

$k \int_o^r (n(s))ds$; this is useful in studying refraction.

The covariance function of phase fluctuations, $C_s(\zeta)$, at two receivers that are both an equal distance L away from the transmitter and are separated from each other by a distance $\zeta = (y^2 + z^2)^{1/2}$ is

$$C_s(\zeta) = 2\pi \int_o^\infty F_s(\kappa) J_o(\kappa\zeta) \kappa d\kappa. \qquad (19.27)$$

$J_o(\kappa\zeta)$ is the Bessel function and $F_s(\kappa)$ is the two dimensional Fourier transforms of $C_s(\zeta)$;

$$F_s(\kappa) = 2\pi\kappa^2 L\, \phi_n(\kappa)\ [cm^2]. \qquad (19.28)$$

The covariance function of amplitude fluctuations, $C_\chi(\zeta)$, for the conditions given above is

$$C_\chi(\zeta) = 2\pi \int_o^\infty F_\chi(\kappa) J_o(\kappa\zeta)\, \kappa d\kappa \qquad (19.29)$$

and

$$F_\chi(\kappa) = \left(\frac{\pi L^3}{6}\right) \kappa^4 \phi_n(\lambda)\ [cm^2]. \qquad (19.30)$$

Under restrictions that $\lambda \ll \ell_o$ and that $L \ll \ell_o^4/\lambda^3$, the smooth perturbation solution of the full wave equation leads to

$$F_s(\kappa) = \pi\kappa^2 L \left(1 + \frac{\sin\alpha}{\alpha}\right) \phi_n(\kappa)\ [cm^2] \qquad (19.31)$$

and

$$F_\chi(\kappa) = \pi\kappa^2 L \left(1 - \frac{\sin\alpha}{\alpha}\right) \phi_n(\kappa)\ [cm^2] \qquad (19.32)$$

where α is $\kappa^2 L/k$. These restrictions limit the validity of Equations (19.31) and (19.32) to wavelengths in the millimeter and optical regions and to relatively short paths, although in certain cases these restrictions may be relaxed, for example for $\lambda \geq \ell_o$. C and F can be measured only over a finite range of values of κ, therefore a complete knowledge of C or of F is not possible.

19.2 ATTENUATION AND BACKSCATTERING

Scattering and attenuation are usually complicated functions of particle size and dielectric properties. The square root of the dielectric constant m is

$$\sqrt{m} \equiv n - i\kappa, \qquad (19.33)$$

where n is the phase *refractive index*, κ is the *absorption index* of the medium, and i is $\sqrt{-1}$.

In describing the properties of the particles, it is convenient to use the parameter K, defined by

$$K \equiv \frac{m - 1}{m + 2} \tag{19.34}$$

When the particles are small in comparison with the transmitted wavelength, the Rayleigh approximation holds, and both the backscatter and the absorption are simple functions of K. In this special case, backscatter is proportional to $|K|^2$, and attenuation is proportional to the imaginary part of minus K or $(\mathrm{Im}(-K))$.

For water, $|K|^2$ is practically constant and equals 0.93 over a wide range of temperatures and wavelengths in the centimeter range. Similarly $|K|^2 = 0.176$ for ice of normal density (0.917 g/cm^3) and centimeter wavelengths. However, the imaginary part of K can vary significantly with temperature and wavelength, and both $(K)^2$ and $(\mathrm{Im}(-K))$ vary with frequency at millimeter and submillimeter wavelengths. Unfortunately, measurements have not been made at every possible combination of temperature and wavelength, and there is no single expression relating all the variables. To obtain the real and imaginary parts of K at the desired temperature and wavelength, the reader is referred to the computer program written by Ray [1972], as corrected by Falcone et al. [1979]. This program interpolates between measured values.

19.2.1 Backscattering and Attenuation Cross Sections

The echo power returned by a scattering particle is proportional to its *backscattering cross section*, σ. The power removed by an attenuating particle is proportional to the *total absorption cross section*, Q_t. *The size parameter* (electrical size) is πD/λ; D is the particle diameter and λ the wavelength of the incident radiation. When the diameter of the scattering or attenuating particle is small with respect to λ, the backscattering and total absorption cross sections may be expressed with sufficient accuracy by the Rayleigh approximation.

For spherical particles, if $D/\lambda < 0.2$,

$$\sigma = \frac{\pi^5 |K|^2 D^6}{\lambda^4} \ [\mathrm{cm}^2] \tag{19.35}$$

For particles with $\pi D/\lambda > 0.2$, σ should be computed from the equations of the Mie theory of scattering [Battan, 1959]. Figures 19-17 and 19-18 show normalized backscattering cross section $(4\sigma/\pi D^2)$ for ice and for water versus the size parameter, computed from the exact Mie equations. The normalized curve for ice is invariant with wavelength in the microwave region; the normalized curve for water is for a temperature of 273 K and a 3.2-cm wavelength. As these figures show, ice spheres equal to or larger than the wavelength may scatter more than an order of magnitude greater than water spheres of the same size. This is confirmed by experimental measurements.

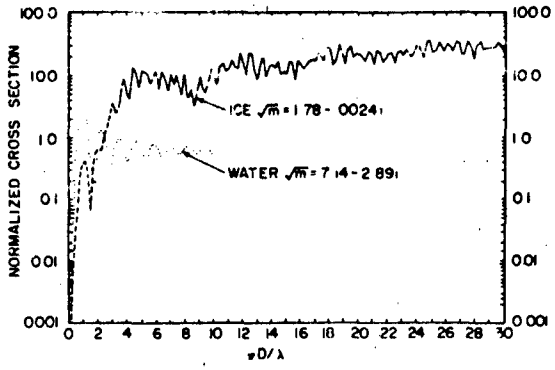


Figure 19-17. Calculated values of the normalized $(4\sigma/\pi D^2)$ backscatter cross section for water at 3.21 cm and 273 K and for ice at wavelengths from 1 to 10 cm.

Measurements at 5 cm wavelength indicate that the back-

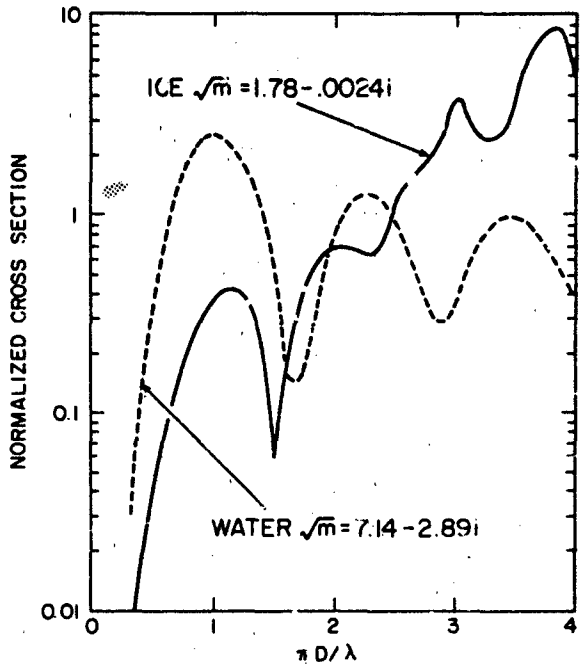


Figure 19-18. Calculated values of the normalized $(4\sigma/\pi D^2)$ backscatter cross section for water at 3.21 cm and 273 K and for ice at wavelengths from 1 to 10 cm. (Detail of Figure 19-17)

scattering of so-called "spongy" hail (a mixture of ice and water) is 3 to 4 dB above that of the equivalent all-water spheres and at least 10 dB above that of the equivalent solid ice spheres [Atlas et al., 1964]. Because of the variabilities of sizes, shapes, and liquid water content of hail, no general rules concerning backscattering and attenuation cross sections for hail can be made. As a first approximation, however, the ice curve of Figures 19-17 and 19-18 may be used.

For spherical particles, if $\pi D/\lambda < 0.1$,

$$Q_t = \frac{\pi^2 D^3}{\lambda} \operatorname{Im}(-K) + \frac{2}{3}\sigma \ [\mathrm{cm}^2]. \qquad (19.36)$$

When $\pi D/\lambda > 0.1$, Q_t must also be computed from the exact Mie equations. Several computer programs are available to compute the total attenuation caused by any distribution of water and/or ice particles. [See, for example, Falcone et al., 1979.]

19.2.2 Reflectivity

The average echo power returned by a group of randomly distributed scattering particles is proportional to their reflectivity η. *Reflectivity* is defined as the summation of the backscatter cross sections of the particles over a unit volume: $\eta \equiv \Sigma\sigma$. When the backscattering particles are spheres and are small enough with respect to wavelength so that the Rayleigh approximation can be used (that is, $\pi D/\lambda < 0.2$), the reflectivity is proportional to the radar reflectivity factor Z which is the summation over a unit volume of the sixth power of the particle diameters, $Z = \Sigma D^6$. Summation over a unit volume of Equation (19.35) gives

$$\Sigma\sigma = \eta = \frac{\pi^5 |K|^2 Z}{\lambda^4} \times 10^{-12} \ [\mathrm{cm}^{-1}] \qquad (19.37)$$

for Z in conventional units of mm^6 and m^{-3} and λ in centimeters.

When the particles are larger than Rayleigh size or composed of ice or water-ice mixtures, it is common practice to measure the radar reflectivity and express it in terms of an equivalent reflectivity factor Z_e. Substituting $|K|^2 = 0.93$ (water at normal atmosphere temperature, wavelength in the centimeter range) in Equation (19.38)

$$Z_e = 3.5 \times 10^9 \lambda^4 \eta \ (\mathrm{mm}^6\ \mathrm{m}^{-3}). \qquad (19.38)$$

Thus, Z_e is simply the ΣD^6 required to obtain the observed signal, if all the drops were acting as Rayleigh scatterers.

Because Z is a meteorological parameter that depends only on the particle size distribution and concentration, it is useful to express Z as a function of either the precipitation rate R or the mass of liquid water (or water equivalent of the ice content) M. The returned radar signal can then be related to Z, and through Z to R or M. Numerous Z-R and Z-M relations have been proposed. They vary geographically, seasonally, and by type of precipitation.

The following Z-R relations are typical of those most often found in the literature:

$Z = 200\ R^{1.6}$	widespread stratiform rain
$Z = 110\ R^{1.47}$	drizzle
$Z = 460\ R^{1.61}$	thunderstorm
$Z = 145\ R^{1.64}$	orographic
$Z = 314\ R^{1.42}$	monsoon

The scattering properties of snow are complicated by the many forms in which snow can occur, either as single ice crystals or aggregates of such crystals. The following relationships are reasonable averages of observations:

$$Z = 500\ R^{1.6} \qquad \text{for single crystals}$$

and

$$Z = 2000\ R^{2.0} \qquad \text{for aggregates}$$

where R is the snowfall rate in millimeters of water per hour.

Measurements by Boucher [1981] relating the reflectivity to the rate of snow accumulation suggest that a relation of the form

$$Z = A\ S^{2.0}$$

where S is the snowfall rate in millimeters of snow per hour, gives good agreement between measured radar reflectivity and measured snowfall accumulation. Boucher found that A varied between 6×10^{-3} to 2×10^{-2} when S was expressed in millimeters per hour. This variability is due to the variability of the density of snow, which ranges over an order of magnitude.

Clouds composed of water particles scatter very poorly at centimeter wavelengths due to the relatively small size of the water droplets. High-power radars operating at millimeter wave lengths can detect water clouds at short ranges. An empirical Z-M relation for water clouds is

$$Z = 0.048M^2$$

where M, the water content, is in grams per cubic meter.

CHAPTER 20

contain H_3O become the dominant species. More recent neutrals involved, only one negative ion composition meas-

Table 19-1 Attenuation (dB/km) at 293 K.

Precipitation Rate (mm/h)	Wavelength (cm)														
	0.03	0.05	0.1	0.15	0.2	0.25	0.3	0.5	0.8	1.0	2.0	3.0	5.0	6.0	15.0
0.25	0.867	0.900	0.874	0.773	0.656	0.539	0.434	0.179	0.0634	0.0381	0.685 $\times 10^{-2}$	0.231 $\times 10^{-2}$	0.657 $\times 10^{-3}$	0.434 $\times 10^{-3}$	0.031 $\times 10^{-4}$
1.25	2.31	2.43	2.51	2.41	2.22	1.99	1.74	0.919	0.374	0.232	0.0449	0.0134	0.304 $\times 10^{-2}$	0.191 $\times 10^{-2}$	0.249 $\times 10^{-3}$
2.50	3.71	3.71	3.90	3.85	3.63	3.34	3.01	1.77	0.783	0.497	0.104	0.0311	0.618 $\times 10^{-2}$	0.374 $\times 10^{-2}$	0.454 $\times 10^{-3}$
5.00	5.35	5.65	6.01	6.02	5.83	5.49	5.08	3.29	1.60	1.05	0.239	0.0750	0.0132	0.758 $\times 10^{-2}$	0.829 $\times 10^{-3}$
12.50	9.35	9.86	10.59	10.80	10.69	10.33	9.81	7.13	3.94	2.70	0.698	0.245	0.0399	0.0209	0.186 $\times 10^{-2}$
25.00	14.27	15.03	16.18	16.67	16.70	16.38	15.81	12.36	7.51	5.38	1.52	0.591	0.100	0.0488	0.348 $\times 10^{-2}$
50.00	21.78	22.90	24.68	25.61	25.89	25.70	25.14	20.89	13.87	10.37	3.23	1.38	0.265	0.124	0.661 $\times 10^{-2}$
100.00	33.22	34.85	37.55	39.18	39.84	39.96	39.50	34.54	24.83	19.40	6.66	3.09	0.706	0.338	0.0128
150.00	42.48	44.51	47.93	50.16	51.10	51.54	51.22	45.94	34.46	27.59	10.06	4.86	1.24	0.613	0.0190

Table 19-2 Temperature correction factor for representative rains.

Precipitation Rate mm/h	Wavelength cm	Temperature (K)				
		273	283	293	303	313
0.25	0.03	1.0	1.0	1.0	1.0	1.0
	0.1	0.99	0.99	1.0	1.01	1.02
	0.5	1.02	1.01	1.0	1.0	1.0
	1.25	1.09	1.02	1.0	1.0	0.99
	3.2	1.55	1.25	1.0	0.81	0.65
	10.0	1.72	1.29	1.0	0.79	0.64
2.5	0.03	1.0	1.0	1.0	1.0	1.0
	0.1	1.0	1.0	1.0	1.0	1.01
	0.5	1.01	1.01	1.0	0.99	0.98
	1.25	0.95	0.96	1.0	1.05	1.10
	3.2	1.28	1.14	1.0	0.86	0.72
	10.0	1.73	1.30	1.0	0.79	0.64
12.5	0.03	1.0	1.0	1.0	1.0	1.0
	0.1	1.0	1.0	1.0	1.0	1.01
	0.5	1.02	1.01	1.0	0.99	0.97
	1.25	0.96	0.97	1.0	1.04	1.07
	3.2	1.04	1.03	1.0	0.95	0.88
	10.0	1.74	1.30	1.0	0.79	0.63
50.0	0.03	1.0	1.0	1.0	1.0	1.0
	0.1	1.0	1.0	1.0	1.0	1.01
	0.5	1.02	1.01	1.0	0.98	0.97
	1.25	0.99	0.99	1.0	1.02	1.04
	3.2	0.91	0.96	1.0	1.01	1.01
	10.0	1.75	1.31	1.0	0.78	0.62
150.0	0.03	1.0	1.0	1.0	1.0	1.0
	0.1	1.0	1.0	1.0	1.0	1.01
	0.5	1.03	1.01	1.0	0.98	0.97
	1.25	1.01	1.0	1.0	1.0	1.01
	3.2	0.88	0.95	1.0	1.04	1.06
	10.0	1.72	1.31	1.0	0.78	0.62

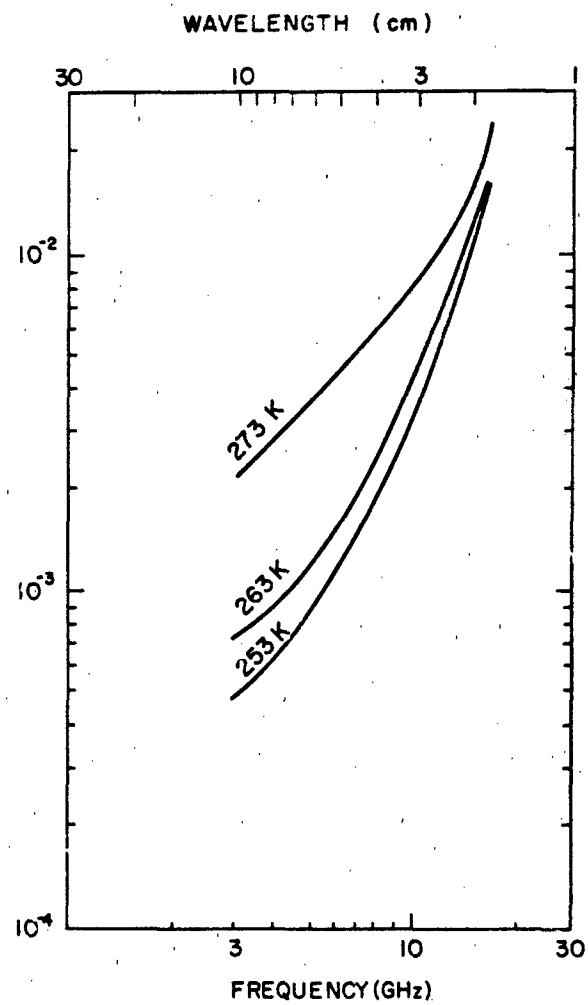


Figure 19-19. Theoretical attenuation due to snowfall of 10 mm of water content per hour as a function of wavelength and temperature.

19.2.3 Attentuation by Precipitation

Attenuation by rain is a function of drop size distribution, temperature and wavelength. Theoretical computations [Dyer and Falcone, 1972] indicate that for a given rainfall rate, wavelength and temperature, the variations in drop size distribution can cause deviations in average attenuation of between 4% and 33%. For comparison, measured attenuations are accurate to no more than ± 20% in general.

Table 19-1 gives the theoretical attenuation, for a wide range of rainfall rates and wavelengths, assuming a constant temperature of 293 K and an exponential distribution of drop sizes. Table 19-2 gives the attenuation correction factor for a range of rain rates and temperatures.

Figure 19-19 shows the theoretical maximum attenuation coefficients assuming a maximum snowfall rate 10 mm of water per hour. Because snowfall rates seldom exceed 3 mm of water per hour, attenuation due to snow should generally be one-third or less the value.

19.2.4 Total Attenuation

In addition to attenuation caused by precipitation particles, microwave and millimeter wave transmission is affected by atmospheric gases, water vapor, and cloud particles. Table 19-3 gives the attenuation as a function of wavelength for molecular oxygen at 293 K, and Table 19-4 gives correction factors for temperature differing from 293 K. Table 19-5 does the same for water vapor attenuation. Computer programs have been written [see, for example, Falcone et al., 1979] to calculate the total attenuation at any frequency, for any input atmospheric condition. Figure 19-20 is an example of the output from one such program.

The solid line (curve A) shows the attenuation at stand-

Table 19-3. Attenuation due to molecular oxygen at a temperature of 293 K and a pressure of 1-atmosphere.

Wavelength (cm)	Attenuation (dB km)
10.0	6.5×10^{-3}
7.5	7.0×10^{-3}
3.2	7.2×10^{-3}
1.8	7.5×10^{-3}
1.5	8.5×10^{-3}
1.25	1.4×10^{-2}
0.8	7.5×10^{-2}
0.7	1.9×10^{0}

Table 19-4. Correction factors for oxygen attenuation.

Temperature (K)	Correction Factor*
293	$1.00\ P^2$
273	$1.19\ P^2$
253	$1.45\ P^2$
233	$1.78\ P^2$

*P is pressure in atmospheres.

Table 19-5. Water-vapor attenuation in dB per kilometer.*

Wavelength (cm)	Temperature			
	293 K	273 K	253 K	233 K
10	$7\ PW \times 10^{-5}$	$8\ PW \times 10^{-5}$	$9\ PW \times 10^{-5}$	$10\ PW \times 10^{-5}$
5.7	$2.4\ PW \times 10^{-4}$	$2.7\ PW \times 10^{-4}$	$3.0\ PW \times 10^{-4}$	$3.4\ PW \times 10^{-4}$
3.2	$7\ PW \times 10^{-4}$	$8\ PW \times 10^{-4}$	$9\ PW \times 10^{-4}$	$10\ PW \times 10^{-4}$
1.8	$4.3\ PW \times 10^{-3}$	$4.8\ PW \times 10^{-3}$	$5.0\ PW \times 10^{-3}$	$5.4\ PW \times 10^{-3}$
1.24	$2.2\ PW \times 10^{-2}$	$2.33\ PW \times 10^{-2}$	$2.46\ PW \times 10^{-2}$	$2.61\ PW \times 10^{-2}$
0.9	$9.5\ PW \times 10^{-3}$	$1.04\ PW \times 10^{-2}$	$1.14\ PW \times 10^{-2}$	$1.26\ PW \times 10^{-2}$

*P is pressure in atmospheres and W is water-vapor content in grams per cubic meter.

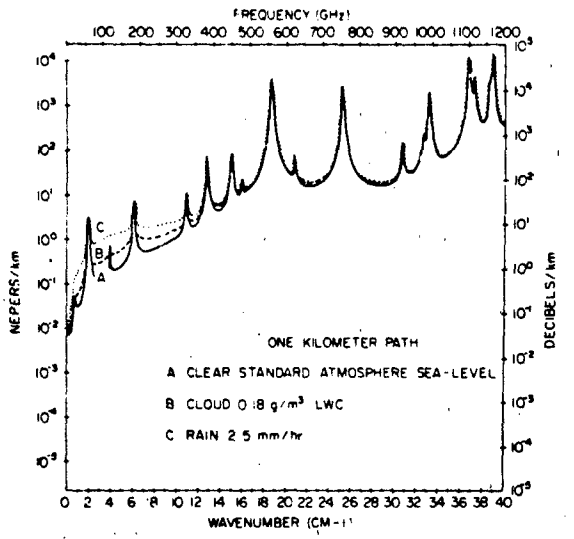


Figure 19-20. Attenuation vs frequency for a clear standard atmosphere, a cloud of 0.18g m^{-3} LWC, and a 2.5 mm/h rain rate for a 1 km horizontal path.

ard sea level conditions on a clear day. The peaks in the curve correspond to absorption lines for oxygen and water vapor. In general, there is a steady increase in attenuation with increasing frequency. Absorption because of atmospheric gases is negligible at centimeter wave lengths (below 30 GHz). The dashed line (curve B) shows the attenuation caused by cloud with a liquid water content of 0.18 g m^{-3}. This is a typical value for non-precipitating stratus or altostratus clouds. The dotted line (curve C) shows the attenuation caused by rain with an intensity of 2.5 mm/hr, assuming a simple exponential drop size distribution. This is a moderate rainfall typical of stratiform situations.

Clouds and precipitation are important factors affecting attenuation of frequencies below 300 GHz. The cloud liquid water content and the rainfall rate are critical parameters to be entered into the computation. Although Figure 19-20 can be useful in giving rough estimates of the attenuation, it is necessary to compute the attenuations for the desired range of atmospheric conditions at the electromagnetic wavelength of interest when predicting the performance of a system.

REFERENCES

Atlas, D., K.R. Hardy, J. Joss, "Radar Reflectivity of Storms Containing Spongy Hail," *J. Geophys. Res.*, **69:** 1955, 1964.

Battan, L.J., *Radar Meteorology*, The University of Chicago Press, p. 33, 1959.

Bean, B.R. and E.J. Dutton, *Radio Meteorology*, Dover, Mineola, New York, 1968.

Boucher, R.J., "Determination by Radar Reflectivity of Short-Term Snowfall Rates During a Snowstorm and Total Storm Snowfall," *Proceedings 20th Conference on Radar Meteorology, Boston, Nov 30-Dec 3, 1981*, American Meteorology Society, Boston, Mass., 1981.

Brown, J.H., R.F. Good, P.M. Bench, and G. Fourher, "Sonde Experiments for Comparative Measurements of Optical Turbulence," AFGL GR-82-0079, ADA118740, 1982.

Crane, R., "Monostatic and Bistatic Scattering from Thin Turbulent Layers in the Atmosphere," Lincoln Lab. Tech. Note 1968-34, ESD-TR-68267, 1968.

Crane, R.K., "A Review of Radar Observations of Turbulence in the Lower Statosphere," *Radio Sci.*, **15:** 177, 1980.

Cunningham, R.M., "Cumulus Climatology and Refractive Index Studies II," *Geophys. Res. Papers*, No. 51, AFCRL, 1962.

Dewan, E.M., "Optical Turbulence Forecasting: A Tutorial," AFGL TR-80-0030, ADA086863, 1980.

Dyer, R.M. and V.J. Falcone, "Variability in Rainfall Rate-Attenuation Relations," Prep, *15 Radar Meteorology Conference*, p. 353, 1972.

Edlen, B., "Dispersion of Standard Air," *J. Opt. Soc. Am.*, **43:**339, 1953.

Falcone, V.J., L.W. Abreu, and E.F. Shettle, "Atmospheric Attenuation of Millimeter and Submillimeter Waves: Model and Computer Code," AFGL TR-79-0253, ADA084485 1979.

Gage, K.S., T.E. VanZandt, and J.L. Green, "Vertical Profiles of Cn^2 in The Free Atmosphere: Comparison of Model Calculations with Radar Observations," *18th Conference on Radar Meteorology, Atlanta, GA, 28-31 March 1978*, American Meteorology Society, Boston, Mass., 1978.

Good, R.E., B. Watkins, A. Quesada, J.H. Brown, and G. Lariot, "Radar and Optical Measurements of Cn^2," *App. Optics*, **21:**2929, 1982.

Gunn, K.L.S. and T.W.R. East, "The Microwave Properties of Precipitation Particles," *Quart. J. Roy. Meteorol. Soc.*, **80:**522, 1954.

Hardy, K.R. and I. Katz, "Probing the Clear Atmosphere with High Power High Resolution Radars," *PROC. IEEE*, **57:**468, 1969.

Hufnagle, R.E. and N.R. Stanley, "Modulation Transfer Function Associated with Image Transmission Through Turbulent Media," *J. Opt. Soc. Am.*, **54:** 52, 1964.

Hufnagle, R.E., "Variations of Atmospheric Turbulence," *Digest of Technical Papers*, Topical Meeting on Optical Propagation Through Turbulence July 9-11, 1974. Optical Society of America, Washington, D.C., 1974.

Liebe, H.J., *Atmospheric Water Vapor* p. 143–202, edited by A. Deepak, T.D. Wilkinson, and A.L. Schemeltekopf, Academic Press, New York, 1980.

Medhurst, R.G., "Rainfall Attenuation of Centimeter Waves: Comparison of Theory and Measurement," *IEEE Trans. Antennas and Propagation*, **AP-13:**550, 1965.

Ottersten, H., "Radar Backscatter from the Turbulent Clear Atmosphere," *Radio Sci.*, **4:**1251, 1969.

Ray, P.S., "Broadband Complex Refractive Indices of Ice and Water," *Appl. Opt.*, **11:** 1836–1844, 1972.

Staras, H. and A.D. Wheelon, "Theoretical Research on Tropospheric Scatter Propagation in the United States 1954–1957," *IEEE Trans. Antennas and Propagation*, **AP-7:**80, 1959.

Strohnbehn, J.W., "Line of Sight Wave Propagation Through the Turbulent Atmosphere," *PROC. IEEE*, **56:**1301, 1968.

Tatarski, V.I., *Wave Propagation in a Turbulent Media*, Chapters 3, 6 and 7, McGraw-Hill, New York, 1961.

Tatarski, V.I., *The Effects of the Turbulent Atmosphere on Wave Propagation*, National Science Foundation, TT-68-50464, 1971.

VanZandt, T.E., K.S. Gage, and J.M. Warnock, "An Improved Model for the Calculation of Profiles of C_n^2 and E in the Free Atmosphere from Background Profiles of Wind, Temperature and Humidity," *Proceedings 20th Conference on Radar Meteorology, Boston, Mass., Nov 30–Dec 3, 1981*, American Meteorology Society, Boston, Mass., 1981.

Chapter 20

ATMOSPHERIC ELECTRICITY

Section 20.1 R.C. Sagalyn and H.K. Burke
Section 20.2 D.R. Fitzgerald

20.1 FAIR WEATHER ELECTRICITY

Electric fields, currents, and conductivities as well as positive and negative ions of greatly varying size and composition constitute the principal electrical properties of the atmosphere in fair weather. Air mass motions, pressure systems, winds, turbulence, temperature, and water vapor distributions have an important influence on the electrical properties in the troposphere through their control over the distributions of charged and uncharged aerosols and radioactive particles of terrestrial origin. These influences are greatest in the atmospheric exchange layer which is generally restricted to 2.5 km above the earth's surface. In the altitude region 30 to 90 km there is a transition from classical concepts of atmospheric electricity to the phenomena of ionospheric physics chiefly because of the changing atmospheric composition and increasing mean free path resulting in an increase in the concentration of free electrons. Recent studies have shown that the ionosphere, with its relatively high conductivity, can no longer be regarded as the upper bound for atmospheric electrical processes. The exact nature of electrical coupling to the ionospheric and magnetospheric regions is currently under investigation.

20.1.1 Electrical Conductivity

The lower atmosphere is a slightly conducting medium due to the presence of positive and negative ions. The principal sources of ionization include (1) cosmic radiation, (2) radiation from radioactive substances in soil, and (3) radioactive gases produced by the decay of (2). Among them (1) is mainly responsible for air conductivity at higher altitudes while (2) and (3) are dominant in the lower atmosphere. Table 20-1 shows average ion pair production rates due to radioactivity and cosmic radiation between 0 and 10 km.

The conductivity σ of air is defined as

$$\sigma = \sum_{i=1}^{m} n_i e_i k_i \qquad (20.1)$$

where m is the number of different ion species present, and n_i, e_i, and k_i are the number density, charge, and mobility of the i^{th} ion species respectively. Small ions because of their low mass have higher mobility than large ions. Over 95% of the total conductivity of the air is contributed by small ions so that

$$\sigma \approx n_+ e k_+ + n_- e k_- \qquad (20.2)$$

where n_+ and n_- are number densities of positive and negative small ions and k_+ and k_- are their mobilities. The mobility of an ion is its velocity per unit electric field strength. Near the earth's surface, $n_+ \approx 600\ cm^{-3}$, $n_- \approx 500\ cm^{-3}$, $k_+ \approx 1.3\ cm^2V^{-1}s^{-1}$ and $k^- \approx 1.6\ cm^2V^{-1}s^{-1}$ yielding a conductivity of about $2.5 \times 10^{-14}\ ohm^{-1}m^{-1}$.

The conductivity of air increases rapidly with altitude for the following reasons: (1) the mobility of ions increases as the neutral number density decreases (Figure 20-1), (2) the presence of numerous large aerosols lowers the conductivity of air near the ground and (3) the cosmic radiation causes an increase in the concentration of small ions with altitude (Table 20-1).

To a first approximation, ion equilibrium exists above the exchange layer (that is, above 2.5 km) so that $q = \alpha n_{\pm}^2$, where q is the net production rate of small ions and α is the volume recombination coefficient. If $k_+ \cong k_-$

Table 20-1. Average ion-pair production rates due to radioactivity and cosmic radiation as a function of altitude [Sagalyn and Fitzgerald, 1965].

Altitude (km)	Ion pairs ($cm^{-3}s^{-1}$)		
	radioactive	cosmic rays	Total
0	7.6	1.5	9.1
0.5	3.8	1.8	5.6
1	2.7	2.6	5.3
2	1.5	5.0	6.5
3	0.9	8.0	8.9
4	0.5	15.0	15.5
5	0.3	23.0	23.3
6	—	37.0	37.0
8	—	75.0	75.0
10	—	125.0	125.0

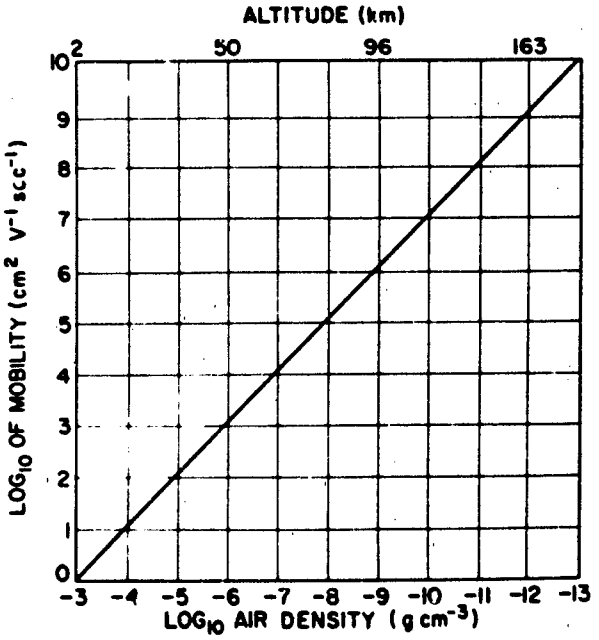


Figure 20-1. Mobility of small positive or negative ions vs air density or altitude, assuming k = 1.4 $cm^2 V^{-1} s^{-1}$ at 273 K and 1013.25 mb pressure.

= k is also assumed, the electrical conductivity can then be given by

$$\sigma = 2ek\sqrt{q/\alpha}\,. \qquad (20.3)$$

Figure 20-2 shows the variation of α with altitude. Figure 20-3 shows the computed variation of volume conductivity with altitude. Values of q are derived from cosmic ray ionization rates; values of k and α are taken from Figures 20-1 and 20-2, respectively.

Variations of electrical conductivity in the atmosphere as a function of altitude have been measured by a number of workers. See Figure 20-4. In this figure, curves 2 through 5 represent earlier experimental results. Curve 1 was calculated by Cole and Pierce [1965] based on variations of

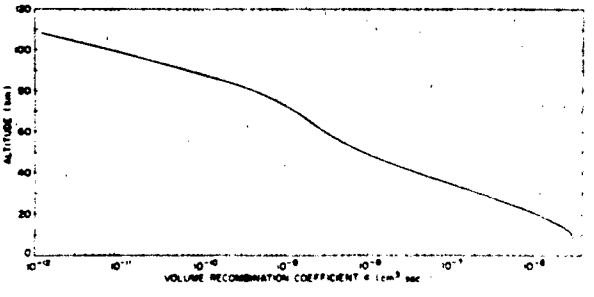


Figure 20-2. Volume recombination coefficient α vs altitude.

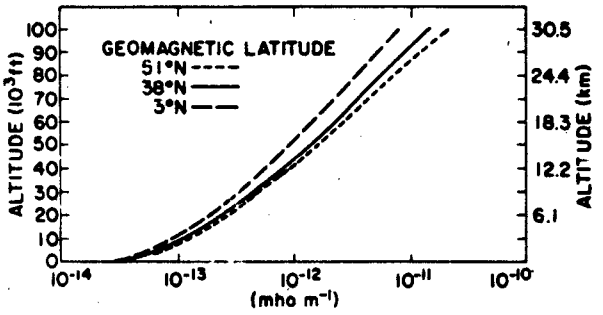


Figure 20-3. Total electrical volume conductivity vs altitude.

electron and ion concentrations and effective collisional frequencies with respect to altitude. The curve 1 values are higher than the others probably because Cole and Pierce neglected large ions. Curve 2 was compiled by Sagalyn [1965] from experimental results prior to 1963. Curves 3, 4, and 5 were obtained from measurements of Paltridge [1965], Mozer and Serlin [1969], and Morita et al. [1971], respectively. Curves 6, 7, and 8 show more recent measurement over Laramie, Wyoming; Hilo, Hawaii; and Sanriku, Japan, respectively. The Laramie measurements were obtained as part of an international effort on Atmospheric Electrical Measurements in 1978 [Rosen et al., 1980]. The 1978 measurements over Laramie are similar to results obtained at the same location in 1974 and 1975 [Takagi et al., 1980]. Positive and negative ion density measurements were combined and smoothed to provide the total conductivity (curve 6). Takagi et al. [1980] also published conductivity profiles for Sanriku, Japan (1973–1975) and Hilo, Hawaii (1975) (curves 7 and 8).

The differences in the conductivity profiles of Figure 20-4 are largely due to the latitude of measurement, envi-

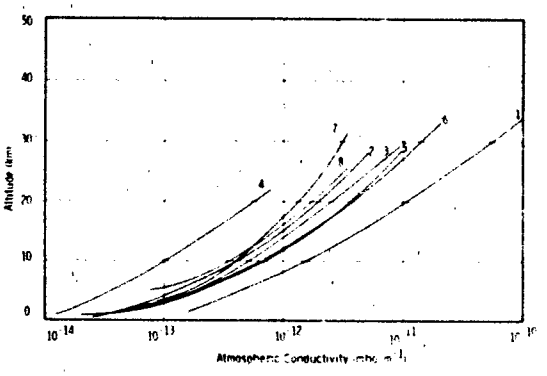


Figure 20-4. Various measurement results of electrical conductivity in the atmosphere as a function of altitude. The numbered curves are from (1) Cole and Pierce [1965], (2) Sagalyn [1965], (3) Paltridge [1965], (4) Mozer and Serlin [1969], (5) Morita et al. [1971], (6) Rosen et al. [1980] for Laramie, Wyoming, (7) Takagi et al. [1980] for Sanriku, Japan, and (8) Takagi et al. [1980] for Hilo, Hawaii.

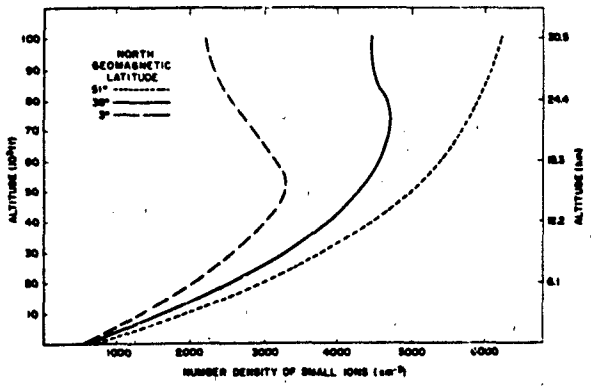


Figure 20-5. Number density of small positive or negative ions vs altitude computed from $q = \alpha n^2$. The variation with latitude illustrates the effect of cosmic rays on production rates.

ronmental conditions at the time of measurement, and the data base upon which the curves were derived. For example, curve 2 [Sagalyn, 1965] is based on the results of approximately 100 aircraft and balloon flights. A large percentage of the flights were conducted in the industrial northeastern part of the U.S. where significant aerosol concentrations exist throughout the troposphere.

The differences in the more recent measurements (curves 6 to 8) can probably be attributed to the latitude variation of cosmic ray intensities which in turn affect the small ion concentration. See, for example, Figure 20-5. Because the conductivity is roughly proportional to the square root of the small ion concentration (Equation (20.3)), the conductivity at Laramie (50° geomagnetic) is expected to be approximately twice as large as that at Sanriku (29° geomagnetic) and Hilo (20° geomagnetic). Figure 20-6 shows the variation with latitude of the average (base) electrical conductivity, the latitude variation following a solar flare, and that expected as a result of Forbush decrease [Hays and Roble, 1979].

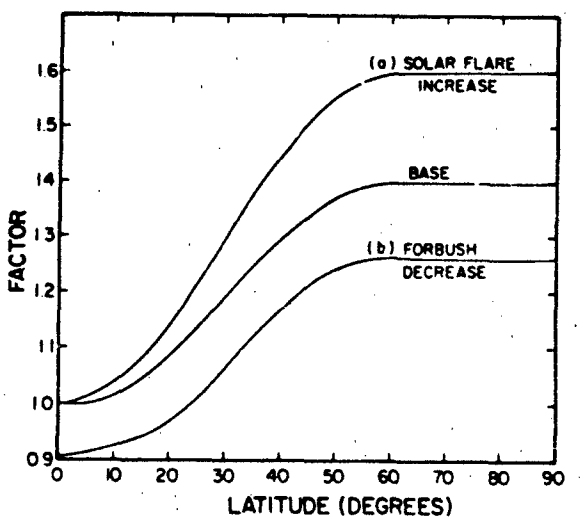


Figure 20-6. Relative variation of electrical conductivity σ/σ_0 with latitude for different environmental conditions: (a) solar flare, (b) Forbush Decrease [Hays and Roble, 1979]

The conductivity variation with altitude is approximated by the relation

$$\sigma(z) = \sigma_0 \exp(2kz), \qquad (20.4)$$

where $(2\ k)^{-1}$ is defined as the scale height of conductivity. It varies typically between 5 and 6 km. The conductivity at the surface, σ_0, varies between 1 and 3 $\times$ 10^{-14} ohm^{-1} m^{-1}.

20.1.2 Electric Field

In fair weather areas the horizontal components of the field are small compared to the vertical component so that the atmospheric electric field E can be defined as

$$E = -\frac{dV}{dZ}, \qquad (20.5)$$

where V is the electric potential and Z is the altitude. The atmospheric electric field is directed vertically, the potential gradient decreases with increasing altitude, and the field by the electrostatic definition is negative (field vector directed downward). Computed from the current density and mobility, Figure 20-7a shows the expected electric field as a function of altitude. Figure 20-7b shows the average value and variation of a measured electric field as a function of altitude [Sagalyn, 1965].

Large scale mappings of the atmospheric electric field which took into account the variational property of air conductivity were carried out by Holzer and Saxon [1952] and Anderson and Freier [1969]. These models do not extend into the ionosphere where the conductivity has a tensor

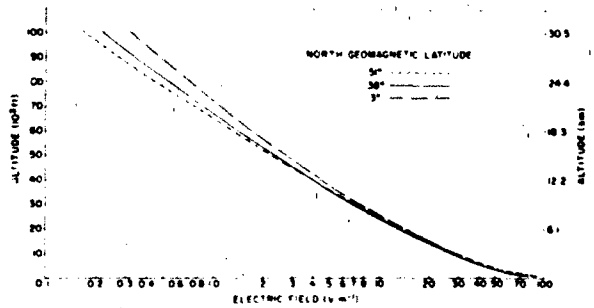


Figure 20-7a. Electric field vs altitude, assuming a conduction current density of 2.7 $\times$ 10^{-12} A m^{-1} and using values of α from Figure 20-2.

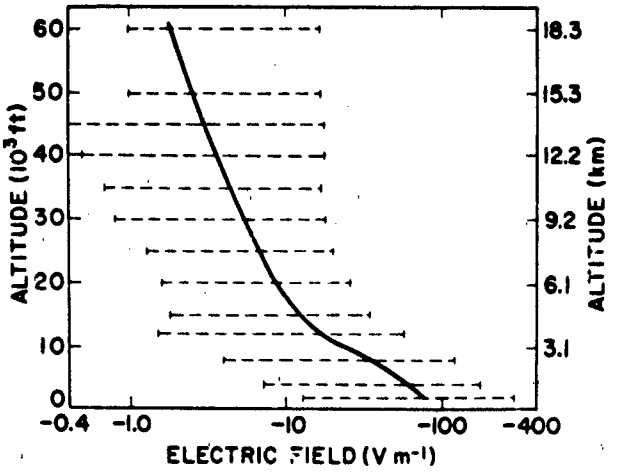


Figure 20-7b. Average value and maximum variation of electric field as a function of altitude in over 80 aircraft and balloon flights in the eastern U.S. [Sagalyn, 1965].

rather than a scalar form. Park and Dejnakarintra [1973] mapped thundercloud electric fields at middle and higher latitudes. The electrical conductivity was represented as an exponential function of altitude, and the anisotropy of air was taken into account above 70 km. The high-latitude magnetic field lines were assumed to be vertical below 150 km and numerical solutions were derived for both the atmosphere and ionosphere.

Assuming an exponential increase in conductivity with altitude, the basic equations for the atmospheric electric field can be expressed:

$$\sigma(z) = \sigma_0 \exp(2kz) \tag{20.4}$$

$$\nabla \times \mathbf{E} = 0, \text{ or } \mathbf{E} = -\nabla\Phi \tag{20.6}$$

where Φ is the electric potential with respect to the earth's surface.

The potential equation is

$$\left[\nabla^2 + 2k\frac{\partial}{\partial z}\right]\Phi = -f\epsilon_0 \tag{20.7}$$

where f is the known source distribution (ρ_{source}). In fair weather there is no charge source and the potential equation reduces to

$$\left[\nabla^2 + 2k\frac{\partial}{\partial z}\right]\Phi = 0. \tag{20.8}$$

Applying appropriate boundary conditions, the fair weather electric field components can be expressed as follows:

$$E_x = 0,$$

$$E_y = 0, \text{ and} \tag{20.9}$$

$$E_z = -\frac{V_0 \cdot 2k}{1 - e^{-2kH}},$$

where V_0 is the total potential difference between the top of the atmosphere and the earth's surface and H is the upper boundary of the atmosphere. In the solution for the fair weather electric field, only the vertical component exists.

Defining $E_0 \equiv -[V_0 \cdot 2k/(1 - e^{-2kH})]$ as the field intensity near the ground, the fair weather electric field can be rewritten as

$$E_z = E_0 e^{-2kz}. \tag{20.10}$$

The average observed values of E_0 over land is about 130 V/m in fair weather. Based on this value, the electric field at the 70 km level is 0.1 mV/m to 1 mV/m. This is the same order of magnitude of electric fields of magnetospheric origin. During disturbed weather conditions, much larger fields have been observed.

20.1.3 Currents and Space Charge

20.1.3.1 Air-Earth Conduction Current. As a result of the existence of predominately vertical electric field in the presence of positive and negative atmospheric ions, a current is constantly flowing to the earth's surface. This current is called the air-earth conduction current. The vertical conduction current density j is defined by

$$j = V/R_c = \sigma E. \tag{20.11}$$

The potential difference between the earth and the upper atmosphere conducting layers V is $-\int_0^z E\,dZ$. The resistance of a vertical air column per unit area R_c is $\int_0^z dZ/\sigma$. Conservation of charge and continuity considerations requires that the current along the column be nearly constant. The conduction current density is, therefore, the net vertical current due to positive ions flowing downward and negative ions flowing upward in the atmosphere. Measurements show that despite large variations in the electric field and atmospheric conductivity, the air-earth conduction current is comparatively stable (Table 20-2). The mean value of air-earth current density from annual observations is 2.3 pA (picoamps)/m^2 over continents and 3.3 pA/m^2 over oceans [Israël and Dolezalek 1973]. The zonal distribution given in Table 20-2 was derived by Hogg [1950]

The mean diurnal variation of the air-earth current density for 10 continental stations derived from data obtained over at least one year is given in Figure 20-8. The diurnal

Table 20-2. Zonal distributions of air-earth current densities

Latitude Zone	Mean Current Density (pA/m^2)
Arctic	3.0
Northern temperate zone	2.1
Tropics	2.6
Southern temperate zone	2.4
Antarctic	3.0

variations are given in percentages to facilitate comparisons [Israël and Dolezalek, 1973]. The early morning increase in current density is attributed to the "sunrise effect" also observed by Burke and Few [1978]. Local effects such as visibility, fog, and humidity can also reduce the current density. Figure 20-9 shows some recent direct measurements of air-earth current density at different altitudes and locations. There are other atmospheric currents such as the convection current that should be distinguished from the conduction current.

20.1.3.2 Convection Current. In regions where a space charge exists (excess of ions of one sign), movement of air produces a transfer of charge that can be defined as a convection current density j_2

$$j_2 = \rho v \approx A_{vr} \frac{\partial \rho}{\partial Z}, \qquad (20.12)$$

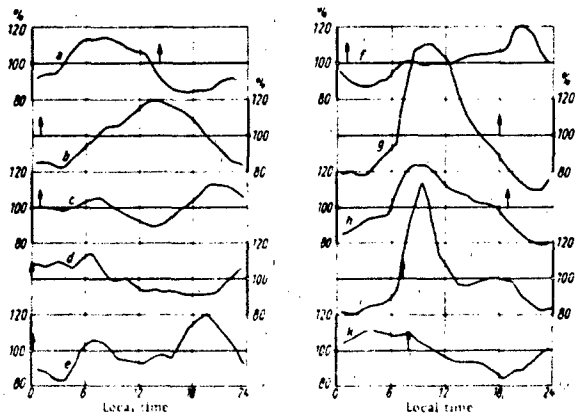


Figure 20-8. Mean diurnal variation of the air-earth current density (according to data over at least one year) at 10 stations in percentage representation. The stations have been arranged according to geographic latitude. The arrows on the time axis represent 0 h GMT: (a) Fairbanks, Alaska; (b) Uppsala, Sweden; (c) Potsdam, E. Germany; (d) Kew, England; (e) Chambon-la-Foret, France; (f) Buchau, Germany; (g) Tucson, Arizona; (h) Huancayo, Peru; (i) Bandung, Indonesia; and (k) Watheroo, W. Australia [Israel and Dolezalek, 1973].

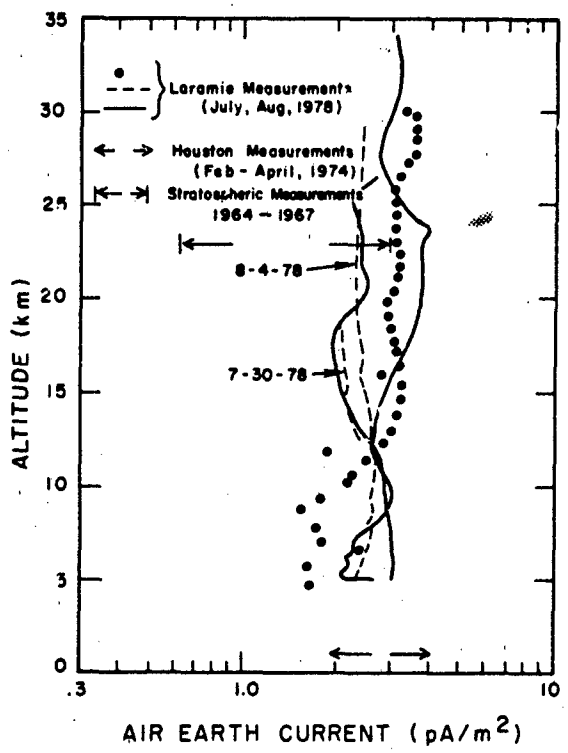


Figure 20-9. Direct measurements of the air-earth-current as reported by several different groups:
1978 Laramie Measurements [Rosen et al., 1980]
1974 Houston Measurements [Burke and Few, 1978]
1964–67 Stratospheric Measurements [Uchikawa, 1972]

where ρ is the space charge per unit volume, v is the vertical air velocity, and A_{vr} is the coefficient of eddy diffusion. Space charge densities and gradients large enough to significantly influence the vertical current occur primarily in the exchange (Austausch) layer, where considerable mixing occurs and greatly varying quantities of charged nuclei are introduced into the atmosphere. Throughout this chapter, nuclei refer to molecular aggregates that may or may not act as condensation nuclei depending on their chemical composition.

Aspinall [1972] showed that the conduction current and mechanical transfer currents, which include the convection current, are about the same order of magnitude in the exchange layer. Convection currents comparable to or greater than the conduction current have been observed below 0.015 km [Smith, 1958].

The space charge density ρ is related to the field and to the potential through Poisson's equation

$$\frac{dE}{dZ} = \frac{d^2V}{dZ^2} = -4\pi\rho \qquad (20.13)$$

in cgs electrostatic units.

20.1.3.3 Displacement Current. A displacement current exists when the electric field (potential gradient) changes with time.

$$j_{displacement} = \varepsilon_o \frac{dE}{dt}$$

Thus, when there exists a significant variation of the electric field with time, instantaneous measurements of the air-earth current will be made up of the sum of the displacement current and the conduction current.

20.1.3.4 "Electrode Effect." As a result of the earth's negative charge, negative ions drift upward near the surface of the earth and unless they are replaced by negative ions from radioactive substances in the ground they will leave a region of net positive charge near the surface. This phenomena is referred to as the "electrode effect." Hoppel [1969], comparing theories of the atmospheric electrode effect with the experimental observations of Crozier [1965], Gathman [1967], and Muhleisen [1961] concluded that within the first quarter of a meter above ground there is a region of positive space charge. Above that height, positive ions are balanced by an upward flow of negative ions from radioactive sources. The reverse of the normal electrode effect is sometimes observed on quiet nights due to the trapping of these radioactive ions. Over water or over the polar caps the electrode effect is observed to extend to higher altitudes due to the absence of ions from radioactive sources.

20.1.3.5 Earth Charge and Worldwide Current System. The total conduction current flowing to the earth at any given time is approximately 1800 amperes. As a result of this current, the bound negative charge of the earth (about 500 000 C) would be neutralized in less than half an hour. The net charge, however, remains nearly constant. Thunderstorms and lightning shower clouds are the most likely mechanisms for the maintenance of the earth's charge. Stergis [1957] found that thunderstorm and shower clouds act as generators, driving current upward in the reverse direction to the current flow in fair weather areas. Lightning discharges are also found to carry negative charges to earth in large amounts for short periods of time. It is estimated that a large portion of the 1800 ampere global conduction current is balanced by the charge of opposite sign transferred by the lightning activity of thunderstorms throughout the world.

It should be noted that the potential gradient in undisturbed fair weather areas has the same diurnal variation over all the earth when referred to universal time [Mauchly, 1923]. This diurnal variation is in phase with the diurnal variation of worldwide thunderstorm activity. The potential gradient in undisturbed areas is proportional to the total potential difference between the earth and ionosphere. The magnitude of this potential difference is 275 ± 50 kV.

20.1.4 Atmospheric Ions

20.1.4.1 Definition and Relations. Atmospheric ions are divided into four main groups according to their size and mobilities: small positive and negative ions, n_+ and n_-, and large positive and negative ions (also referred to as charged nuclei), N_1 and N_2. Table 20-3 gives the range of mobilities and radii of atmospheric ions.

The continuity equations for the production and destruction of ions may be written

$$\frac{dn_+}{dt} = q_1 - \alpha n_+ n_- - \eta_{12} n_+ n_- - \eta_{10} n_+ N_o, \quad (20.14)$$

$$\frac{dn_-}{dt} = q_2 - \alpha n_+ n_- - \eta_{21} n_- N_1 - \eta_{20} n_- N_o, \quad (20.15)$$

$$\frac{dN_1}{dt} = Q_1 - \eta_{10} n_+ N_o - \eta_{21} n_- N_1 - \gamma N_1 N_2, \quad (20.16)$$

$$\frac{dN_2}{dt} = Q_2 - \eta n_- N_o - \eta_{12} n_+ N_2 - \gamma N_1 N_2, \quad (20.17)$$

where q_1 and q_2 are net production rates of small ions due to radioactive emanations from the earth's surface, cosmic radiation and diffusion, and Q_1 and Q_2 are net production rates of large ions due to ionizing sources and diffusion ($q_1 \approx q_2$ and $Q_1 \approx Q_2$). N_o is the concentration of neutral nuclei. The volume recombination coefficient for small ions is, in $cm^3 s^{-1}$,

$$\alpha = 1.75 \times 10^{-6} \left(\frac{273}{T}\right)^{2.5} (2M)^{-1.2} f(Y), \quad (20.18)$$

Table 20-3. Mobility and size range of atmospheric ions.

Ions	Mobility ($cm^2 s^{-1} V^{-1}$)	Radius (cm)
Small	1 to 2	6.6×10^{-8} to 7.8×10^{-7}
Average large	10^{-3} to 10^{-2}	7.8×10^{-7} to 2.5×10^{-6}
Langevin	2.5×10^{-4} to 10^{-3}	2.5×10^{-6} to 5.7×10^{-6}
Ultra large	$< 2.5 \times 10^{-4}$	$> 5.7 \times 10^{-6}$

where T is the temperature in K, M is the molecular weight of the ions in atomic mass units, and f(Y) is the probability function.

$$f(Y) = 1 - 4Y^{-4}\left[1 - (Y + 1)\,e^{-Y}\right]^2, \quad (20.19)$$

and

$$Y = 0.81\left(\frac{273}{T}\right)^2\left(\frac{P}{760}\right)\left(\frac{L_A}{L}\right). \quad (20.20)$$

P is the pressure in mm H_g and L_A/L is the ratio of the mean free path of a molecule to that of an ion at normal temperature and pressure; for air, $L_A/L = 3$. Table 20-4 gives values of f(Y) for Y between 0 and 2.5. The variation of α with altitude was shown previously in Figure 20-2.

The attachment coefficients are η_{12} for collisions between small positive and large negative ions, η_{21} for collisions between small negative and large positive ions, η_{10} for collisions between small positive ion and a neutral nucleus, and η_{20} for collisions between a small negative ion and a neutral nucleus. Bricard [1949] derived expressions for these attachment coefficients as a function of the charge and radius of the nuclei and mobility and diffusion coefficient of the small ion. The diffusion coefficients of small positive and negative ions are approximately equal, therefore

$$\eta_{12} \approx \eta_{21} = 4\,\pi D\,a/I\,(\zeta, p), \quad (20.21)$$

$$\eta_{20} \approx \eta_{10} = 4\,\pi D\,a/I\,(\zeta, o), \quad (20.22)$$

where D is the average value of the small diffusion coefficient, a is the radius of the interacting charged or uncharged nucleus, and I (ζ, p) is a dimensionless parameter that contains the dependence of the attachment coefficient on the radius and charge of the nucleus.

$$I(\zeta, p) = \int_1^\infty X^{-2} \exp\left[-\zeta\left(\frac{p}{X} + \frac{1}{2X^2(X^2 - 1)}\right)\right] dX \quad (20.23)$$

where p is the number of elementary charges on the large ion, $\zeta = ek/Da$ in cgs electrostatic units, or $\zeta = ek/4\pi\varepsilon_o Da$ in mks units (k is small-ion mobility), and X = r/a, where r is the radius of a sphere of influence centered on the nucleus. Figure 20-10 gives values of I (ζ, p).

In Equations (20.16) and (20.17), γ is the combination

Table 20-4. Probability function f (Y) from Equation (20.20).

Y	0	1	2	3	4	5	6	7	8	9
0.03	0.03916	0.04044	0.04172	0.04299	0.04426	0.04553	0.04680	0.04806	0.04933	0.05059
0.04	0.05185	0.05311	0.05437	0.05562	0.05688	0.05813	0.05938	0.06063	0.06187	0.06312
0.05	0.06436	0.06560	0.06684	0.6808	0.06932	0.07055	0.07179	0.07302	0.07425	0.07547
0.06	0.07670	0.07792	0.07914	0.08037	0.08158	0.08280	0.08402	0.08523	0.08644	0.08765
0.07	0.08886	0.09007	0.09127	0.09248	0.09368	0.09488	0.09608	0.09727	0.09847	0.09966
0.08	0.1009	0.1020	0.1032	0.1044	0.1056	0.1068	0.1080	0.1091	0.1103	0.1115
0.09	0.1127	0.1138	0.1150	0.1162	0.1174	0.1185	0.1197	0.1209	0.1220	0.1232
0.1	0.1243	0.1358	0.1472	0.1583	0.1694	0.1802	0.1910	0.2015	0.2119	0.2222
0.2	0.2324	0.2423	0.2522	0.2619	0.2715	0.2809	0.2903	0.2994	0.3085	0.3175
0.3	0.3263	0.3350	0.3435	0.3520	0.3603	0.3686	0.3767	0.3847	0.3926	0.4004
0.4	0.4080	0.4156	0.4231	0.4304	0.4377	0.4449	0.4519	0.4589	0.4658	0.4726
0.5	0.4793	0.4859	0.4924	0.4988	0.5051	0.5114	0.5175	0.5236	0.5296	0.5355
0.6	0.5414	0.5471	0.5528	0.5584	0.5639	0.5694	0.5748	0.5801	0.5853	0.5905
0.7	0.5956	0.6006	0.6056	0.6105	0.6153	0.6201	0.6248	0.6294	0.6340	0.6385
0.8	0.6430	0.6474	0.6517	0.6560	0.6602	0.6644	0.6685	0.6726	0.6766	0.6805
0.9	0.6844	0.6883	0.6921	0.6958	0.6995	0.7032	0.7068	0.7103	0.7138	0.7173
1.0	0.7207	0.7241	0.7274	0.7307	0.7339	0.7371	0.7403	0.7434	0.7465	0.7495
1.1	0.7525	0.7555	0.7584	0.7613	0.7641	0.7669	0.7697	0.7724	0.7751	0.7778
1.2	0.7804	0.7830	0.7856	0.7881	0.7906	0.7931	0.7955	0.7979	0.8003	0.8027
1.3	0.8050	0.8072	0.8095	0.8117	0.8139	0.8161	0.8182	0.8203	0.8224	0.8245
1.4	0.8265	0.8285	0.8305	0.8325	0.8344	0.8363	0.8382	0.8401	0.8419	0.8437
1.5	0.8455	0.8473	0.8490	0.8508	0.8525	0.8541	0.8558	0.8574	0.8591	0.8607
1.6	0.8623	0.8638	0.8654	0.8669	0.8684	0.8699	0.8713	0.8728	0.8742	0.8756
1.7	0.8770	0.8784	0.8796	0.8811	0.8824	0.8837	0.8850	0.8863	0.8876	0.8888
1.8	0.8901	0.8913	0.8925	0.8937	0.8948	0.8960	0.8971	0.8983	0.8994	0.9005
1.9	0.9016	0.9027	0.9037	0.9048	0.9058	0.9068	0.9079	0.9089	0.9098	0.9108
2.0	0.9118	0.9127	0.9137	0.9146	0.9154	0.9165	0.9174	0.9182	0.9191	0.9200
2.1	0.9208	0.9217	0.9225	0.9233	0.9242	0.9250	0.9258	0.9266	0.9273	0.9281
2.2	0.9289	0.9296	0.9304	0.9311	0.9318	0.9325	0.9332	0.9339	0.9346	0.9353
2.3	0.9360	0.9367	0.9373	0.9380	0.9386	0.9393	0.9399	0.9405	0.9411	0.9417
2.4	0.9423	0.9429	0.9435	0.9441	0.9447	0.9452	0.9458	0.9464	0.9469	0.9475

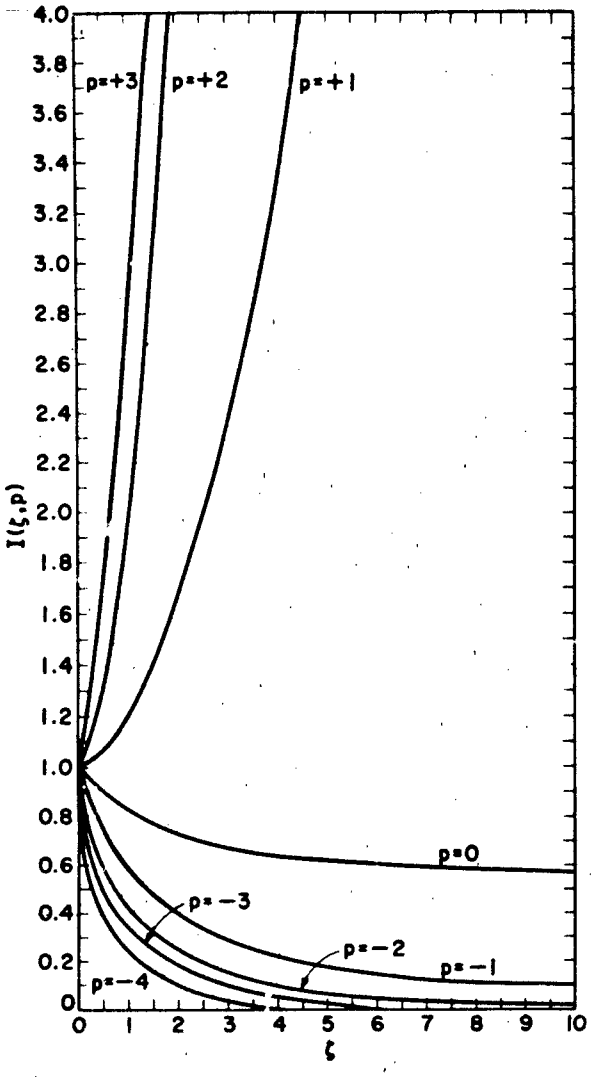


Figure 20-10. Parameter I (ζ,p) for Equations (20.21) and (20.22); p is the number of positive or negative elementary charges on the large ion; and ζ is a dimensionless parameter.

coefficient for collisions between large ions of opposite sign with a value of approximately 10^{-9} cm^3s^{-1}. This is three orders of magnitude less than the value of the attachment coefficients so that terms containing γ, therefore, can usually be neglected.

20.1.4.2 Positive Ion Chemistry and Composition in the Stratosphere and Troposphere. A positive ion reaction scheme for the stratosphere and troposphere is given in Figure 20-11a and b [Ferguson, 1979]. It is essentially the O_2^+ reaction sequence of the D-region positive ion chemistry augmented by reactions that involve some of the minor constituents of the lower atmosphere. See also Chapter 21.

Variations of typical ionization production rates in the troposphere are shown in Figure 20-12. Values of less than

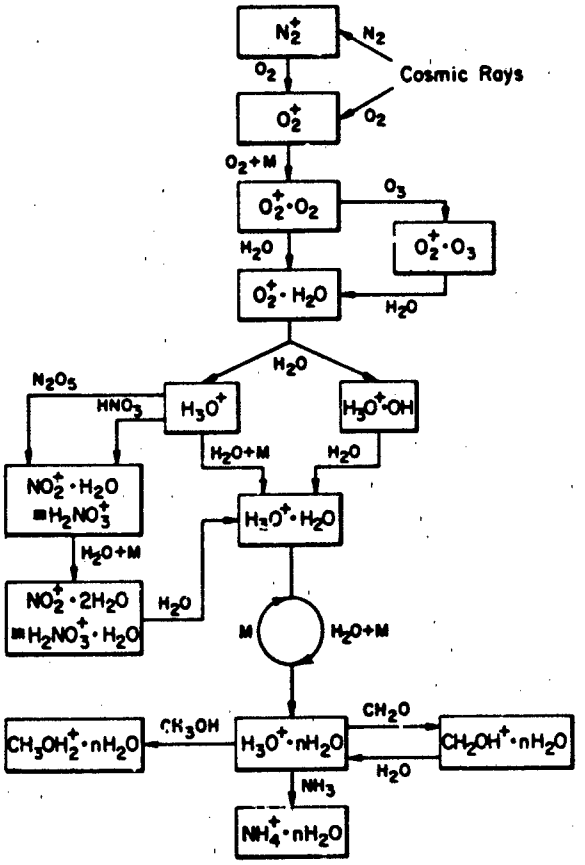


Figure 20-11a. Stratospheric positive ion chemistry [Ferguson et al., 1979].

1 to more than 20 ion pairs $cm^{-3}s^{-1}$ occur in the stratosphere. The ambient concentration of positive ions is determined by the recombination rate with negative ions. The ion concentration in the lower stratosphere has been found to have a value of about 5×10^3 cm^{-3}. Positive ion lifetimes are on the order of a few thousand seconds.

As shown in Figure 20-11a and b, the initial products of the ionization below 60 km are predominantly N_2^+ and O_2^+, with lesser amounts of O^+ and N^+. These ions are rapidly converted to O_2^+, as well as an inconsequential amount of NO^+, by well-established reactions. Once formed, the O_2^+ ions associate with O_2

$$O_2^+ + O_2 + M \rightleftarrows O_2 \cdot O_2^+ + M. \quad (20.24)$$

The formation of $O_2^+ \cdot O_2$ begins a series of fast switching reactions (see Figure 20-11b) that involve H_2O and leads to the formation of the water cluster ions, $H_3O^+ \cdot nH_2O$.

In the troposphere, where the H_2O mixing ratio is about 10^{-2}, the conversion of $O_2^+ \cdot O_2$ to $O_2^+ \cdot H_2O$ proceeds so rapidly that the reacting rates outlined must lead to the water cluster ions. However, in the stratosphere, where the H_2O mixing ratio is of the order of 10^{-6}, other neutral con-

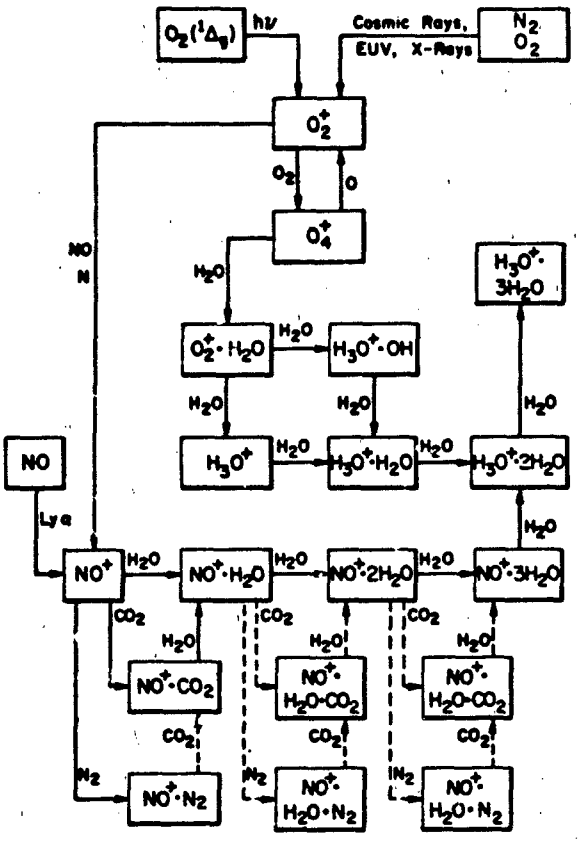


Figure 20-11b. D-region positive ion chemistry [Ferguson et al., 1979].

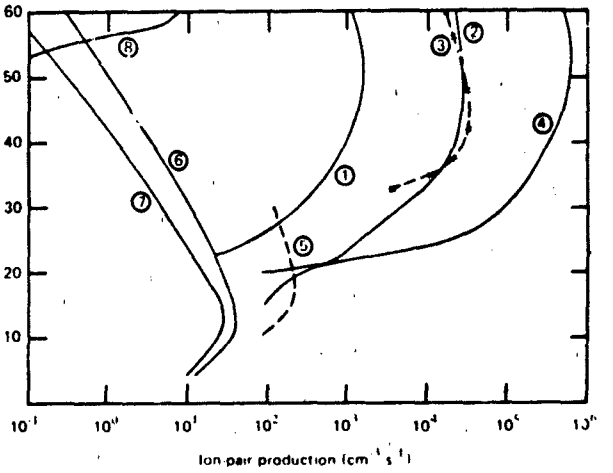


Figure 20-12. Ion-pair production rates due to solar protons: (1) Polar cap absorption (PCA), 11/2/69; (2) PCA, 8/4/72; 1500–1600 UT; (3) PCA, 8/4/72, 1508 UT; (4) PCA, 8/4/72, 2200 UT; (5) PCA 9/29/61; (6) SSMIN (sunspot minimum) galactic cosmic rays; (7) SSMAX (sunspot maximum) galactic cosmic rays; and (8) precipitating electrons in a hard aurora [Herman and Goldberg, 1978a].

stituents have comparable abundances. CO_2 has a much larger concentration and O_3 and CH_4 concentrations are comparable with that of H_2O. If $O_2^+ \cdot O_2$ reacts with any of these neutrals, then the ion chemistry outlined in Figure 20-11 might be significantly altered.

These $O_2^+ \cdot O_2$ reactions have been examined and the results are given in Table 20-5. The first entry is the fast reaction against which the alternative paths must compete. In contrast, the reaction of $O_2^+ \cdot O_2$ with CH_4 is very slow. The reactions of $O_2^+ \cdot O_2$ with CO_2 and O_3 were studied in equilibrium and the thermochemical constants are given. However, the exothermicity for the reaction involving O_3 is much larger; therefore, the $O_2^+ \cdot CO_2$ and $O_2^+ \cdot O_3$ ion concentrations will be more comparable. The $O_2^+ \cdot O_3$ concentrations are estimated to never drop below one tenth of the $O_2^+ \cdot O_2$ concentration and even equal it at 30 km. This means that the $O_2^+ \cdot O_3$ chemistry must also be considered.

Recently, the first stratospheric ion composition measurements were made in rocket-borne mass-spectrometer flights. Three high-latitude studies sampled the positive ions during descent through the altitude range 55 to 35 km.

Above 45 km, the substantial signal levels permitted high resolution identification of the water cluster ions $H_3O^+ \cdot nH_2O$ with n = 9, 1, 2, and 3 as the dominant species in this region. A rather sharp transition at about 45 km was reported from the water cluster ions to ions with e/m ratios of 29 ± 2, 42 ± 2, 60 ± 2, and 80 ± 2. In the region below this altitude, the cluster ions that do not

Table 20-5. $O_2^+ \cdot O_2$ reactions with H_2O, CH_4, CO_2, and O_3 (300 K).

Reactions	Result[a]	Source
1. $O_2^+ \cdot O_2 + H_2O \rightarrow O_2^+ \cdot H_2O + O_2$	$k = 1.5(-9)$ cm^3 s^{-1}	Howard et al. (1972)
2. $O_2^+ \cdot O_2 + CH_4 \rightarrow$ products	$k < 3(-12)$ cm^3 s^{-1}	Dotan et al. (1978)
3. $O_2^+ \cdot O_2 + CO_2 \rightleftarrows O_2^+ \cdot CO_2 + O_2$	$\Delta H = 0.3 \pm 1.0$ kcal mole^{-1} $\Delta S = 4.3 \pm 2.6$ cal mole^{-1}K^{-1}	Dotan et al. (1978)
4. $O_2^+ \cdot O_2 + O_3 \rightleftarrows O_2 \cdot O_2^+ \cdot O_3 + O_2$	$\Delta H = -3.7 \pm 1.0$ kcal mole^{-1} $\Delta S = 4.5 \pm 2.6$ cal mole^{-1} K^{-1}	Dotan et al. (1978)

[a] 1.5(−9) implies 1.5×10^{-9}.

contain H_2O become the dominant species. More recent balloon flights indicate that water cluster ions are the major species at 35 km. Other ions were observed in appreciable concentration, in particular ions with an e/m ratio of 96 ± 2.

The observation of water cluster ions as the dominant species down to altitudes of 45 km or below, have been formed undoubtedly by the chemistry described in Figure 20-11a. Those observations place upper limits on the neutral molecules that react rapidly with $H_3O^+ \cdot nH_2O$ ions.

There is a clear need for more detailed measurements of stratospheric and tropospheric positive ions with instruments of higher mass resolution. There is also a need for measurements of the critical neutral constituents for the ion chemistry, such as NaOH.

20.1.4.3 Negative Ion Chemistry and Composition in the Stratosphere and Troposphere. The negative ion chemistry of the stratosphere and troposphere is even more speculative than the positive ion chemistry of these regions. In addition to there being few measurements of the trace neutrals involved, only one negative ion composition measurement has so far been reported. Furthermore, the D-region measurements from which one could could draw guidance are relatively sparse and somewhat ambiguous at present. Thus, our present understanding of the negative ion processes in the stratosphere stems largely from laboratory studies.

The approach of Ferguson [1979] has been to start with the D-region negative ion chemistry given in Figure 20-13 and modify it to be in accordance with the expected differences in neutral composition of the two regions. Then one must look first for possible changes in the chemistry that leads to the NO_3^- ions and then consider whether these ions would react with trace neutral species that are suspected in the lower atmosphere to form even more stable ions.

The first necessary modification of the D-region scheme is to disregard the reactions of atomic oxygen, whose concentration below 50 km is negligible in comparison to that of O_3. In addition, NO will no longer play a role. Furthermore, the rapid formation of cluster ions can strongly alter the evolution of ion chemistry. For example, the reaction of O_3^- with CO_2 is known to decrease rapidly with O_3^- hydration. Because H_2O bonds more strongly to O_3^- than to CO_3^- this reaction may become endothermic when O_3^- becomes heavily hydrated. In the lower atmosphere, O_3^- hydration is likely to occur before the reaction of O_3^- with CO_2.

The reaction scheme based on these considerations is shown in Figure 20-14. The dashed lines represent places of considerable uncertainty. For example, it is not clear whether $O_2^- \cdot nH_2O$ ions will react with O_3, as do the unclustered O_2^- ions. The situation with $O_3^- \cdot nH_2O$, O_3^-, and CO_2 is the same; namely, the hydrated ions may not follow the same reaction paths as the unhydrated ions.

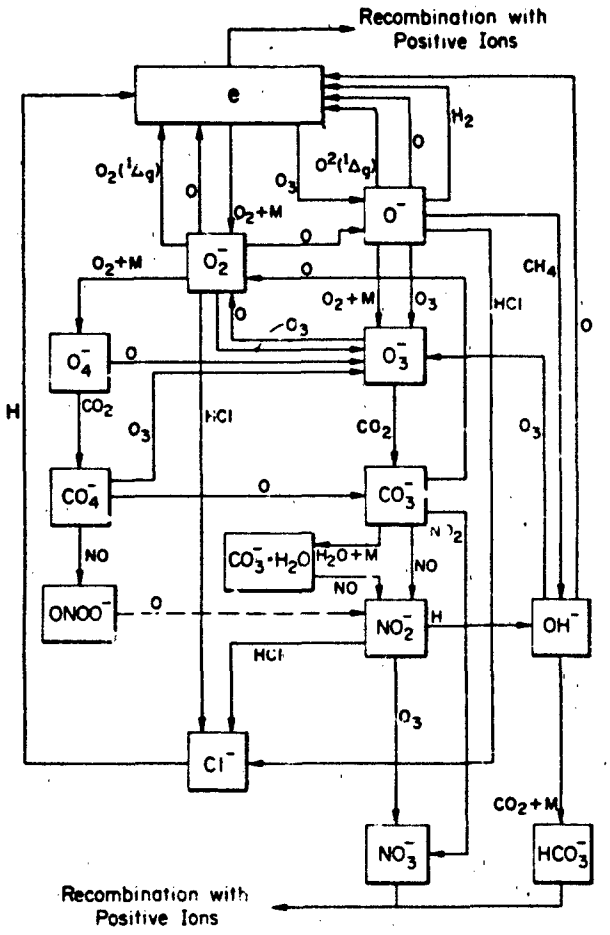


Figure 20-13. Reaction scheme of D-region negative ion chemistry [Ferguson, 1979].

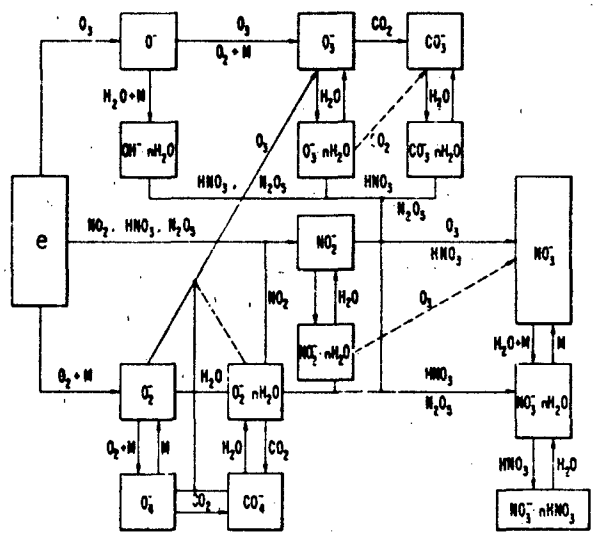


Figure 20-14. Stratospheric and tropospheric negative ion chemistry [Ferguson et al., 1979].

Table 20-7. Summary of lightning charges, altitudes, and moment charges in various geographical locations. [Jacobson and Krider, 1976].

However, even if the paths that lead to NO_3^- are somewhat uncertain, its eventual formation does not appear to be in doubt. Both HNO_3 and N_2O_5 provide effective NO_3^- production routes. It is highly likely that the terminal negative ions of the lower atmosphere are complex cluster ions such as $NO_3^- \cdot \ell H_2O \cdot mSO_2 \cdot nHNO_3$, with NO_3^- as the core ion. Recently, middle atmosphere ions have been tentatively identified as $NO_3^-(HNO_3)_n$ and $NO_3^-(HCl)(HNO_3)_n$ with $n = 1, 2, 3$ in each case, as well as some ions which may possibly involve HSO_4^- cores and H_2SO_4 neutrals. Our understanding of the negative ion chemistry of the middle atmosphere is clearly in a rather elementary state.

20.1.5 Electrical Equilibrium and Variations

20.1.5.1 Electrical Equilibrium. Above the exchange layer in fair weather areas, the charged and uncharged nucleus content is low. To a first approximation, ion equilibrium exists. Assuming charge neutrality, Equations (20.14) and (20.15) reduce to

$$q = \alpha n_\pm^2. \qquad (20.25)$$

Measured values of production rates due to cosmic rays at various atmospheric depths and geomagnetic latitudes are given in Chapter 6.

Under equilibrium conditions in regions of significant nucleus concentration, Equations (20.14) and (20.15) reduce to

$$q = \alpha n_\pm^2 + n_\pm (\eta_{12} N_\pm + \eta_{10} N_0), \qquad (20.26)$$

or

$$n_\pm = -\frac{[\beta N_\pm \pm (\beta^2 N_\pm^2 + 4\alpha q)^{1/2}]}{2\alpha} \qquad (20.27)$$

where $\beta = \eta_{12} + \eta_{10}(N_0/N_\pm)$, and $N_\pm$ is the charged nucleus concentration of either sign. The dependence of small-ion density and conductivity on charged nucleus concentration computed for several altitudes is given in Figure 20-16. Figure 20-15 compares theoretical relations [Equation (20.27)] with experimental data for small ion content and charge nuclei concentration at 1.5 km. The good agreement between the theoretical curves and experimental data indicates the relation between small ions and nuclei is well understood.

20.1.5.2 Variations in the Exchange Layer. The exchange layer varies in depth between 0.15 and 3.0 km above the surface. Large variations in the electrical properties oc-

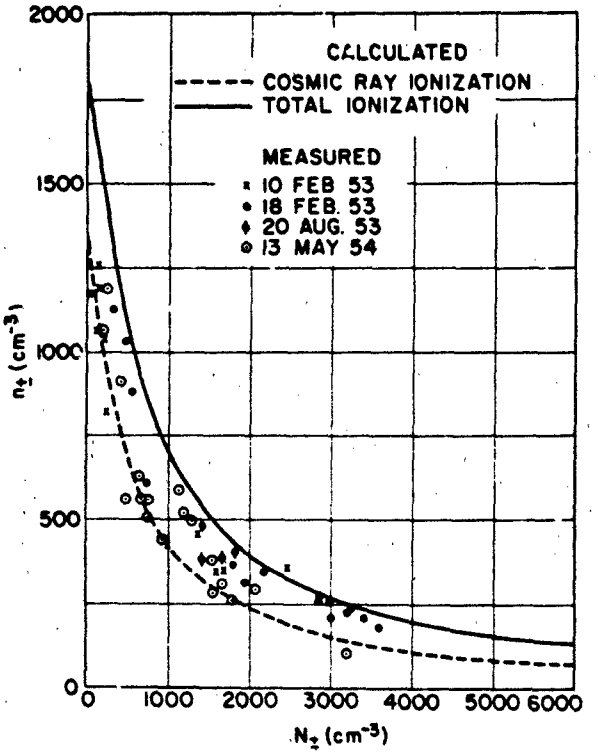


Figure 20-15. Positive or negative small ion number density n ± vs positive or negative charged nucleus concentration, N ± at 1.5 km altitude [Sagalyn and Faucher, 1956].

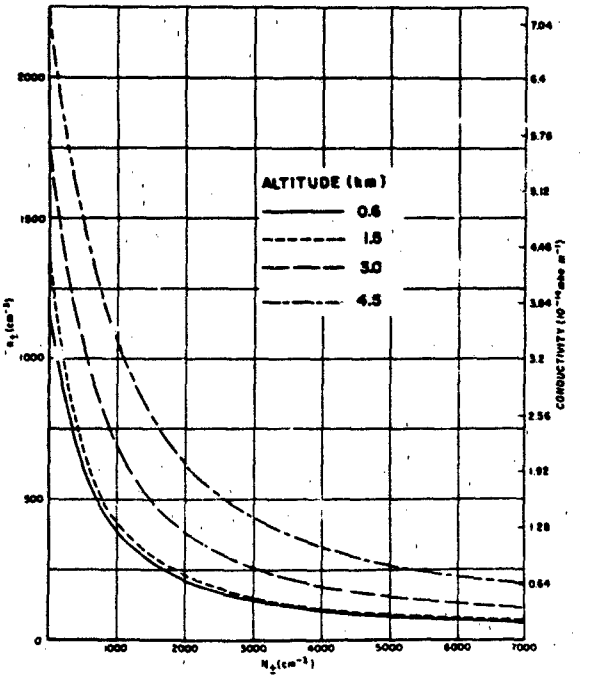


Figure 20-16. Positive or negative small ion number density η ± vs positive or negative charged nucleus concentration, N ± at 1.5 km altitude [Sagalyn and Faucher, 1956].

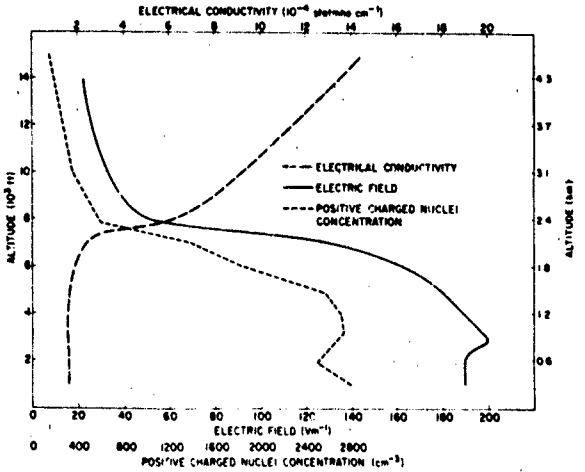


Figure 20-17a. Simultaneous values of electrical conductivity, electric field, and positive charged nuclei concentration vs altitude observed on 20 August 1953, 1311 to 1416 Eastern Standard Time [Sagalyn and Faucher, 1954].

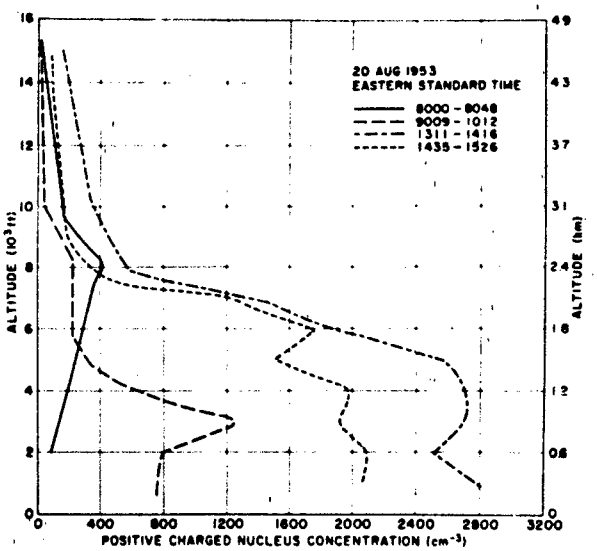


Figure 20-18a. Altitude and time variation of positive charged nucleus concentration during a period of low advection over southern New Hampshire [Sagalyn and Faucher, 1956].

cur in this layer. Meteorologically, the layer is characterized by uniform specific humidity and a nearly adiabatic temperature gradient rate at the upper boundary. The horizontal wind velocity differs from the wind velocity in the general circulation due to the frictional influence of the earth's surface. Figures 20-17a and b illustrate the variations of electrical properties with altitude in a well developed exchange layer.

The distribution and magnitude of the electrical properties vary from day to day and with time of day. During periods of low advection, regular daily variations can be observed. Figures 20-18a, b, and c demonstrate the daily variations observed during a series of aircraft flights over southern New Hampshire. The diurnal variation of the electrical properties illustrated in these figures results from the daily variation of the intensity of atmospheric turbulence and the influx of nuclei from the surface. Figure 20-19 shows the average daily variation in the columnar resistance of the atmosphere.

Over land areas, the most complicated distribution of the electrical properties is found between the surface and

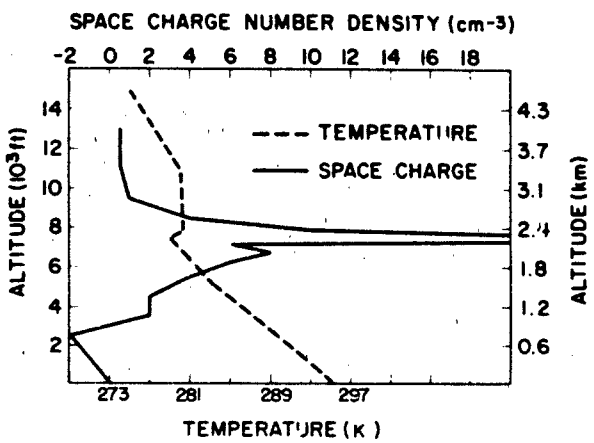


Figure 20-17b. Simultaneous temperature and space charge density vs altitude on 20 August 1953, 1311 to 1416 Eastern Standard Time [Sagalyn and Faucher, 1954].

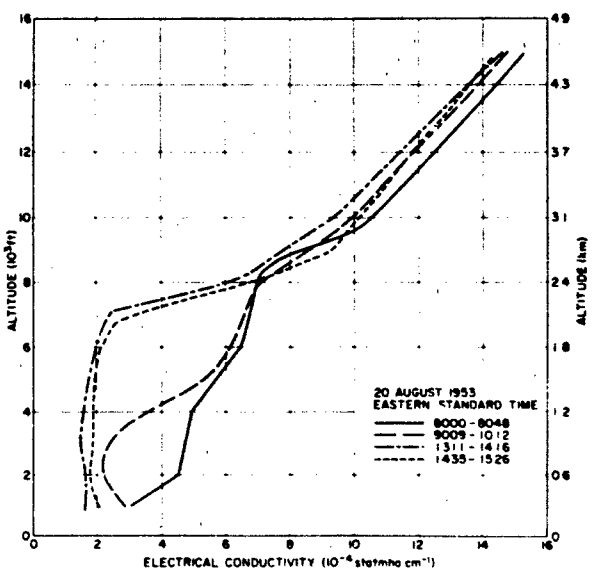


Figure 20-18b. Altitude and time variation of total volume conductivity during a period of low advection over southern New Hampshire [Sagalyn and Faucher, 1956].

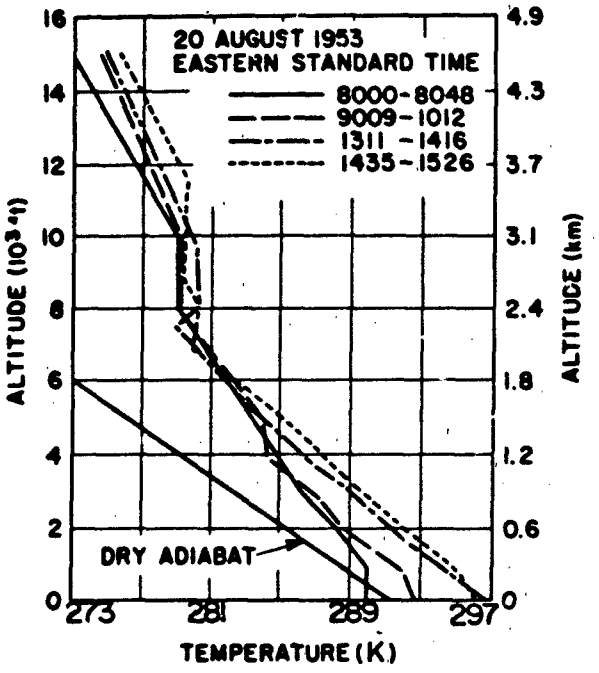


Figure 20-18c. Altitude and time variation of temperature during a period of low advection over southern New Hampshire [Sagalyn and Faucher, 1956].

approximately 0.3 km. In this region ionization produced by radioactive emanations from the surface of the earth can, particularly in the early morning hours, cause an initial decrease of conductivity with height. This causes the potential gradient to increase with altitude and produces a negative space charge region near the ground. With the onset of turbulence, the concentration of ions in the region between the surface and 0.3 km is greatly reduced, and the conductivity increases or remains constant with altitude. In accordance with Poisson's equation, the potential gradient then decreases with altitude, and a positive or zero space

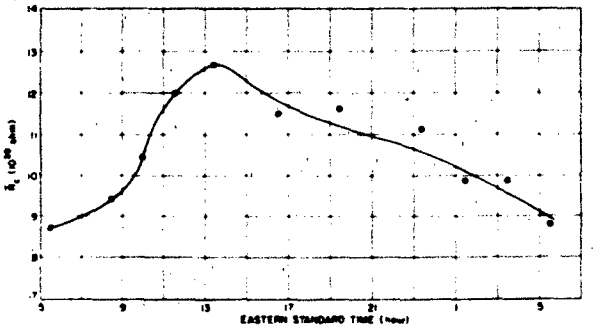


Figure 20-19. Diurnal variation of average resistance air column R_c of 1-cm² cross section and 4.6 km height [Sagalyn and Faucher 1956].

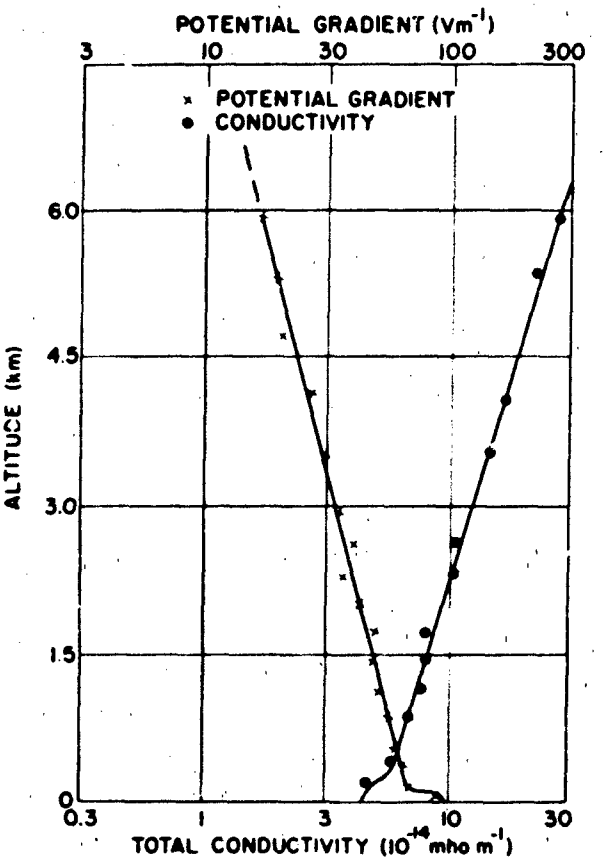


Figure 20-20. Electrical conductivity and electric field as a function of altitude over Greenland. (After J.F. Clark and J.M. Kraakevik [Smith, 1958].)

charge exists in this region. The exact nature of the diurnal variation depends upon the radioactive content and porosity of the surface and on the intensity of atmospheric turbulence.

The influence of the exchange layer on the electrical properties is minimal in regions where air masses are forming, such as the arctic and antarctic. This is illustrated in Figure 20-20 where the vertical distribution of electrical conductivity and the electric field measured on an early morning flight over Greenland are shown. Over oceans in high-pressure areas where very stable exchange layers are formed, negligible diurnal variation of the electrical properties are found. This is partly due to the absence of radioactive emanations from the surface and to a stable nuclei source. Examples of the electrical properties in a stable exchange layer over the oceans are shown in Figure 20-21a,b,c and d. Figure 20-22 shows the influence of the exchange layer on conductivities observed in mountainous regions.

20.1.5.3 Variations in the Free Atmosphere. The variability of the electrical quantities in fair weather in the troposphere is largely due to influences discussed throughout this

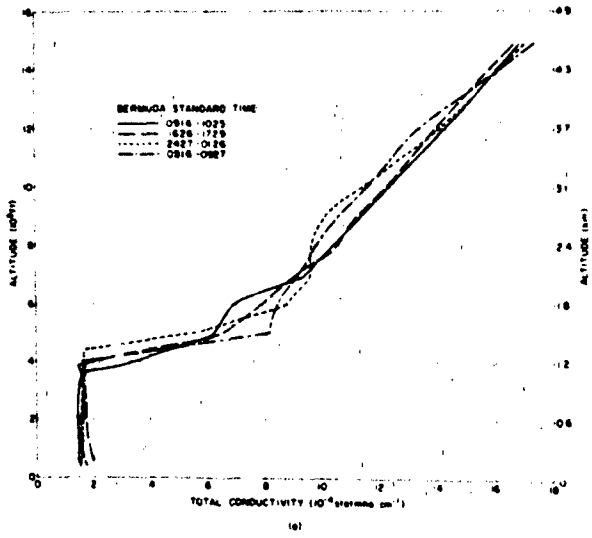


Figure 20-21a. Time variation of total electrical conductivity, positive charged nucleus concentration, temperature, and absolute humidity over the Atlantic Ocean at latitude 33°, longitude 64° on 13–14 April 1954. (From R.C. Sagalyn [Smith, 1958].)

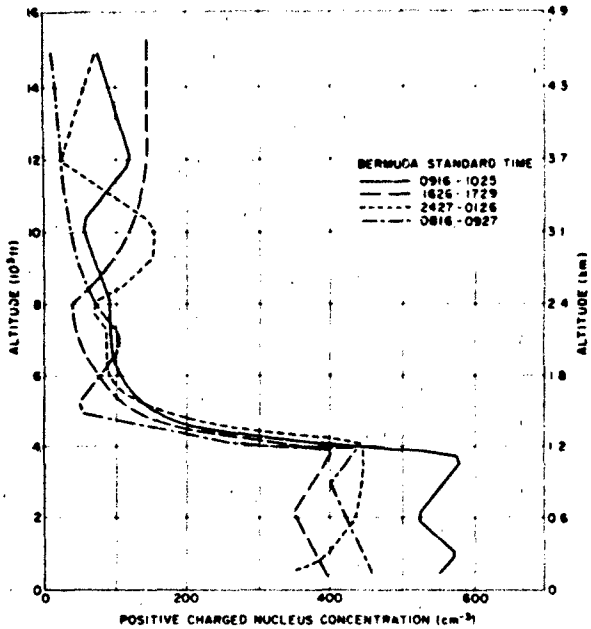


Figure 20-21b. Time variation of total electrical conductivity, positive charged nucleus concentration, temperature, and absolute humidity over the Atlantic Ocean at latitude 33°, longitude 64° on 13–14 April 1954. (From R.C. Sagalyn [Smith, 1958].)

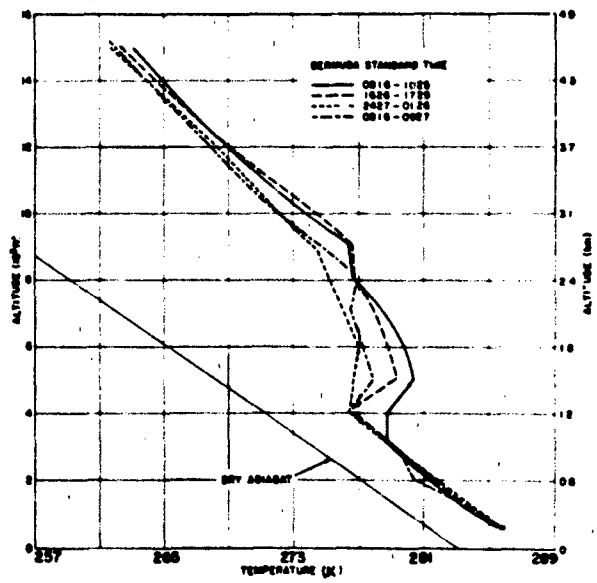


Figure 20-21c. Time variation of total electrical conductivity, positive charged nucleus concentration, temperature, and absolute humidity over the Atlantic Ocean at latitude 33°, longitude 64° on 13–14 April 1954. (From R.C. Sagalyn [Smith, 1958].)

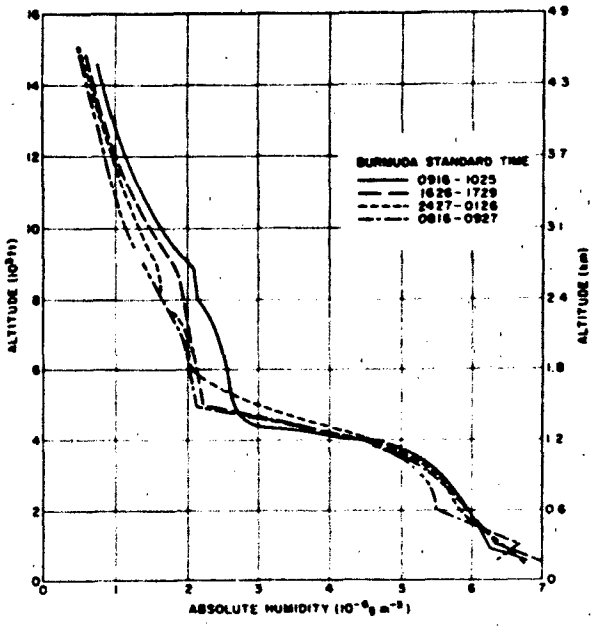


Figure 20-21d. Time variation of total electrical conductivity, positive charged nucleus concentration, temperature, and absolute humidity over the Atlantic Ocean at latitude 33°, longitude 64° on 13–14 April 1954. (From R.C. Sagalyn [Smith, 1958].)

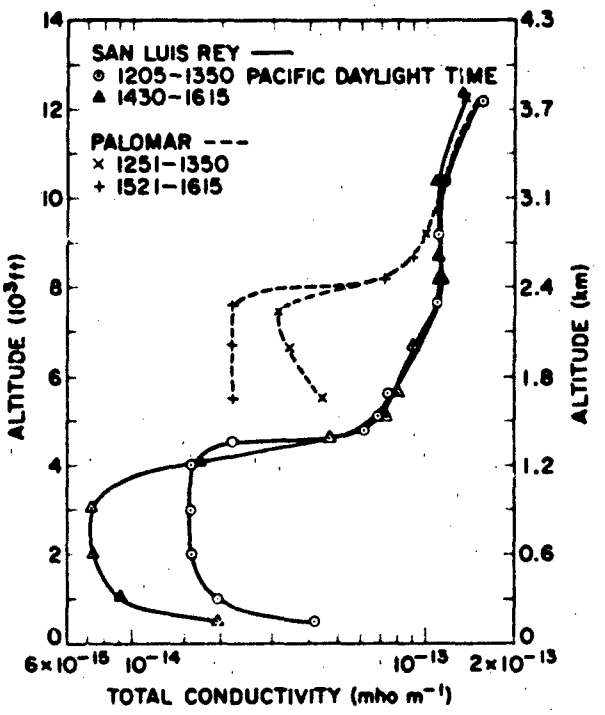


Figure 20-22. Electrical conductivity vs altitude over San Luis Rey station at 0.15 km and Palomar station at 1.7 km on 8 June 1953.

chapter. For example, the variations in electrical conductivity both from model computation and measurements have been shown in Figures 20-3, 20-4, 20-18b, and 20-21. Those for the electrical field were demonstrated in Figures 20-7a, 20-7b, 20-17a, and 20-20. For conduction current, the variations were shown in Figures 20-8 and 20-9. Figure 20-5 was an example of the variations with altitude of small ion density. It is further demonstrated in Figure 20-23 [Sagalyn, 1965]. The variability of large nucleus concentration with altitude is shown

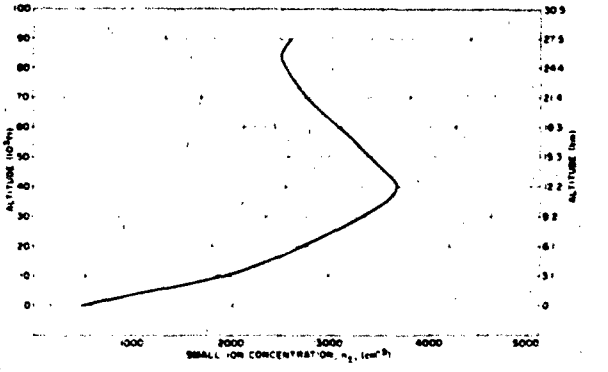


Figure 20-23. Average value and maximum variation of a small-ion density as a function of altitude. Data is based on over 60 aircraft flights throughout the eastern U.S.

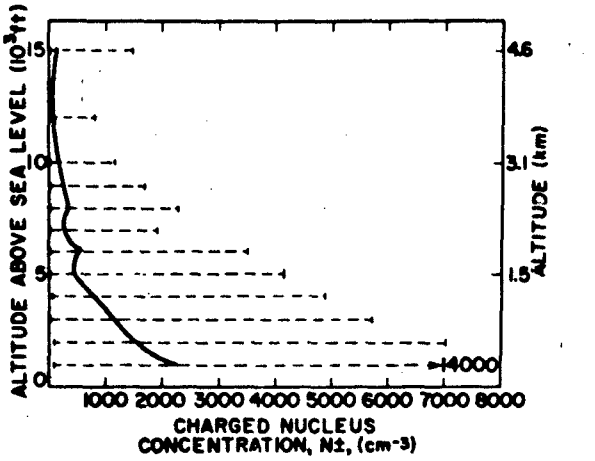


Figure 20-24 Average value and maximum variation of charged nucleus concentration as a function of altitude. Data is based on over 60 aircraft flights throughout the eastern U.S.

in Figure 20-24 [Sagalyn, 1965]. These results were obtained from over 80 aircraft and balloon flights throughout the U.S.

20.1.6 Solar Influence on Earth's Atmospheric Electrical Parameters

The connection between solar activity and the earth's atmospheric electrical properties and meteorological conditions is not yet understood. Investigations are underway to determine whether solar emission can act as a trigger mechanism for climatic changes over decades or millennia. Existing evidence strongly suggests a genuine link between transient energy-generating processes on the sun and electrical-meteorological responses in the near earth environment. The strongest effects are observed in the Northern Hemisphere winter when solar insolation is least effective, at middle to high latitudes [Herman and Goldberg, 1978a].

One way in which solar activity can couple into the atmospheric electrical system is through changes in the ion-pair production rate. Variations in atmospheric ionization and therefore in conductivity (except very near the surface) are produced by variations in the galactic cosmic ray flux intensity and by solar proton events.

Variations of the average ground-level cosmic ray intensity with the 11-year sunspot cycle are well documented and show a clear inverse correlation [Forbush, 1954]. An example of this is given in Figure 20-25 in which the normalized neutron monitor count rate from Climax, Colorado, is compared with average monthly Zurich sunspot numbers over two solar cycles. There is an approximate 20% decrease in count rate from solar maximum year 1957 to minimum year 1968. Data from Kent and Pomerantz [1971] indicate that at middle latitudes (55° geomagnetic) the count rate is

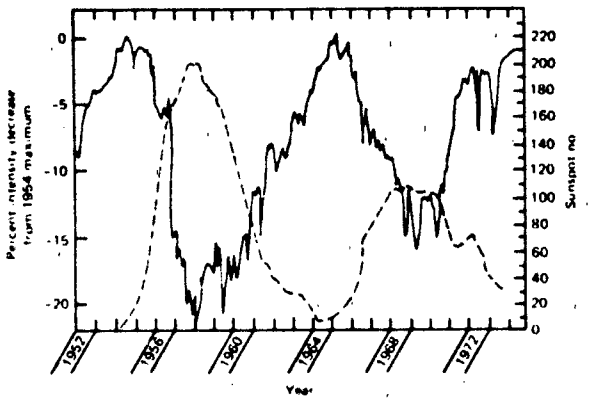


Figure 20-25. Solar cycle variation of cosmic ray intensity (solid curve) and sunspot number (dashed curve) over two solar cycles [Fulks, 1975]

about 21% lower at solar maximum, but near the equator is only 7% lower. The Climax cosmic ray intensity peaks approximately two-thirds of a year later than the time of sunspot minimum, but the amount of lag seems to be slightly different at different stations [Fulks, 1975].

The secular trend of potential gradient measured at stations in the British Isles, France, and Spain over a 20-year period (1902–1922) was investigated by Bauer [1926]. The average potential gradient for all stations combined, expressed in percentage of the 20-year mean value, was clearly above the mean during sunspot maximum years and below the mean in minimum years (Figure 20-26). For the potential gradient averages, days with local thunderstorms were excluded. Both the diurnal amplitude and the annual amplitude show a strong positive correlation with sunspot number over the two solar cycles. More recently, Muhleisen [1971] and Fischer and Muhleisen [1972] investigated the 11-year cycle influence on ionospheric potential. Their results also indicate a positive correlation with annual mean sunspot number.

Figure 20-12 shows the ion-pair production rate as a function of height for solar radiation (curves 1 to 5) and galactic radiation (curves 6 and 7). Although galactic cosmic rays carry more energy per particle and penetrate deeper into the lower atmosphere, episodic solar emissions and the related interplanetary magnetic field (dependent on flare activity and solar rotation) can produce higher ion-pair production rates in the lower atmosphere. This is demonstrated in curves 1 to 5 in which injections of solar protons during solar flares produce ionization above 20 km considerably in excess of the normal quiet background production rate provided by galactic cosmic rays in both sunspot minimum (curve 6) and sunspot maximum (curve 7) years. Curves (6 and 7) from Webber [1962] demonstrate the solar cycle variation in cosmic ray ion production at a magnetic latitude of about 70°.

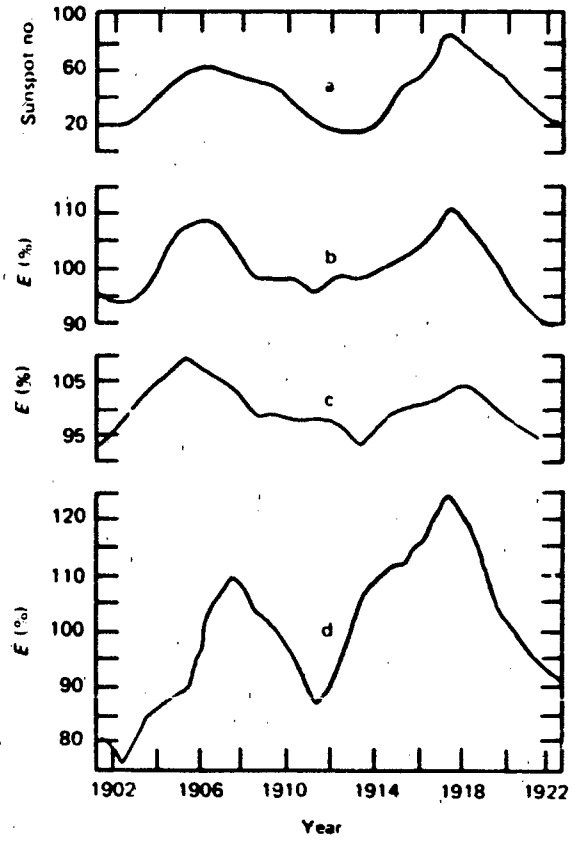


Figure 20-26. Cyclical variations in potential gradient compared to 11-year sunspot cycle. Curve a, relative sunspot number; b, potential gradient; c, daily amplitude of potential gradient; d, potential gradient annual amplitude. Curves b, c, and d are in percentage of mean values [Bauer, 1926 and Israel and Dolezalek, 1973]

On shorter time scales, solar flare eruptions and solar magnetic sector boundary crossings that result in increased numbers of solar protons entering the lower atmosphere also cause a decrease in galactic cosmic rays. The decrease in galactic cosmic rays following a solar flare is termed a Forbush decrease. From long term investigations, it appears that air-earth current, potential gradient, ionospheric potential, and thunderstorm activity all respond to solar flares. Enhancements in these atmospheric electrical quantities occur 1 to 4 days after the eruption of a major solar flare, with thunderstorm occurrence responding slowest. The increases range from 12% [Cobb, 1967] to 50% [Reiter, 1969] in air-earth current density, 30% to 60% in potential gradient [Reiter, 1969], and 20% [Flohn, 1950] to 70% [Bossolasco et al., 1972] in thunderstorm occurrence frequency.

The responses of atmospheric electrical quantities to solar flare occurrence are probably due to a combination of factors. High energy solar protons emitted during flares produce enhanced atmospheric ionization and thus increased

conductivity above ~ 20 km altitude. The Forbush decrease in cosmic ray intensity following flares results in a decreased conductivity below 20 km altitude. These changes could lead to an increase in the electric field at levels below 20 km and a decrease above 20 km. This in turn can produce the potential increase observed during thunderstorm activity [Herman and Goldberg, 1978b].

20.1.7 Global Model of Atmospheric Electricity

Significant advances in the understanding of atmospheric electrical processes during the past decade have resulted from the development of global modeling and from inclusion of coupling with the ionosphere and outer atmosphere. Previously, atmospheric electrical phenomena were assumed to take place within a "concentric spherical capacitor" [Chalmers, 1967]. The lower boundary of this capacitor was the earth's surface and the electrosphere (a highly conducting layer at about 50–70 km) was assumed to be the upper boundary. The electrosphere is defined as the height of an equipotential that acts as a perfect electrostatic shield to physical processes taking place outside this region. Thunderstorms and other electrical processes in the lower atmosphere would then have no effect outside the electrosphere. Conversely, processes taking place beyond the ionosphere were assumed to have a negligible effect on electrical conditions in the lower atmosphere. In this view, thunderstorm activity was the major source of charge generation within the atmospheric electrical system and was balanced by the fair weather conduction current.

This "close-in" theory of atmospheric electricity is correct to first order in most cases. However, it does not allow for horizontal electric fields that are known to exist in disturbed weather areas and for variation in ionosphere characteristics. Recent studies indicate that the electrosphere is not a perfect equipotential layer. The "capacitor" is in fact leaky, with the result that atmospheric electrical processes interact with ionospheric and magnetospheric phenomena.

Hays and Roble [1979] derived a quasi-static numerical model of global atmospheric electricity in which thunderstorms are the main source of electric current. Thunderstorms are assumed to be distributed geographically in accordance with the known statistical distribution of thunderstorm frequency. In the model, the electrical conductivity increases exponentially with altitude and electrical effects are coupled into the magnetosphere along geomagnetic field lines. The electrical conductivity is assumed to vary with latitude, simulating the latitudinal variation of known cosmic ray production. The global distribution of electric potential and current is then calculated. The results show that large positive electric potentials are generated over thunderstorms and that these perturbations penetrate upward to ionospheric altitudes. The effect of a thunderstorm in one

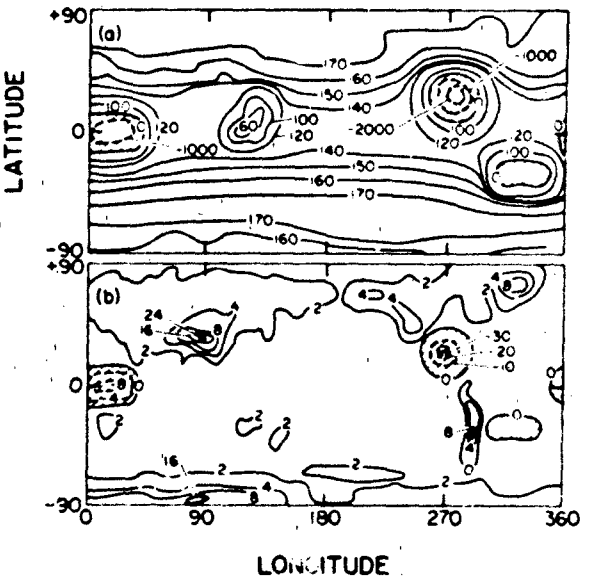


Figure 20-27. Results from Hays-Roble global atmospheric electricity model (a) latitude and longitude variation of the vertical component of electric field at the earth's surface (b) latitude and longitude variation of conduction current at the earth's surface [Hays and Roble, 1979]

hemisphere can be transmitted along geomagnetic field lines into the conjugate hemisphere; however, the potential perturbation in the conjugate hemisphere is damped below stratospheric altitudes. Electric fields over thunderstorm regions may approach 0.25–0.50 mV/m at ionospheric heights for nighttime conditions. The return current at the earth's surface in the fair weather region is greater over mountainous regions than at sea level. The perturbation of the calculated electric potential and current distributions due to an increase in cosmic rays during a solar flare increase and the subsequent Forbush decrease in cosmic ray ionization can also be carried out using the model. An example of the output of the model is shown in Figure 20-27. The results in Figure 20-27 were derived assuming Northern Hemisphere summertime background with 2000 individual thunderstorms acting at 1900 UT. The vertical component of the electric field and the conduction current at the earth's surface are shown. Equipotential curves and constant current density contours are plotted in panels a and b, respectively.

20.1.8 Recent Advances in the Middle Atmosphere

20.1.8.1 Middle Atmosphere. The term "middle atmosphere" is used to describe the region between the tropopause and the mesopause, with possibly a slight extension into the lowest part of the thermosphere, that is, the altitude region 10 to 100 km. It includes the stratosphere and the meso-

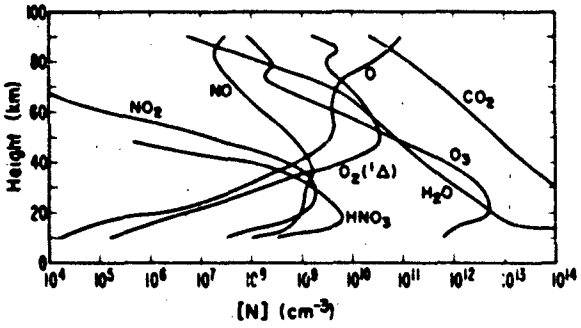


Figure 20-28 Typical daytime concentration profiles of some minor constituents of the middle atmosphere [Reid et al., 1979]

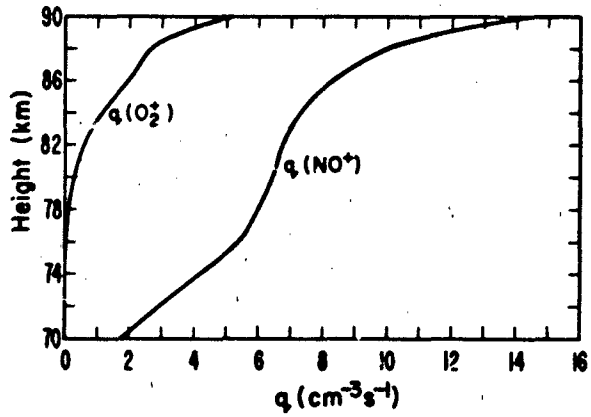


Figure 20-29 Rate of ion production for a solar zenith angle of 45° and undisturbed conditions [Reid et al., 1[illegible]]

sphere, knowledge of which has increased greatly as a result of recent studies of the ozone layer (Chapter 21).

Knowledge of the chemical composition is fundamental to the understanding of the electrical characteristics of the middle atmosphere. Several minor constituents, typical profiles of which are shown in Figure 20-28, play prominent roles. The composition of the middle atmosphere was discussed in Section 20.1.4. The temperature structure and the wind systems of the middle atmosphere are determined by a balance between the heating and cooling rates. The heating is due primarily to the absorption of solar radiation by ozone, and the cooling to infrared radiation, mainly by carbon dioxide. Thus, these two minor constituents determine the structure of the middle atmosphere. The other constituents shown in Figure 20-28 are the chief participants in the production of ionization, or in the ion chemistry that determines the steady-state ion concentration and composition.

During the winter, there is a pronounced equatorward-directed temperature gradient in the middle atmosphere, and the resultant geostrophic winds are westerly as in the troposphere. Wind speeds reach their maximum near the stratopause, forming the so-called polar-night jet, surrounding a strong polar vortex. During the summer the situation tends to be reversed, with a poleward-directed temperature gradient caused by the combination of a high-latitude ozone maximum and essentially continuous insolation, giving rise to prevailing easterlies throughout most of the middle atmosphere. The spring and fall transitions between these two average conditions and major transient stratospheric warming events of the late winter are periods of particular interest in the middle atmosphere.

Vertical transport in the middle atmosphere is usually parametrized through the use of an eddy-diffusion coefficient whose profile is based on observed vertical diffusion rates. Characteristic vertical mixing times are long, varying from years in the lower stratosphere, to months in the upper stratosphere and lower mesosphere, and to weeks in the upper mesosphere. Since these are much longer than the characteristic lifetimes of ions, the ions are not themselves directly affected by vertical transport. Many of the minor species that determine the ion composition, however, are dominated by transport effects rather than by photochemistry in much of the middle atmosphere.

The major source of mesospheric ionization is the NO molecule, which can be ionized by solar Lyman-alpha radiation at 1216 Å. The Lyman-alpha line coincides with a window in the O_2 absorption spectrum and penetrates most of the mesosphere reaching unit optical depth at about 75 km for an overhead sun. NO^+ is the primary ion species produced by this source. Figure 20-29 shows the rates of production of both O_2^+ (from the $O_2(^1\Delta)$ source as well as from x-rays during average solar conditions) and NO^+ for a solar zenith angle of 45° and for currently accepted profiles for NO and $O_2(^1\Delta)$.

In the lower mesosphere and throughout the stratosphere, the chief source of ionization under normal conditions is galactic cosmic rays. The rate of ion production due to cosmic rays was shown in Figure 20-12 for different latitudes and solar conditions. The cosmic ray induced variation of small ion density with latitude is given in Figure 20-5.

The solar magnetic field and solar wind act to exclude cosmic-ray particles from the near earth environment through mechanisms that are not yet fully understood. The chief result is that the cosmic-ray flux at the earth is lower during solar maximum, when the solar magnetic field is more intense on the average, than during solar minimum. The inverse relation between solar activity and cosmic ray intensity near the earth's surface is illustrated in Figure 20-25. This produces a significant solar-cycle variation in the electrical properties of the middle atmosphere.

20.1.8.2 Middle Atmosphere Ion Concentration, Mobility, and Conductivity. The fundamental relations discussed in Sections 20.1.1 through 20.1.6 are valid throughout the middle atmosphere. The steady-state ion concentra-

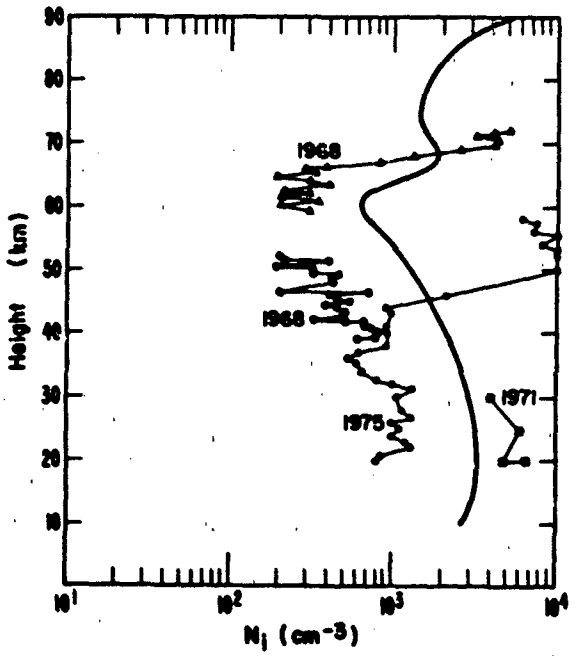


Figure 20-30. Positive-ion concentration profile from model calculation (solid line) and observed [Reid et al., 1979].

tion profile of the middle atmosphere is illustrated in Figure 20-30 which shows the results of a recent model calculation of positive-ion concentrations for solar-maximum year at geomagnetic latitude 40°, and for a solar zenith angle of 45°. Also shown in the figure are the results of direct measurements of small-ion concentrations for similar geomagnetic latitude but for different phases of the solar cycle. Obviously there is a general agreement in the shape of the profiles, but the sharp enhancements in ion concentration that appeared at 65 km in 1968 and at 45 km in 1975 have

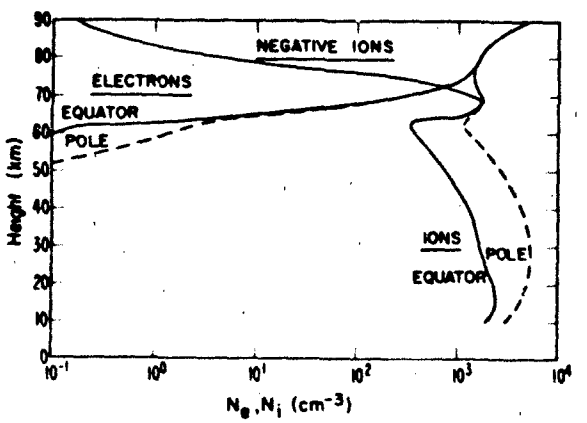


Figure 20-31. Negative ion and electron concentration profiles at the pole and equator from model calculation [Reid et al., 1979].

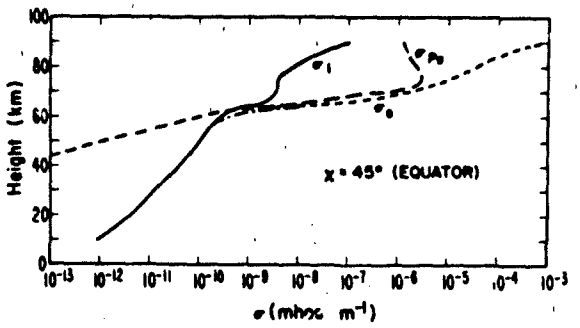


Figure 20-32. Calculated profiles of electron and ion conductivity at the equator. [Reid et al., 1979].

no counterpart in the theoretical results. The model results also appear to be higher by about a factor of three than the actual observations over most of the range. Since our knowledge of the ion production rate is unlikely to be seriously in error, at least below 60 km, the discrepancies must reflect our lack of knowledge of the full details of the ion chemistry of the stratosphere.

Figure 20-31 shows the results of a similar model calculation of the concentrations of the negatively charged species—negative ions and electrons—for the same conditions of solar illumination but for polar and equatorial latitudes. The very sharp transition between the overlying electron-dominated region and the underlying negative/positive-ion dominated region is the most obvious feature. There is considerable observational confirmation consistent with the model that the transition from negative ion to electron occurs at about 70 km during quiet daytime conditions.

Figures 20-32 and 20-33 show theoretical profiles of conductivity for equatorial and polar latitudes, again for solar maximum conditions and for a solar zenith angle of 45°. The solid lines represent the ionic component and the broken lines the electronic component of the conductivity, with the crossing points marking a sudden increase in the gradient of total conductivity. Below this altitude, substan-

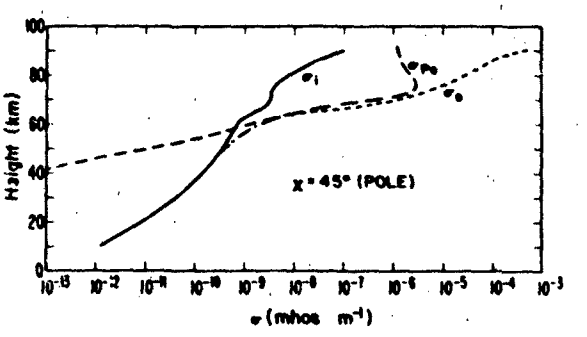


Figure 20-33. Calculated profiles of electron and ion conductivity at the poles. [Reid et al., 1979].

tial horizontal potential differences can be maintained, but the conductivity at higher altitudes is large enough to allow only relatively small potential variations on a global scale.

In the upper mesosphere the geomagnetic field begins to have an important effect on the electron conductivity. The electrons are forced to spiral about the magnetic field, and if the collision frequency is low enough they can no longer move freely along the direction of an applied electric field that has a component perpendicular to the magnetic field. Collisions with gas molecules, however, break up this organized spiraling motion, and allow some movement along the electric field direction, thereby preventing the electron conductivity along the direction of the electric field from vanishing entirely. The residual electron conductivity perpendicular to the magnetic field is referred to as the Pedersen conductivity, and is shown in Figures 20-32 and 20-33 by the broken line labeled σ_p. The conductivity along the direction of the magnetic field is unaffected by the spiralling motion of the electrons, and is identical to the uniform conductivity that would exist if there were no magnetic field. It is shown by the broken line labeled σ_c. Since the electrons drift along a direction perpendicular to both the magnetic and electric fields between collisions, there is a third component of the conductivity known as the Hall conductivity, but it differs in character from the other components by virtue of its vanishing power dissipation (since $j_H \cdot E = 0$). Throughout the middle atmosphere the ions remain collision dominated and are just as free to move perpendicular to the magnetic field as they would be in the absence of a magnetic field.

The various sources of middle-atmosphere ionization vary on a wide range of time scales, and a corresponding variability in the electrical parameters must exist. In the mesosphere, the solar x-ray flux varies with the 27-day solar rotation period, with the 11-year solar cycle, and with individual flares when its intensity can increase by at least three orders of magnitude above the normal value. Except during intense solar flares, x rays remain a fairly negligible ionization source.

The major mesospheric source of ionization of NO, Lyman-alpha, increases by nearly a factor of two from solar minimum to solar maximum and by about 30% in the course of a solar rotation. The concentration of NO will vary with the production of NO in the lower thermosphere by energetic particles. Auroral ionization is known to be accompanied by the production of large quantities of NO near 100 km altitude, but downward diffusion of NO into the mesosphere is still unknown.

The electrical properties of the middle atmosphere are highly complex and variable. The middle atmosphere responds to driving forces from above and from below, and is subject to electric fields of both tropospheric and magnetospheric origin. It must play a significant role in any solar terrestrial coupling mechanisms that involve the lower atmosphere.

20.2 THUNDERSTORM ACTIVITY

20.2.1 Thunderstorm Charge Distribution and Electric Field Pattern

The thunderstorm is the final product of the growth of a series of clouds with vertical development known in order of increasing size as cumulus, cumulocongestus, and cumulonimbus. This class of cumuliform clouds is formed through the upward transfer of energy from surface heating effects leading to buoyant vertical air motions or through the lifting effects associated with the motion of atmospheric frontal weather zones. The cumuliform clouds derive their energy from the latent heat of condensation and sublimation of atmospheric water vapor drawn into the cloud circulation. The intensity of the resulting convective activity depends on the available moisture and the stability of the lower atmosphere as expressed by the air temperature variation with height. Weather services technically distinguish the thunderstorm from other convective clouds by the observation of thunder. These storms normally are associated with extensive lightning activity, heavy rain showers, possible hail or snow pellets, and strong subcloud wind gusts due to the spreading out of cold downdrafts from the areas of precipitation.

The storm electrical activity is spectacular and dangerous; however, it represents an energy expenditure of less than 1% of the thermal, gravitational, and kinetic energy associated with the condensation of moisture and the development of drafts and precipitation [Braham, 1952]. The storm progresses through three stages of its life cycle. The growth stage is characterized primarily by in-cloud updrafts. The active stage is accompanied by cloud top glaciation or conversion to ice forms, and the presence of strong up and downdrafts, lightning, and precipitation. The dissipating stage is reached when a lifting or breaking up of the lower cloud and a layering and spreading out of the upper anvil or ice form cloud material occurs. At this point the vertical air motions are greatly reduced and the precipitation is reduced to widespread fairly uniform small droplets or small snowflakes falling from the anvil [Byers and Braham, 1949]. By this time the electrical activity has subsided. However, considerable residual electric charge remains stored on the cloud particles and there is a possibility of occasional strong cloud to ground lightning or of a strike to an aircraft or rocket entering the cloud.

Details of the electric charge generation process in storms are not fully established. Numerous observations indicate that strong electric fields and lightning are usually found in convective clouds that contain mixtures of water droplets and various forms of snow and ice particles. The charge is probably generated through a combination of processes involving differential small ion capture by water droplets, induction effects of collisions of different sized particles in

an existing electric field, and physical effects. These include a tendency of water droplets to form an easily shattered electrical double layer in the droplet surface, possible proton migration in the ice crystal lattice structure due to thermal gradients, and other effects associated with the liquid to solid and solid to liquid phase changes which occur during the formation and fallout of precipitation.

Separation of the charge generated as a result of an enormous number of these microscale events is brought about primarily by the differential motion of variously sized precipitation and cloud particles due to gravitational forces and by the draft velocity production of aerodynamic drag forces on the particles. On the order of 1000 coulombs of bound space charge are produced and maintained during the late growth and active period of the storm. Observations indicate certain regions of the storm characteristically have an excess of negative or positive charge. These zones of unbalanced charge create electrostatic fields that extend to the surface of the earth and for a considerable distance into the clear atmosphere surrounding the storm. Measurements of these fields and of the field changes due to the neutralization of portions of the charge centers during lightning flashes were originated by Wilson [1916]. Numerous subsequent observations confirm that the electrical structure of the storm appears to resemble a bipolar charge distribution with an upper positive and lower more concentrated negative charge center. Some storms have an additional small positive center located near the cloud base. This center seems to be closely associated with the major rain shafts falling from the storm. A typical midlatitude storm is depicted in Figure 20-34. The charge center locations are more closely related to the vertical temperature structure than to the geometric differences in height of storms in various geographic areas of the world. The approximate temperature dependence is shown by the scale along the vertical axis. The estimated percentage frequency of lightning channel paths is given by the circled numbers adjacent to the various forms of lightning events depicted.

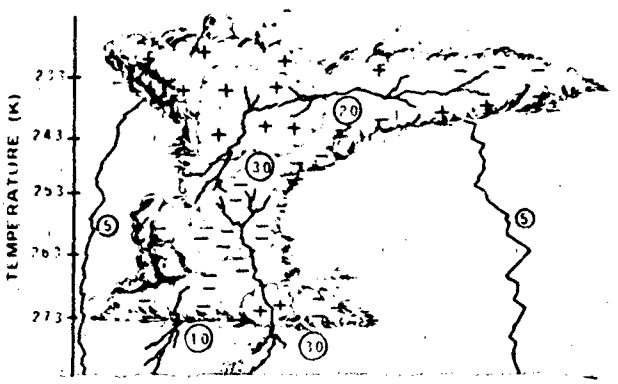


Figure 20-34 Typical charge distribution and lightning patterns of a mid-latitude thunderstorm

20.2.1.1 Surface Electric Field. The magnitude of the electric field of a thunderstorm as measured at ground level is strongly dependent upon local topography and the degree to which vegetation and structures are present to supply sources of point discharge current from the earth in response to the storm-initiated field. Storms over water and in mountainous terrain above timberline may produce maximum surface fields of about 100 kV/m. Storms moving over wooded, brushy or other terrain, which can provide numerous small pointed sources for development of a point discharge current, normally form a layer of space charge about 200 m thick near the ground under the storm. This layer acts to reduce the field at the ground to about 10 kV/m. Moving storms typically have some wind induced tilt to the main bipolar charge axis. This results in a wave like pattern in the time series observation of the surface field as a storm passes. Figure 20-35 indicates the general form and dimensions of the electrostatic disturbance

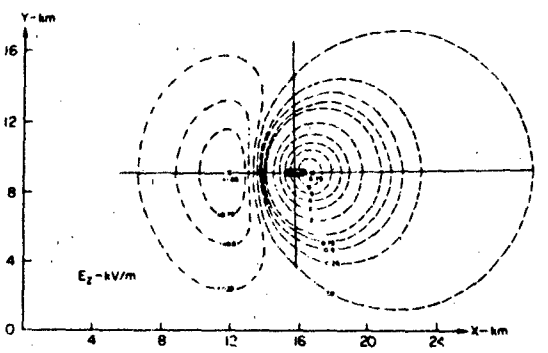


Figure 20-35. Electric field produced at surface of earth by the tilted bipole model [Fitzgerald, 1957].

associated with these storms. The figure is drawn for a lower 10 coulomb negative center at 3.5 km height and an upper 10 coulomb positive center at 4.5 km which is displaced down wind by 1 km. This figure indicates that the electric field at a particular site may exhibit a strong initial increase of field in either a positive or negative direction if a thunderstorm develops or moves into the vicinity. Interpretation of the time series field record for lightning warning is strongly dependent upon station location with respect to the cloud, the state of development of the cloud, and the structure of the winds aloft.

A typical pattern of the electric field associated with the passage from west to east of a squall line thunderstorm complex in the midwestern United States is shown in Figure 20-36. The initial rise in field associated with the approaching anvil charge is seen to the right of the figure. This is followed by a field reversal as the main-storm mass approaches the station. Numerous lightning field changes are indicated by the abrupt discontinuities in the basic wave

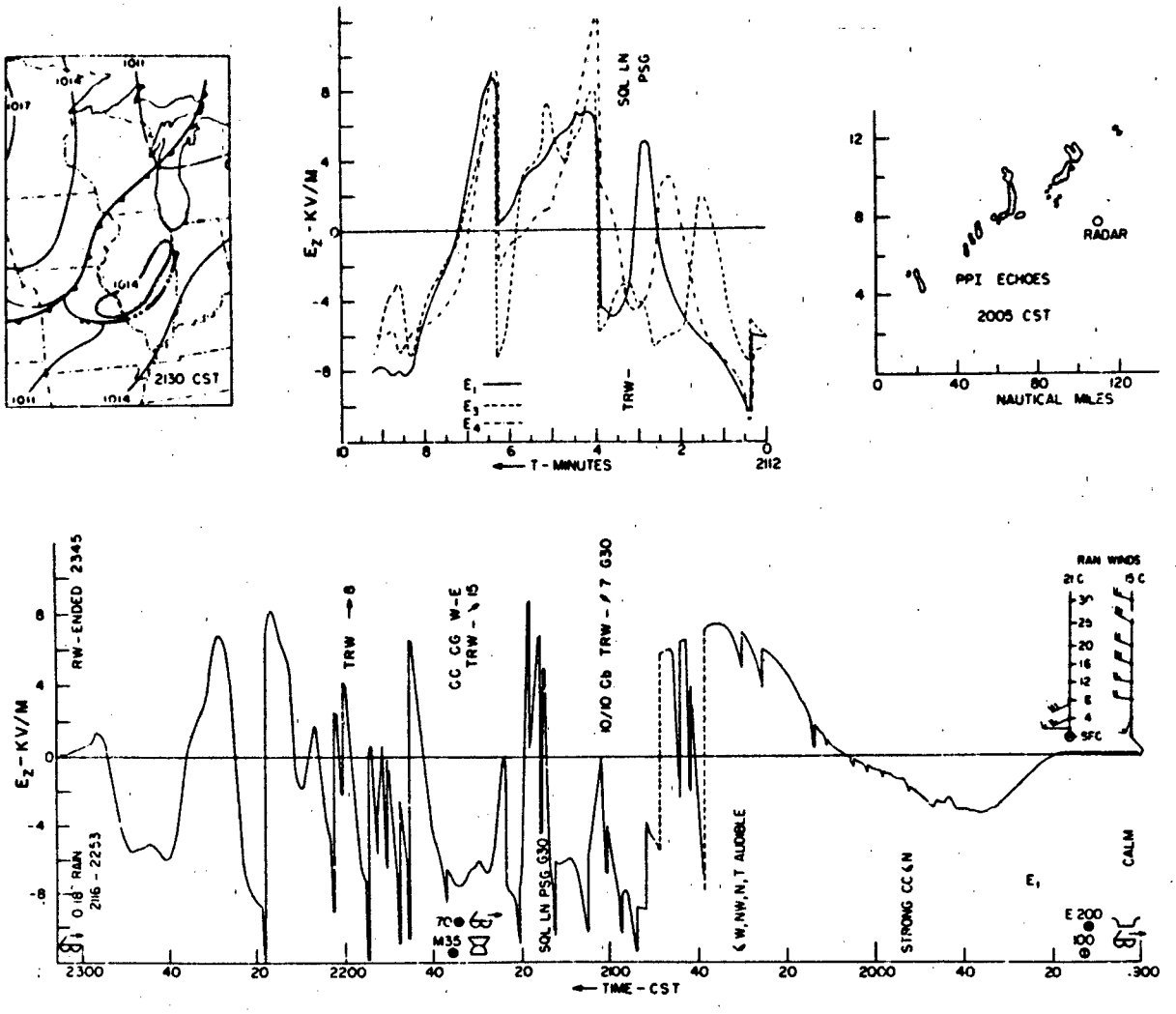


Figure 20-36. Surface field with active squall line passage, Illinois [Fitzgerald, 1956].

pattern. The squall line passage from 2115 to 2120 was accompanied by a lower positive charge region in the main part of the storm. Several subsequent alterations of lower negative and positive regions are noted as the storm passed the station. The final positive rise before return to the fair weather field was associated with the receding anvil structure. Some characteristics on the duration and lightning activity of typical slowly moving summer storms at Kennedy Space Center, Florida are given in Table 20-6 from a study by Livingston and Krider [1978]. These storms are frequently of long duration and large dimensions. They are accompanied by considerable lightning activity and heavy precipitation.

Jacobson and Krider [1976] have summarized magnitudes and altitudes of charge centers destroyed in lightning events studied in various geographical regions. Their summary is given in Table 20-7.

20.2.1.2. Field Patterns Aloft. Aircraft measurements indicate that electrostatic field components within storms are usually in the range of from 30 to 150 kV/m. Frequent alterations of polarity are found. These are usually associated with areas of differing precipitation forms, free air temperature, and vertical air motion. The scale of these complex field regions ranges from about 300 m to several km. Small intensely charged regions are occasionally found. These may have field components as high as 400 kV/m. These small regions are considered likely sources of lightning initiation.

The general pattern of the vertical component of the electric field in thunderstorms occurring in England was determined by Simpson, Scrase, and Robinson through a number of balloon soundings. Figure 20-37 adapted from Simpson and Robinson [1941] represents a summary of their results. It shows isolines of constant vertical field for a

Table 20-6. Summary of electrical behavior of 1975 and selected 1976 air mass storms at the NASA Kennedy Space Center [Livingston and Krider, 1978].

Date	Start/Stop Times, UT	Storm Duration, min	Total Number of Discharges	Maximum Flashing Rate, min^{-1}	Average Flashing Rate, min^{-1}	Maximum Radar Top, km
			1975			
May 15	1455/1608	73	533	20.0	7.3	14.0
June 2	1739/2004	145	866	17.0	6.0	13.7
	2002/2325	203	503	6.0	2.5	14.9
June 3	1702/1915	133	434	10.6	3.3	15.2
June 7	2011/2114	63	361	17.2	5.7	
	2202/2243	41	105	9.8	2.6	
June 9	1930/2350	260	879	10.4	3.4	15.5
June 10	2053/2127	34	25	1.2	0.7	10.1
June 16–17	2006/0105	299	1853	24.8	6.2	14.3
June 17	1514/1615	61	31	1.2	0.5	12.2
	1617/1745	88	79	1.4	0.9	12.2
	1753/1855	62	126	3.6	2.0	12.2
	1850/1926	36	74	4.0	2.1	12.2
June 18	1533/1709	96	122	3.6	1.3	12.2
June 20	1541/1605	24	8	0.4	0.3	8.2
June 26	2012/2230	138	579	9.2	4.2	15.8
June 27	1930/2021	51	178	6.8	3.5	
	2031/2300	149	1392	21.0	9.3	17.7
June 28	0000/0120	80	31	1.0	0.4	12.2
	0115/0145	30	18	1.0	0.6	
July 8	1831/2020	109	303	9.0	2.8	11.3
July 9	0000/0220	140	185	5.2	1.3	9.4
July 9–10	2037/0010	213	1943	24.0	9.1	16.5
July 10	1940/2325	225	1987	26.0	8.8	16.5
July 11	1714/1829	75	97	3.4	1.3	11.0
July 13	2001/2032	31	12	1.0	0.4	14.0
	2020/2232	132	96	2.0	0.7	14.0
TOTAL		2991	12820		4.2	
			1976			
July 8	1513/1543	30	47	7.0	1.6	
July 13	1725/1835	70	359	15.4	5.1	15.8

model containing +24 coulombs of upper charge, −20 coulombs of intermediate level charge and +4 coulombs of lower positive charge. The lower trace represents the field at the surface due to this model. The height and distance scale shown is typical of most European thunderstorm situations. United States and tropical area storms may have much larger dimensions. Later balloon soundings by Chapman [1958], Winn et al. [1981], Weber et al. [1982], and Rust and Moore [1974], for example, have been made in New Mexico storms. Parker and Kasemir [1982] have summarized the electrical structure of the lower regions of Colorado and Wyoming storms as measured by Kasemir and Holitza. Two aircraft equipped with electric field sensors were used to acquire data at two altitudes in the storm. A typical vertical field profile is shown in Figure 20-38. Chapman's [1958] balloon sounding had a generally similar shape. However, positions of the field maxima and zero crossover points were located about 500 m lower and the maximum fields inferred were about ±90 kV/m.

Combined cloud physics and electrical measurements of thunderstorm properties were conducted as a joint Air Force-National Severe Storms Project experimental program during the 1962–1966 time period. Figure 20-39 depicts the vertical electric field, voltage on the F-100F penetration air

Table 20-7. Summary of lightning charges, altitudes, and moment charges in various geographical locations. [Jacobson and Krider, 1976].

Geographical location	Charge (C)	Altitudes of lightning charges above local terrain† (km)	Range of air temperatures (K)	Moment changes (C · km)	Investigators
Florida	−10 to −40	6 to 9.5 (0)	263 to 239	100 to 600 400 (av)	present study
England	−11.5 to −46	7 (1)	239*	33 to 430	Wilson (1916)
England	−20	2 (1)	272.5*	100 (av)	Wilson (1920)
England	−10 to −40	4.5 to 5 (1)	257 to 254*	220 (av)	Wormell (1939)
England	—	—	—	150 (av)	Pierce (1955)
South Africa	−15	3 (1.8)	260*	93 (av)	Schonland (1928)
South Africa	−4 to −40	2.5 to 8.7 (1.8)	263 to 225*	41 to 495	Barnard (1951)
South Africa		4 to 8.5 (1.8)	266 to 235	—	Malan and Schonland (1951)
New Mexico	−24 (av)	4 to 7 (1.6)	268 to 248	—	Workman et al. (1942)
New Mexico	−5 to −20	4.3 to 7.2 (2.1)	266 to 240	—	Reynolds and Neill (1955)
New Mexico	−5 to −60	3 to 8 (1.8)	260 to 237*	249 (av)	Brook et al. (1962)
New Mexico	−30 to −48	4.5 to 6 (1.8)	270 to 250*	—	Krehbiel et al. (1974)
Japan	−50 to −150	4 to 8 (1)	269 to 248	—	Hatakayama (1958)
Japan	−6 to −55	6 to 8 (1)	262 to 249*	—	Tamura (1954)
Japan	−20 (av)	3.5 to 5.5 (1)	276 to 265*	—	Takeuti (1966)
Hong Kong	−25 (av)	4 (1)	272*	210 (av)	Wang (1963)
Australia	−17 (median)	3 (median) (1)	278*	150 (av)	Mackerras (1968)

*Estimated temperatures using climatological averages for the month of interest [U.S. Department of Commerce, 1971].
†The numbers in parentheses indicate the estimated height of the local terrain above sea level.

craft, and the storm draft and temperature structure during a typical penetration of an active thunderstorm at 9 km altitude. Traces on the left of the figure represent the complete penetration path; those on the right are an expanded depiction of the electrical structure bracketing the time of an aircraft lightning strike. Aircraft potential with respect to the environment at several wingspan distances is seen to be as large as 400 kV. The vertical electrostatic field change associated with this strike was in excess of 200 kV/m. The general magnitude of the vertical field at 9 km in this large Florida storm is similar to that shown in the 4 to 6 km range in the smaller Colorado storms (Figure 20-38).

Electric fields extend a considerable distance into the clear air surrounding the storm. The finite electrical conductivity of air due to small ion pair production by the cosmic radiation permits a weak current flow of ions to the cloud boundary in response to the storm-generated field. Effects of the development of boundary layer sheaths of space charge of opposite polarity to the main storm charges due to this current flow may complicate the interpretation of measurements of field and field changes taken external to the storm. Theoretical analyses of these problems are given, for example, by Brown et al. [1971], Hoppel and Phillips [1971], Illingsworth [1971], and Klett

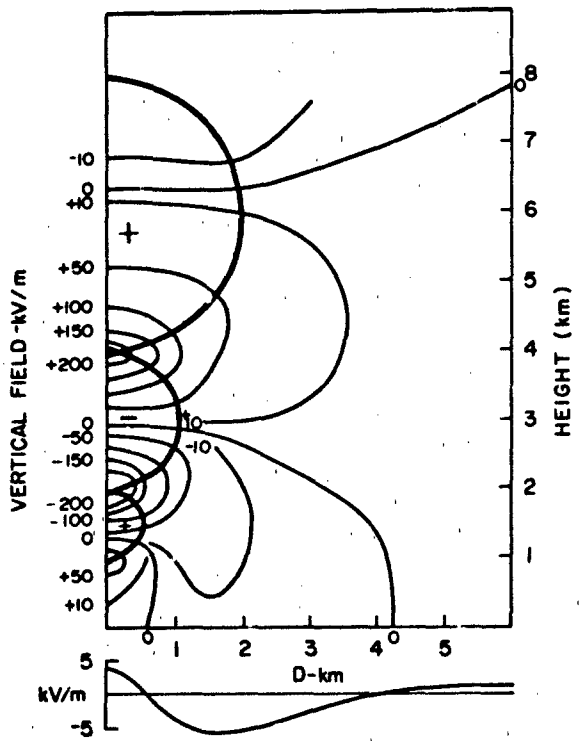


Figure 20-37. Electric field structure of representative European thunderstorms [Simpson and Robinson, 1941].

[1972]. These studies indicate external measurements, due to a field screening effect by the boundary sheath, may yield underestimates of the major cloud charges and of the true charge center regeneration rates that follow a lightning event. The electrical structure of real storm boundaries is further complicated by the mechanically

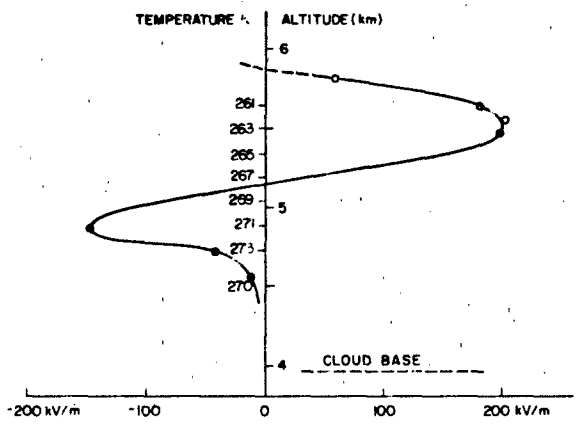


Figure 20-38. Vertical field pattern in Colorado thunderstorms [Parker and Kasemir, 1982].(Reprinted with permission from *IEEE*, © 1982.)

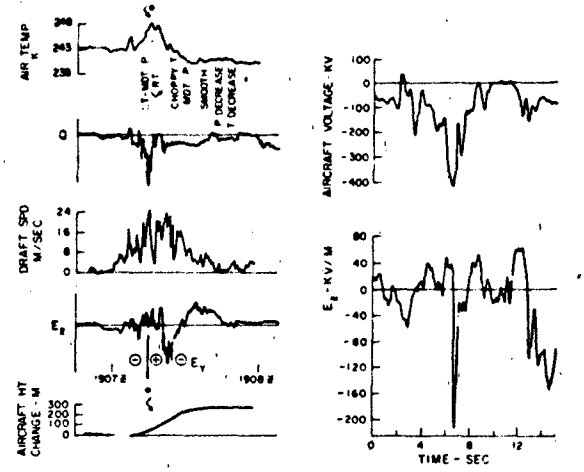


Figure 20-39. Electrical, thermal, and draft structure associated with aircraft lightning strike at 9 km altitude.

forced motions of the cloud particles to which the moving small ions attach. These motions are due to the combined horizontal wind and vertical up and down draft velocities associated with the turbulent cloud motions. Given these possible complications in the interpretation of the data, two examples of the field measured in the clear air at high altitudes by instrumented aircraft are shown in Figures 20-40 and 20-41. The first indicates how three components of the electric field incident on an aircraft vary during a track along the western edge of a large Florida storm. The aircraft radar was used to map several contours of storm reflectivity of the storm's precipitation horizontal cross section structure in the vicinity of 9 km altitude. The three field components, E_x oriented along the flight track, E_y directed along the wing, and E_z the vertical component, undergo systematic amplitud

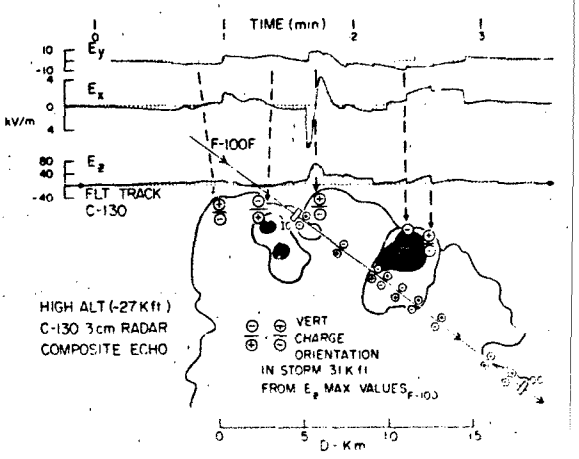


Figure 20-40. Clear air electric field components and radar cross section of a Florida thunderstorm.

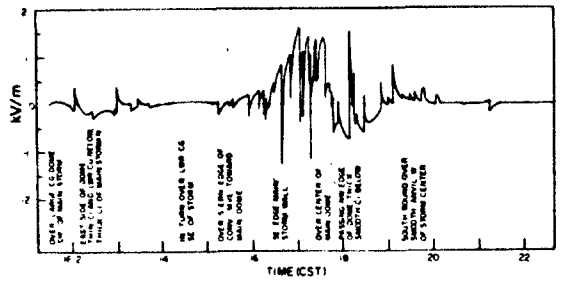


Figure 20-41. Vertical field at 20 km altitude on overflight of an Oklahoma thunderstorm [Fitzgerald and Cunningham, 1965].

variations as the aircraft moves past the major charge concentrations. In general, maxima in E_y and E_z occur when the aircraft is at the closest radial distance to charge centers. The two maxima in E_x are related to boundaries of charge volumes and the zero crossover in E_x occurs near the maxima in E_y and E_z, again indicating the direction to a centroid of the charge center. The charge centers are usually located in the vicinity of the higher values of radar reflectivity but no exact coincidence is found. Field components at midstorm heights may have values to about 50 kV/m in the vicinity of the cloud. Figure 20-41 shows a complex field pattern associated with a U-2 overflight of an Oklahoma thunderstorm. The cloud tops were near 13 km altitude; the aircraft was at 20 km. Field values 7 km above the storm were in the range of 1 kV/m. The sharp discontinuities in the trace are due to the very active lightning activity that was in progress. Other overflight data indicate fields about 3 km above the storm are about 5 kV/m with lightning field changes ranging to 10 kV/m. These fields produce a significant current flow. Gish and Wait [1950] in the first overflight electrical measurements reported the average upward directed current flowing through the upper regions of storms to be about one ampere. The range of currents for a data base of 65 traverses was from 0.1 A to 6 A.

20.2.2 Lightning Characteristics

20.2.2.1 General Phenomenology of the Discharge. The lightning discharge is a complex propagating gas breakdown process. It is thought to originate when a locally strong electric field has developed in the vicinity of small volumes of intense space charge. These differential charge concentrations are trapped on cloud and precipitation particles in the thunderstorm. In midlatitudes lightning usually begins after storm tops have risen to 7 to 9 km and when a radar echo has been present for 10 to 20 minutes. Vertical drafts are normally in the range of 10 to 20 m/s by this time. In high latitudes or winter storm conditions, lightning may occur with cloud tops in the 3 to 4 km range.

Characteristics of the discharge channel are deduced from high-speed photography, optical spectra, and measurement of the acoustic and electromagnetic signals associated with the flash. Further understanding of the physical properties of lightning channels has been greatly advanced through measurements on sparks of several meter lengths as produced by very nigh voltage generators. Cloud ground strike characteristics have been analyzed from numerous measurements taken at instrumented towers. Test aircraft are in use to study properties of intracloud strokes and their electrical interactions with the aircraft.

The terminology which has evolved to aid in description of the very complex gas electrical breakdown phenomena is rather specialized. The definitions of terms given here is adapted from extensive descriptions of lightning characteristics by Schonland [1956], Uman [1984], and Uman and Krider [1982]. The overall cloud to ground lightning event is defined as a flash. It normally has a duration in the range of 0.1 to 1 s. A frequent value is 0.5 s. A majority of these events neutralize 10s of coulombs of negative charge. However, a significant number of long duration high energy flashes neutralize a positive cloud charge. The common negative flashes frequently are made up of three or four discharge components of about 1 ms duration which are separated in time by 40 to greater than 100 ms. These events are defined as the individual strokes. The time sequence of events starts as a preliminary breakdown process, which may begin in the vicinity of the lower positive and negative charge regions in the cloud as shown in Figures 20-34 and 20-37. The channel extension below cloud base for the first stroke in a flash is called a stepped leader based on the appearance of the chain of bursts of luminosity accompanying the channel motion toward the earth. The initial leader breaks down the dielectric strength of air through a sequence of luminous connecting links of about 50 m lengths. Each segment is formed in about 1 μs, with an average 50 μs delay before formation of the next segment. The return stroke begins when this leader approaches the ground and contacts an upward directed discharge from the ground. This discharge is initiated by the very large potential difference of about 10^8 V between the tip of the approaching leader and the earth's surface. A very rapid equalization of charge in the channel then occurs at a speed of about one third the velocity of light. The most spectacular lightning events such as intense luminosity, high peak current and rate of change of current, rapid electric and magnetic field changes, and production of thunder due to strong heating and expansion of the lightning channel are associated with the return stroke. Subsequent strokes usually follow the existing partially ionized channel. Recharging of this channel from new regions of cloud charge is accomplished by a fast continuous process known as the dart leader. Subsequent return strokes usually do not display the branching structure of the first stroke. The time sequence of events for a three-stroke lightning flash as it would be photographed with a moving film and a stationary camera is summarized in Figure 20-42 adapted from Uman [1984].

Intervals between strokes are variable. A representative

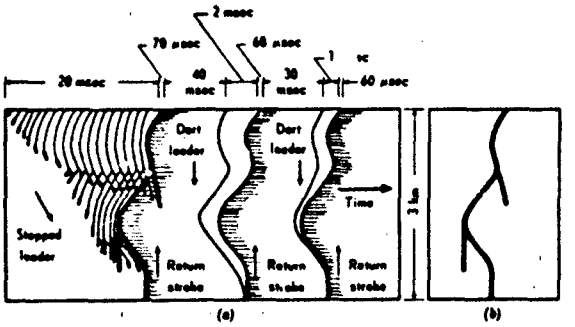


Figure 20-42. Luminous features of cloud ground lightning flash: (a) moving film camera and (b) stationary film camera [Uman, 1984].

value is 50 ms. These are usually preceded by an impulsive dart leader utilizing the existing ionized air to recharge the channel at about a 10^6 m/s rate. The overall strike sequence may involve a time period ranging up to about one second. During the time intervals between the leader-return stroke sequences, the channel can remain weakly conducting for about 100 ms thus permitting a long duration current of the order of 100 A to flow. Most of the actual charge transfer is accomplished by this current. The large peak currents and fast rise time events are associated with the initial portion of the return stroke. Rise times of 0.1 to 2 μs are common. Peak currents may exceed 100 kA although 20 to 30 kA are typical. The statistical distribution of lightning parameters is given in Table 20-8 as prepared by Cianos and Pierce [1972].

20.2.2.2 Characteristics of Radio Frequency Signals due to Lightning. The return stroke current surge generates the most energetic natural radio signals found on earth. The spectral energy has a maximum in the 4 to 8 khz portion of the very low frequency (VLF) band. The stepped leader pulses and a somewhat similar in-cloud channel propagation mode involving current surges known as "K" changes also radiate at VLF with spectral amplitudes about 10% that of return strokes. Long duration strokes also produce extremely low frequency (ELF) signals in the 30 hz to 3 khz band. These VLF and ELF signals can easily be received at distances of many thousand kilometers since they propagate as a guided wave within the spherical shell formed by the earths surface and the lower ionosphere. Very extensive studies of ionospheric properties, radio wave propagation, and long distance direction finding for locating zones of disturbed weather have been conducted using lightning as the signal source. A comprehensive treatment of ELF and VLF propagation and detailed models of natural radio noise sources can be found in Galejs [1972].

The general nature of the radiation produced by a lightning return stroke at a range of some 40 to 100 km from a surface observation point is calculated as a solution to the electromagnetic problem of a current monopole above a flat conducting plane or of the equivalent free-space dipole.

The vertical electric field for this model is given by

$$E_z(t) = \frac{1}{4\pi\varepsilon_o}\left(\frac{M}{d^3} + \frac{1}{cd^2}\times\frac{dM}{dt} + \frac{1}{c^2d}\times\frac{d^2M}{dt^2}\right) \qquad (20.28)$$

Here $M = 2\Sigma qh$ at retarded time $(t - d/c)$ is the charge distribution involved in the flash where q is the charge and h its height, c is the velocity of light, d is the range in meters, and ε_o is the permittivity of free space. The terms in the equation are generally known as the "electrostatic," "induction," and "radiation" components of the field change.

Table 20-8. Statistical distributions for lightning parameters [Cianos and Pierce, 1972].

Parameter[a]	Percentage of Occurrence				
	2%	10%	50%	90%	98%
Number of return strokes	10 to 11	5 to 6	2 to 3	—	—
Duration of flash, ms	850	480	180	68	36
Time between strokes, ms	320	170	60	20	11
Return strokes current,[b] kA	140	65	20	6.2	3.1
Charge transfer per flash, C	200	75	15	2.7	1
Time to peak current, μs	12	5.8	1.8	0.66	0.25
Rates of current rise, kA/μs	100	58	22	9.5	5.5
Current half-value time, μs	170	100	45	17	10.5
Duration of continuing current, ms	400	260	160	84	58
Continuing current, A	520	310	140	60	33
Charge of continuing current, C	110	64	26	12	7

[a] Note that all of the parameters are independent. Some judgment must be made in using the values for consistency.
[b] Values for first stroke.

The relative importance of the terms varies as a function of frequency. The radiation term starts to predominate at range $d = c/2\pi f$.

The associated magnetic field is

$$B_\phi(t) = \frac{\mu_0}{4\pi}\left(\frac{I}{d^2} \times \frac{dM}{dt} + \frac{I}{cd} \times \frac{d^2M}{dt^2}\right) \quad (20.29)$$

The terms are known as the "induction" and "radiation" components for this equation. McLain and Uman [1971] have derived exact equations based on more realistic current variations of the "M" term. At ranges beyond 50 km, Equations (20.28) and (20.29) were shown to be accurate. At distances where the channel length is comparable to the range, the assumptions for derivation of the above equations are invalid. Wait [1959] has shown that the $d \ll \ell$, the vertical field is approximately

$$E_z \simeq \frac{2\mu_0 I}{\pi} \times \ln\left(\frac{2\pi d}{\lambda}\right) \quad (20.30)$$

with a corresponding magnetic field

$$B_\phi = \mu_0 I/2\pi d, \quad (20.31)$$

where f is the frequency, λ the wavelength, and I the channel current.

A variety of more complex models of the lightning radiation source current shape and of propagation velocity in the channel have been developed to permit more realistic calculations of the electromagnetic field waveforms. Details of the recent status of this work can be found in Lin et al. [1980] and Master et al. [1981].

The VLF-LF electric field spectra associated with Florida thunderstorms has recently been obtained for source distances ranging from 1.5 km to 200 km. Table 20-9 taken

Table 20-9. Mean and standard deviations (s.d.) of electric field spectra in decibels for various distances [Serhan et al., 1980].

		First strokes			Subsequent strokes		
Frequency kHz	Distance, km	Number of strokes	Mean, dB	s.d., dB	Number of strokes	Mean, dB	s.d., dB
2	1.5	6	−15.8	4.0			
5	1.5	6	−19.2	3.0			
10	1.5	6	−24.2	3.2			
100	1.5	6	−43.6	3.4			
300	1.5	6	−52.0	3.4			
2	4.0	4	−35.0	2.4	4	−39.8	2.8
5	4.0	4	−38.8	2.2	4	−41.6	2.8
10	4.0	4	−44.2	2.6	4	−47.0	2.0
100	4.0	4	−64.0	1.8	4	−67.2	3.2
300	4.0	4	−71.2	3.4	4	−72.0	2.2
2	7.0	9	−44.6	3.2	8	−50.0	2.0
5	7.0	9	−46.8	1.8	8	−51.0	2.2
10	7.0	9	−53.4	2.2	8	−56.6	3.2
100	7.0	9	−72.8	4.0	8	−72.8	4.2
300	7.0	9	−77.0	3.2	8	−82.2	4.0
2	10.0	6	−51.4	3.8	15	−56.0	4.8
5	10.0	6	−53.4	3.2	15	−57.5	4.2
10	10.0	6	−58.8	3.6	15	−64.6	4.0
100	10.0	6	−75.2	6.0	15	−80.2	7.0
300	10.0	6	−81.4	2.6	15	−85.6	3.4
2	50.0	13	−71.4	6.4	21	−77.0	4.5
5	50.0	13	−69.2	6.4	21	−75.2	4.5
10	50.0	13	−76.2	4.6	21	−79.4	3.0
100	50.0	13	−92.0	6.6	21	−95.0	4.0
300	50.0	13	−103.6	6.6	21	−108.6	4.8
2	200.0	22	−86.0	5.6	34	−89.3	5.0
5	200.0	22	−83.6	5.1	34	−87.0	3.6
10	200.0	22	−88.2	4.1	34	−91.2	3.6
100	200.0	22	−105.6	5.4	34	−108.0	4.0
300	200.0	22	−124.5	6.1	34	−127.0	6.2

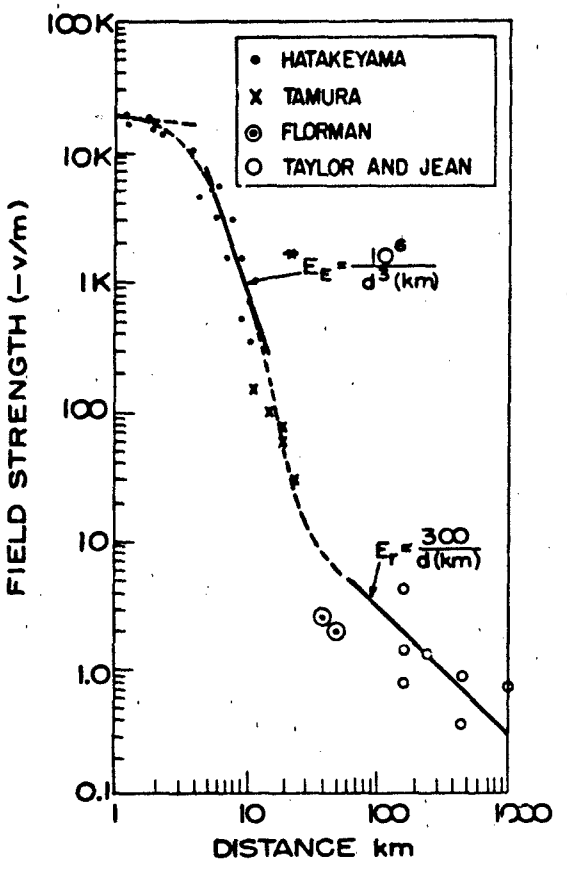


Figure 20-43. Peak electric field variation vs distance from lightning discharges [Watt, 1969].

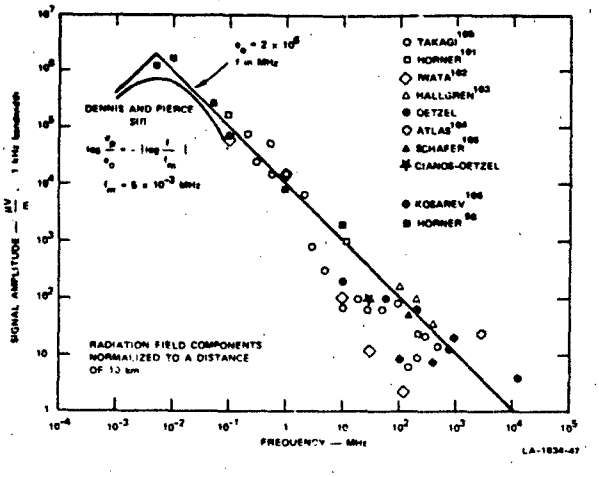


Figure 20-44. Peak received amplitude as a function of frequency [Cianos and Pierce, 1972].

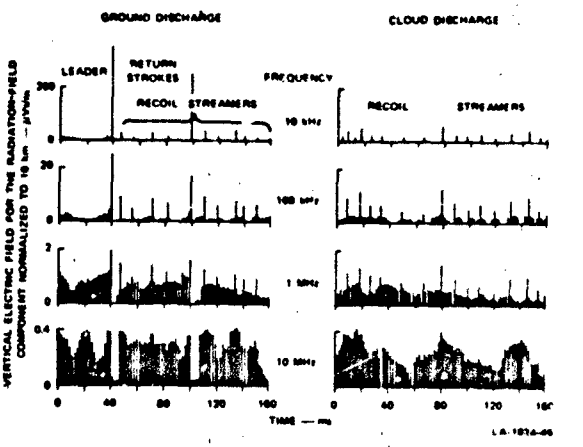


Figure 20-45. The structure of the fields radiated by lightning as a function of time and frequency [Cianos and Pierce, 1972].

from Serhan et al. [1980] gives the db values as 20 times the magnitude (Vs/m) of the frequency spectrum for frequencies ranging from 2 khz to 300 khz.

Earlier data on the peak electric field vs source distance as summarized by Watt [1969] is given in Figure 20-43. The transitions of slope from a nearly constant field associated with a nearby flash to the $1/d^3$ electrostatic field regime and final $1/d$ radiation field are clearly depicted. The peak signal amplitude in a 1 khz bandwidth as a function of received signal frequency is given in Figure 20-44. Data of the investigators listed in the figures have been normalized

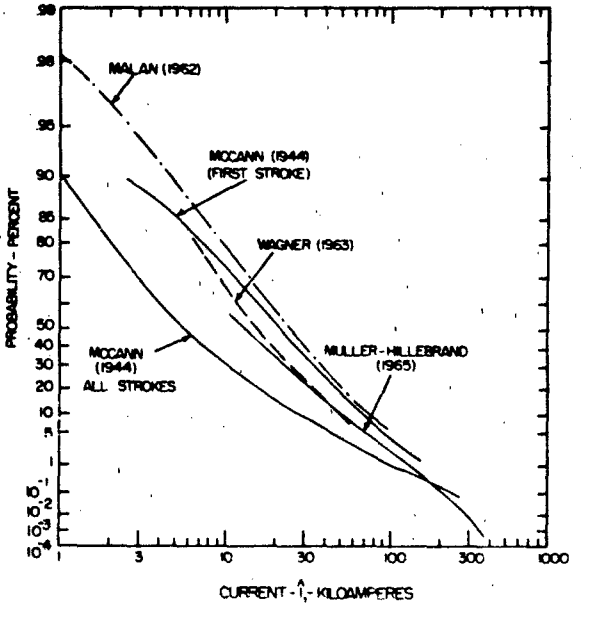


Figure 20-46. Probability distribution of current peaks in a lightning flash [Gaiys, 1972]

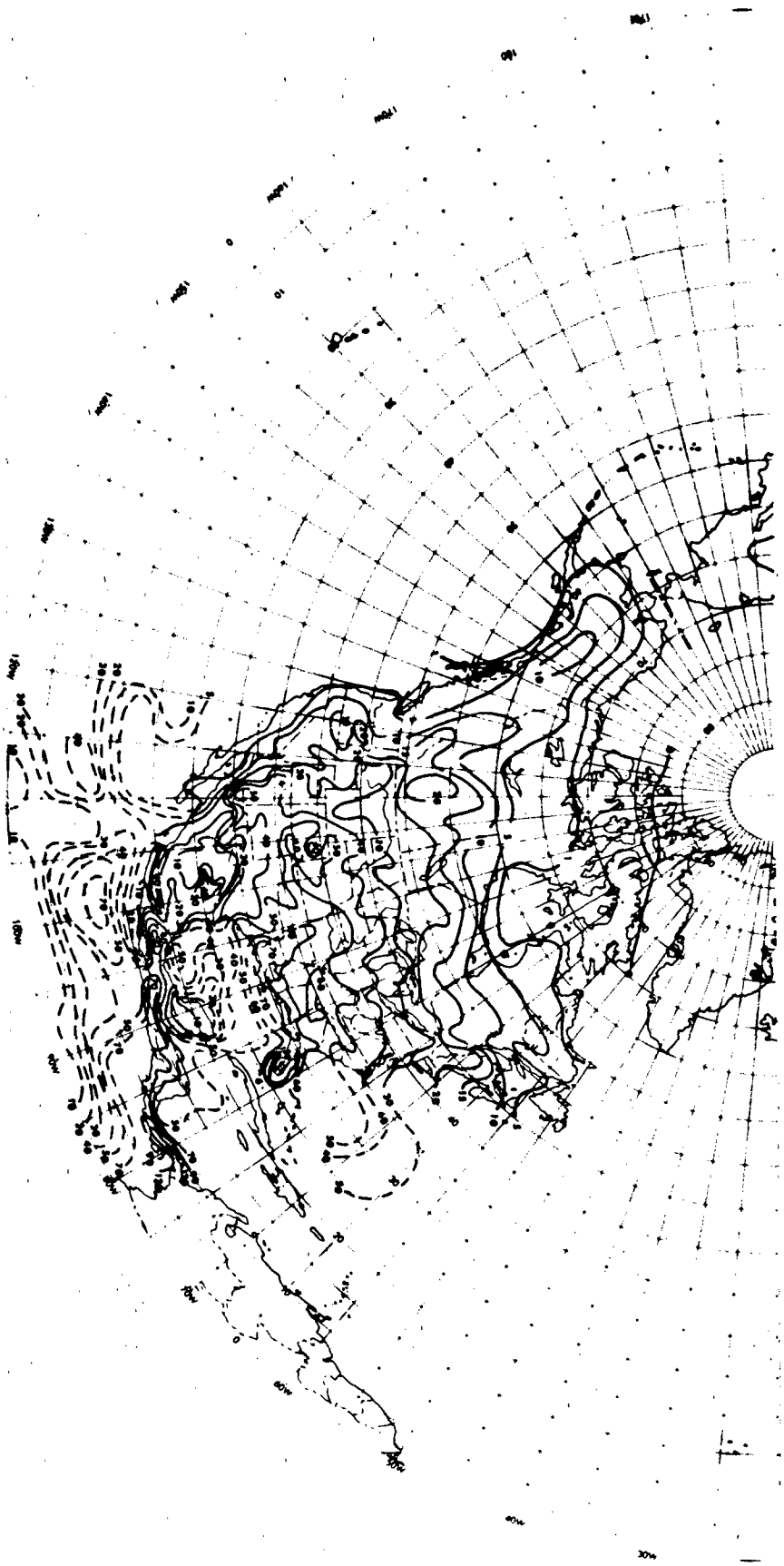


Figure 20-47. Annual distribution of thunderstorm days for North America.

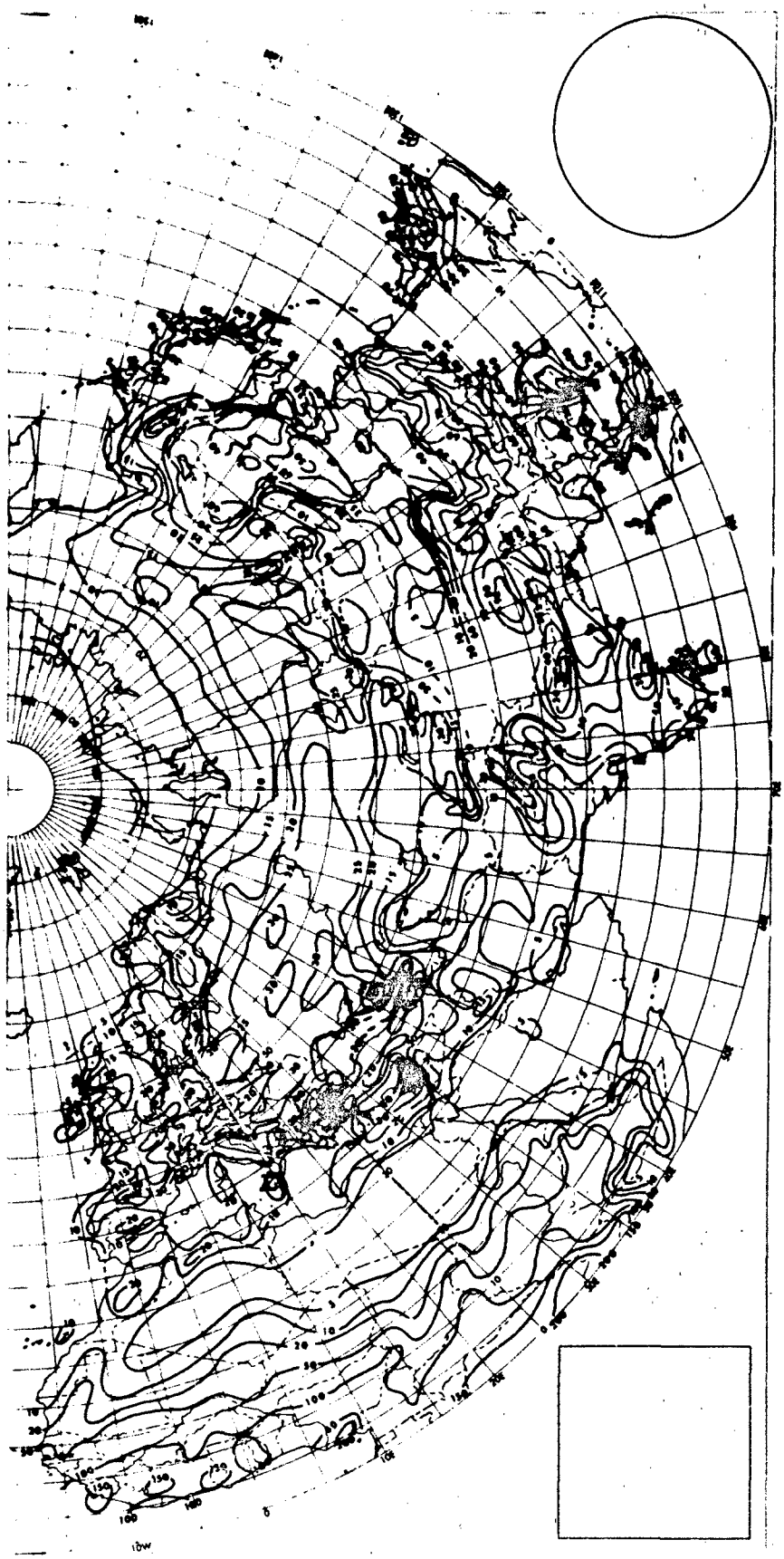


Figure 20-48. Annual distribution of thunderstorm days for Asia, Europe, North Africa, and the Middle East.

to a source distance of 10 km and presented by Cianos and Pierce [1972].

The amplitude variation patterns of the radio noise burst due to a lightning event are complex and have a considerable variation of structure at different frequencies. Figure 20-45, also due to Cianos and Pierce [1972] depicts the general nature of the received fields due to cloud-ground and intracloud flashes at four frequencies. The terminology "recoil streamers" refers to within cloud current surges associated with the channel contacting small highly charged regions. These are the same as the "K" changes previously described.

The magnitude of the various parameters of lightning phenomena vary over a wide range. It is common to assume that the logarithm of the parameter of interest can be fit to the normal or Gaussian statistical distribution. For example the probability that a peak current I exceeds a reference current I_o can be written as

$$P(I > I_o) = \frac{1}{\sigma\sqrt{2\pi}} \int_{I_o}^{\infty} \frac{dI}{I} \exp\left[\frac{-(\ln I - m)^2}{2\sigma^2}\right]$$

$$= \frac{1}{2}\left[1 - \operatorname{erf}\left(\frac{\ln I_o - m}{\sigma\sqrt{2}}\right)\right],$$

where m is the median value of ln I and σ is the standard deviation. Galejs [1972] has summarized data on peak currents involved in power transmission line flashes as measured in several countries. The probability distribution is shown in Figure 20-46.

In recent years the relation of VHF radio noise to channel characteristics has become much better known as a result of studies of incipient lightning streamers and subsequent channel development. As suggested by Figure 20-45, the noise structure is very complex, but it appears as indicated by Proctor [1981] that irregular patterns of short individual pulses and of longer bursts of pulses occur at VHF. Pulse rates range from 10^3 to greater than 10^5/s with individual pulse durations from about 0.1 to 1 μs. The stronger pulses are thought to be associated with ionization produced at the tip of the channel extension into previously neutral air while weaker noise bursts are associated with reintensification of ionization in a decaying channel due to dart leader or recoil streamer reconnections with the main charge sources. Streamer propagation velocities range from 10^4 to 10^7 m/s with current in the channel of the order of 100 A. A charge density of around 10^{-3} C/m of channel length is established in the process.

Kachurin et al. [1974] demonstrated systematic changes in the RF spectra over a range 0.1 to 300 Mhz as convective clouds developed to the thunderstorm stage. Short bursts of noise, usually of 10 to 15 ms duration, occurred for 5 to 10 min preceding lightning activity. At 100 Khz, the burst amplitudes were about 100 μV/m in 1 Khz bandwidth as normalized to 10 km distance. With storm development, a secondary maximum in the pulse duration statistics was found to occur in the 100 to 150 ms interval. Concurrently the 100 Khz signal amplitude increased to about 50 000 μV/m Khz normalized as above.

20.2.3 Precipitation Static Characteristics

Flight through various haze conditions, dust or sand storms, rain, and particularly through snow and ice crystal clouds, causes a considerable accumulation of charge and an aircraft voltage of the order of 100 kV with respect to the environment. Charge accumulation results from collisions of the cloud and precipitation particles with the aircraft. The magnitude of the charging current is related to the mean net charge transferred per impact and the number of particles intercepted per unit time. Thus it is proportional to the frontal area of the aircraft. Discharge takes place through the engine exhaust and by corona from the high curvature portions of the airframe. The corona pulses have a serious effect on radio reception below about 30 Mhz and at times interfere with VHF and UHF reception. An early discussion of this phenomenon and of the development of corona-resistant antenna sheaths and low noise dischargers is given in a series of articles by Gunn et al. [1946]. Development of later dischargers for jet aircraft is described by Nanevicz and Tanner [1964].

High-speed aircraft have additional noise problems. At speeds above 130 m/s the noise level may increase at rates up to the sixth power of the speed [Couch, 1960]. Three noise generating mechanisms are corona pulses, an effect called streamering (flashover to the airframe of charge built up on dieleetric surfaces), and an effect due to antenna pickup of sudden changes in particle-impact voltage. The latter noise is at least 54 dB below the streamering noise level. The spectrum of impact noise decreases inversely with frequency. The streamering spectrum is nearly flat to 100 Khz, rolling off to 3 dB down at 480 Khz. Corona pulses have rise times of a few nanoseconds with decay to 50% in 30 ns at sea level pressure. At 10 km altitude the rise and decay times shift to 20 and 100 ns with the reduced pressure. Wide band RF noise is easily generated by these pulses. It must be decoupled from the aircraft antennas. This is accomplished through use of very sharply pointed static discharge points which act to reduce the amplitude of the corona pulses and by the use of high resistance rods extending from trailing outboard locations on aircraft. These rods provide a mechanical connection for the points and maintain them at the same DC potential as the aircraft. The resistance and distributed capacitance of the installation act as a filter to greatly reduce the radio frequency content of the discharge current.

20.2.4 Distribution and Duration of Thunderstorms

About 50 000 thunderstorms occur every day throughout the world and, on the average, about 2000 storms are in

progress at any time. Large scale charts of the frequency of occurrence of storms over the earth as a function of location and month were published by the World Meteorological Organization, Geneva, Switzerland and were reprinted in the 1960 edition of the Handbook of Geophysics. A thunderstorm day is internationally defined as a day on which thunder is heard. No account is taken of the duration or number of storms on a given day in the tabulation of these statistics. As illustrations of the data available, new more detailed charts have been prepared taking into account airfield data compiled by the USAF Environmental Technical Applications Center (ETAC), some topographic features, and the probable distribution across national boundaries. These are shown as Figures 20-47 and 20-48. Data from Stekol'nikov [1955] and from Arkhipova [1957] have been used for assistance in constructing the maps for the USSR. The accuracy of charts such as these could be considerably improved through addition of satellite and RF direction finder information. In general about one half the storms last for one hour, a quarter for two hours. Complex systems may produce activity for many hours. A comprehensive discussion and monthly summaries of thunderstorm days at a number of stations in the USSR can be found in Arkhipova [1957].

CHAPTER 20

REFERENCES

Anderson, F.J. and G.D. Freier, "Interactions of the Thunderstorms with a Conducting Atmosphere," *J. Geophys. Res.*, **74:** 5390, 1969.

Arkipova, E.P., "Maps of the Geographic Distribution of the Number of Thunderstorm Days in the USSR" (Karty Geograficheskogo Raspredeleniia Chisla Dnei s Grozoi na Territorii SSSR), *Leningrad. Glavnaia Geofizicheskaia Observatoria. Trudy*, **74:** 41–60 1957 (Engl. transl. USWB), 1960.

Aspinall, W.P., "Mechanical Transfer Currents of Atmospheric Electricity," *J. Geophys. Res.*, **77:** 3196, 1972.

Bauer, L.A., "Sunspots and Annual Variations of Atmospheric Electricity with Special Reference to Carnegie Observations 1915–1921," *Res. Dep. Terr. Magn. Carnegie Inst. Wash.*, **175**(5): 359, 1926.

Berger, K., "Novel Observations on Lightning Discharges, Results of Research on Mount San Salvatore," *J. Franklin Inst.*, **283:** 478, 1967.

Bossolasco, M., I. Dagnino, A. Elena, and G. Flocchini, "Solar Flare Control of Thunderstorm Activity," in *Studi in onore di G. Aliverti*, Instituto Universitario Navale Di Napoli, 213, 1972.

Braham, Jr., R.R., "The Water and Energy Budgets of the Thunderstorm and Their Relation to Thunderstorm Development," *J. Meteor.*, **9:** 227, 1952.

Bricard, P.J., "L'equilibre Ionique de la Basse Atmosphere," *J. Geophys. Res.*, **54:** 39, 1949.

Brown, K.A., P.R. Krehbiel, C.B. Moore, and G.N. Sargent, "Electrical Screening Layers Around Charged Clouds," *J. Geophys. Res.*, **76:** 2825, 1971.

Bruce, C.E.R. and R.H. Golde, "The Lightning Discharge," *J. Inst. Elect. Engrs. London*, **88:** 487, 1941.

Burke, H.K. and A.A. Few, "Direct Measurement of the Atmospheric Conduction Current," *J. Geophys. Res.*, **83:** 3093, 1978.

Byers, H.R. and R.R. Braham, *The Thunderstorm*, U.S. Government Printing Office, Washington, D.C., 1949.

Chalmers, J.A., *Atmospheric Electricity, 2nd Edition*, Pergamon Press, New York, 1967.

Chapman, S., "Corona-Point-Discharge in Wind and Application to Thunderclouds," in *Recent Advances in Atmospheric Electricity*, edited by L.G. Smith, Pergamon Press, New York, 1958.

Christian, H., C.R. Holmes, J.W. Bullock, W. Gaskell, A.J. Illingworth, and J. Latham, "Airborne and Ground-Based Studies of Thunderstorms in the Vicinity of Langmuir Laboratory," *Quart. J. Roy. Meteor. Soc.*, **106:** 159, 1980.

Cianos, N. and E.T. Pierce, "A Ground-Lightning Environment for Engineering Usage," Stanford Research Institute, Report 1, Contract LJS-2817-A3, SRI Project 1834, 1972.

Clarence, N.D. and D.J. Malan, "Preliminary Discharge Processes in Lightning Flashes to Ground," *Quart. J. Roy. Meteorol. Soc.*, **83:** 161, 1957.

Clifford, D.W., "Another Look at Aircraft-Triggered Lightning," FAA/NASA/FIT Symposium on Lightning Technology, Hampton, Va., 1980.

Cobb, W.E., "Evidence of Solar Influence on the Atmospheric Electric Elements at Mauna Loa Observatory," *Monthly Weather Rev.*, **95:** 12, 1967.

Cole, R.K. and E.T. Pierce, "Electrification in the Earth's Atmosphere for Altitudes between 0 and 100 km," *J. Geophys. Res.*, **70:** 2735, 1965.

Couch, P.W., "Precipitation Static on Modern Aircraft," in *Electromagnetic Wave Propagation*, edited by M. Desirant and J.L. Michiels, Academic Press, New York, 1960.

Crozier, W.D., "Atmospheric Electrical Profiles below Three Meters," *J. Geophys. Res.*, **70:** 2785, 1965.

Dennis, A.S., and E.T. Pierce, "The Return Stroke of the Lightning Flash to Earth as a Source of VLF Atmospherics," *Radio Sci.*, **68D:** 777, 1964.

DuBro, G.A. (ed.), *Atmospheric Electricity-Aircraft Interaction*, AGARD Lecture Series 110, Technical Editing and Reproduction Ltd., London (Avail. NTIS), 1980.

Fergusor, E.E., [including all references therein], "Ion Chemistry of the Middle Atmosphere," in *NASA CP-2090*, National Aeronautics and Space Administration, p. 71–88, 1979.

Fischer, H.J. and R. Mühleisen, "Variationen des Ionospharen-potentials und der Weltgewitteraatigkeit im elfjahrigen solaren Zyklus," *Meteorologische Rundschau*, **25:**6., 1972.

Fisher, F.A. and J.A. Plumer, *Lightning Protection of Aircraft*, NASA Reference Publication 1008, U.S. Government Printing Office, Washington, D.C., 1977.

Fitzgerald, D.R., "Theoretical and Experimental Studies of Convective Cloud Electrification," Final Report, Contract AF 19(604)-1552, AD 98761, 1956.

Fitzgerald, D.R., "Some Theoretical Aspects of the Relation of Surface Electric Field Observations to Cloud Charge Distributions," *J. Meteor.*, **14:** 505, 1957.

Fitzgerald, D.R., "Probable Aircraft Triggering of Lightning in Certain Thunderstorms," *Monthly Weather Rev.*, **95:** 835, 1967.

Fitzgerald, D.R. and H.R. Byers, "Aircraft Observations of Convective Cloud Electrification," in *Recent Advances in Atmospheric Electricity*, edited by L.G. Smith, Pergamon Press, New York, 1958.

Fitzgerald, D.R. and R.M. Cunningham, "Multiple Aircraft Studies of the Electrical Properties of Thunderstorms," *Proceedings International Conference on Cloud Physics*, Tokyo and Sapporo, Japan, 24 May–1 June, p. 157, 1965.

Flohn, H. "Solare Korpuskularausbruche und Gewitterfrequenz, (Solar Corpuscular Eruptions and Thunderstorm Frequency)", in *Das Gewitter*, edited by H. Israel, Akad. Verlag Gesel., Leipzig, 143, 1950.

Forbush, S.E., "World-wide Cosmic Ray Variations, 1937–1952," *J. Geophys. Res.*, **59:** 525, 1954.

Fulks, G.J., "Solar Modulation of Galactic Cosmic Ray Electrons, Protons, and Alphas," *J. Geophys. Res.*, **80:** 1701, 1975.

Galejs, J., *Terrestrial Propagation of Long Electromagnetic Waves*, Pergamon Press, New York, 1972.

Gathman, S.C., "The Effects of Local Atmospheric Space Charge on Surface Electric Fields," *Trans. (AGU)* **48:** 107, 1967.

Gish, O.H. and G.R. Wait, "Thunderstorms and the Earth's General Electrification," *J. Geophys. Res.*, **55:** 473, 1950.

Golde, R.H. (ed.), *Lightning*, Vol. 1, Physics of Lightning, Vol. 2, Lightning Protection Academic Press, New York, 1977.

Gunn, R., "Electric Field Intensity Inside of Natural Clouds," *J. Appl. Phys.*, **19:** 481, 1948.

Gunn, R., W.C. Hall, and G.D. Kinzer, "Army-Navy Precipitation Static Project," *Proc. Inst. Radio Engrs.*, **34** (4&5), 1946.

Harrison, H.T., "UAL Turbojet Experience with Electrical Discharges," *UAL Meteorology Circular No. 57*, United Airlines, Chicago, 1967.

Harrison, L.P., "Lightning Discharges to Aircraft and Associated Meteorological Conditions," NACA Tech. Note 1001, Washington, D.C., 1946.

Hays, P.B. and R.G. Roble, "A Quasi-Static Model of Global Atmospheric Electricity, 1. The Lower Atmosphere," *J. Geophys. Res.*, **84,** 3291, 1979.

Herman, J.R. and R.A. Goldberg, "Sun, Weather and Climate," *NASA SP-426*, National Aeronautics and Space Administration, Washington, D.C, 1978a.

Herman, J.R. and R.A. Goldberg, "Initiation of Non-Tropical Thunderstorms by Solar Activity," *J. Atmos. Terr. Phys.*, **40,** 121, 1978b.

Hogg, A.R., "Air-Earth Observations in Various Localities—Arch. Meteor.," Wien A., 3.40, 1950.

Holmes, C.R., E.W. Szymanski, S.J. Szymanski, and C.B. Moore, "Radar and Acoustic Study of Lightning," *J. Geophys. Res.*, **85:** 7517, 1980.

Holzer, R.E. and D.S. Saxon, "Distribution of Electrical Conduction Currents in the Vicinity of Thunderstorms," *J. Geophys. Res.*, **57:** 207, 1952.

Hoppel, W.A., "Electrode Effect: Comparison of Theory and Measurement," *Proceedings 4th International Conference on the Universal Aspects of Atmospheric Electricity, II*, 167, 1969.

Hoppel, W.A. and B.B. Phillips, "The Electrical Shielding Layer Around Charged Clouds and Its Role in Thunderstorm Electricity," *J. Atmos. Sci.*, **28:** 1258, 1971.

Illingsworth, A.J., "The Variation of the Electric Field after Lightning and the Conductivity within Thunderclouds," *Quart. J. Roy. Met. Soc.*, **97:** 440–456, 1971.

Israël, H. (ed.), *Das Gewitter*, Akad. Verlag Gesel., Leipzig, 1950.

Israël, H., and H. Dolezalek, *Atmospheric Electricity, Vol II*, Israel Program for Scientific Translations, Jerusalem, 1973.

Jacobson, E.A. and E.P. Krider, "Electrostatic Field Changes Produced by Florida Lightning," *J. Atmos. Sci.*, **33:** 103, 1976.

Kachurin, L.G., M.I. Karmov, and Kh.Kh. Medaliyev, "The Principal Characteristics of the Radio Emission of Convective Clouds," Isv., *Atmos. Oceanic Phys.*, **10:** 1163–1169, 1974.

Kent, D.W. and M.A. Pomerantz, "Cosmic Ray Intensity Variations in the Lower Atmosphere," *J. Geophys. Res.*, **76:** 1652, 1971.

Kitagawa, N. and M. Brook, "A Comparison of Intracloud and Cloud-to-Ground Lightning Discharges," *J. Geophys. Res.*, **65:** 1189, 1960.

Klett, J.D., "Charge Screening Layers Around Electrified Clouds," *J. Geophys. Res.*, **77:** 3187, 1972.

Krider, E.P., G.A. Dawson, and M.A. Uman, "The Peak Power and Energy Dissipation in a Single-Stroke Lightning Flash," *J. Geophys. Res.*, **73:** 3335, 1968.

Krider, E.P., R.C. Noggle, and M.A. Uman, "A Gated, Wideband Magnetic Direction Finder for Lightning Return Strokes," *J. Appl. Meteor.*, **15:** 301, 1976.

Lhermitte, R. and P.R. Krehbiel, "Doppler Radar and Radio Observations of Thunderstorms," *IEEE Trans. Geosci. Electron.*, **GE-17:** 162, 1979.

Lin, Y.T., M.A. Uman, and R.B. Standler, "Lightning Return Stroke Models," *J. Geophys. Res.*, **85:** 1571, 1980.

Livingston, J.M. and E.P. Krider, "Electric Fields Produced by Florida Thunderstorms," *J. Geophys. Res.*, **83:** 385, 1978.

Loeb, L.B., "The Mechanisms of Stepped and Dart Leaders in Cloud-to-Ground Lightning Strokes," *J. Geophys. Res.*, **71:** 4711, 1966.

Magono, C., *Thunderstorms*, Elsevier Press, Amsterdam, 1980.

Malan, D.J., *Physics of Lightning*, English Universities Press, London, 1963.

Master, M.J., M.A. Uman, Y.T. Lin, and R.B. Standler, "Calculations of Lightning Return Stroke Electric and Magnetic Fields Above Ground," *J. Geophys. Res.*, **86:** 12, 127, 1981.

Mauchly, S.J., "On the Diurnal Variation of the Potential Gradient of Atmospheric Electricity," *Terr. Mag. Atmos. Elec.*, **28:** 61, 1923.

McLain, D.K. and M.A. Uman, "Exact Expression and Moment Approximation for the Electric Field Intensity of the Lightning Return Stroke," *J. Geophys. Res.*, **76:** 2101, 1971.

Morita, Y., H. Ishikawa, and M. Kanada, "The Vertical Profiles of the Small Ion Density and the Electric Conductivity in the Atmosphere up to 19 km," *J. Geophys. Res.*, **76:** 3431–3436, 1971.

Mozer, F.S. and R. Serlin, "Magnetospheric Electric Field Measurements with Balloons," *J. Geophys. Res.*, **74:** 4739, 1969.

Muhleisen, R., "Electrode Effect Measurements Above the Sea," *J. Atmos. Terr. Phys.*, **20:** 79, 1961.

Muhleisen, R., "Neue Ergebnisse und Probleme" in *der Luftelektrizitat. Zs. Geophysik*, **37:** 759, 1971.

Naneviez, J.E., and R.L. Tanner, "Some Techniques for the Elimination of Corona Discharge Noise in Aircraft Antennas," *Proc. IEEE*, **52:** 53, 1964.

Ogawa, T. and M. Brook, "Charge Distribution in Thunderstorm Clouds," *Quart. J. Roy. Meteor. Soc.*, **95:** 513, 1969.

Paltridge, G.W., "Experimental Measurements of the Small Ion Density and Electrical Conductivity of the Stratosphere," *J. Geophys. Res.*, **70:** 2751, 1965.

Park, C.G. and M. Dejnakarintra, "Penetration of Thundercloud Electric Fields into the Ionosphere and Magnetosphere, I. Middle and Subauroral Latitudes," *J. Geophys. Res.*, **78:** 6623, 1973.

Parker, L.W., and H.W. Kasemir, "Airborne Warning Systems for Natural and Aircraft-Initiated Lightning," *IEEE Trans. Electr. Compatib.*, **EMC-24:** 137, 1982.

Pierce, E.T., "Electrostatic Field-Changes Due to Lightning Discharges," *Quart. J. Roy. Meteor. Soc.*, p. 211, 1955.

Proctor, D.E., "VHF Radio Pictures of Cloud Flashes," *J. Geophys. Res.*, **86:** 4041, 1981.

Reid, G.C. [including all references therein], "The Middle Atmosphere," in *NASA CP-2090*, National Aeronautics and Space Administration, 27–42, 1979.

Reiter, R., "Solar Flares and Their Impact on Potential Gradient and Air-Earth Current Characteristics at High Mountain Stations," *Pure Appl. Geophys.*, **72:** 259, 1969.

Rosen, J.M., D.J. Hofmann, W. Gringel, J. Berlinski, S. Michnowski, Y. Morita, T. Ogawa, and D. Olson, *Results of an International Workshop on Atmospheric Electrical Measurements*, Dept. of Physics and Astronomy, University of Wyoming, Laramie, 1980.

Rosen, J.M. and D.J. Hoffman, "Balloon-Borne Measurements of the Small Ion Concentration," *J. Geophys Res.*, **86:** 7399, 1981a.

Rosen, J.M. and D.J. Hoffman, "Balloon-Borne Measurements of Electrical Conductivity, Mobility and the Recombination Coefficient," *J. Geophys. Res.*, **86:** 7406, 1981b.

Rust, W.D. and C.B. Moore, "Electrical Conditions Near the Bases of Thunderclouds Over New Mexico," *Quart. J. Roy. Met. Soc.*, **100:** 450, 1974.

Rustan, P.L., M.A. Uman, D.G. C' ders, and W.H. Beasley, "Lightning Source Locations from VHF Radiation Data for a Flash at Kennedy Space Center," *J. Geophys. Res.*, **85:** 4893, 1980.

Sagalyn, R.C. and D.R. Fitzgerald, "Atmospheric Electricity," in *Handbook of Geophysics and Space Environments*, Office of Aerospace Research, U.S. Air Force, 1965.

Sagalyn, R.C. and G.A. Faucher, "Aircraft Investigation of the Large Ion Content and Conductivity of the Atmosphere and Their Relation to Meteorological Factors," *J. Atmos. Terr. Phys.*, **5:**253, 1954.

Sagalyn, R.C. and G.A. Faucher, "Space and Time Variations of Nuclei and Electrical Conductivity of the Atmosphere," *Quart. J. Roy. Meteorol. Soc.*, **82:**428, 1956.

Schonland, B.F.J., "The Lightning Discharge," *Handbuch der Physik*, **22:** 576–628, Springer Verlag OHG, Berlin, 1956.

Schonland, B.F.J., *The Flight of Thunderbolts*, 2nd ed., Clarendon Press, Oxford, U.K., 1964.

Serhan, G.I., M.A. Uman, D.G. Childers, and Y.T. Lin, "The RF Spectra of First and Subsequent Lightning Return Strokes in the 1- to 200-km Range," *Radio Sci.*, **15:** 1089, 1980.

Simpson, G.D., and G.D. Robinson, "The Distribution of Electricity in Thunderclouds, II," *Proc. Roy. Soc.*, **A 177:** 281, 1941.

Smith, L.G. (ed.), *Recent Advances in Atmospheric Electricity*, Pergamon Press, New York, 1958.

Stekol'nikov, I.S., *Study of Lightning and Lightning Protection*, **JPRS 29:** 407, trans. of *Izucheniye Molnii i Grozozashchita*, Moscow, 1955.

Stergis, C.G., G.C. Rein, and T. Kangas, "Electric Field Measurements in the Stratosphere" and "Electric Field Measurements above Thunderstorms," *J. Atmos. Terr. Phys.*, **11:** 77–83, 1957.

Takagi, M., Y. Morita, Y. Kondo, and H. Isnikawa, "Balloon Observations of Atmospheric Electrical Elements and Aerosols in the Lower Stratosphere," *Proceedings of VI International Conference on Atmospheric Electricity*, University Manchester Inst. of Science and Technology, England, 1978.

Takagi, M., Y. Morita, Y. Kondo, and H. Ishikawa, "Balloon Observations of Atmospheric Electrical Elements and Aerosols in the Lower Stratosphere," *5th International Conference on Atmospheric Electricity*, Garmisch-Partenkirchen Germany, 1980.

Takeuti, T., M. Nakano, M. Nagatani, and H. Nakada, "On Lightning Discharges in Winter Thunderstorms," *J. Meteorol. Soc. Japan*, **51:** 494, 1973.

Taylor, W.L., "Radiation Field Characteristics of Lightning Discharges in the Band 1 kc/s to 100 kc/s," *J. Res. Natl. Bur. Stds.*, **67D:** 539, 1963.

Teer, T.L. and A.A. Few, "Horizontal Lightning," *J. Geophys. Res.*, **79:** 3436, 1974.

Tiller, J.A., M.A. Uman, Y.T. Lin, and R.D. Brantley, "Electric Field Statistics for Close Lightning Return Strokes Near Gainesville, Florida," *J. Geophys. Res.*, **81:** 4430, 1976.

Uchikawa, K., "On the Seasonal Variations of the Atmospheric Electric Elements," *Pure Appl. Geophys.*, **100**(8): 54, 1972

Uman, M.A., *Lightning*, Dover, New York, 1984.

Uman, M.A. and E.P. Krider, "A Review of Natural Lightning: Experimental Data and Modeling," *IEEE Trans. on Electromagnetic Capability*, **EMC-24:** 79–112, 1982.

Wait, J.R., *Electromagnetic Radiation from Cylindrical Structures*, Pergamon Press, New York, 1959.

Watt, A.D., *V.L.F. Radio Engineering*, Pergamon Press, New York, 1967.

Watt, A.D., "ELF Electric Fields from Thunderstorms," *J. Res. Natl. Bur. Stds.*, **64D:** 425, 1969.

Webber, W., "The Production of Free Electrons in the Ionospheric D-Layer by Solar and Galactic Cosmic Rays and the Resultant Absorption of Radio Waves," *J. Geophys. Res.*, **67:** 5091, 1962.

Weber, M.E., H.J. Christian, A.A. Few, and M.F. Stewart, "A Thundercloud Electric Field Sounding: Charge Distribution and Lightning," *J. Geophys. Res.*, **87:** 7158, 1982.

Weidman, C.D. and E. P. Krider, "The Radiation Field Waveforms Produced by Intracloud Lightning Discharge Processes," *J. Geophys. Res.*, **84:** 3159, 1979.

Williams, J.C., "Some Properties of the Lower Positive Charge in Thunderclouds," in *Recent Advances in Atmospheric Electricity*, edited by L.G. Smith, Pergamon Press, New York, 1958.

Wilson, C.T.R., "On Some Determinations of the Sign and Magnitude of Electric Discharges in Lightning Flashes," *Proc. Roy. Soc.*, **A 92:** 555, 1916.

Winn, W.P., C.B. Moore, and C.R. Holmes, "Electric Field Structure in an Active Part of a Small, Isolated Thundercloud," *J. Geophys. Res.*, **86:** 1187, 1981.

Winn, W.P., G.W. Schwede, and C.B. Moore, "Measurements of Electric Fields in Thunderclouds," *J. Geophys. Res.*, **79:** 1761, 1974.

World Meteorological Organization, "World Distribution of Thunderstorm Days" Geneva, WMO No. 21, 1956.

Chapter 21

ATMOSPHERIC COMPOSITION

Section 21.1 G.P. Anderson, H.S. Muench
Section 21.2 R.E. Good
Section 21.3 C.R. Philbrick
Section 21.4 W. Swider

The thin envelope of gas that surrounds the earth has evolved into the present state routinely accepted as our atmosphere. While the evolutionary history of the atmosphere has been studied by several scientists, this chapter deals only with our current understanding of the gas species present. The density and temperature structure and the nomenclature adopted to describe it have been discussed in Chapters 14 through 16.

The atmosphere can be viewed as a giant photochemical and dynamical experiment where aeronomers try, with limited snapshot glimpses, to understand and model its complexities. Nitrogen and oxygen in molecular and atomic forms contribute 99% of the total atmospheric composition to altitudes near 500 km. However, it is the minor species which comprise less than 1% of the atmosphere that are most important in establishing the chemical and dynamical behavior of the atmosphere. The minor species of interest include all forms of the molecules that result following dissociation by sunlight of not only the major species but also the less abundant molecules containing hydrogen and carbon. The noble gases also contribute to the sum of the minor species and provide inert tracers for examining the dynamical processes in the upper atmosphere. Sources such as meteorites, interplanetary dust, and the solar wind contribute additional minor species to the top of the atmosphere. Tropospheric clouds, aerosols, ocean spray, anthropogenic sources, and industrial pollution contribute sources of minor species from below. Within the past 10 to 15 years, the potential impact of manmade pollution in modifying the bulk properties of the atmosphere has been realized. The excited states of several significant minor species must be examined separately because of the special importance they have on particular chemical processes.

The role of the minor species in the atmosphere is most important to the following topics:

1. Absorbing and shielding the earth's surface from solar ultraviolet radiation
2. Establishing the radiation balance by absorption and emission of infrared radiation which establishes the thermal structure of the atmosphere
3. Absorption of the extreme ultraviolet radiation which leads, through ion chemistry, to the formation of the D, E, and F regions of the ionosphere

During the past 10 years significant advances in mass spectrometry, optical spectrometry, radiometry, and interferometry have rapidly improved our knowledge of the atmospheric composition. The region between 50 and 140 km can be directly sampled only by rocket borne instruments, but recent advances in remote sounding techniques are beginning to provide additional data of a more synoptic nature in this region. The lower altitudes can be probed by aircraft and balloons while the higher altitudes have been examined during the past 20 years by satellite instruments.

21.1 ATMOSPHERIC OZONE, BELOW 50 KM

Ozone, a minor, variable constituent of the atmosphere, is recognized as a dominant influence on both the present and past character of the earth's environment. It contributes less than 0.0001 % the total atmospheric mass, yet is the critical blocking filter for dangerous high energy solar radiation between 0.2 and 0.3 μm. The ozone redistributes this absorbed ultraviolet energy into the stratospheric thermal structure through collisional heating. Coincident with its absorption properties, ozone also emits radiation at longer wavelengths, serving as an effective cooling agent. Because ozone is chemically responsive to the concentrations of other minor constituents, some of which may be manmade, it has received much scientific and public attention.

Ozone (O_3) is composed of three atoms of oxygen and is gaseous under conditions found in the earth's atmosphere. The bent triatomic structure leads directly to its rich rotation-vibration infrared signature. Actual concentrations vary con-

siderably in time and space; the vertical distribution is dominated by the occurrence of a deep stratospheric layer with maximum values of 5×10^{12} m^{-3} [Krueger and Minzner, 1976]. Since this layer is embedded within the exponential decrease in total atmospheric concentration, ozone is often reported in units of mixing ratio. Typical mixing ratios of 0.03 to 0.10 parts per million by mass (ppmm) are found throughout the troposphere, with an abrupt increase beginning only at the tropopause. A stratospheric maximum mixing ratio of approximately 12 ppmm occurs near 35 km (Figure 21-1 [Dütsch, 1978]). (Note: Data referenced in Dütsch [1978] have been graciously supplied to AFGL by Professor Dütsch and were the basis for Figures 21-1, 21-10 and 21-11.)

The gas is a very powerful oxidant, readily attacking organic compounds (particularly double-bonded structures such as natural rubber) and will harm sensitive plant and animal tissue. Most life forms are tolerant to normal surface ozone amounts, but unusually high doses, even for brief periods, will produce damage [CRC, 1969; Parent, 1978]. For example, humans begin to smell ozone as concentrations increase above levels of 0.03 to 0.06 ppmm, and they might notice some discomfort at 0.05 to 0.1 ppmm. When doses exceed 0.6 to 0.8 ppmm, as might occur under extreme conditions near the ground or at jet aircraft altitudes, a two-hour exposure will produce strong irritation (reversible) to the pulmonary system and will lower resistance to infection. Doses of 15 ppmm are fatal to laboratory animals and will produce severe pneumonia-like illness in humans. Such high concentrations, although frequent in the middle atmosphere, will not occur naturally in the habitable environment.

Units for Atmospheric Ozone: Published reports on atmospheric ozone vary in their units for describing ozone content. The choice depends on the measuring technique as well as the intended user. Vertically integrated or total ozone is normally presented in "Dobson Units;" 1 DU equals 1 milli-atmosphere-centimeter (m-atm-cm), the equivalent vertical depth of pure ozone under standard temperature and pressure. The vertical ozone distribution is sometimes expressed in DU/km, but more frequently is presented as a mixing ratio, partial pressure, mass density, or number density. Units such as micrograms (μg), parts-per-million (ppm), and nanobars (nbar) are used to avoid including powers of ten. Some definitions and conversions are shown in Table 21-1.

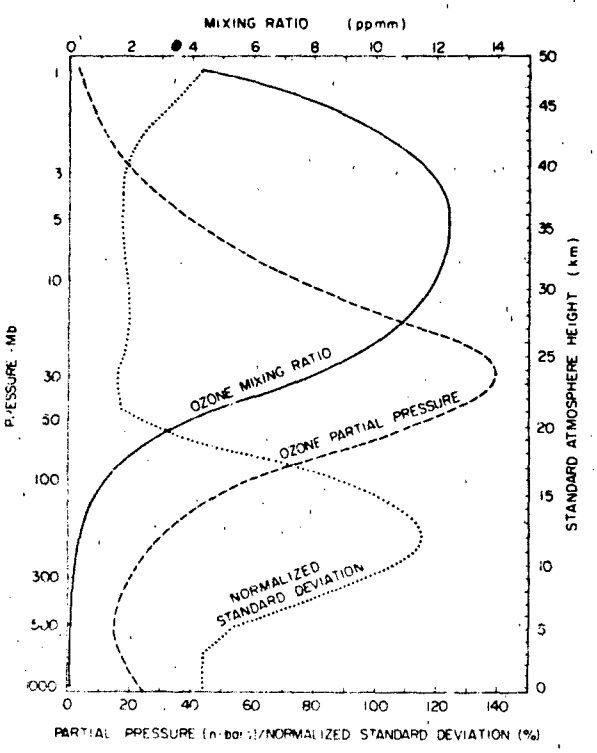


Figure 21-1. Global annual mean vertical distribution of ozone mixing ratio, ozone partial pressure, and normalized standard deviation.

21.1.1 Production

Atmospheric ozone originates primarily from photochemical reactions in the stratosphere, although some minor local tropospheric sources, both natural and anthropogenic, exist. The simplest stratospheric model for ozone production (first postulated by Sidney Chapman in 1930 [Chapman, 1930]) hypothesizes an equilibrium among the major forms of oxygen: O_2, O and O_3; the latter two are referred to collectively as "odd" oxygen. Molecular oxygen provides a reservoir of O atoms through photodissociation:

$$O_2 + h\nu \rightarrow O + O$$

where $h\nu$ is the quantized representation of the available solar ultraviolet energy; (h = Planck's constant and ν = frequency of the photon.* The atomic oxygen rapidly combines with the abundant O_2 in a three-body collision to form ozone:

$$O + O_2 + M \rightarrow O_3 + M.$$

The third reactant (M), most likely N_2 or O_2, is necessary for energy and momentum balance. The primary control to this ozone formation sequence is destruction through photodissociation:

*The frequency of solar energy required for photodissociation is specifically dependent on the structure of the molecule. To bring about O_2 dissociation, a photon has sufficient energy only at wavelengths less than 0.24 μm; for O_3, energy associated with wavelengths between 0.2 and 0.3 μm is most efficient, although other narrow wavelength intervals contribute. For more detail on the molecular physics governing absorption and photodissociation in the atmosphere, see, for example, Goody [1964].

Table 21-1. Ozone units definitions and conversions.

Ozone Parameter	Symbol	Typical Values
Total Ozone	Ω	305 DU or 0.305 atm-cm NTP
Mass Mixing Ratio	r_{3m}	5 ppmm or 5×10^{-6} g O_3/g air
Volume Mixing Ratio	r_{3v}	3 ppmv or 3×10^{-6} cm^3 O_3/cm^3 air
Partial Pressure	P_3	80 nbars or 80 μmb
Number Density	N_3	26×10^{17} molecules O_3/m^3
Mass Density	ρ_3	210 μg/m^3 or 2.1×10^{-7} kg/m^3

$$r_{3v}\ (\text{ppmv}) = 0.604\ r_{3m}\ (\mu\text{g/g}) = P_3\ (\text{nbar})/P(\text{mb}) = \frac{N_3\ T(\text{K})}{7.248 \times 10^{18}\ P(\text{mb})}$$

$$\rho\ (\mu\text{g/m}^3) = 7.97 \times 10^{17}\ N_3\ (\text{molecules/m}^3) = 21.41\ \Delta\Omega/\Delta Z(\text{DU/km})$$

$$P_3\ (\text{nbar}) = 1.732 \times 10^{-3}\ \rho_3\ (\mu\text{g/m}^3)\ T\ (\text{Kelvin})$$

$$\Delta\Omega = 0.790\ P_3 \ln(P_a/P_b) \text{ for a layer bounded by } P_a \text{ and } P_b$$

$$O_3 + h\nu \rightarrow O + O_2.$$

Direct recombination of O and O_3 or O and O is slow.

The observed diurnal modulation in ozone densities at and above the stratopause is a direct response to the loss of the photodissociative Chapman reactions at sunset. The photochemical balance between the odd oxygen species is abruptly shifted toward O_3, producing nighttime increases of approximately 10% at 50 km and 50% at 60 km [Herman, 1979; Lean, 1982; Allen et al., 1984].

The simple Chapman scheme predicts an ozone distribution with a properly located stratospheric maximum; however, the magnitudes of the calculated concentrations are generally higher than the measured amounts. Reactions with and among other trace constituents, including oxides of nitrogen, hydrogen, and chlorine, significantly alter the ozone photochemical balance [Crutzen, 1971; Molina and Rowland, 1974; Nicolet, 1975]. The full chemical description is so complex that the number of recognized reactions now exceeds one hundred and is continually being updated. (See for example NASA/JPL [1982] or Anderson et al. [1985].) The catalytic gas phase nature of many of these reactions has the net effect of accelerating the conversion of O and O_3 to O_2. For example, in the upper stratosphere the hydroxyl (OH) radical can react with ozone to form hydrogen dioxide (HO_2) plus molecular oxygen [Thrush, 1980]. The HO_2 rapidly combines with atomic oxygen, reforming OH and molecular oxygen. This cycle is written in chemical notation as

$$\begin{aligned} OH + O_3 &\rightarrow HO_2 + O_2 \\ HO_2 + O &\rightarrow OH + O_2 \\ \text{NET:}\quad O + O_3 &\rightarrow 2O_2 \end{aligned}$$

Similar catalytic reactions (where OH may be replaced by nitrogen or chlorine radicals) strongly modify stratospheric O_3 concentrations and are dominant in the tropospheric ozone chemistry [Fishman, 1985].

External modifications to the steady-state distribution of any of the catalytic constituents can perturb the momentary or longer-lived equilibrium ozone concentrations. For instance, mixing ratios near the stratopause respond to propagating planetary waves because of the strong temperature dependence of the hydrogen reactions [Barnett et al., 1975]. In addition, the occurrence of solar proton events has been associated with a marked, though temporary, reduction in ozone near the polar stratopause, initiated by downward transport of catalytically active nitric oxide from the mesosphere [Crutzen et al., 1975; Heath et al., 1977; Peters et al., 1981; Solomon et al., 1983]. It is, of course, the effects of long term anthropogenic deposition of catalytic reactants that remains a major concern. Photolysis of chlorofluorocarbons (for example, CCL_2F_2, CCl_2F), used primarily as refrigerants and aerosol spray carriers, provides an increasing stratospheric source of ClO [Stolarski and Cicerone, 1974; Rowland and Molina, 1975]. Since no other effective removal mechanism has been established for these compounds [NAS, 1979], their release at the surface is followed by slow but eventual (~10 years) diffusion and transport into the stratosphere. Ultraviolet solar energy then frees the ClO, initiating the potential reaction schemes with ozone and other constituents. Additional anthropogenic modification to the ozone balance, both in the troposphere and stratosphere, can be expected from the release of nitrogen and carbon compounds related to engine exhausts, fertilizers, biomass burning, and atmospheric nuclear testing [WMO, 1981; Angell and Korshover, 1983]. The extent and magnitude of ozone modulation directly attributed to man has not yet been firmly established theoretically nor experimentally. See Sections 21.1.4 and 21.1.7 for further discussion.

21.1.2 Transport

In the upper stratosphere the photochemical reaction rates are sufficiently fast that ozone equilibrium will usually be approached within hours. However, lower in the atmosphere dynamic processes begin to compete and eventually

dominate [Hartmann, 1978]. A simplified explanation for the loss of photochemical control of O_3 lower in the stratosphere is the disappearance of the ultraviolet sunlight necessary for dissociation of both O_2 and O_3. The ozone and molecular oxygen above 30 km remove almost 90% of the selective energy, reducing both the instantaneous source ($O_2 + h\nu$) and the sink ($O_3 + h\nu$) of odd oxygen. Photochemical reactions affecting ozone, including diurnal modulations, do occur in the troposphere; however, the source molecules are primarily those capable of responding to visible and near-ultraviolet light.

The relatively complete conversion to dynamic control, with ozone acting as a conserved tracer for the atmospheric motions, occurs by approximately 25 km [Cunnold et al., 1980]. The altitude and thickness of the transition zone between the photochemical and dynamic regimes varies with latitude and season, depending largely on the solar geometry and the strength and frequency of planetary wave disturbances. The transition zone is important because much of the poleward transport of ozone is thought to occur in this region [Garcia and Hartmann, 1980].

Eventually, ozone formed photochemically in the upper stratosphere reaches the tropopause and passes into the troposphere. The major exchange mechanism occurs in the vicinity of the midlatitude jet stream [Danielsen et al, 1970; Danielsen and Hipskind, 1980], although its efficiency is not well established. So little immediate blending occurs that measurements of ozone within a complex tropopause fold (Figures 21-2a and 21-2b) can determine whether air samples are of stratospheric or tropospheric origins [Shapiro et al., 1980; Roe, 1981]. Once in the troposphere, the lifetime for ozone is only a few days, since it is rapidly destroyed upon contact with the earth's surface. The removal is most efficient on vegetation, taking tens of minutes, but significant destruction also occurs over oceans, lakes, and even snow [Calbally, 1980].

In addition to its buffer role between the major stratospheric ozone sources and the ultimate surface destruction,

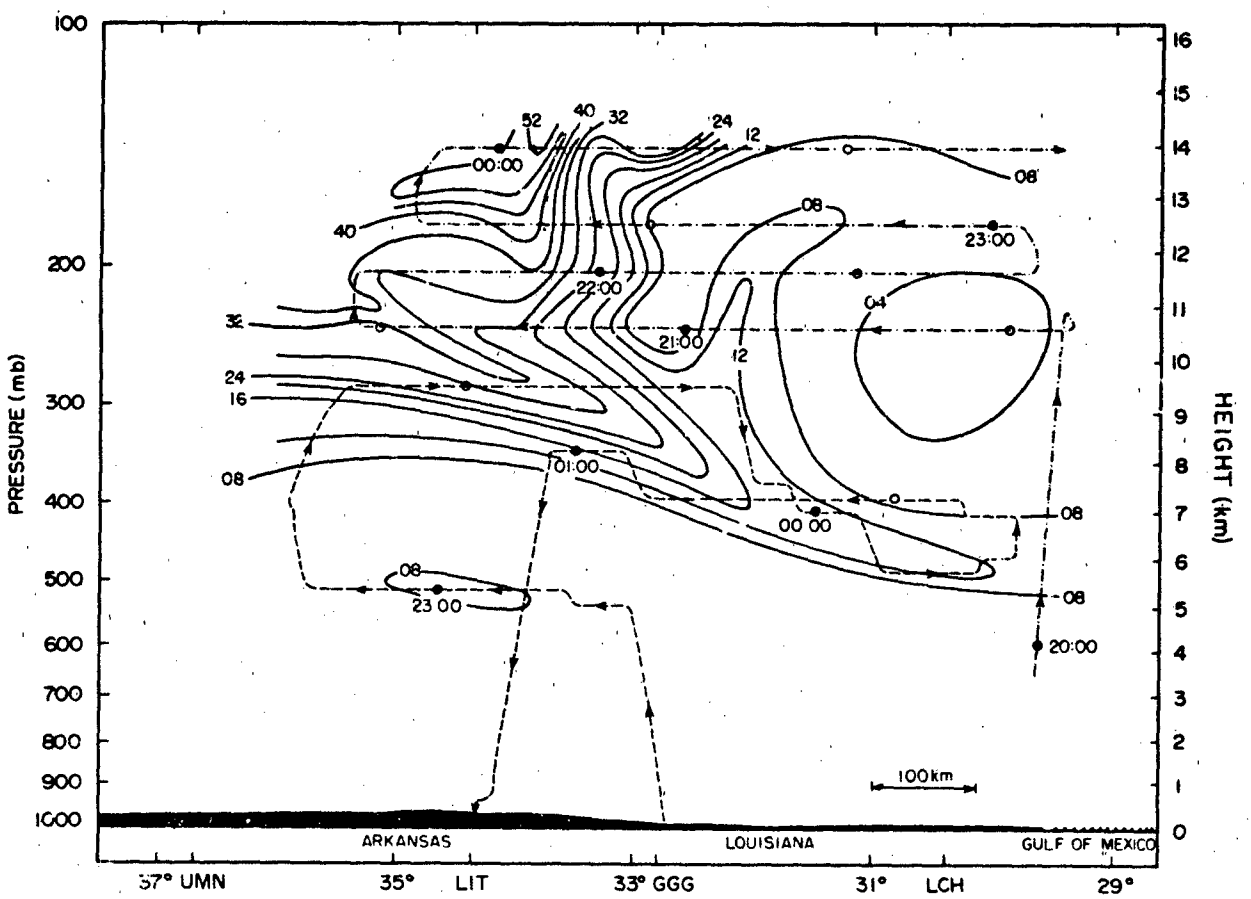


Figure 21-2a. Vertical cross section of ozone concentration (parts per hundred million) from western Missouri to western Louisiana. Analysis based on aircraft observations from 2000 GMT, 16 March 1978 to 0100 GMT, 17 March 1978 illustrates sloping "tongue" of ozone-rich air between 5 and 10 km [after Shapiro et al., 1980].

the troposphere is also an *in situ* chemically active environment. High temperature combustion in power plants, automobiles, and forest fires releases oxides of nitrogen, carbon, and other gases that near-ultraviolet sunlight can dissociate [Crutzen et al., 1975]. If gaseous organic compounds (such as petroleum fumes) are also present, the ozone equilibrium levels are significantly enhanced. These locally-increased concentrations contribute to the photochemical smogs found downwind of major cities. In addition, non-urban tropospheric chemistry can provide a globally-varying, non-anthropogenic *in situ* ozone source [Fishman et al., 1979; Liu et al., 1980]. The magnitude of this natural O_3 production is modified by the availability of such minor constituents as methane (from anaerobic metabolism [Ehhalt and Schmidt, 1978]), CO (from natural combustion sources [Persson, 1974], vegetation [Zimmerman et al., 1978], and NO (from lightning discharges [Noxon, 1976]). Direct O_3 photolysis is also enhanced by cloud and molecular scattering [Thompson, 1984]. Logan et al., [1981] have estimated some of the global pollutant budgets as almost evenly divided between anthropogenic and natural contributions. The balance varies with hemisphere, season, local weather, and chemical environments.

Because much research effort is being expended in understanding the tropospheric/stratospheric ozone cycle and its relation to man, comprehensive reviews of its chemistry and dynamics are published frequently. See, for instance, Brasseur [1982]; Brasseur and Solomon [1984]; *The Stratosphere, 1981* [WMO, 1981]; Logan et al., [1981]; *The NASA Assessment Report* [1984]; Whitten [1985]; Bojkov [1984].

21.1.3 Ozone Measurement

The complex control mechanisms—dynamic, radiative, and photochemical—that govern the global distribution of ozone prohibit a simple prediction of any local ozone en-

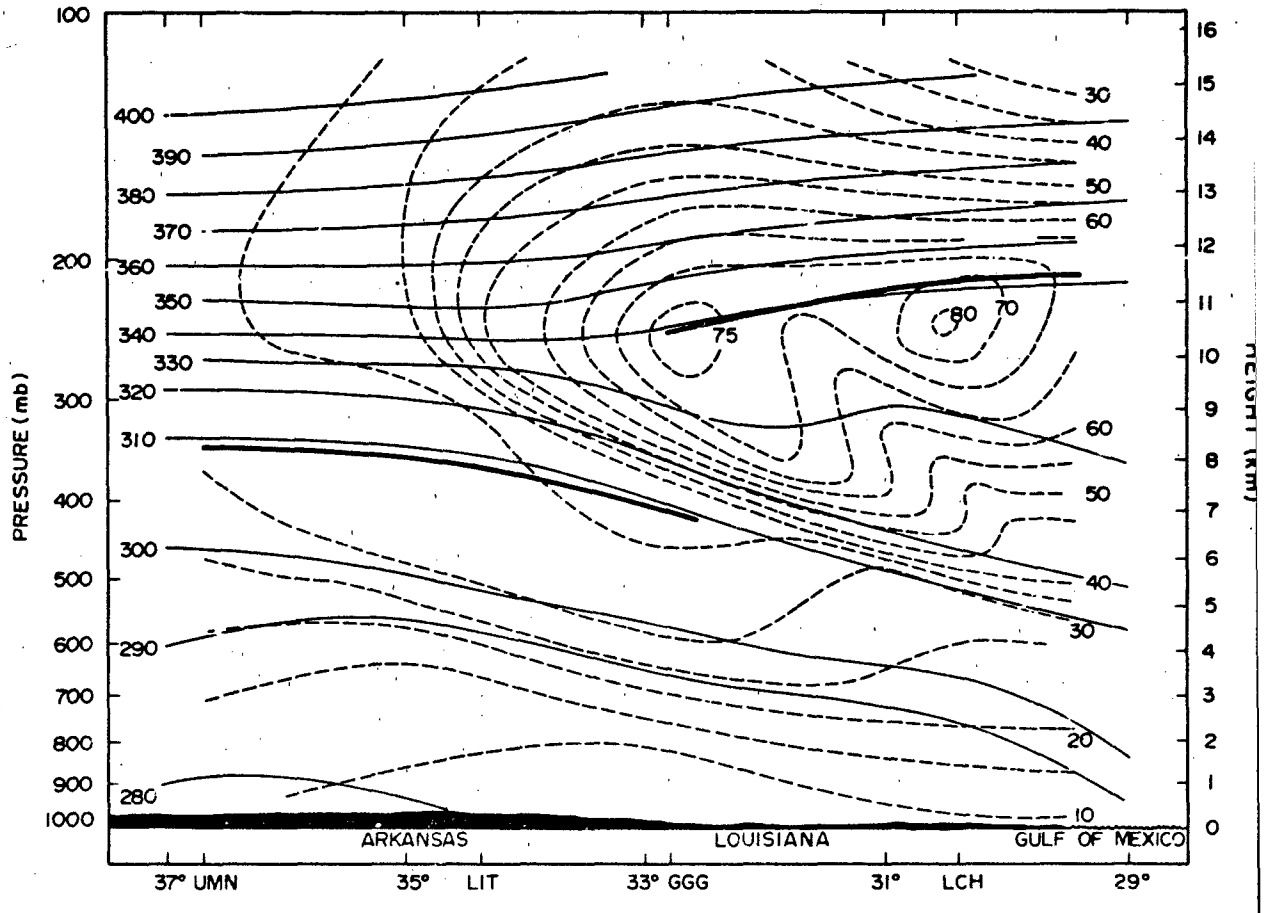


Figure 21-2b. Vertical cross section of potential temperature (thin solid lines), wind speed (dashed lines), and tropopause (heavy line) corresponding to previous figure. Potential temperature and static stability of air in the ozone tongue is characteristic of lower stratosphere air normally found north of the jet stream [after Shapiro et al., 1980].

vironment (see Section 21.1.7 on modeling efforts). A large complement of measurement techniques has been developed for both integrated column and vertical profile definition. The distinction is important because the total ozone overburden serves as the absorbing buffer for the solar radiation, while the vertical distribution governs energy deposition. Instrumentation based on the molecular, optical, and/or chemical properties of ozone has been adapted to ground and space platforms, including aircraft, balloons, rockets, and satellites.

In particular, the optical properties of ozone, both in the ultraviolet and infrared, provide rich instrumental applications. Spectral measurements of the differential absorption in the UV Hartley (0.2 to 0.3 μm) and Huggins (0.3 to 0.35 μm) bands are employed routinely. Beginning in the 1930s, ground station networks [Dobson, 1930; Dütsch, 1973] have been inferring atmospheric total ozone from observed variations in the measured solar irradiance within these wavelength intervals ("Unkehr/Dobson" and "M83" measurements). Observations of the zenith angle modulation of that same irradiance provide broad resolution (5–10 km) vertical structure information [Mateer, 1965].

Satellite instrumentation has created the opportunity for global coverage, a particular advantage when examining a constituent as variable as O_3 [Krueger et al., 1980]. Both total column and/or vertical distributions have been inferred from UV Hartley and Huggins band measurements [Anderson et al., 1969; Heath et al., 1973; Rusch et al., 1983]. These same absorption techniques, when used on balloons and rockets [Krueger, 1973; Mentall et al., 1980; Lean, 1982] provide an invaluable calibration test for the sounding networks. (For information on general mathematical formulations and algorithms for inversion of remotely-sensed atmospheric data, see for example Twomey [1977], Deepak [1977], Rodgers [1976], or Gordley and Russell [1981]).

Another satellite-based ultraviolet system infers ozone distributions by observing the occultation of a light source (solar or stellar) crossing the earth's horizon [Hays and Roble, 1973; Aiken et al., 1982; Millier et al., 1981]. For this method the vertical resolution is 4 to 6 km. Ultraviolet lasers have also been used to remotely probe the ozone layer from the ground (with potential satellite applications [Remsberg and Gordley, 1978]); the synchronous, atmospherically reflected laser signal [Pelon and Megie, 1982] provides both altitude (that is, distance from the source) and density information. An *in situ* application of ozone's ultraviolet absorption properties, used primarily on large balloons and rockets, includes both source and detector separated by a chamber filled with a local air sample [Ainsworth and Hagemeyer, 1980; Profitt and McLaughlin, 1983; National Academy Report, 1982; Robbins 1980]. Pumping efficiency then governs the vertical resolution (< 20 m, see Figure 21-12).

Infrared systems to measure atmospheric ozone employ the thermal emission of the triatomic molecule, usually in the strong vibration-rotation band near 9.6 μm. Infrared rocket and satellite instruments have determined both total ozone [Prabhakara et al., 1970; Lovill et al., 1978; Crosby et al., 1980] and vertical profiles [Gille et al., 1980b; Remsberg et al., 1984; Stair et al., 1984] using nadir (vertical) or limb viewing, respectively. The thermal dependence of the emission requires simultaneous temperature sounding; however, it is independent of solar insolation, providing full diurnal coverage. Vertical resolution is approximately 5 km throughout the stratosphere and lower mesosphere (12–65 km).

The longer wavelength (infrared, millimeter, and microwave) emission/absorption properties of the ozone molecule (see Harries [1980] for a general discussion) are also atmospherically useful. The shape and equivalent width of the spectrally-isolated lines provide independent measures of both density and pressure. In particular, the emission spectrum in the wavelength interval 50–300 Ghz (1–6 mm), as measured from the ground [Parrish, 1981] has been proposed for inclusion in the standard global sounding network. Similarly, molecular line parameters have been measured in absorption against the solar background spectrum [Farmer et al., 1980; Louisnard et al., 1983; Weinrab et al., 1984].

Fluorescence radiation at 1.27 μm, emanating from ozone photolysis ($O_3 + h\nu \rightarrow O_2 + O$), has been measured on a limited number of high-altitude rocket experiments [Weeks et al., 1978]. More recently this experiment, as part of an instrument complement on a spinning satellite, has measured global O_3 mixing ratio profiles up to 90 km [Thomas et al., 1983].

Small *in situ* chemical detectors (which record iodide or chemiluminescent reactions to the presence of ozone) are frequently used as portable ground and aircraft sensors or as expendable balloon and rocket-borne ozonesondes [Regener, 1960; Komhyr, 1971; Hilsenrath, 1980]. Their calibration is demanding, often requiring normalization of the profile to some other measurement; however, fine scale vertical structure (<50 m) measurements can be achieved at low cost. Mass spectrometers, because of their size and complexity, are not used routinely as sondes; they can be mounted on either aircraft or balloons and provide detailed *in situ* observations, including altitude profiles of isotopic abundances [Mauersberger, 1981].

A summary of the instrument systems currently being used for climatological analyses of ozone in the stratosphere and troposphere is presented in Table 21-2. The values for relative accuracy are based on instrument intercomparisons as well as observational consistency. They do not include the systematic differences (6%–8%) found between the total column ozone measured by Dobson ground-based instruments, which are considered the standard, and the ozone measured both by the back-scattered UV and infrared satellite instruments. The accuracies stated for the electrochemical and chemiluminescent sondes have been enhanced by an independent measurement of total ozone as part of

Table 21-2. Major data sources for atmospheric ozone below 50 km.

Type of Observation	Period	Obs./Mo.	Vertical Range	Vertical Resolution	Horizontal Resolution	Relative Accuracy	Comments
1. World Ozone Network (a)							
Dobson Spectrophotometer	1930–84 +	700	0–50 km	(Total Ozone)	Point	± 3%	Cross-calibration of instruments difficult
M83 Instruments (filter)	1960–84 +	250	0–00 km	(Total Ozone)	Point	± 8%	Problems with filter stability
Dobson Spectrophotometer-Umkehr	1935–84 +	60	0–60 km	6 km	Point	± 6%	Requires clear sunrise or sunset
Electro-chemical Ozonesondes	1965–84 +	30	0–30 km	0.3	Point	± 3%	Pump efficiency problems above 25 km
2. AFGL (North America Network (b)							
Chemiluminescent Ozonesonde	1962–64	40	0–30 km	0 km	Point	± 8%	Problems with chemiluminescent materials
3. NASA-OGO4 (BUV) (c)	1967–69	1000	30–50 km	7 km		± 10%	Monthly means (daytime only)
4. NASA Nimbus 4 Satelite System (c)							
Back-Scattered Ultraviolet (BUV)	1970–77	18 000	0–60 km	(Total Ozone)	200 km	± 3%	Excellent global coverage (daytime only)
Back-Scattered Ultraviolet (BUV)	1970–77	1800	22–60 km	7 km		± 6%	Excellent global coverage (daytime only)
IR Interferometer Spectrometer (IRIB)	1971–72	40 000	0–60 km	(Total Ozone)	100 km	± 3%	Excellent global coverage (day and night)
5. USAF DMSP Satellites, F1-F4 (d)							
Infra-Red Spectrometer	1977–80	3 000 000	0–60 km	(Total Ozone)	100 km	± 7%	Cross-path scanning, excellent coverage, day and night
6. NASA Nimbus 7 Satellite System (c)							
Solar Back-Scattered UV (SBUV)	1978–84 +	32 000	0–60 km	(Total Ozone)	200 km	± 3%	Excellent global coverage (daytime only)
Solar Back-Scattered UV (SBUV)	1978–84 +	32 000	22-60 km	7 km	200 km	± 6%	Excellent global coverage (daytime only)
Limb IR Monitor of Stratosphere (LIMS)	1978–79	32 000	12-60 km	4 km	200 km	± 7%	Excellent day-night coverage, requires cryogenics
Total Ozone Mapping Spectrometer (TOMS)	1978–84 +	1 200 000	0-60 km	(Total Ozone)	60 km	± 3%	Cross-path scanning, daily global coverage
7. Solar Mesospheric Explorer (SME) (c)							
Back Scattered Ultraviolet (BUV)	1982–84 +	30 000	30–60 km	7 km	200 km	± 5%	Global covering,
Near Infrared	1982–84 +	30 000	30–60 km	7 km	200 km	± 5%	limb scanning (daytime only) Global covering, limb scanning (daytime only)

Data Sources

(a) Archived at World Ozone Data Centre, Atmosphere Environment Service, 4905 Dufferin St., Downsview, Ontario, Canada M3H5T4a (Data published bimonthly in *Ozone Data For The World*)

(b) Archived at NOAA Environmental Data and Information Center, National Climatic Center, Federal Building, Asheville NC 28801

(c) Archived at National Space Science Data Center, Goddard Space Flight Center, Code 601, Greenbelt MD 20771

(d) For information, Dr. J. Lovell, PO Box 808, Lawrence Livermore Laboratories, Livermore CA 94550

the calibration. Further details on ozone instrumentation may be found in *The Stratosphere 1981* [WMO, 1981] or more recent reviews.

21.1.4 Total Ozone: Its Global Distribution and Variability

The most complete ozone data set is that for total ozone. Ground-based measurements have been made for over 40 years and a global network (albeit coarse) has existed for 25 years [London et al., 1976]. With the introduction of satellite sensors (see Table 21-2), the frequency, spatial coverage, and relative accuracy of such soundings have all been improved [Hilsenrath and Schlesinger, 1979]. Combined ground and satellite determinations allow the formulation of climatologies that describe the distribution and variability of total ozone on scales ranging from hours to years.

Hemispheric contour plots of the annually averaged column amounts of ozone and their normalized standard deviations are shown in Figures 21-3a and 21-3b. The distributions graphically demonstrate the impact of global dynamics on ozone in the troposphere and lower stratosphere (the contributions from these lowest altitudes to the total column is approximately 70%). Subsequently, the features of any total ozone distribution correlate well with those of similarly formulated maps of standard meteorological data near the tropopause [Miller et al., 1979]. For example, the structure in Figure 21-3 is negatively correlated with the annual mean height of the tropopause [Dobson, 1963]. The hemispheric asymmetries that persist in these annual means are primarily

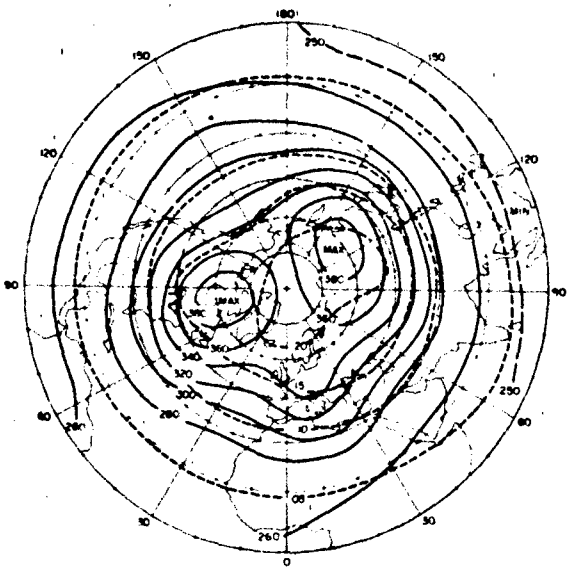


Figure 21-3a. Total ozone northern hemisphere, annual mean (m atm cm). Dashed lines indicate normalized standard deviation.

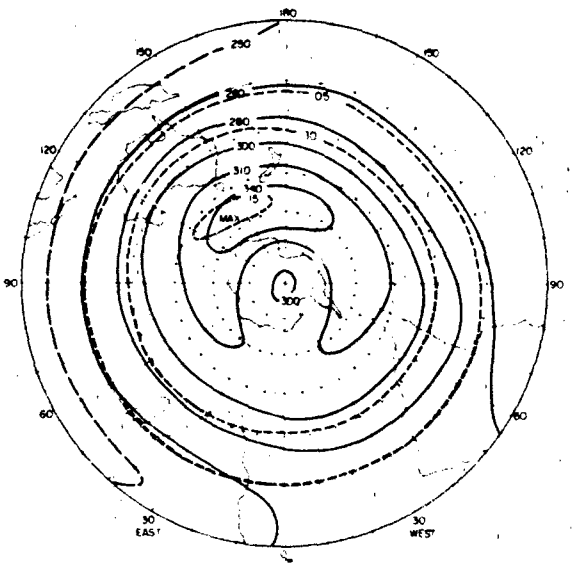


Figure 21-3b. Total ozone southern hemisphere, annual mean (m atm cm). Dashed lines indicate normalized standard deviation.

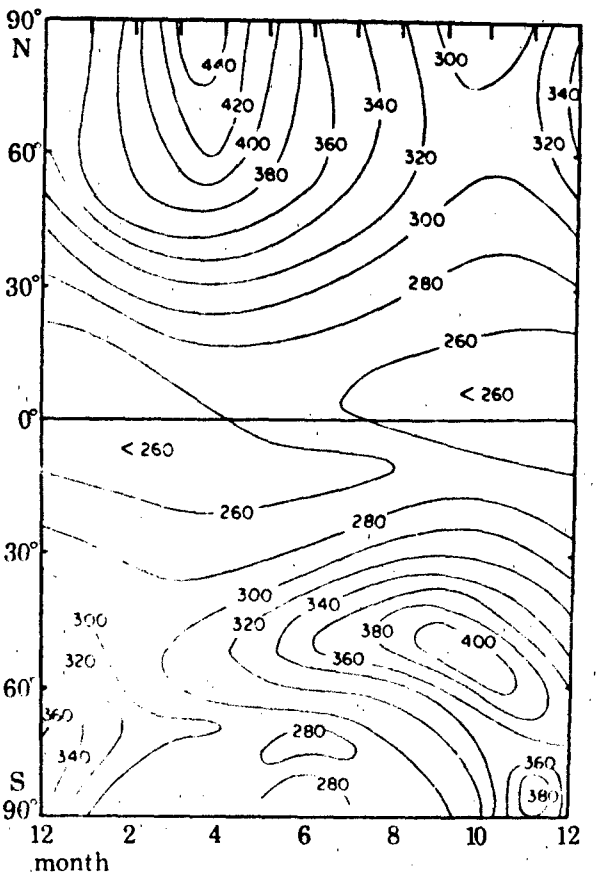


Figure 21-4. The observed zonally averaged distribution of total ozone (m atm cm). The low latitude minima and the hemispheric differences immediately reveal the importance of the dynamics of transport processes.

a consequence of geographically-stable, planetary-scale winter weather patterns [Hartmann, 1977]. See Holton [1980] for a more complete description of the controling dynamic processes.

The local variability in total ozone can be very large and, again, highly correlated with the instantaneous (synoptic) global weather system. Similarities between the features of the daily, satellite-derived total ozone measurements and the standard synoptic height maps of the 200 mb pressure surfaces have been observed [Shapiro et al., 1982]. If instantaneous O_3 column amounts can be rigorously related to the global tracking of planetary waves, the pertinent satellite data could be added to the meteorological inputs for weather prediction.

Total ozone climatology exhibits natural variations in both amplitude and period, ranging from daily to interannual, and perhaps even solar cycle [Wilcox et al., 1977; London and Reber, 1979]. The dominant influence on the variability is seasonal, as depicted in Figure 21-4. A harmonic analysis of ground (or satellite) data bases provides insight into the latitudinal dependence (Figure 21-5). At middle and high latitudes the annual period dominates, with additional power in the aperiodic component. Near the equator the total variability is small, allowing the statistical identification of semi-annual, annual, and quasi-biennial (26 month) waves [Angell and Korshover, 1983]. Anthropogenic modification of total ozone has yet to be established [Reinsel, 1981; Bloomfield et al., 1983]; its magnitude, as predicted by theoretical calculations, would be very small (a decrease of <0.1%/year, Logan et al., 1978; Penner, 1982; Wuebbles et al., 1983; NASA Academy Report, 1984), but over long time scales could be dramatic. Given ozone's large variability plus current instrumental limitations (both precision and intercalibration: see Table 21-2) man's influence is not yet detectable.

21.1.5 Vertical Profiles

The modified "Chapman" photochemical mechanism is sufficient to describe the basic vertical distribution of ozone; its dominant feature, as discussed previously, is the pronounced ozone layer, a region in the stratosphere of maximum concentrations (actually resulting from complex photochemical production and subsequent transport). However, measurable amounts of O_3 also occur throughout the troposphere and into the mesosphere. The climatological variability within this framework is similar to that for the total column, with additional complications from the altitude-dependent chemistry and dynamics.

An annual, globally-averaged vertical distribution (as in Figure 21-1) has a broad peak in mixing ratio (ppmm) near 35 km (note that the choice of units governs the definition of the maximum). The normalized standard deviation of the composite profile suggests that the variability is greatest near 12 km and 45 km (related respectively to dynamics at the tropopause and chemical responses to thermal and/or solar disturbances below the stratopause). The annual variances in Figure 21-1 are minimized by the geographical dominance

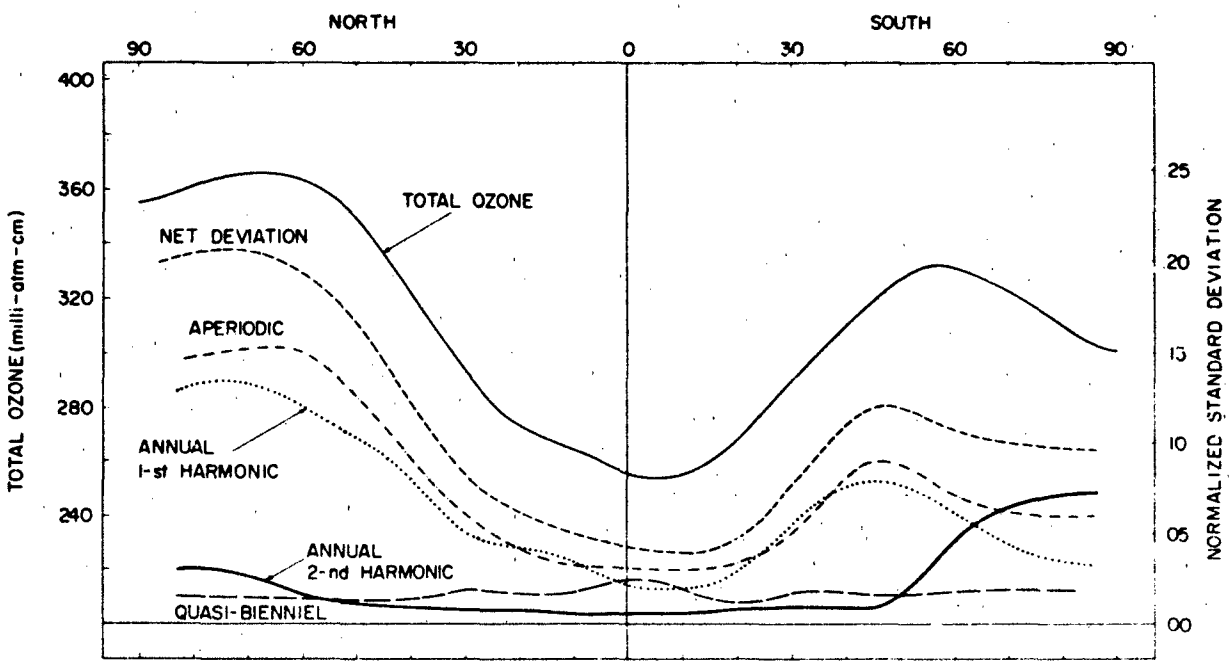


Figure 21-5. Meridional profile of total ozone and the components of temporal variability.

of large areas of relatively stable ozone profiles in the summer and low latitude; winter and equinoctial profiles actually exhibit much greater natural excursions over both long (seasonal) and short (hours to days) time scales.

As with total ozone, the first order variations about a basic vertical distribution can be ascribed to seasonal effects [Dütsch, 1978; Klenk et al., 1983]. Systematic satellite measurements provide the opportunity to examine these annual progressions as a function of latitude, particularly at and above 10 mb (approximately 30 km). For instance, sample one-month zonal mean distributions for 1979 (Figures 21-6a, b, c, d), based on satellite-derived, backscattered ultraviolet data [McPeters et al., 1984], encompass the seasonal features of the middle and upper stratosphere. Towards the winter pole the mixing ratio contours maximize at increasingly higher altitudes (often above 40–45 km, that is, 2–3 mb), in response to weaker photodissociation and stronger transport. The summer hemisphere, however, exhibits a stably located maximum. This imbalance reverses hemispheres with the seasons and is symmetric at the equinoxes. A statistical analysis of a similar data set for the years 1970–71, in combination with balloon measurements [Klenk et al., 1983], suggests that the zonal mean variance can be typically reduced to a four-parameter function (including latitude and day of the year) for selected pressure levels. The 1971 seasonally-dependent profiles for high,

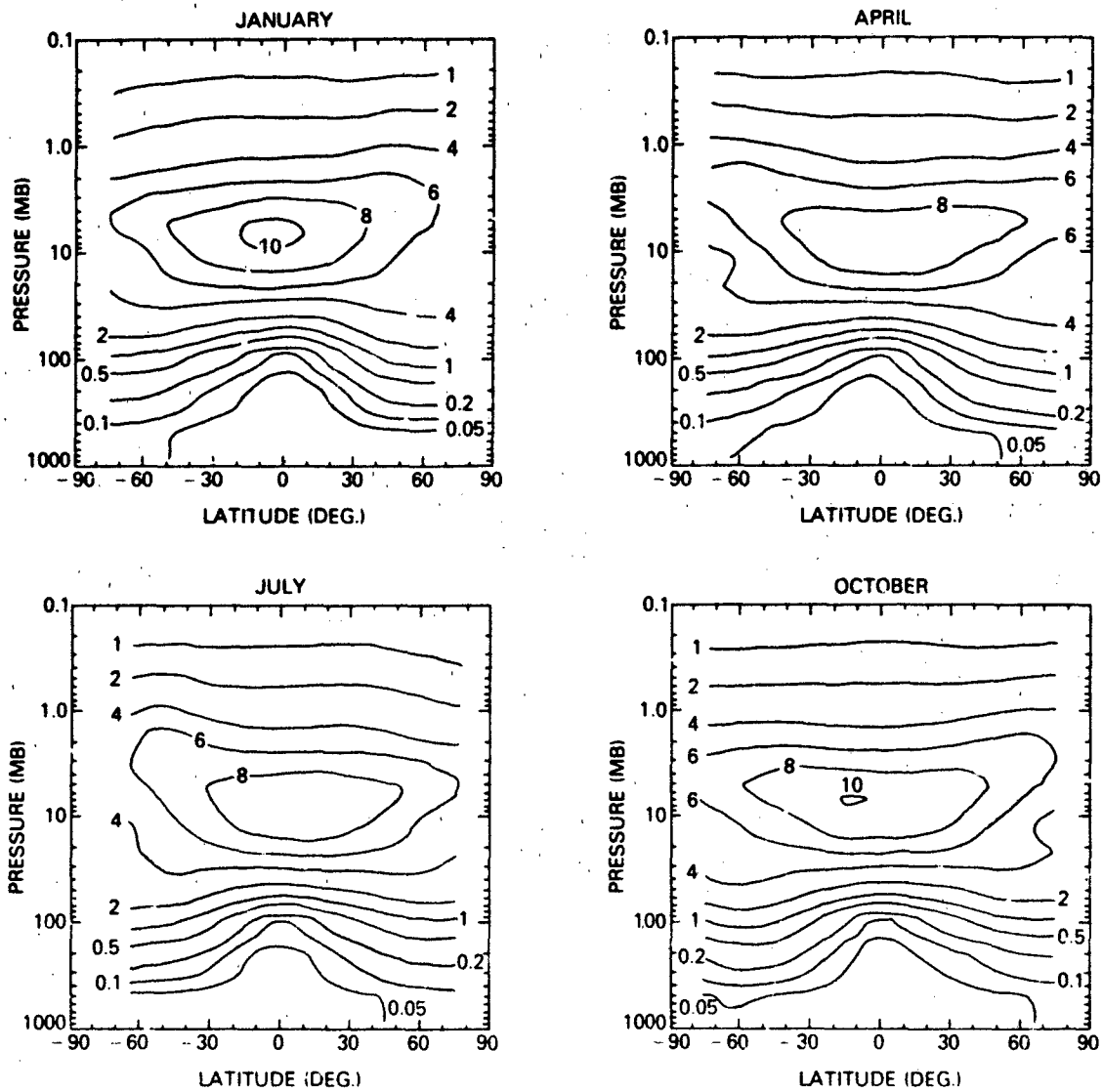


Figure 21-6. Latitude-height cross sections of the ozone mixing ratio (ppmv) for the months of (a) January, (b) April, (c) July, and (d) October. All profile measurements within 10° latitude zones were averaged each month. [McPeters et al., 1984].

middle, and low latitude intervals appear in Figures 21-7, 21-8, and 21-9.

A meridional cross section of annually-averaged ozone partial pressure was computed from approximately 15 years of ozonesonde ascents and Umkehr observations [Dütsch, 1978] and is presented in Figure 21-10. The pattern of partial pressure (solid lines) closely parallels the location of the tropopause, and the percent variability about the mean (dashed lines) is greatest where the slope of the tropopause is greatest—the region of midlatitude storm activity. The amplitude of the annual cycle (expressed as percent) is presented in Figure 21-11, together with the month of maximum value. In general, largest seasonal changes are near the tropopause with a peak in the spring season. Large seasonal changes are also found above 45 km at polar latitudes due to the effects of the annual temperature cycle on ozone photochemistry.

In addition to the seasonally-controlled annual cycle, other natural periods can also be extracted from the climatological profiles. The quasibiennial oscillation appears at low latitudes (mirroring the total ozone behavior) and is most apparent near the mixing ratio maximum [Angell and Korshover, 1983]. At 40 to 45 km there is a weak indication of a solar rotation response [Gille et al., 1984] which is again statistically separable only at low latitudes where the profiles exhibit the most stability. A semiannual oscillation in ozone mixing ratio has been identified over a broad range of pressures (1–30 mb) and latitudes [Maeda, 1984].

The day-to-day longitudinal distribution is not well represented by the zonal means. In general, the hemispheric weather patterns that influence total ozone can be identified in the vertical profiles; the details, however, depend on the atmospheric level being sounded. In the troposphere, O_3 is usually a reliable tracer for weather conditions, but it is also subject to *in situ* chemical reactions [Fishman, 1985]. Near the tropopause and in the lower stratosphere, planetary scale wave patterns are accurately replicated in the ozone field

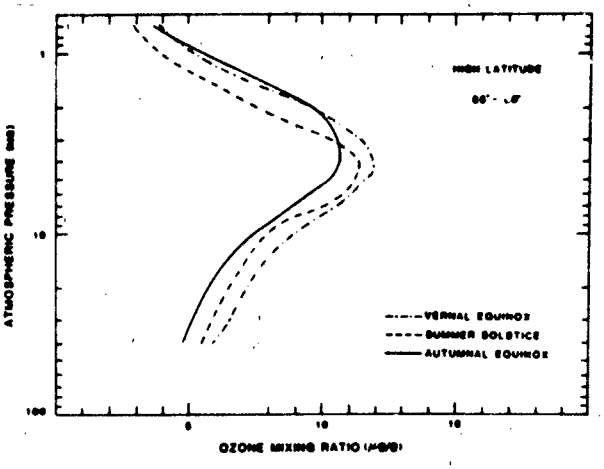


Figure 21-8. As in Figure 21-7 but for high latitude region 65°–80°.

[Gille et al., 1980b; Remsberg et al., 1984]. In the upper stratosphere the ozone concentrations begin to be modulated directly by the temperature sensitivity of the controlling chemistry [Barnett et al., 1975].

Preliminary climatological averages of ozone mixing ratios on constant pressure surfaces have been compiled for a range of time periods, from daily to seasonal, annual and interannual (in the stratosphere, see Remsberg et al., [1984], Barth et al., [1983], Nagatani and Miller [1984]; in the troposphere, see Bojkov [1983]). The complex structure of these climatologies is an integral response to the total atmospheric environment, including the instantaneous and gradient behavior (temporal and geographic) of temperature, transport parameters, and other minor constituent concentrations. Satellite platforms and field programs are currently being designed to address these complexities with complementary instrumentation (for instance, the Upper Atmosphere Research Satellite mission scheduled for the late 1980s [NASA Research Summaries, 1984]).

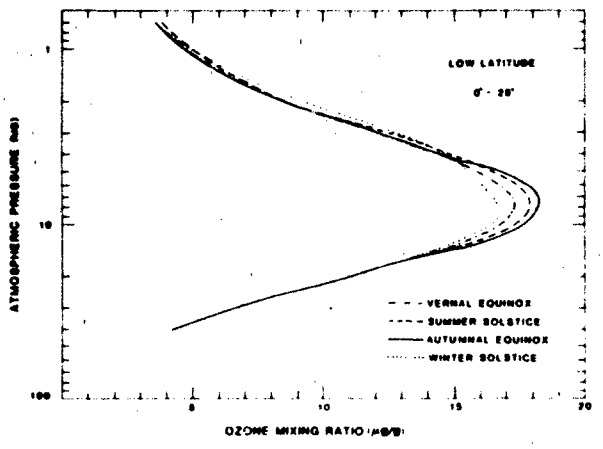


Figure 21-7. Ozone mixing ratio versus atmospheric pressure for the low latitude region ±25° [Klenk et al., 1983].

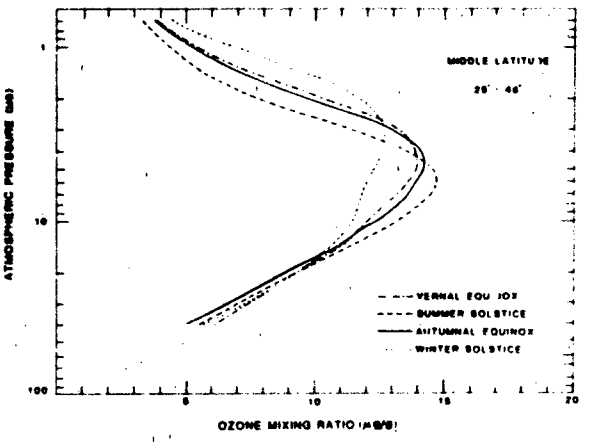


Figure 21-9. As in Figure 21-7 but for the middle latitude region 25°–45°.

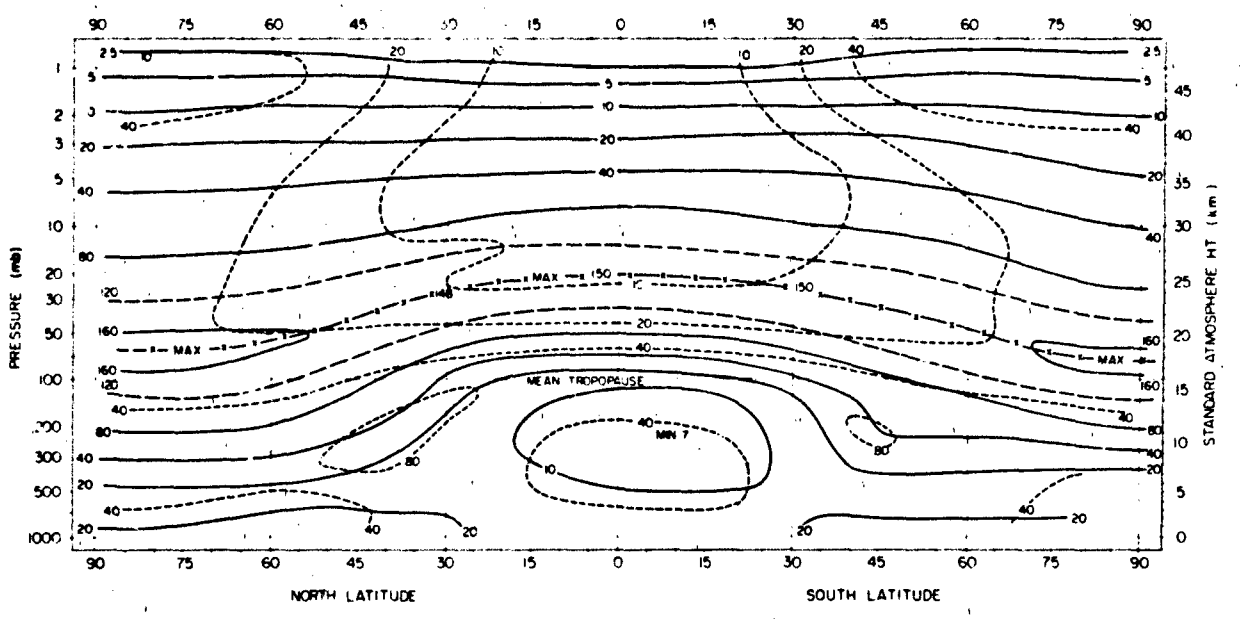


Figure 21-10. Pole-to-pole cross section of annual zonally averaged ozone (nbar) and normalized standard deviation (%).

21.1.6 Spatial and Temporal Scales of Ozone Variability

Satellite instrumentation cannot provide sufficient resolution (temporal, vertical, or horizontal) to describe modulations brought about by small-scale thermal waves and turbulence. This smaller scale ozone variability (Figure 21-12) as measured by *in situ* aircraft, balloon, and rocket data, appears as narrow horizontal layers (a kilometer or less in thickness) often extending for hundreds of kilometers along and tens of kilometers perpendicular to the flow [Dütsch, 1969]. These features can sometimes be directly related to tropopause folding, but even minor dynamic perturbations can lead to such turbulent stratification in a hydrostatically stable atmosphere [Dewan, 1981]. This structure in the ozone distribution, although prevalent, is not detectable from re-

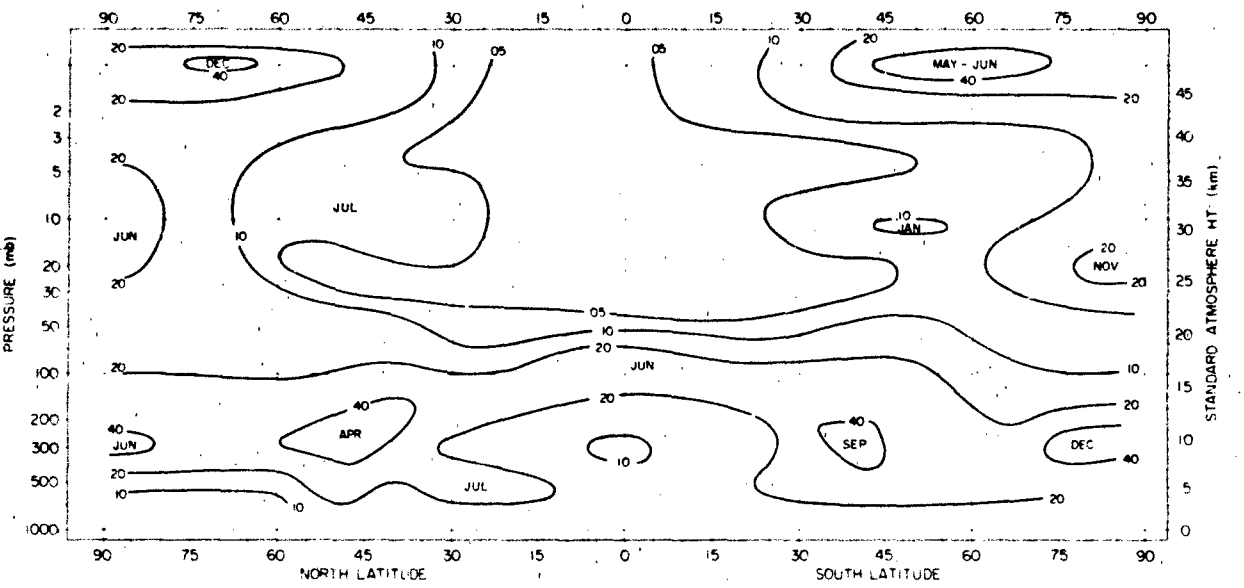


Figure 21-11. Pole-to-pole cross section of seasonal ozone variability (%) and time of maximum

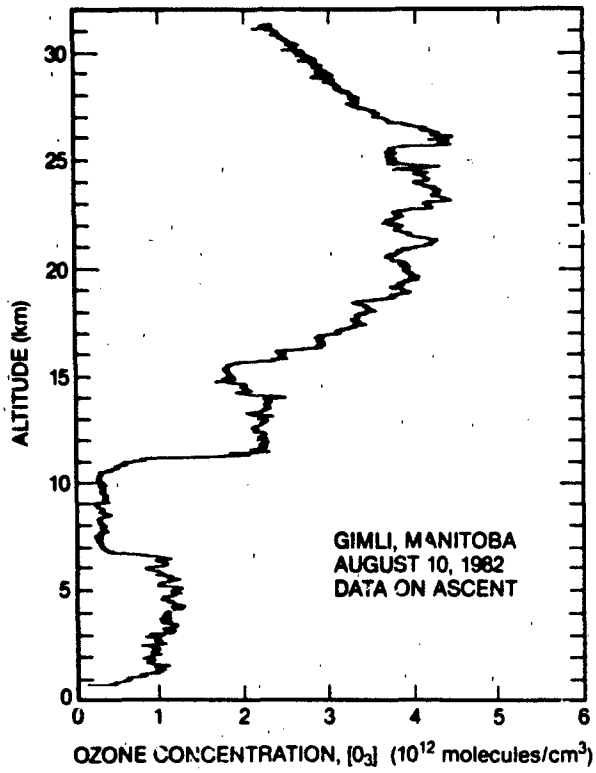


Figure 21-12. Vertical ozone ascent on 10 Aug 1982 over Gimli, Manitoba (3200 m) based on sampling at 5 m (1 s) intervals, illustrating thin layers of relatively high ozone concentration [Profitt and McLaughlin, 1983].

mote platforms (see Section 21.1.3 on measurement techniques). Measurements of both total ozone (ground-based) and local concentrations (*in situ* aircraft sampling at 11 km) have been used to estimate their respective degree of autocorrelation with east-west distance (Figure 21-13). The 11 km autocorrelation decays more rapidly, indicating a greater sensitivity to small scale motions; total ozone is

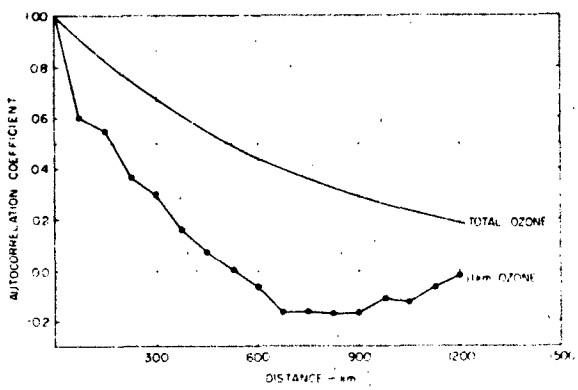


Figure 12-13. Variation of autocorrelation of total ozone with distance (E-W) in mid latitudes.

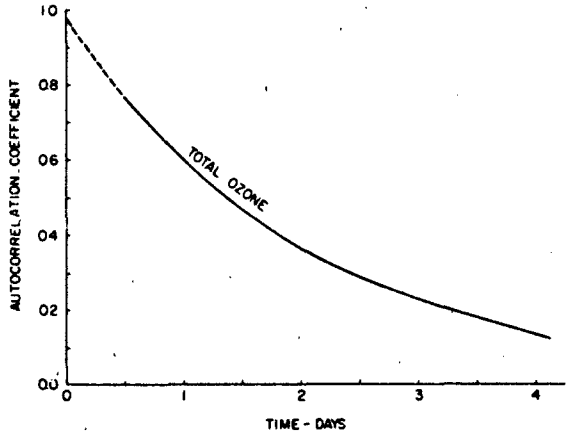


Figure 21-14. Decay of autocorrelation of total ozone with time.

associated with thicker atmospheric layers controlled by large scale dynamics. The time variation of the autocorrelation for total ozone (Figure 21-14) shows significant contributions from both long (seasonal) and short periods [Wilcox, 1978].

21.1.7 Mathematical Models

Mathematical descriptions of the atmospheric ozone layer date back to the original scheme of Chapman [1930]. However, as the natural variability in both total column abundance and vertical distribution were recognized, the need for more extensive modeling efforts becomes apparent. This requires balancing complex chemical schemes with an understanding of the feedbacks between ozone and the radiative-dynamic environments. The current formulations encompassing one, two, and three dimensional models, do not focus strictly on the ozone layer, but rather attempt to parametrize broader aspects of the surrounding atmosphere. [See the NASA Assessment Report, 1984 for a more complete discussion.]

The simplest conceptual models (other than simulations of local volume or laboratory environments) are one dimensional, with altitude or pressure as the vertical coordinate [Sze et al., 1980]. Because of the large number of individual chemical reactions necessary for atmospheric description [NASA/JPL, 1982], models typically divide the minor constituents into "families" [Wofsy and McElroy, 1974]. For example, molecular concentrations of NO or NO_2 are not calculated separately within the original mathematical formulation. Instead, a solution for the whole group of nitrogen-oxygen compounds (known as NO_x's) is found first; other standard families include the O_x (oxygen-oxygen), HO_x (hydrogen-oxygen), and ClO_x (chlorine-oxygen) compounds. Temporal modulations (with scales ranging from minutes to years) can be accomodated in 1-D models; broader applications include sophisticated parametrization of clouds,

radiative transfer, vertical transport, and exchange mechanisms.

Two-dimensional models can now include most of the photochemical and some of the temporal precision of the 1-D cases [Gidel et al., 1983]. The second dimension parametrizes the latitudinal gradients that exist on a global scale. The zonal mean temperature structure is incorporated either as input or internal calculation. There are various mathematical approaches adopted in 2-D models, including boundary conditions, radiative balance, integration timesteps, solution methods, etc. (see Table 2.14 in *The Stratosphere 1981* [WMO, 1981]). One of the most critical assumptions, however, is the method of approximating the eddy (non-zonal) motions, both in the horizontal and vertical directions. Pragmatic solutions have centered on "diffusion" coefficients which take the form of three dimensional tensors (K-theory, as proposed by Reed and German [1965]). An alternative mechanistic description, where at least part of eddy modulation is considered quasi-harmonic, has also been formulated [Garcia and Solomon, 1982].

The value of two dimensional models lies in their ability to replicate zonally-averaged measurements. They successfully reproduce the seasonal excursions of both the mean "total" ozone (Figure 21-5) and the mixing ratio maximum (Figure 21-6). Because they can incorporate time-dependent mechanisms (such as proposed solar rotation and solar cycle variability or minor constituent fluctuations [Garcia et al., 1984]), they have been used in conjunction with one dimensional models to investigate possible trends in ozone concentration. Prediction capability, however, is tied to the reliability of the input assumptions (including dynamic parametrization, reaction rates, and the atmospheric measurements against which they are evaluated). As an example of "model-related" sensitivities, the one dimensional models models formulated in 1979 forecast a potential 16%–18% anthropogenically-caused decrease in column O_3 by the year 2020. Models in 1982–83 predict a possible overall decrease of 2%–4% but with a strong altitude dependence and possible nonlinear response [Maugh, 1984; NASA Assessment Report, 1984]. The differences are related to updated determinations of critical chemical "constants", including the temperature and pressure dependence of reaction rates and absorption coefficients. Obviously, a definitive answer on the future of the ozone layer cannot yet be established, but 1 and 2-D models will facilitate the mathematical exploration of the possibilities.

Three dimensional models attempt to realistically simulate the fluid-dynamic, radiative, and photochemical characteristics of the atmosphere. The two general types of grid models (with variables proportional to latitude, longitude, and height) are similar to those used for large-scale meteorological forecasting. The global circulation models (GCMs) employ a comprehensive set of equations governing the radiation field, 3-D motions on a sphere, and at least part of the chemistry [Fels et al., 1980]. The lower boundary conditions can extend down to the solid-earth orography (including driving regons such as the Tibetan plateau) or are predicated on a modulated tropopause, initiating global stratospheric motions wth tropospheric forcing. The mechanistic models (MMs) parametrize general planetary characteristics (artificial ocean-continent boundaries, for instance), sacrificing realistic accuracy for increased flexibility. They have been used to simulate stratospheric warmings and their perturbing effect upon the whole global ozone distribution [Lordi et al., 1980; Hsu, 1981].

The horizontal and vertical resolution requirements for adequate three-dimensional representation of the physical processes affecting ozone are less rigid than those for tropospheric meteorological predictions. However, the grid sizes still serve as a limiting factor for computation speed and flexibility. These models are costly for trend predictions, but do provide the forum for physical understanding of climatological behavior and transport. For a fuller discussion, see *The Stratosphere 1981* [WMO, 1981] or more recent reviews.

21.2 MINOR CONSTITUENTS IN THE STRATOSPHERE

The existence of the stratosphere was discovered at the turn of the century by unmanned balloon measurements which found the temperature increased with altitude. The first understanding of the stratosphere occurred when Dobson suggested the heating was the result of UV absorption by ozone in the late 1920s. Chapman's pioneering photochemical explanation of the stratosphere in the 1930s remained the definitive description into the 1950s. The explosive research and recognition of the complex chemical nature of the stratosphere began inauspiciously in the late 1960s with the fear that supersonic transports (SSTs) would cause irreparable damage to the earth's environment. It is now recognized that the stratosphere has a very complex chemical composition generated from natural and man made sources. Furthermore, anthropogenic sources may lead to future changes of the stratosphere. The description that follows is not complete and must be updated and corrected as new information becomes available. A series of reports by government agencies and committees indicates the world wide interest in environmental protection and on-going stratospheric research - SCEP [1970], CIAP [1974], COMESA [1975], COVOS [1976], NAS [1975, 1979], DOE [1979], UNEP [1979, 1980], NASA [1977, 1979], FAA [1979], and WMO [1982]. These repons contain excellent summaries of minor constituents and theoretical and model explanations. Each succeeding years report contains new or revised information. Journal reviews of the stratosphere [Cicerone, 1975; Hudson, 1979; Murgatroyd, 1982; and Solomon, 1983] contain succinct descriptions of the minor con-

stituents. Emphasis has now shifted from measurements to developing a better understanding of the mechanisms and chemical processes identified earlier.

Neutral minor constituents of the stratosphere can be grouped into large categories of radicals, sources of radicals, sinks of radicals, and sources for the formation of aerosols. The radicals are conveniently divided into groups of odd oxygen—$O(^3P)$, $O(^1D)$, $O_2(^1\delta)$, $O_2(^1\Sigma)$, O_3; odd hydrogen —OH, HO_2, H_2O_2; odd nitrogen—NO, NO_2, HNO_3, N_2O5, HO_2NO_2; and odd halogens—Cl, ClO, F, Br, BrO, HCl, HF, $ClONO_2$. The precursors or sources of the radicals are constituents such as N_2O, CH_4, H_2O SF_6, CO_2, H_2, CCl_4, CH_3CCl_3, CH_3Cl, C_2HCl_3, CCl_3F, CCl_2F_2 and other chlorofluoromethanes (CFMs), some of which are introduced into the troposphere at mixing ratios greater than 100 pptv; these represent the major source of chlorine in the stratosphere. Stratospheric NO_x derives from nitrous dioxide (N_2O) transported from the ground. Large solar proton events could contribute appreciable amounts of NO in the upper stratosphere. The products of radical–radical reactions, which are generally inert, serve as sinks for the radicals. They also form shortlived storage from which radicals can be released by photolysis or oxidation. The major sinks are HCl, HNO_3, HNO_4, $ClONO_2$, H_2O_2, O_3, HO_2, NO_2, HOCl, N_2O_5, $COCl_2$, COFCl and HF. The minor constituents that are believed important as sources for aerosol growth are H_2SO_4, SO_2, COS and NH_3.

The following description of the minor constituents (Sections 21.2.1 through 21.2.9) has been excerpted from the World Meteorological Organization publication "The Stratosphere 1981 Theory and Measurements" with the permission of NASA. Even though this information can be found in a more detailed form in the original publication, it has been included in this chapter to provide the reader with the basic facts of stratospheric composition and a more complete picture of "Atmospheric Composition" as a whole.

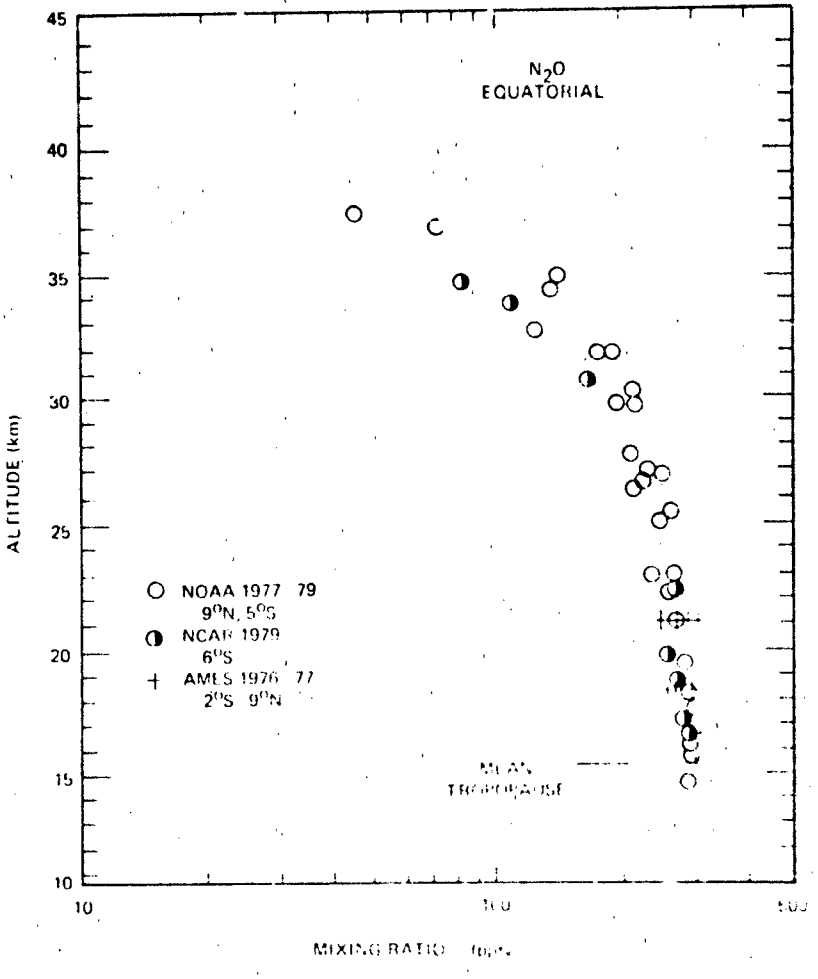


Figure 21-15. Measurements of N_2O as a function of altitude in the equatorial region

21.2.1 Nitrous Oxide, FC-11, FC-12, and Methyl Chloride (N_2O, CCl_3F, CCl_2F_2 and CH_3Cl)

Since 1975 a number of measurements have been made of the stratospheric concentrations of N_2O, CCl_3F and CCl_2F_2. Four field programs have been particularly extensive, all involving laboratory analysis by gas chromatography of air samples collected by balloon-borne samplers. The few measurements made by *in situ* balloon-borne infrared techniques [Farmer et al., 1980] are consistent with these data. The four research groups are

NOAA Goldan et al. [1980, 1981]. Balloon-borne evacuated grab samples.

NCAR Heidt et al. [1975]. Balloon-borne cryosampler.

KFA Volz et al. [1981]. Balloon-borne cryosampler.

Ames Tyson et al. [1978]; Vedder et al. [1978]; Inn et al. [1979]; Vedder et al. [1981]. Balloon and aircraft-borne cryosampler.

The results from three of these groups, compared over narrow latitude bands, are shown for equatorial latitudes in Figures 21-15, 21-16, 21-17 and mid-latitudes in Figures 21-18 through 21-23. They are in good agreement with one another. The two single altitude profiles obtained by infrared techniques at 32°N and 30°S also fall in between the trends shown in Figures 21-18 through 21-23. The data for CCl_3F and CCl_2F_2 have been corrected for the well-established secular increase in each using tropospheric measurements.

The vertical profile of CH_3Cl as obtained from gas chromatographic measurements on cryogenic samples from balloon flights over Southern France (44°N) is shown in Figure 21-24. The data, although largely scattered, clearly show a very rapid decrease from 600 ppt at the tropopause to about 20 to 40 ppt at 30 km indicating the short lifetime of CH_3Cl.

21.2.2 Carbon-Containing Species

21.2.2.1 Carbon Dioxide (CO_2). Vertical profiles of CO_2 have been measured routinely at NCAR using a volumetric technique, achieving a precision of ±3 ppm. About the same precision has been achieved at the KFA using gas-chromatography. The precision of the gas-chromatographic measurements has recently been improved to ±0.5 ppm [Volz et al., 1981]. Additional measurements have been made by Bischof et al. [1980] using the infrared absorption technique, and by Mauersberger and Finstad (1980) using a balloon-borne mass spectrometric method.

The individual profiles of CO_2 are plotted in Figure 21-25. As can be seen, the older volumetric and GC-data show a large scatter but no significant gradients of the CO_2 mixing ratio in the stratosphere. However, from the more recent data [Volz et al., 1981] and from the measurements by Bischof et al. [1980], a weak but significant gradient is observed in the lower stratosphere, namely the CO_2 mixing ratio is found to decrease by about 6 to 7 ppm between the tropopause and 20 km altitude.

21.2.2.2 Carbon Monoxide (CO). Although the experimental data are extremely limited, the measurements cover

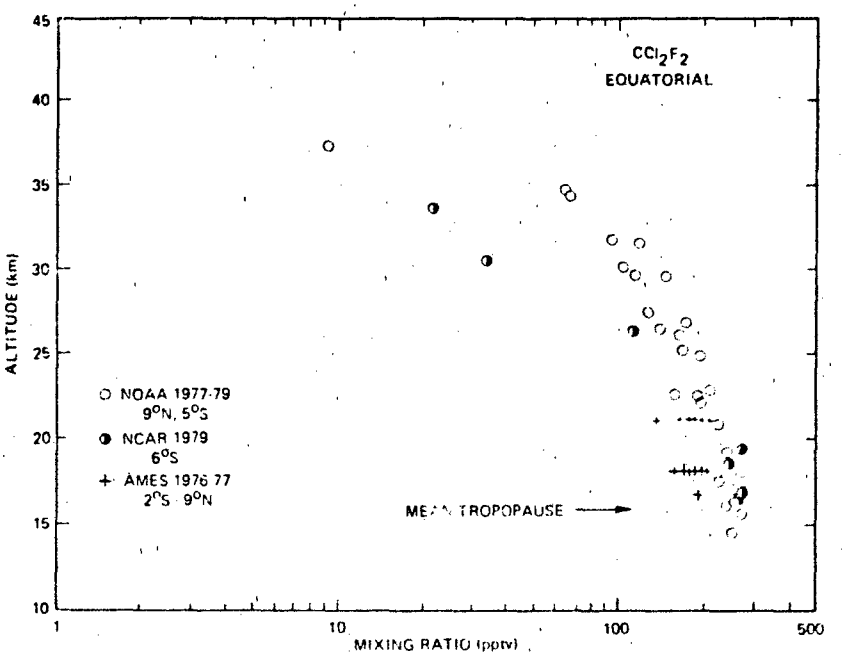


Figure 21-16. Measurements of CF_2Cl_2 (FC-12) as a function of altitude in the equatorial region.

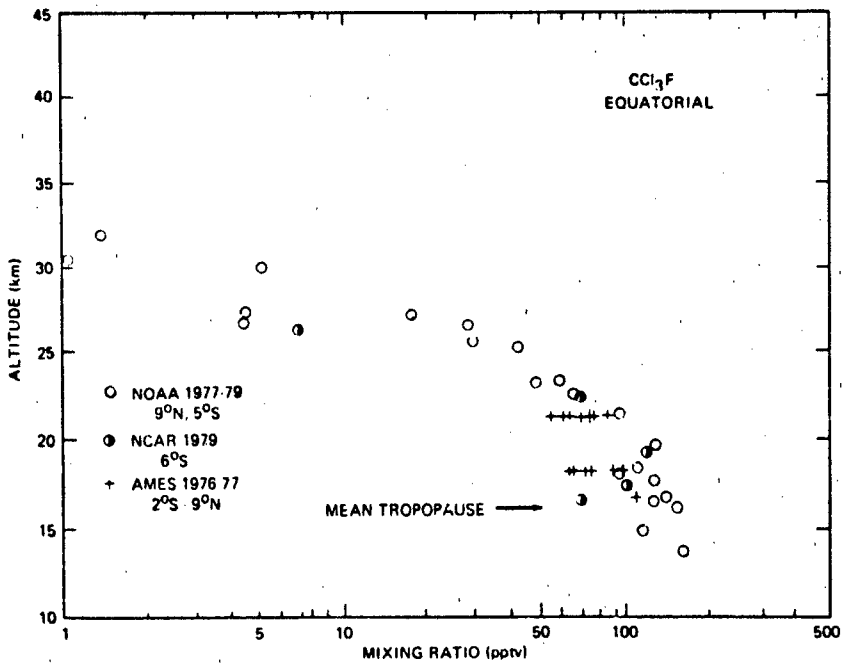


Figure 21-17. Measurements of CCl_3F (FC-11) as a function of altitude in the equatorial region.

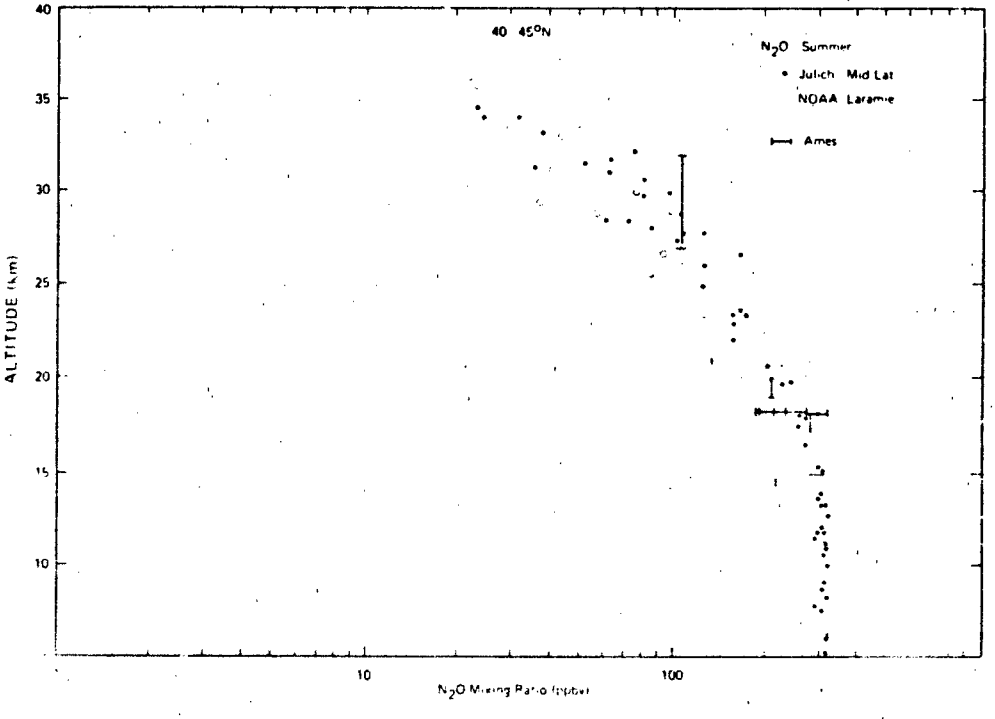


Figure 21-18. Midlatitude summer vertical profile of N_2O.

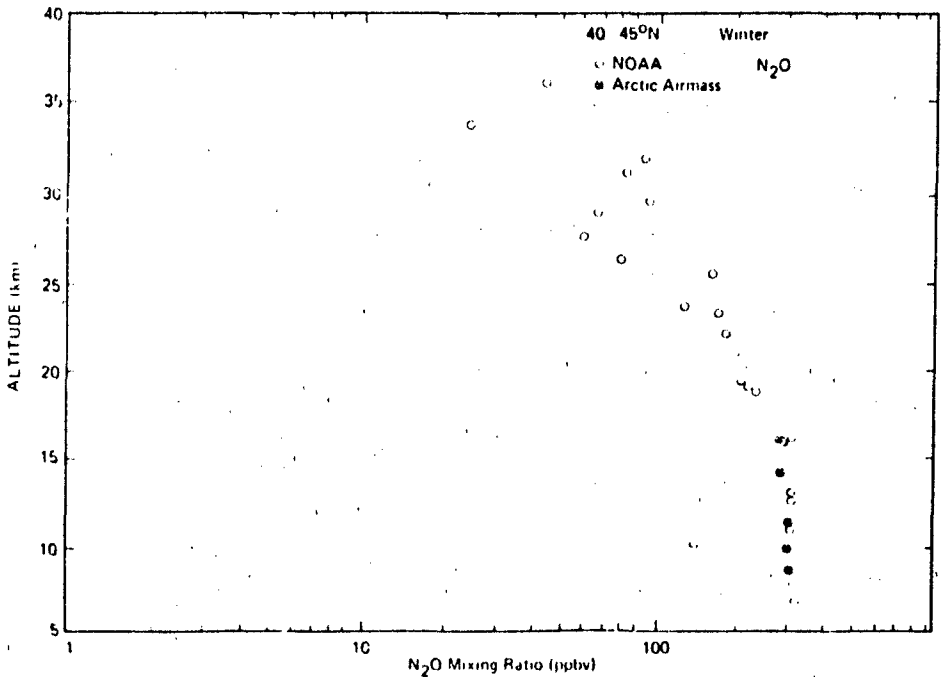


Figure 21-19. Midlatitude winter vertical profile of N_2O.

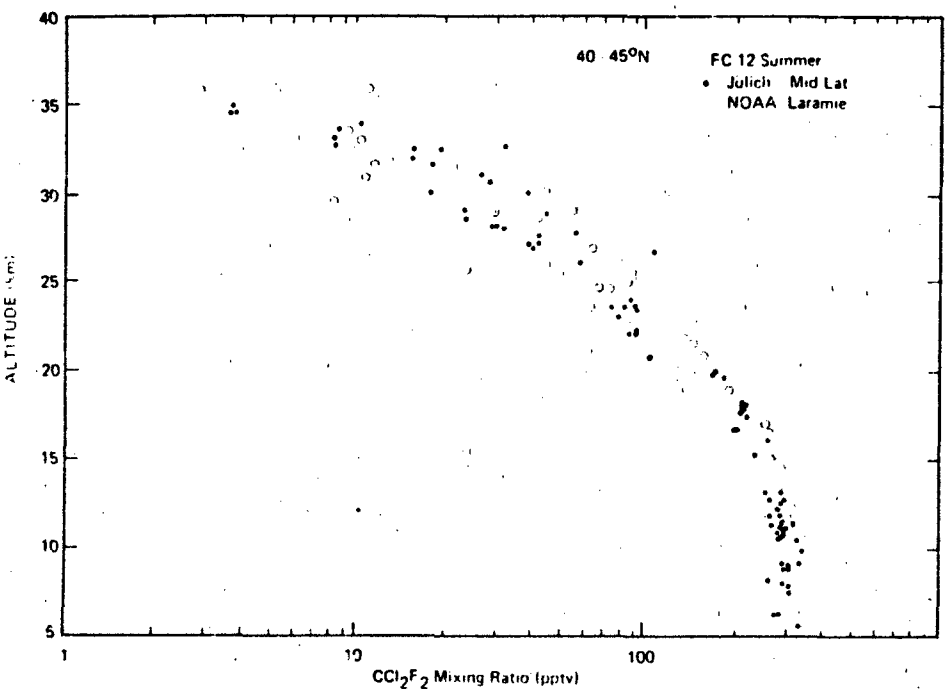


Figure 21-20. Midlatitude summer vertical profile of CCl_2F_2 (FC-12).

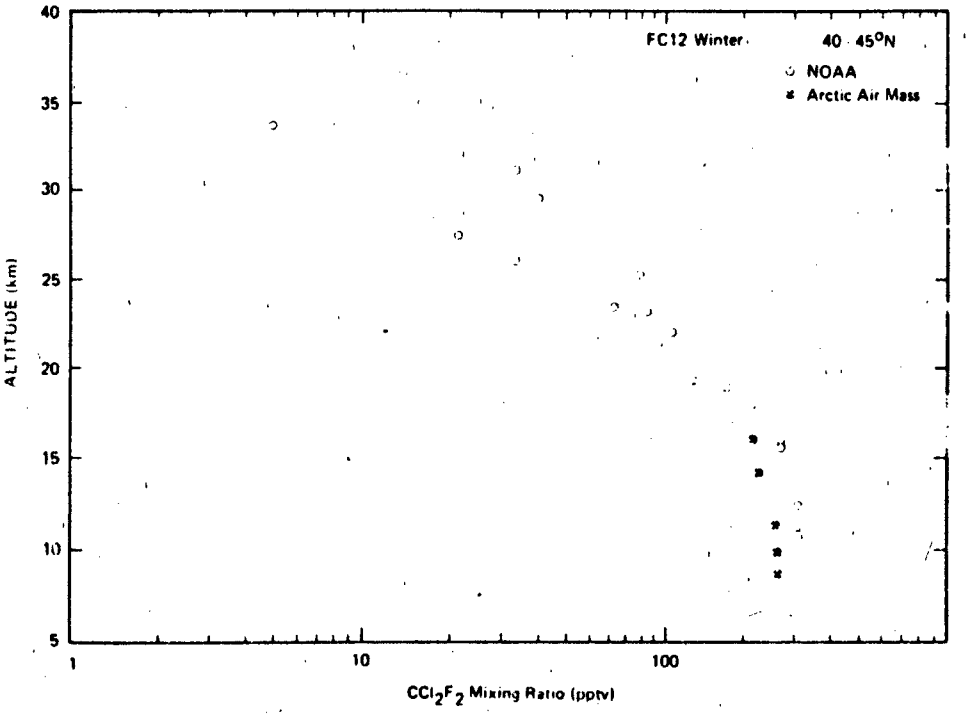


Figure 21-21. Midlatitude winter vertical profile of CCl_2F_2 (FC-12).

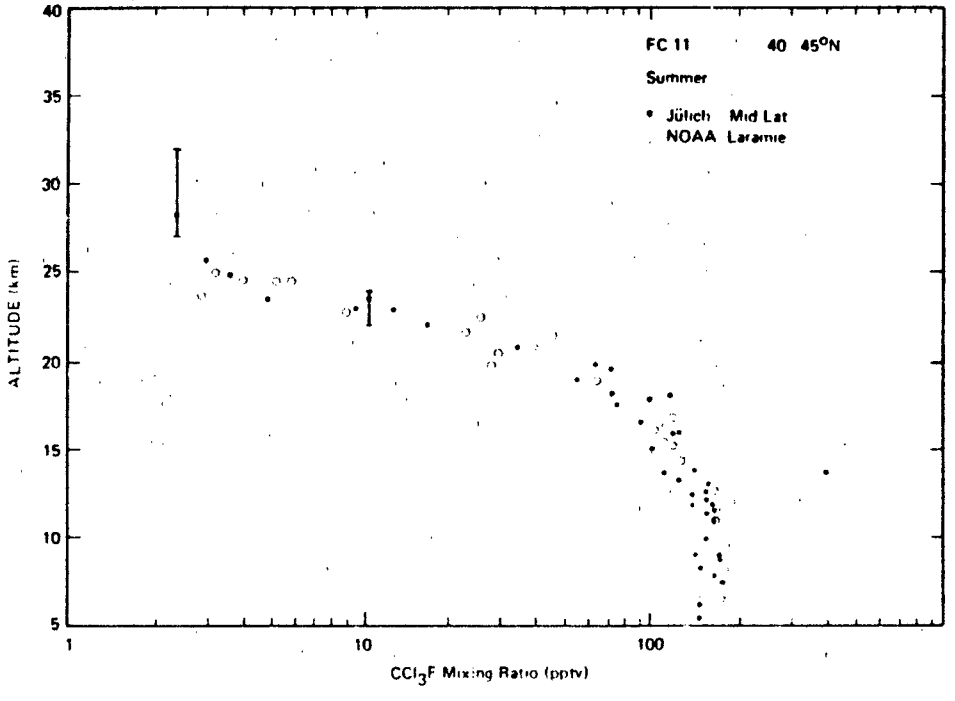


Figure 21-22. Midlatitude summer vertical profile CCl_3F (FC-11).

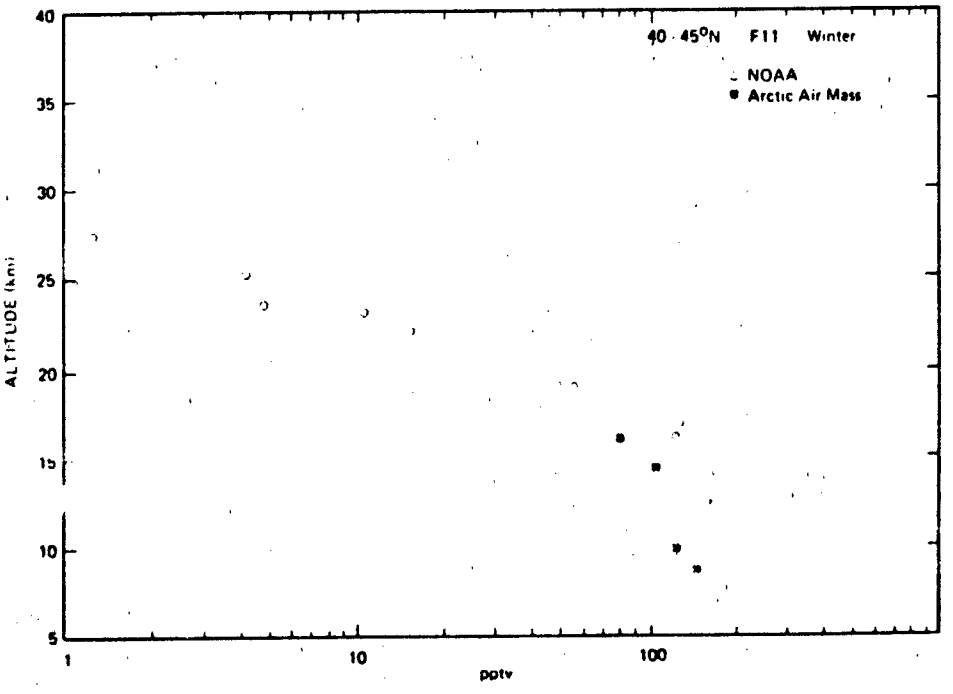


Figure 21-23. Midlatitude winter vertical profile of CCl_3F (FC-11).

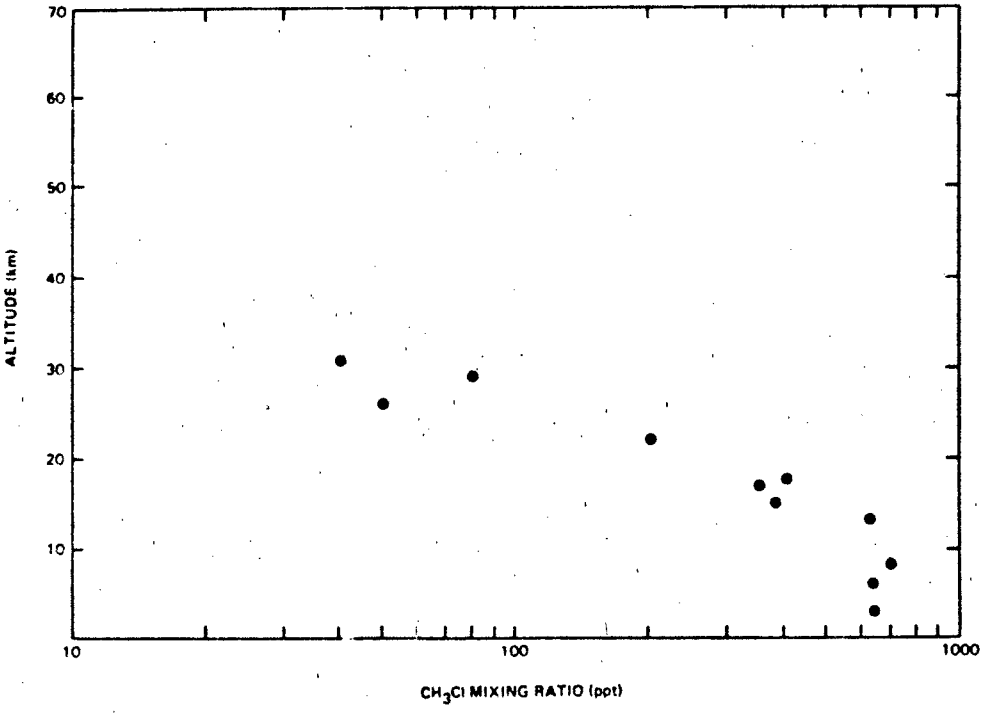


Figure 21-24. Vertical profile of CH_3Cl at 44° northern latitude. The data were measured by GC-FID and GC-MS [Penkett et al., 1979] from samples collected cryogenically during a joint balloon program [Fabian et al., 1979].

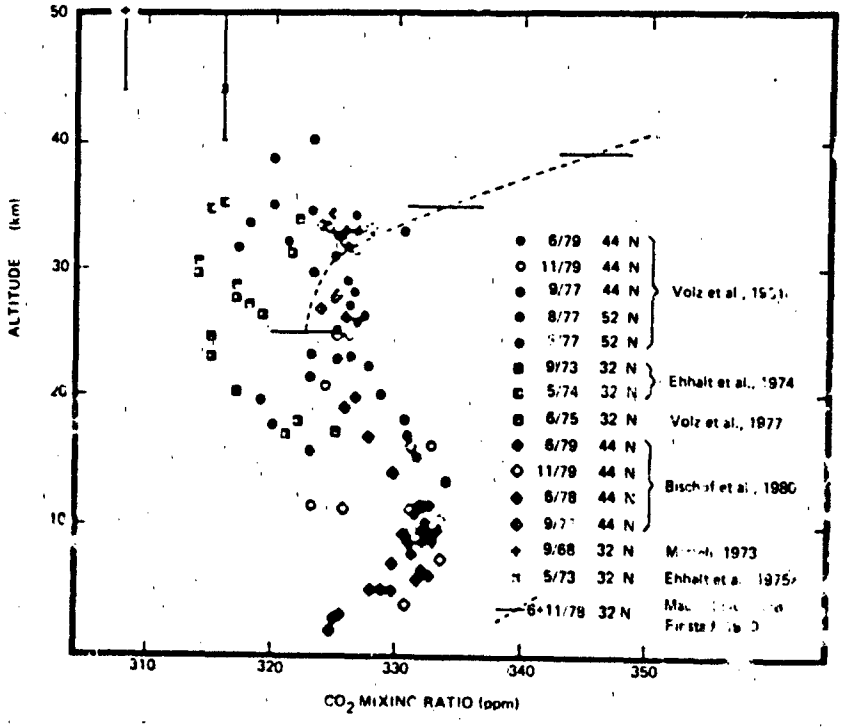


Figure 21-25. Vertical profiles of CO_2 at midlatitudes. Some of the scatter is due to the observed secular increase of CO_2 in the stratosphere [Volz et al. 1981] during the time the different profiles were measured.

the vertical profile of CO between the tropopause and 50 km. The combined data in Figure 21-26 suggest a decrease of CO across the tropopause and in the lower stratosphere up to 20 km. Above this altitude the CO mixing ratio is constant, at 10 ppb, up to 30 km. It then increases to 40 ppb around 40 to 50 km altitude. More reliable measurements of CO in the middle and upper stratosphere are required to establish the vertical profile.

21.2.3 Hydrogen-Containing Species

21.2.3.1 Molecular Hydrogen (H_2). Recent stratospheric measurements of H_2 are summarized in Figure 21-27. The profiles obtained at three latitudes show little vertical and latitudinal variation.

The data at 40 to 60°N show a clear decrease with altitude, from 0.55 ppm at the tropopause to 0.45 ppm at 35 km altitude [Ehhalt, 1978; Fabian et al., 1979]. A similar trend although with a much larger uncertainty can be deduced from the data at 60°N. No significant trend is found for the data at 32°N over Palestine, Texas. It should be noted that the absolute calibration of the NCAR data is about 10% lower than that of the KFA. It has been shown by Ehhalt and Tonnissen [1980] in a qualitative way that elevated levels of stratospheric H_2 are associated with, and probably caused by, increased concentrations of CH_4.

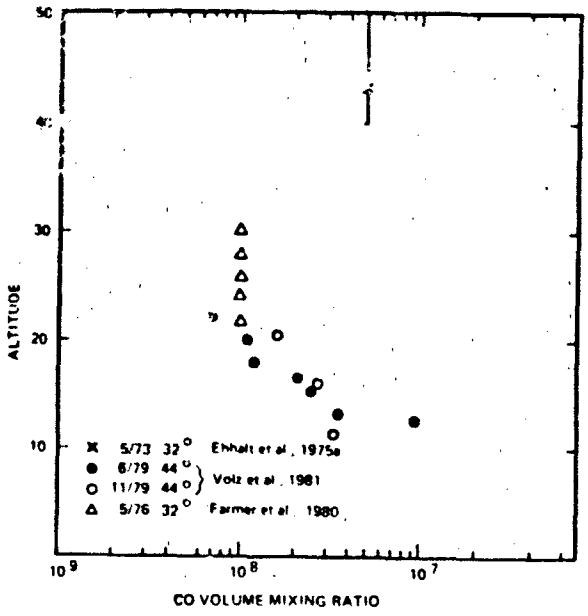


Figure 21-26. Vertical profile of CO at midlatitudes.

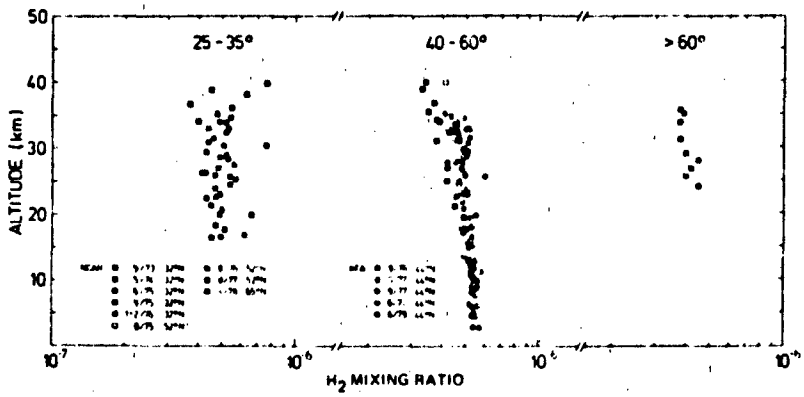


Figure 21-27. Vertical profiles of H_2 at different latitudes. The NCAR data are published by Pollock et al. [1980] and Ehhalt et al. [1975a,b] by the KFA data by Volz et al. [1981] and Fabian et al. [1979, 1981].

21.2.3.2 Methane (CH_4). Vertical profiles of CH_4 have been measured since 1965. Most of the sampling flights were performed at 32°N, 44°N, and 52°N. Only data from two flights exist for latitudes >60° (NAS) and only one profile for the tropics. Measurements were performed either by gas chromatography on grab samples and cryogenic samples collected *in situ* during balloon and aircraft flights or by using long path infrared absorption from balloons. In Figure 21-28 the results are plotted separately in four latitude bands. All NCAR measurements made prior to 1974 were multiplied by a factor of 1.2 [Heidt and Ehhalt, 1980].

At first glance, only the tropical profile deviates significantly from the others, showing a much weaker gradient in the stratosphere. This behavior, which is also confirmed by the profiles of other long-lived trace gases such as N_2O and FC-12, signifies a considerably stronger upward transport in the tropics than in mid-latitudes. A closer investigation shows some minor but still significant differences among the midlatitude profiles. The average profile at 32°N shows only a very weak gradient between 20 and 30 km, and on several occasions, profiles with a well-mixed layer in this altitude range were observed. This behavior can be explained by the stratospheric branch of the tropical Hadley circulation displacing air from the tropical mid-stratosphere with a weak CH_4 gradient into the lower stratosphere at 30°N, [Ehhalt and Tonnissen, 1980]. On some occasions, layers of almost constant mixing ratio were also observed at 44°N. In addition, from the individual profiles collected at 44°N, there is a slight hint of a seasonal variation of the stratospheric CH_4 concentration, especially above 25 km, where the average profile shows a relatively large variability compared to lower altitudes. At higher latitudes, the data though sparse indicate a more or less linear decrease of the CH_4 mixing ratio with altitude above the tropopause.

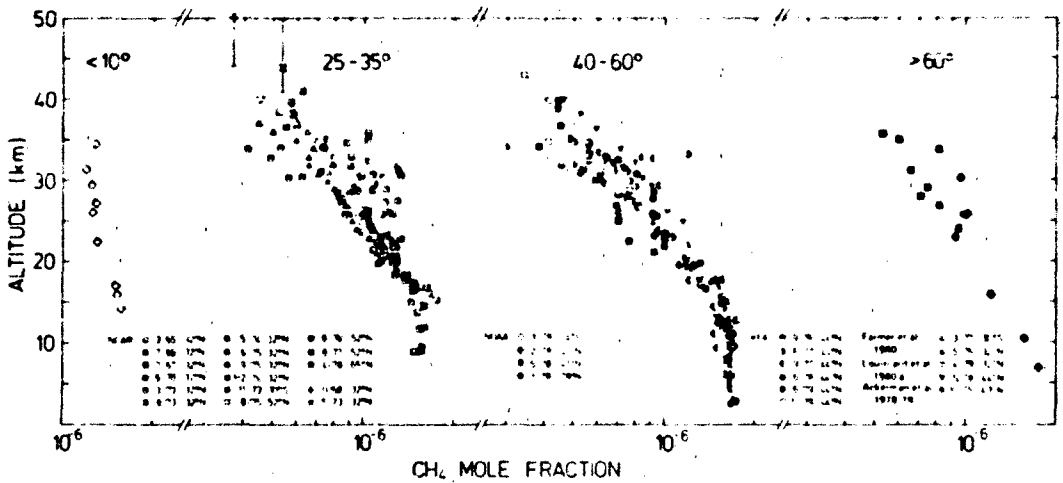


Figure 21-28. Vertical profiles of CH_4 at different latitudes. The NCAR data by Pollock et al. [1980], Ehhalt and Heidt [1973a, b], Ehhalt et al. [1974, 1975a,b]; the NOAA data by Bush et al. [1978]; and the KFA data by Volz et al. [1981], Fabian et al. [1979, 1981]. The NCAR data prior to 1975 were corrected by a factor of 1.2 [Heidt and Ehhalt, 1980].

21.2.3.3 Ethane, Propane, and Acetylene (C_2H_6, C_3H_8, and C_2H_2). Only three of the hydrocarbons present in the troposphere have been observed in the stratosphere: ethane, C_2H_6, propane, C_3H_8, and acetylene, C_2H_2. The tropospheric background mixing ratios of these species are quite low; around 1 to 2 ppb for C_2H_6 and up to several hundred ppt for C_3H_8. In addition, they show a strong latitudinal gradient with even lower values at the equator. The C_2H_6 mixing ratio at the equator is lower by about a factor of five; C_3H_8 and C_2H_2 decrease by a factor of ten [Rudolph et al., 1979; Singh et al., 1979; Harrison et al., 1979; Cronn and Robinson, 1979]. Because of their low tropospheric concentrations the fluxes of these gases into the stratosphere are small and their impact on the stratospheric carbon and hydrogen budgets is negligible. Singh and Hanst [1981] have proposed that oxidation products of ethane and propane are important carriers of reactive nitrogen.

C_2H_6 and C_3H_8 react rapidly with atomic chlorine, Cl, and can decrease the Cl concentration significantly in the lower stratosphere [Aiken and Maier, 1978; Rudolph et al., 1981]. Measured profiles of C_2H_6 have been used to deduce the vertical profile of Cl atoms in the lower stratosphere, where direct observation of Cl atoms is not yet feasible [Rudolph et al., 1981].

The measured vertical profiles of C_2H_6, C_2H_2, C_3H_8 are shown in Figures 21-29 through 21-32. All of these gases exhibit a strong decrease in the mixing ratio with altitude. The measured profiles of C_2H_2 and C_3H_8 agree reasonably well with those predicted from a one-dimensional steady-state model. In contrast C_2H_6, which is destroyed by reaction with Cl, decreases less steeply than predicted. This has been interpreted to indicate substantially lower Cl-atom concentrations in the lower stratosphere than predicted by models [Rudolph et al., 1981].

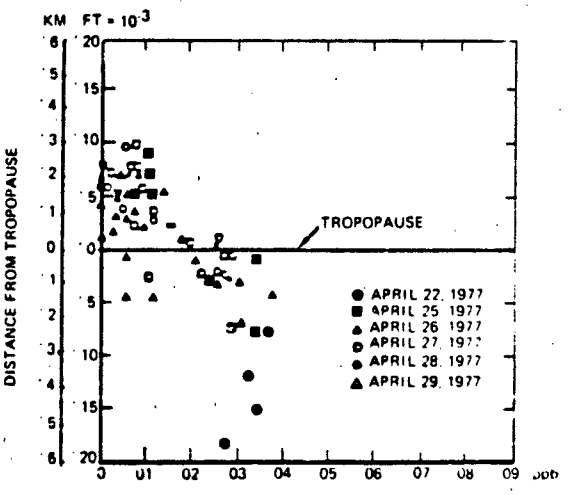


Figure 21-30. Distribution of acetylene relative to tropopause height. The latitude range of the samples from the first flight were 36 to 47°N and 36 to 38 N for the remaining flights.

21.2.4 Stratospheric Water Vapor (H_2O)

Ellsaesser et al. [1980] discussed the knowledge of the physical and chemical properties of stratospheric H_2O in 1979. In that work, a compilation of measured profiles was given in graphical form and a number of conclusions were drawn. In general, the basic Brewer theory of tropical "freeze drying" within the rising equatorial branch of the Hadley

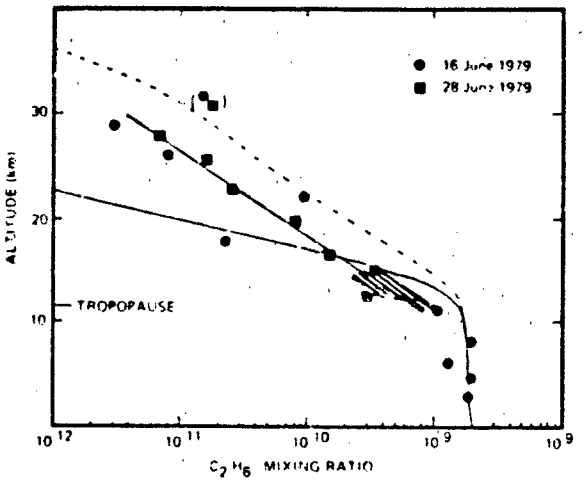


Figure 21-29. Vertical profiles of the C_2H_6 mixing ratio in the stratosphere over Southern France 44° N latitude [Rudolph et al., 1981]. The stippled area shows the range of data by Cronn and Robinson [1979] over San Francisco Bay area (37°N) in April 1977.

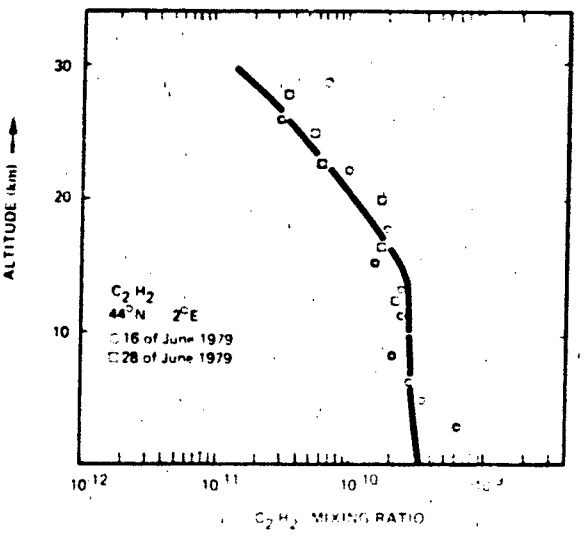


Figure 21-31. Vertical profiles of the C_2H_2 mixing ratio in the stratosphere over Southern France 44 N latitude [Rudolph et al., 1981].

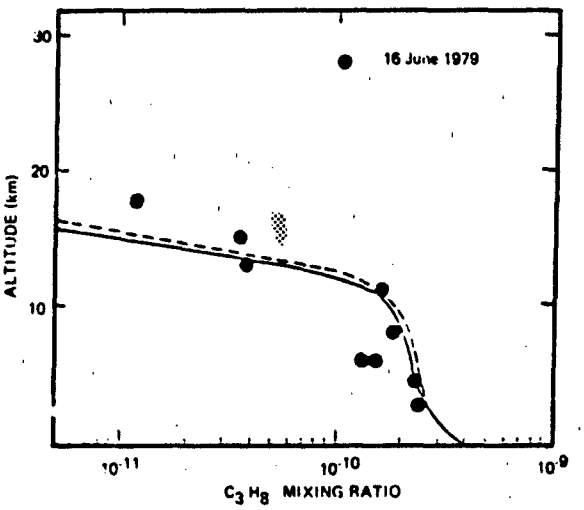


Figure 21-32. Vertical profiles of the C_3H_8 mixing ratio in the stratosphere over Southern France, 44°N latitude [Rudolph et al., 1981].

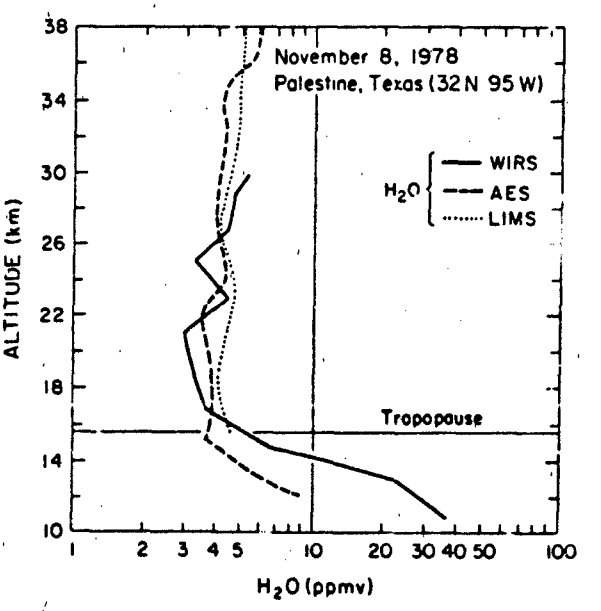


Figure 21-33. Measurements of H_2O mixing ratio over Palestine, Texas on 8 Nov 1978 by two balloon-borne *in situ* instruments (WIRS and AES) compared to the LIMS profile retrieved with the operational algorithm [Gille and Russell, personal communication].

cell was thought to be valid. However, a number of questions remained.

- Are there significant long-period trends in lower stratospheric mixing ratio, suggesting changes in circulation of tropical tropopause temperatures?
- Is the decrease in mixing ratio with height just above the tropical tropopause as identified by Kley et al. [1979] during their only tropical sounding a regular feature of the stratosphere?
- Are there latitudinal gradients of the mixing ratio? If so, are they poleward-directed or equator-directed?
- Are there increases in mixing ratio with height? If so, can the increase be fully accounted for by CH_4 oxidation?

21.2.4.1 Satellite Measurements Water vapor profiles have been obtained by the LIMS instrument, extending from the tropopause up to ~50 km, for a number of specific occasions, and generally show a fairly gradual increase of mixing ratio with height over this range. An example of this data is compared in Figure 21-33 with profiles obtained simultaneously by two balloon-borne instruments: an infrared radiometer from the Atmospheric Environment Service (labelled AES), and the WIRS instrument of the National Physical Laboratory. The data agree closely giving encouragement that at least in the lower stratosphere, the LIMS data appears to agree well with independent observations.

21.2.4.2 Other Measurements. The observations shown in Figure 21-33 show an increase in H_2O mixing ratio between the tropopause, or in some cases a somewhat higher level, and 30 or 40 km altitude. The satellite observations seem to indicate a marked peak in the H_2O mixing ratio at the 55 to 60 km level. However, Rogers et al. [1977] and Waters et al. [1980] have reported constant mixing ratios, centered around 4 ppmv at these altitudes. Since there are no theoretical arguments for large mixing ratios peaking around 55 to 60 km, confirming measurements are needed before those results can be considered real.

A new *in situ* device, developed by Kley and his as-

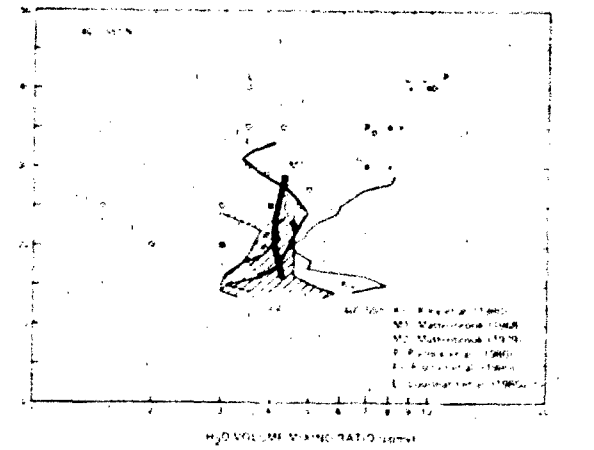


Figure 21-34. Altitude profile of H_2O between 40 and 55°N latitude. If more than one flight was made by a certain group, the banded area represents the range of their measurement.

sociates [Kley et al., 1979; Kley and McFarland, 1980] has now evolved as a powerful technique. These authors have built a sensitive instrument with a fast response time for balloons and aircraft which can resolve tens of meters in the vertical. Measurements of H_2O and temperature in mid-latitudes (Wyoming) generally show undersaturation in the upper troposphere although on one flight in Brazil [Kley et al., 1979] saturation was observed at the tropical tropopause. The results of the Brazil flight are shown in Figure 21-35. However, previous observations by Dobson and co-workers [Brewer, 1949; Dobson et al., 1946] over southern England (~52°N) have shown cases of both saturation *and* undersaturation below the tropopause. Saturation would indicate a contradiction to the Brewer model. Even supersaturation has been observed [Dobson et al., 1946; Kley et al., 1980]. It was pointed out by Kley et al. [1979] that tropical stratospheric air has a minimum mixing ratio at about 60 mb or 19 km (see Figure 21-35). This is well above the tropopause. Robinson [1980] has found more examples from literature studies. These observations indicate that the simple Brewer mechanism needs to be refined.

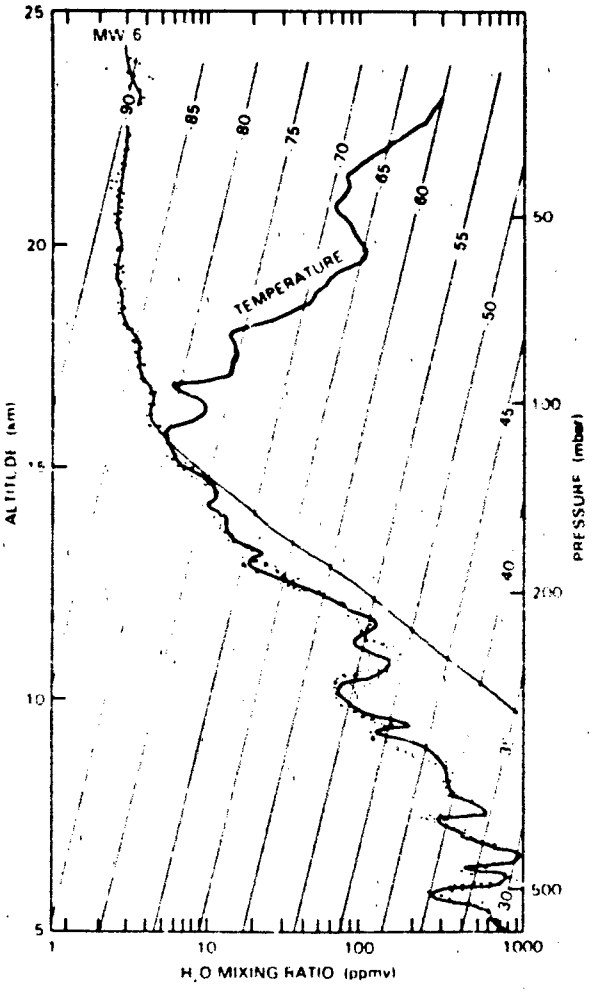


Figure 21-35. The water vapor mixing ratio measured over Quixeramobim, Brazil on 27 Sept 1978 [Kley et al., 1979].

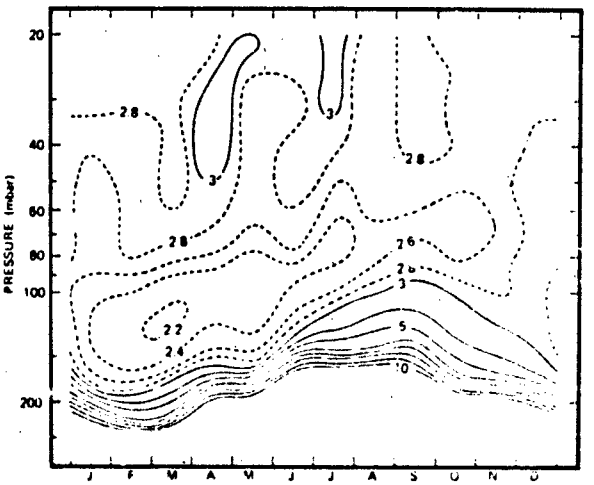


Figure 21-36. The mean annual variation of stratospheric water vapor at Washington, D.C., based on 12 years of data by Mastenbrook and Daniels [1980].

Another set of water vapor data are the frost point measurements of Mastenbrook that are being continued by NOAA. In a recent paper [Mastenbrook and Daniels, 1980] data from four flights over Washington, D.C. during the early part of 1979 are reported. A map of annual variations of stratospheric H_2O over Washington, D.C., based on 12 years of data, is presented here in Figure 21-36.

The composite water vapor profile of Figure 21-36 is recommended for chemical modeling purposes in the lower stratosphere at mid-latitudes ($z < 20$ km). Mastenbrook's instrument tends to produce altitude independent mixing ratios whereas Kley's instrument normally gives a moderate increase of 1 to 3 ppmv between tropopause and 32 km at mid-latitudes. For the equatorial lower stratosphere it is suggested that Figure 21-35 be used. This profile is similar to earlier ones by Mastenbrook [1968] but shows the hygropause clearly. It should be noted that it is the minimum in water vapor some 3 km above *tropical* tropopause that was referred to as hygropause by Kley et al. [1979].

21.2.5 Odd Oxygen

The principal oxygen radicals are $O(^1P)$, $O(^1D)$, $O_2(^1\delta)$, $O_2(^1\Sigma)$, O_2 (other excited states) and O_x. While it is feasible in this section to critically analyze all of the available data on the first five species, a thorough discussion of O_x is obviously of such magnitude that it warrants special treatment (see Section 21.1). Atomic oxygen in the ^1D level is of critical importance for establishing the oxidation rate of source molecules which enter the stratosphere such as CH_4,

N_2O etc., but there currently are no observations of (O^1D) in the stratosphere. The electronically excited states, $O_2(^1\delta)$ and $O_2(^1\Sigma)$, have been observed but the data base is small. The remaining electronically excited states of O_2 ($A^3\Sigma_u^+$, $C^3\delta_u$, $C\Sigma_u$) have not been observed.

21.2.5.1 Atomic Oxygen ($O(^3P)$). There are six observations of $O(^3P)$ in the stratosphere, all obtained using a parachute-borne, *in situ* atomic resonance fluorescence instrument, the results of which are shown in Figure 21-37. Experimental accuracy is ± 30% and experimental precision ± 10% for each measurement [Anderson, 1975].

Several points are apparent from Figure 21-37. First, there is both local structure within and absolute displacement among observed distributions which exceed respectively the precision and accuracy of the measurements. It should also be noted that the local structure does not appear consistently. For example the profiles observed on October 25, 1977 and December 2, 1977 display a small degree of local structure, typically less than ± 20% variation over an interval of ± 1 km above approximately 34 km. Below that altitude significantly greater local structure is apparent, though seldom more than ± 50%. On the other hand, the remaining four observations exhibit at least one example of major (factor of two) variation over ± 2 km interval with an increasing structural development below the 33 to 35 km interval.

21.2.6 Odd Nitrogen

21.2.6.1 Nitric Oxide (NO) NO has been the most extensively studied stratospheric odd-nitrogen species. The last decade has seen numerous measurement programs apply a variety of experimental techniques at many different altitudes, locations, and seasons. Thus, the problem in establishing a picture of our best current experimental understanding of stratospheric NO is not a scarcity of data, but rather of making a proper assessment and selection of the data.

Altitude Profile

There have been a large number of measurements by many different groups of the NO mixing ratio as a function of altitude. The techniques employed in these measurements fall into two classes: *in situ* and remote. While the long-path, vertical-column measurements from the latter technique have provided some of the best information regarding the seasonal and latitudinal variations for NO, the determination of a detailed height profile from long-path data involves deconvolution at solar zenith angles near 90° in order to obtain the maximum number of absorbing molecules along the sight path The rapid NO → NO_2 and NO_2 → NO conversions at sunset or sunrise complicate the comparison of these data with model predictions.

As a consequence, *in situ* measurements are used here to establish the NO height profile. A limited number of

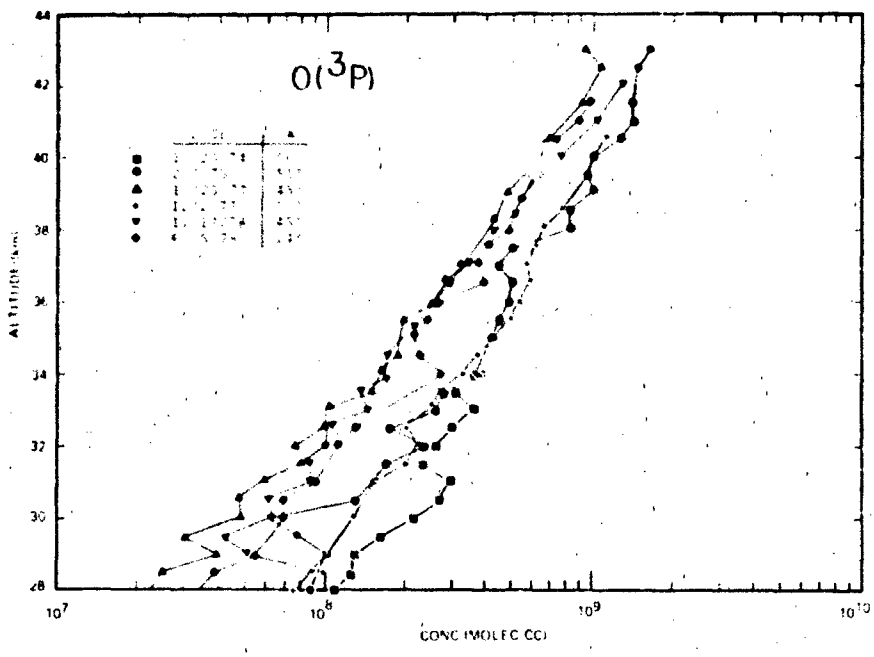


Figure 21-37. Observed concentration of $O(^3P)$ between 28 and 43 km in the stratosphere. All determined *in situ* using atomic resonance fluorescence.

midday long-path measurements are also included. The result is a more nearly homogeneous, relatively high sun profile that should afford a better defined comparison between observations and theory. The exclusion of the results of the remote techniques does not unduly restrict the size of the data set on which the present height profiles can be based, since there are a sizable number of *in situ* measurements made by several groups, using a variety of techniques: chemiluminescence deployed with balloons [Ridley and Howlett, 1974; Drummond et al., 1977], aircraft [Loewenstein et al., 1975], and rockets [Mason and Horvath, 1976]; photoionization mass spectrometer [Aikin and Maier, 1978], spin-flip laser absorption [Patel et al., 1974]. A balloon-borne pressure-modulated radiometer [Chaloner et al., 1978] has been used to obtain midday long-path information.

Assuming instrumental reproducibility, the variations in the results from a series of flights employing the same instrument are the best indicator of natural seasonal and geographic trends, since it is likely that differences between the results of different research groups and/or instrument packages can be dominated by unknown systematic instrumental discrepancies. Consequently the approach used is (a) to examine separately the data of each group that has accrued sizeable sets for possible seasonal and geographic effects, and (b) then to combine the data of all of the groups into an appropriate profile.

Figure 21-38 shows the balloon-borne measurements of Ridley and coworkers [Roy et al., 1980; Ridley and Schiff, 1981; Ridley and Hastie, 1981] made in October, 1977 and thereafter. All six of these flights were made with a chemiluminescence instrument that incorporated an improved inlet and in-flight calibration procedure [Ridley and Schiff, 1981]. The internal consistency of the data set in Figure 21-38 strongly suggests that a sizable part of the much larger variation that this group observed earlier from flight-to-flight was instrumental. For example, the nearly coincident, half-filled symbols represent data gathered from three flights at the same place and season (32°N, fall), but in two different years. In addition, the data taken on two flights at the corresponding Southern Hemispheric latitude, but different equivalent season (SH, summer), are only slightly lower than the NH results. Lastly, measurements from the summer flight at 51°N lie wholly within the 32° data set. Since the differences between the results of the flights are very nearly equal to the variations within any one of the flights, these data present no evidence of systematic patterns over the given parameter ranges: summer and fall, 50° to 30°N and 30°S, solar zenith angle 37 to 75°. Therefore, they present no reason not to take a factor-of-three-wide band of NO mixing ratios as representative for these parameter ranges and for the indicated altitudes.

Figure 21-39 shows the rocket-borne chemiluminescence measurements made by Horvath and Mason over a

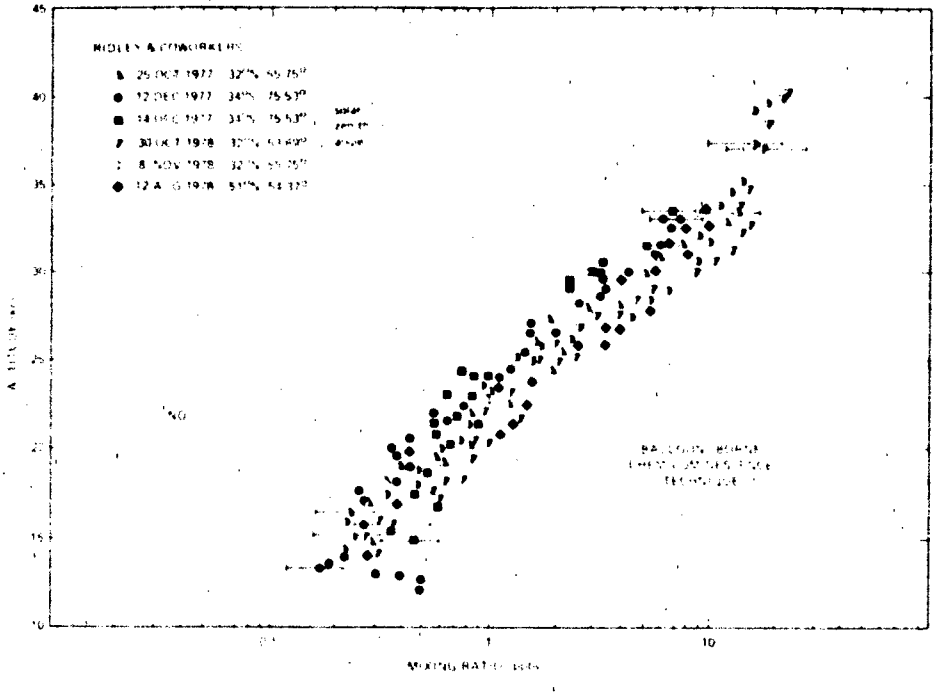


Figure 21-38 *In situ* NO mixing ratio measurements of Ridley and coworkers. All of the flights were made with instrumentation that incorporated a new inlet and flight calibration procedures

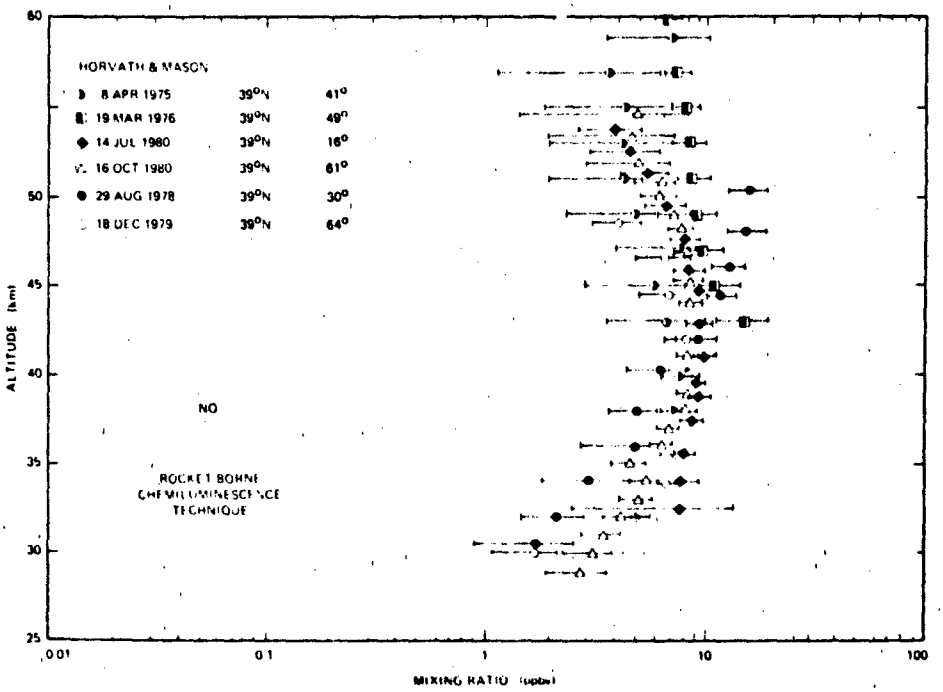


Figure 21-39. *In situ* NO mixing ratio measurements of Horvath and Mason. The 1980 measurements were made on parachute descent.

five-year period at one location, 39°N, and four seasons, [Mason and Horvath, 1976; Horvath and Mason, 1978; Horvath, personal communication]. There is good agreement within the data set. The later measurements (1980 and 1981) have better precision because they were taken on parachute descent, rather than rocket ascent. No clear seasonal trend emerges from this data set. The data set's homogeneity does, however, establish a factor-of-three band of NO mixing ratios at these altitudes, which are generally higher than those reached by Ridley and coworkers.

The measurement bands corresponding to the two extensive NO mixing ratio data sets are reproduced in Figure 21-40. The width of the indicated ranges includes essentially all of the individual measurements and their estimated uncertainties. The results of four other measurement series are shown. The first is that of Loewenstein and coworkers, whose chemiluminescence instrument has been flown extensively at two altitudes aboard U-2 aircraft [Loewenstein and Savage, 1975; Loewenstein et al., 1975, 1977, 1978a, b]. The rectangles in Figure 21-40 represent the range of NO mixing ratios found in spring, summer, and fall and over latitudes from 5° to 50°N. Measurements were also made in winter at mid-latitudes and above 50°N. A pronounced winter variation was found and is discussed in detail below. The winter variation is excluded from the data in Figure 21-40. The second measurement series represented in Figure 21-40 is a short one, namely, two flights by Patel and coworkers, who used a spin-flip Raman laser to detect NO in absorption in a multipass cell [Patel et al., 1974; Burkhardt et al., 1975]. Although only two flights were performed, the results, taken in the fall of 1973 and the spring of 1974, are in remarkably good agreement for several hours at float altitude. A single data point, representing the noontime mean, is given in Figure 21-40.

As Figure 21-40 shows, the data sets are in good agreement. Even though there are some discrepancies, there is nevertheless substantial overlap. There are, of course, some *in situ* and remote measurements that have been excluded from this comparison and that conflict with the data in Figure 21-40. It is worth stressing here that these exclusions were based on the reasons given above and not on the fact that they conflict. Furthermore, some of the excluded measurements agree with those in Figure 21-40.

Diurnal Variation

Observations have confirmed all of the major NO diurnal variations expected from the stratospheric odd-nitrogen chemistry: essentially no NO at night, a rapid increase at sunrise, a slow increase during the daytime, and a rapid decrease at sunset. All of these features have been observed in detail with both *in situ* and remote techniques.

The most extensive set of observations are from the *in situ* studies of Ridley and coworkers, who used their balloon-borne chemiluminescence instrument. Figures 21-41 and 21-42 show sunrise and sunset data, respectively. The sunrise 1975 flight employed two separate chemiluminescence instruments [Ridley et al., 1977], neither of which had the improved inlet calibration procedures adopted in 1977 [Ridley and Schiff, 1981]. Although it might be for-

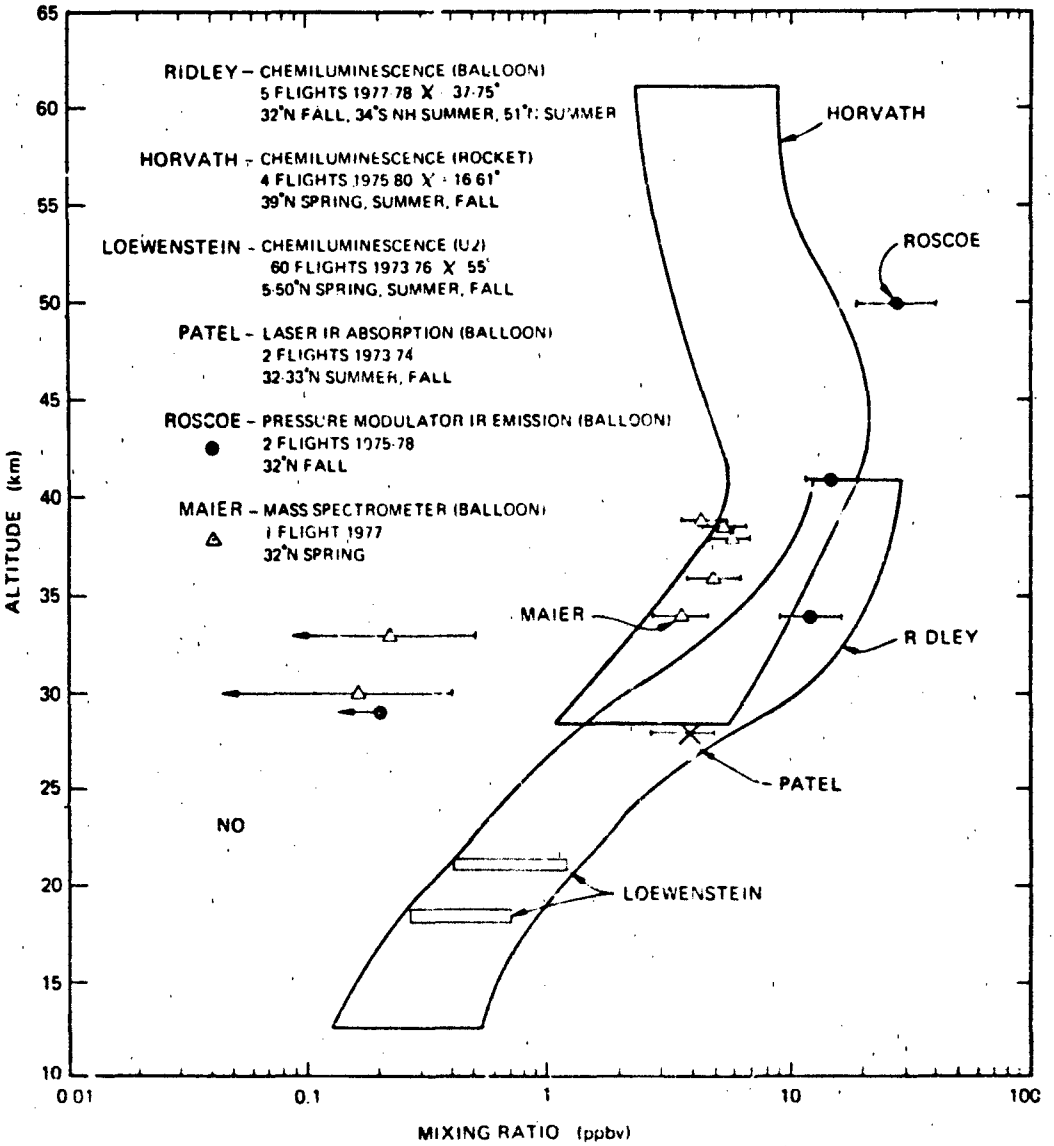


Figure 21-40. *In situ* NO mixing ratio measurements of several series of flight yielding self-consistent results

tuitous, the NO mixing ratios are, however, in accord with those from their later flights (see Figure 21-38). Figure 21-42 shows this group's sunset measurements [Ridley and Schiff, 1981]. Both sets of data reveal these rapidly changing events in remarkable detail and Figure 21-41 clearly shows a slow NO increase during the daytime.

Seasonal Variation

The variation of NO with season necessarily requires an extensive measurement program. Fortunately, several such studies have been conducted with different techniques. The various results are in reasonable harmony.

The most extensive of such investigations are those of Loewenstein and coworkers using a U-2 chemiluminescence instrument. The studies have revealed two major seasonal effects. The first of these stems from a 4-year flight series at 21.3 km during all months of the year. The results are shown in Figure 21-43 [Loewenstein et al., 1977]. A rather sharp winter minimum and a broader summer maximum is apparent. The ratio of the maximum and minimum concentrations is about six. The reproducibility of the pattern over 4 years makes it difficult to doubt its reality. Furthermore, the same trend, although defined by less data, has been found at 18.3 km.

The second striking seasonal variation discovered by Loewenstein and coworkers is shown in Figure 21-44. The data are from several summer and fall flights at 18 km

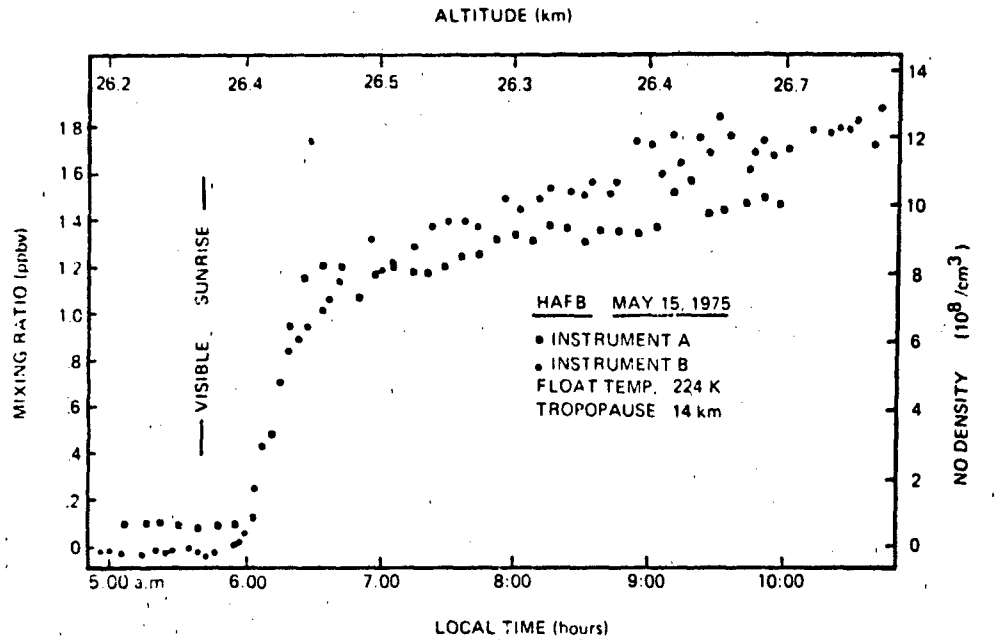


Figure 21-41. The sunrise *in situ* NO measurements of Ridley et al. [1977] obtained using two chemiluminescence instruments simultaneously flown on 15 May 1979 near 33°N, 106°W.

altitude and from 5°N to 80°N latitude. North of about 50° or 60°, the NO concentration exhibits a marked seasonal variation; the summer values are an order of magnitude larger than the fall values. There were no winter or spring high-latitude flights, so only a summer-fall comparison is possible. The fall high-latitude values are nearly zero; therefore, the winter values cannot be much lower. At 21 km, the high-latitude data are sparse; hence, the spring-fall difference seen at 18 km could not be investigated.

Latitudinal Variations

Between 5°N and about 50°N, there is substantial evidence that the latitudinal variation of NO is not large. Figure 21-47 shows the north-south variation for summer at 18 and 21 km altitude measured by Loewenstein et al. [1978a]. These data are typical of those obtained from their other north-south transects. The largest 5°N-to-50°N difference is about 2.5, the values increasing in a northern direction.

21.2.6.2 Nitrogen Dioxide (NO_2). Several extensive flight and ground-based measurement programs have established a better picture of stratospheric NO_2 in recent years. NO_2 is a rather variable constituent, both in time and in location, and some of its now recognized variations, such as the winter minimum at high latitudes, are an unanswered challenge to theory.

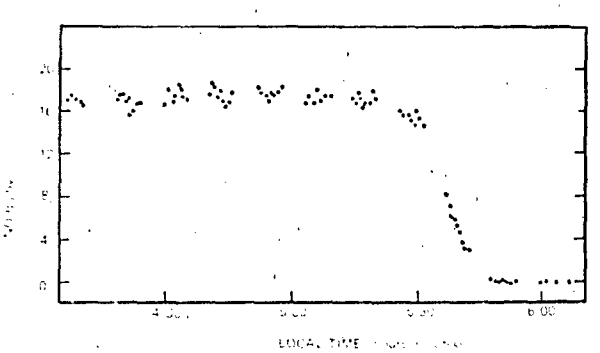


Figure 21-42. The sunset *in situ* measurements of Ridley and Schiff [1981] obtained with a chemiluminescence instrument flown on 8 Nov 1978 near 32°N, 96°W.

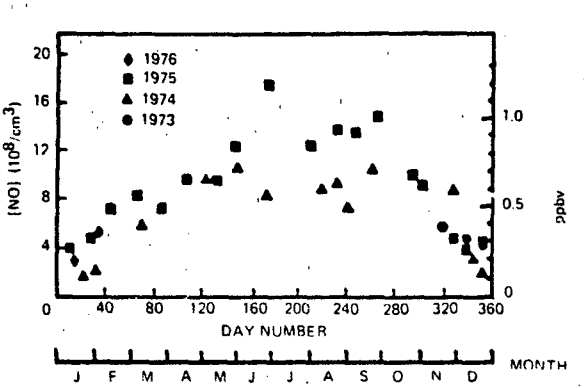


Figure 21-43. Nitric oxide seasonal data (122°W, 40°N) summary at 21.3 km. The in situ NO measurements of Loewenstein et al. [1977] obtained with a chemiluminescence instrument.

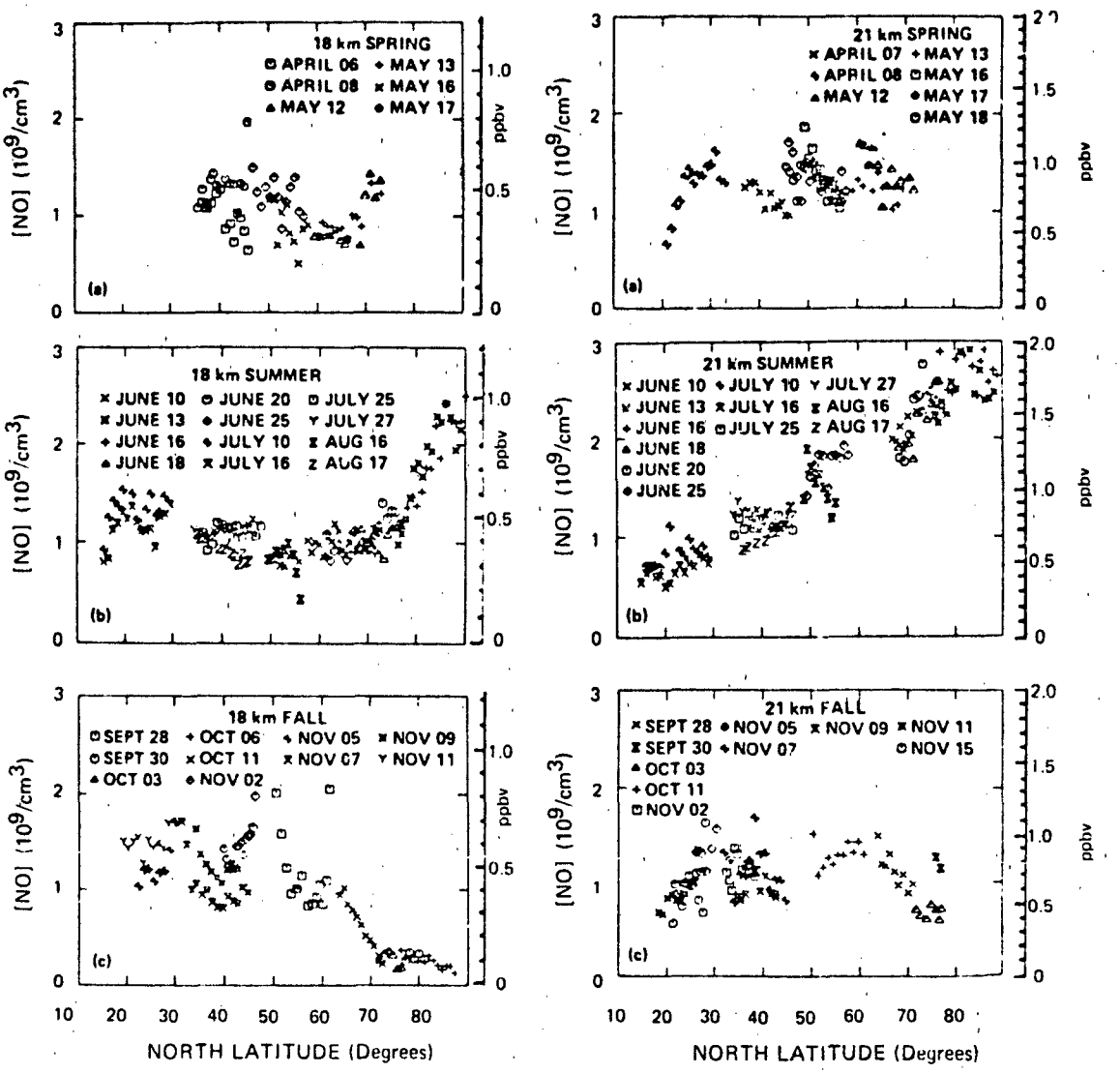


Figure 21-44. The *in situ* NO measurements of Loewenstein et al. [1978a] obtained with a chemiluminescence instrument in 1975 and April–May 1976.

Altitude Profile

There are virtually no *in situ* measurements of NO_2 in the stratosphere; hence, essentially all of the information about the NO_2 altitude profile has come from remote techniques. Most of these methods have required the long path associated with the rising or setting Sun, obtaining the vertical variation of the NO_2 mixing ratio by unfolding it from the change of the slant column density as a function of the viewing angle. Thus, the majority of the NO_2 altitude profile measurements are grouped into either sunrise or sunset profiles. The sunset mixing ratios of NO_2 are larger than those at sunrise, the increased NO_2 having been formed by the photolysis of N_2O_5. In addition to this sunrise/sunset difference, there are marked seasonal and latitudinal variations in the vertical column of NO_2. As noted in detail below, the NO_2 vertical-column values increase with increasing latitude and are larger in the summer than winter. Therefore, the large number of profile observations must be gathered into subgroups of certain times, places, and seasons, in order to have a well-defined, homogeneous profile. The largest of these is the subset of sunset profiles. This subset is examined here to see if they also show latitudinal or seasonal trends.

Figures 21-45 21-46, and 21-47 are observed sunset profiles at approximately 32°, 48°, and 55°N, respectively. Each profile is made up of at least two flights. The error bars reflect the reported uncertainties.

Figure 21-45 contains the results of the extensive flight series of Murcray and coworkers, who employed infrared [Murcray et al., 1974; Blatherwick et al., 1980] and visible

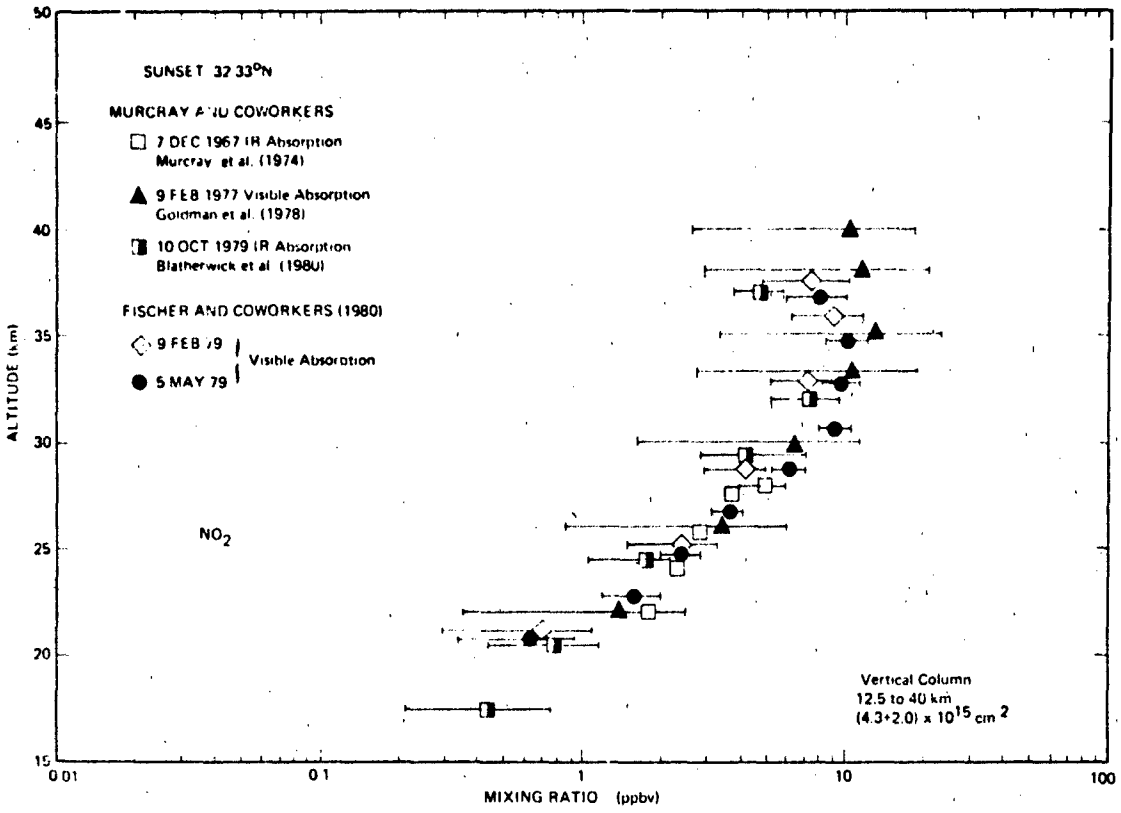


Figure 21-45. Remote measurements of the sunset altitude profile of NO_2 made at 32 to 33°N latitude.

[Goldan et al., 1978] absorption techniques from a balloon platform launched in the southern United States. Accompanying these data are the results of the two recent flights of Fischer and coworkers [personal communication, 1981]. The data set includes winter, spring, and fall flights, but there are no obvious seasonal differences between the profiles. This lack of variation is in accord with the vertical-column observations, which find little seasonal variation at latitudes below 35°N. However, the lack of a summer flight precludes the most sensitive test, namely, a summer/winter comparison.

Figure 21-46 gives the results of flights that were made in the early 1970s using infrared absorption on balloon [Ackerman and Muller, 1973; Ackerman et al., 1975] and aircraft [Fontanella et al., 1975] platforms in France. The 16×10^{-9} value of Rigaud et al. [1977] at 37 km in May of 1976 seemed too large to warrant inclusion. Three seasons are represented (winter is missing), but the small data set and the experimental uncertainties complicate the examination for seasonal trends.

Figure 21-47 shows the results from flights made over a 3-year period in Canada using visible absorption techniques [Kerr and McElroy, 1976; Evans et al., 1978]. Two sets of results are shown. The data of Evans et al. [1978] summarize the results of four flights. All five flights occurred in the summer.

The data in Figures 21-45, 21-46, and 21-47 are in good agreement within each latitude range. The high-latitude mixing ratios are up to a factor of two larger than the mid-latitude values in the range from 15 to 30 km, as a comparison of Figures 21-45 and 21-47 shows. Since the concentrations associated with such profiles reach a maximum value at about 25 km, these larger high-latitude values cause the associated vertical-column value to be larger. Although the uncertainties are relatively large, this 2.4×10^{15} cm^{-2} difference, which is a 50% increase corresponding to a 20° change in latitude in the summer, is in good agreement with the change that has been observed in vertical-column studies of the latitudinal dependence of NO_2 in the summer. Unfortunately, no one research group has data represented in two or more of Figures 21-45 through 21-47. Therefore, it remains possible that the agreement may be fortuitous, particularly since the 47° to 50°N vertical-column datum does not fit very well into the trend. It is nevertheless reassuring that the associated vertical-column values determined from these profiles agree fairly well with those directly measured.

Figure 21-48 contains the results of measurements that

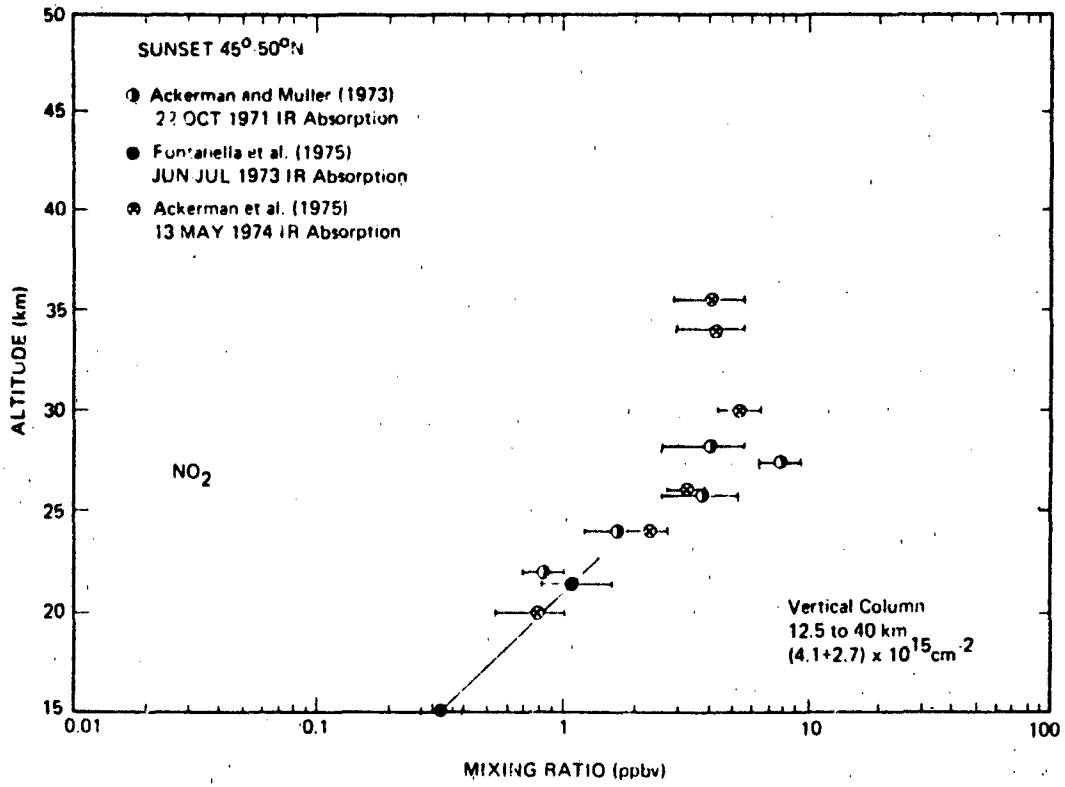


Figure 21-46. Remote measurements of the sunset altitude profile of NO_2 made at 45 to 50°N latitudes.

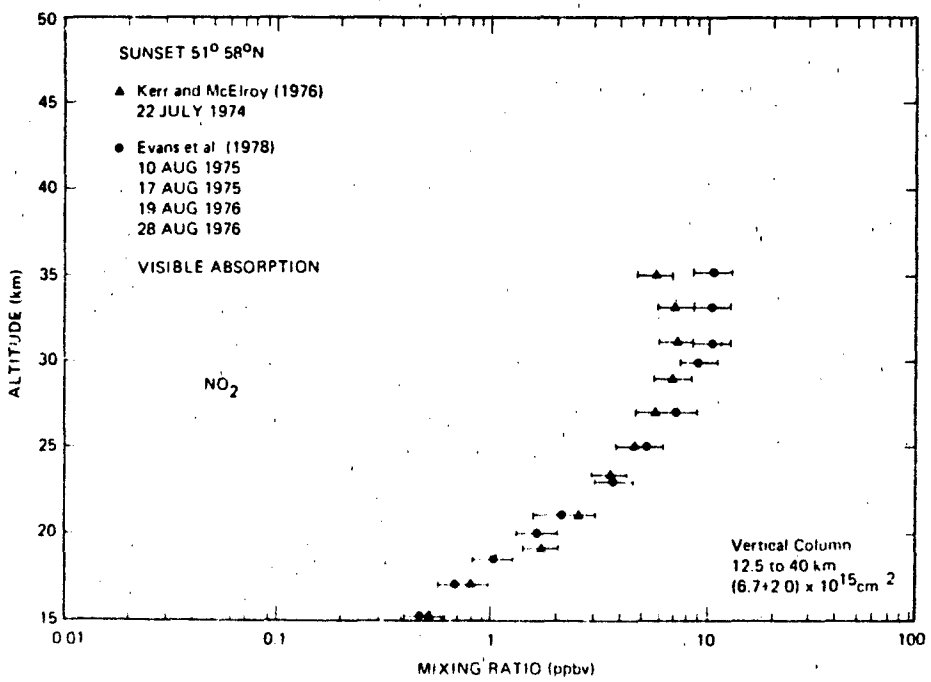


Figure 21-47. Remote measurements of the sunset altitude profile of NO_2 made at 51 to 58°N latitudes.

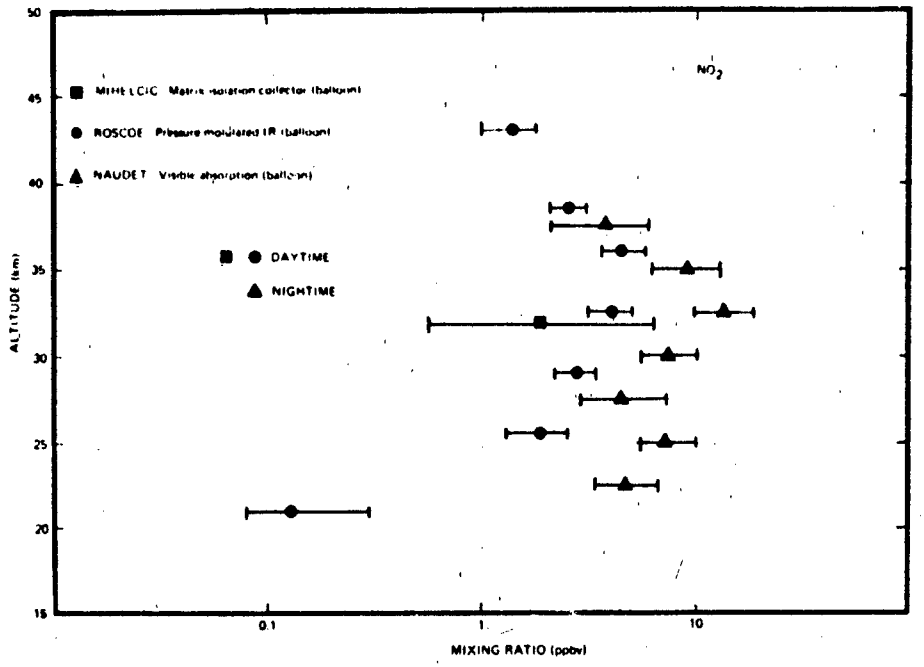


Figure 21-48. Results of full daytime and full nighttime measurements of NO_2. The data of Roscoe et al. [1981] (daytime) and Naudet et al. [1980] (nighttime) were obtained using long pathlength techniques, while the daytime data of Mihelcic et al. [1978] were collected *in situ*.

were made in full daytime and nighttime. The daytime studies are from the long-path pressure-modulated infrared radiometer of Roscoe et al. [1981] and the *in situ* matrix isolation collector of Mihelcic et al. [1978]. Earlier daytime long-path studies [Harries et al., 1976; Drummond and Jarnot, 1978] had considerably less precision and are not included here. The nighttime study is the long-path absorption investigation of Naudet et al. [1980] who used a star as light source.

Diurnal Variations

Both vertical-profile and vertical-column measurements have defined the diurnal variation of NO_2. Figure 21-49 shows the morning-evening difference reported by Evans et al. [1978], who averaged the results of four flights. The decrease from evening to morning is about a factor of two, which agrees with vertical-column measurements. Figure 21-50 shows the results of Mankin and coworkers [Coffey et al., 1981], who used infrared absorption spectroscopy

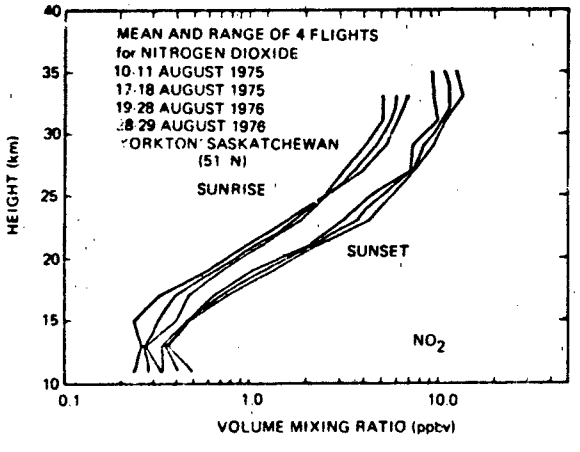


Figure 21-49. The sunrise and sunset altitude profiles of NO_2 reported by Evans et al. [1978] from the Canadian stratoprobe flight series. The upper and lower limits indicate the maximum observed deviations from the mean. The measurements were made using a balloon-borne visible absorption apparatus.

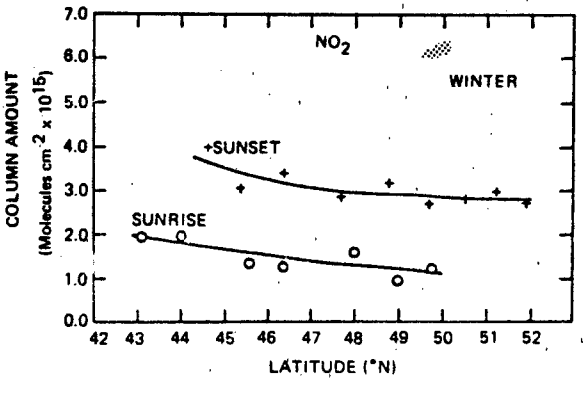


Figure 21-50. Sunrise and sunset vertical-column measurements of Mankin and coworkers, who used an infrared absorption apparatus on an aircraft platform [Coffey et al., 1981].

aboard an aircraft. Noxon's ground-based absorption spectroscopic technique [Noxon et al., 1979; Noxon, 1980] shows a factor of two larger NO_2 vertical column density at night when compared to daytime values, which is consistent with the above studies. Girard et al. [1978/1979], using similar techniques, did not initially find a sunrise-sunset difference; however, recent, more precise measurements (Girard, personal communication) have found this difference.

Seasonal Variations

The seasonal variations of NO_2 are best illustrated by the extensive ground-based measurements of Noxon [1979]. Figure 21-51 shows the results of 4 years of vertical-column measurements at various northern latitudes. The winter minimum and summer maximum is extremely regular and the ratio is as large as a factor of five at the higher latitude. Girard et al. [1978/79] report the same seasonal trend, but it is less well defined in their smaller data set. Other vertical column measurements have also been less extensive and some unusually large values have been reported [Pommereau and Hauchecorne, 1979].

21.2.6.3 Nitric Acid (HNO_3). HNO_3 has been studied with a variety of *in situ* and remote techniques. The main features of the height profile have been established and much of the latitudinal variation is now well established. The current status is summarized in the section below and details are contained in the remaining sections.

Altitude Profile

The altitude profile of nitric acid in the stratosphere has been established experimentally by *in situ* and remote techniques at Northern Hemispheric midlatitudes and is shown in Figure 21-52.

Two *in situ* methods are represented. The first is that of Lazrus and Gandrud [1974], who used a filter collection technique on balloon and aircraft platforms. The second *in situ* method is the rocket-borne ion-sampling technique of Arnold and coworkers [1980], with which the HNO_3 mixing ratio is deduced from the observed ion concentrations and the ion chemistry leading to their formation from the ambient HNO_3.

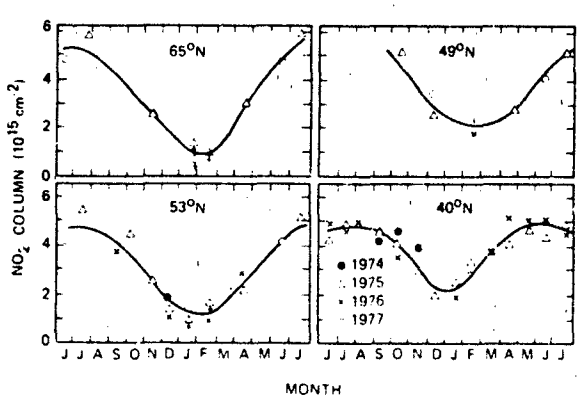


Figure 21-51 Seasonal variation of late-afternoon NO_2 at four latitudes, as given by the ground based visible absorption spectroscopic measurements by Noxon [1979]. The abundance should be multiplied by 1.25 [Noxon, 1980].

The remote measurements have employed the long path associated with the rising or setting sun. Since HNO_3 has a long lifetime, the time of the day at which the measurement was made is not as critical a parameter as it is for NO and NO_2. Thus, both *in situ* and remote measurements can be meaningfully intercompared in Figure 21-52. Both infrared emission [Evans et al., 1978; Harries et al., 1976; Murcray, personal communication, 1980] and absorption [Fontanella et al. 1975] have been used. The data of Evans et al. [1978] represent the mean of the results of four flights. All of the data in Figure 21-52 are in good agreement, with the exception of the lower-altitude results of Lazrus and Gandrud [1974] which tend to be lower than the other measurements.

There is no obvious seasonal trend in these profile data. Lazrus and Gandrud [1974] reported that their winter and spring measurements showed higher HNO_3 concentrations than did their summer and fall measurements, but the data were too sparse to be able to make a more positive statement. The vertical column density associated with these profile measurements in Figure 21-52 is $(8.6 \pm 4.0) \times 10^{15}$ cm^{-2}, which is in fair agreement with the vertical-column measurements discussed below.

Diurnal Variation

No change in the HNO_3 abundance as a function of the time of the day has been observed.

Seasonal Variation

At latitudes less than about 40°N, there is no strong evidence to support a large seasonal variation. The vertical-column measurements of Mankin and coworkers, who used an aircraft-borne infrared absorption instrument [Coffey et al., 1981], found essentially no summer-to-winter change, as Figure 21-53 shows. Lippens and Muller [1980] found the slightly higher value $(1.8 \pm 0.4) \times 10^{16}$ cm^{-2} at 40°N in April.

Latitudinal Variation

The latitudinal variation of the vertical column of HNO_3 is well documented [Murcray et al., 1975, Coffey et al., 1981]. There is a strong increase in the vertical-column density with increasing latitude, both in the Northern and Southern Hemispheres. Figure 21-54 shows the data of Murcray et al. [1975]. Above about 50 to 60°N latitude, there appears to be a pronounced seasonal variation. At these high latitudes, the winter HNO_3 concentration is high (see Figure 21-53) and in the early summer it seems to be much lower, as the data of Murcray et al. [1978] in Figure 21-55 show.

21.2.6.4 Nitrogen Trioxide (NO_3). A single height profile is available for nighttime NO_3 [Naudet et al., 1981]. The measurements were made at 43°N in September from a balloon. They are based on absorption in the visible region using Venus as the light source. The derived profile and stated uncertainties are shown in Figure 21-56.

The corresponding column abundance between 20 and

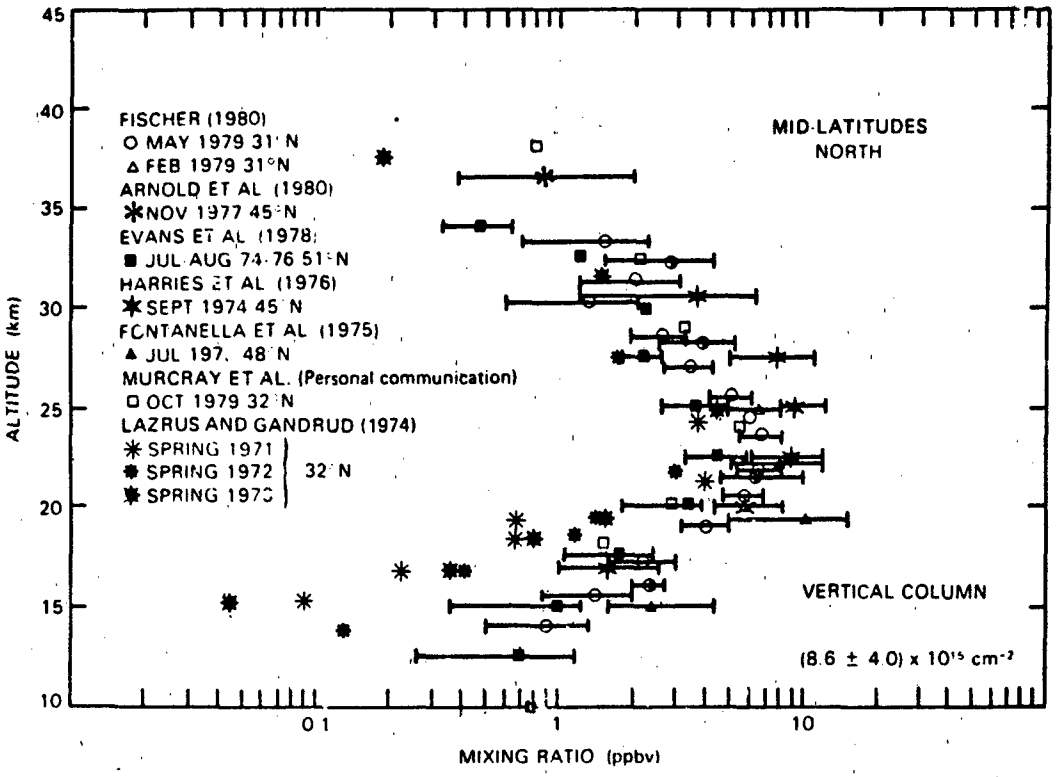


Figure 21-52. *In situ* and remote measurements of the HNO_3 mixing ratio at northern midlatitudes.

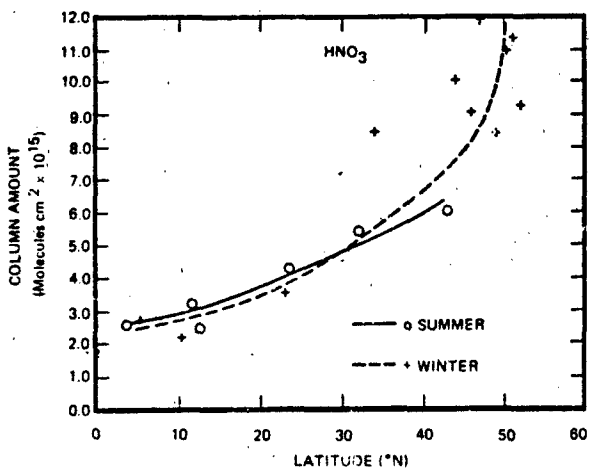


Figure 21-53. Evidence for the lack of seasonal variation in the vertical column density of HNO_3 at latitudes less than 40° as measured by Coffey et al. [1981] using infrared absorption.

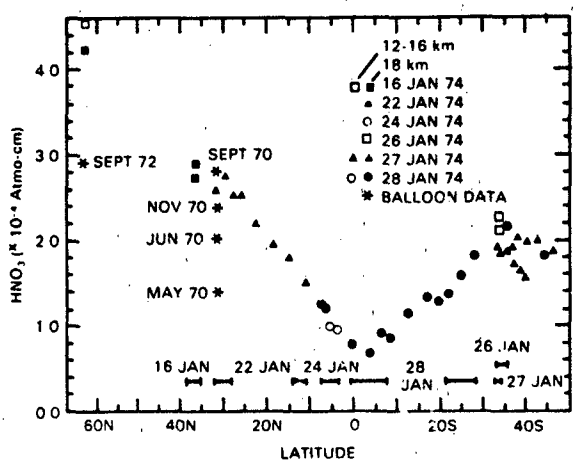


Figure 21-54. The latitudinal variation in the vertical column density of HNO_3, as measured by Murcray et al. [1975] using infrared emission (1 atm cm = 2.7×10^{19} molecules/cm^2).

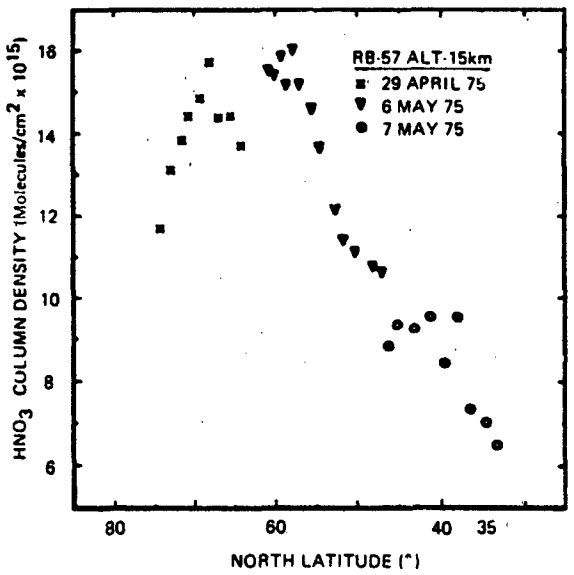


Figure 21-55. The decline of the HNO_3 vertical column density at high latitudes in the spring north of 70° as seen by Murcray et al., [1978] using infrared emission.

39 km is (3.5 ± 1.2) × 10^{13} cm^{-2}. This appears to be consistent with Noxon's estimate of about 10^{14} cm^{-2} in the spring and an upper limit of 4 × 10^{13} cm^{-2} in the summer.

21.2.6.5 Nitrogen Pentoxide (N_2O_5). There have been no new measurements of N_2O_5. This situation remains as reported in NASA RP 1049: A tentative detection of 2 ppbv at 30 km a few hours after sunrise by Evans et al. and an upper limit of 1.2 × 10^{15} cm^{-2} above 18 km in February by Murcray.

21.2.6.6 Peroxynitric Acid (HO_2NO_2). Despite the recent interest in this species no detection of its presence has been reported. The upper limit remains at 0.4 ppbv as reported in NASA RP 1049.

21.2.7 Odd Hydrogen

21.2.7.1 Hydroxyl Radical (HO). Hydroxyl has been observed in the stratosphere by four independent techniques:

- Solar flux induced resonance fluorescence observed by a rocket-borne spectrophotometer [Anderson, 1971a; Anderson, 1971b] which provides a local concentration measurement by determining the change in total column emission rate as a function of altitude.

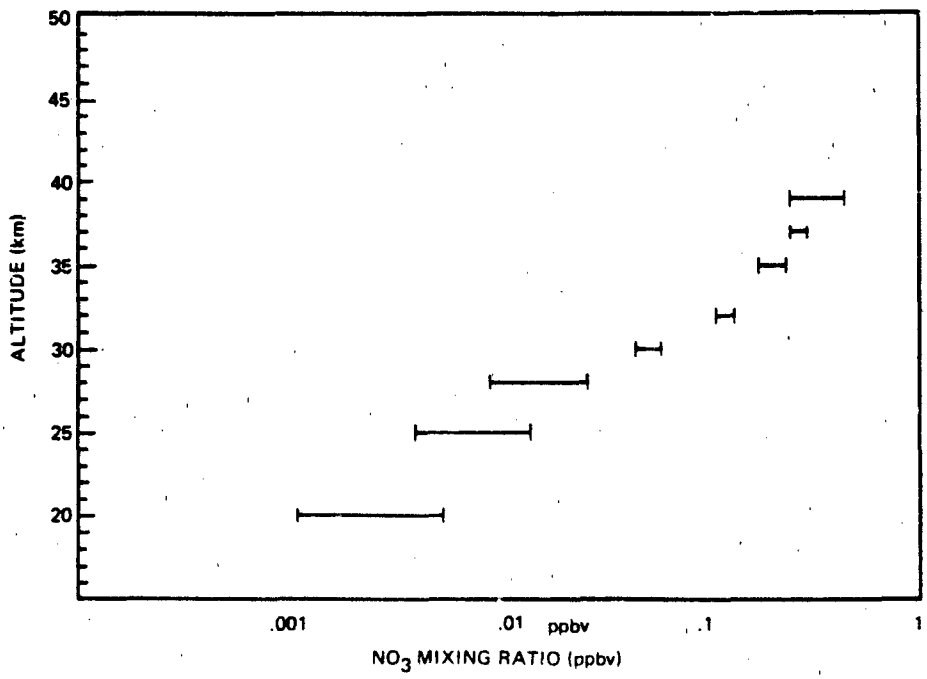


Figure 21-56. Vertical nighttime profile of NO_3 [Naudet et al., 1981] measured using visible absorption with Venus as a light source.

- Balloon-borne *in situ* molecular resonance fluorescence using a plasma discharge resonance lamp to induce fluorescence. The fluorescence chamber is lowered through the stratosphere on a parachute to control the altitude and velocity of the probe [Anderson, 1976; Anderson, 1980].
- Ground-based high resolution solar absorption by a PEPSIOS (Poly-Etalon Pressure Scanned Interferometer) instrument which resolves a single rotational line in the (O-O) band of HO at 309 nm. The total column density of terrestrial HO between the instrument and the sun is observed, dominated by the altitude interval 25 to 65 km [Burnett, 1976, 1977; Burnett and Burnett, 1981].
- Balloon-borne laser induced detection and ranging (LIDAR) in which a pulsed laser system coupled to a telescope is used to observe the backscattered fluorescence from HO. The laser is tuned to the 1-0 band of the A-X transition at 282 nm and the fluorescence at 309 nm (the 0-0 band) is observed as a function of time following the laser pulse [Heaps and McGee, 1982].

Data from the upper stratosphere rocket data from Anderson [1975] and the stratosphere balloon data using *in situ* resonance fluorescence [Anderson, 1980] and LIDAR [Heaps and McGee, 1982] are used to form a composite HO profile (Figure 21-57) with an upper limit of the mean tropospheric HO concentration taken from the methyl chloroform lifetime studies and tropospheric laser experiments.

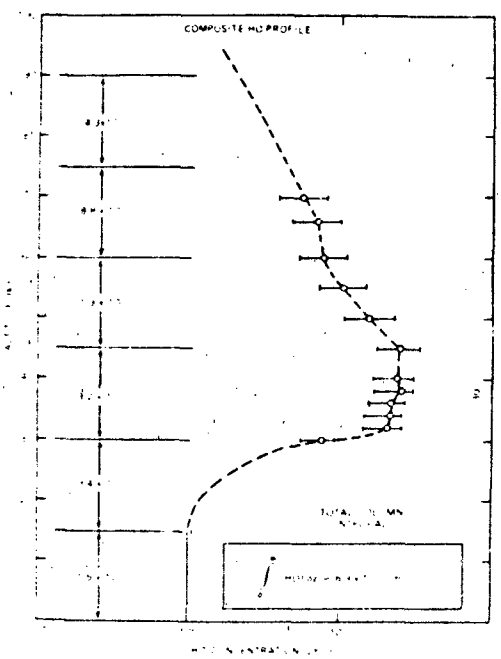


Figure 21-57. A composite HO profile based on the rocket and balloon data. The tropospheric concentration is estimated from methyl chloroform studies and the tropospheric laser experiments.

In summary, an analysis of balloon and rocket data of HO in the stratosphere and ground-based total column observations leads to the following conclusions:

1. There is substantial agreement among the three techniques: the *in situ* data provides a consistent picture of the altitude dependence of HO between 30 and 70 km implying a peak concentration at 40 km of 2.4×10^7 cm^{-3} and a total column density at midday of 6.9×10^{13} cm^{-2}. The midday total column abundance determined from the ground is 5.7×10^{13} cm^{-2}.
2. There is a systematic increase of approximately 1×10^{13} cm^{-2} per year between December 1976 and December 1979 and a suggestion of a yearly spring maximum and fall minimum. The spring to fall decrease is approximately 30%.

21.2.7.2 Hydroperoxyl Radical (HO_2). A total of four HO_2 observations have appeared in the literature, one by the matrix isolation technique and three by the resonance fluorescence method. Those observations are summarized in Figure 21-58. There is significant scatter evident in those observations which should not be attributed to atmospheric variability until:

- The signal-to-noise ratio of the observations is improved;
- Simultaneous observations of photochemically related species such as HO or H_2O demonstrates a correlation in concentration fluctuations.

21.2.7.3 Atomic Hydrogen (H). There are no reported observations of atomic hydrogen in the stratosphere either direct or indirect, nor have any upper limits been reported.

21.2.7.4 Hydrogen Peroxide (H_2O_2). The only reported observation of H_2O_2 is the tentative result reported by Waters et al. [1981] using the Balloon-Borne Microwave Limb Sounder (BMLS) to observe the purely rotational emission of H_2O_2 at 204 GHz. All aspects of the experimental hardware and uncertainties analysis are identical to that discussed in the BMLS section of the ClO discussion.

21.2.8. Odd Chlorine

21.2.8.1 Chlorine Oxide (CO).

Seasonal Variation

Although there is an indication of some seasonal dependence in ClO, the data base is clearly inadequate to draw any clear conclusions. This question should be addressed both by an improvement in the accuracy and precision of the balloon-borne methods and by the more extensive deployment of ground-based methods. Only when those are done in concert will an adequate definition of this important point emerge.

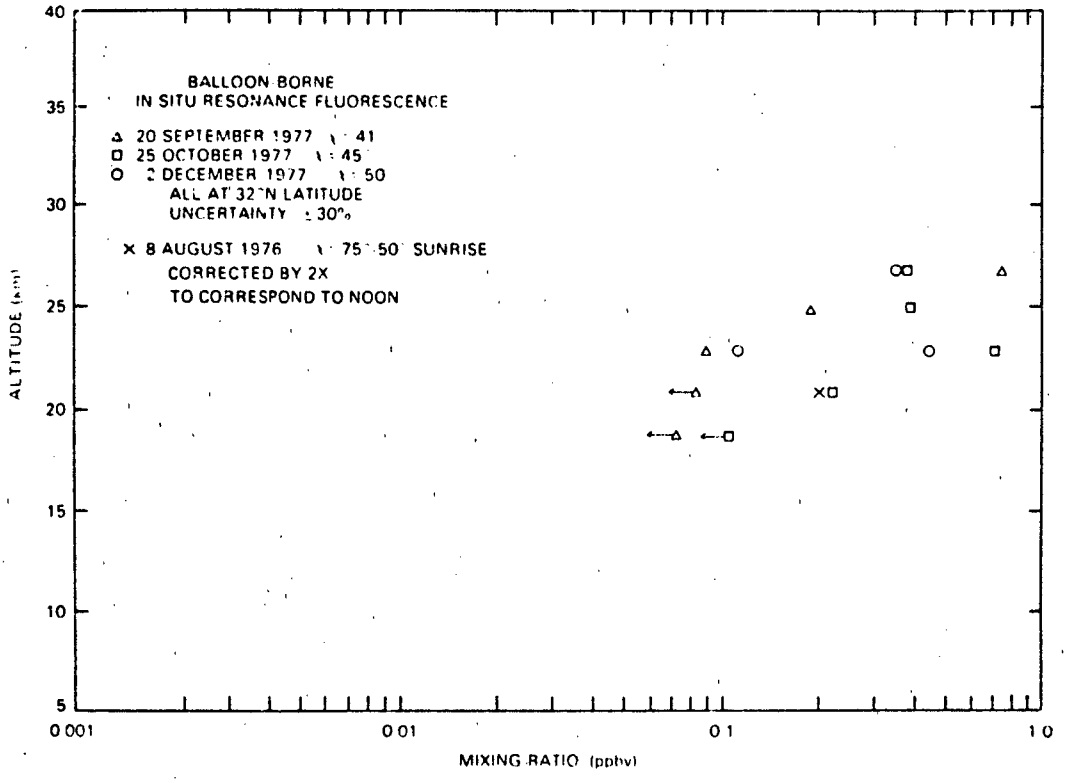


Figure 21-58. Vertical profiles of HO_2 from balloon-borne resonance fluorescence measurements [Anderson et al., 1981].

Latitude Variation

There is virtually no information on the critical question of latitude variation. The balloon-borne observations are all carried out at 32°N and the ground based data at 42°N. As will be noted in the subsequent sections, if the steep gradient in ClO below the peak is characteristic of both mid- and low latitude conditions, the interpretation of chlorine induced depletion will be significantly simpler.

21.2.8.2 Hydrogen Chloride (HCl). HCl in the stratosphere has been observed by three different remote sensing techniques and one *in situ* method. Most of the presently available data on the vertical profile of concentration come from balloon-borne observations made at ~32°N latitude (Texas and New Mexico); there are, in addition, single profiles taken at 30°S (Australia) and 65°N latitude (Alaska). The profile measurements cover the altitude range from 14 to 40 km, and are supplemented by values for the total column abundance in the upper stratosphere. There is insufficient data from which to discern any seasonal variability, and the location of the altitude of peak relative abundance is not clearly established. The available data cover the period from 1975 to 1980.

Altitude Profile

Balloon-borne near-infrared absorption spectroscopy has been used to obtain vertical profiles covering the 14 to 40 km altitude range (Figure 21-59). Several groups of investigators have made measurements by this method, and obtained results which are in fairly good agreement. In addition there have been observations made by pressure-modulation radiometry, by emission spectroscopy and by absorption spectroscopy from the ground (Figure 21-60).

Seasonal Variation

Since many of the IR absorption measurements of HCl were made at the same latitude (32° and 33°N), it might be expected that any seasonal trend would be seen in this subset of the data. The results, however, do not show any variation greater than the quoted uncertainties associated with each measurement. (It should be mentioned also that the same conclusion regarding the absence of a seasonal variation was reached by Lazrus et al. [1977] from the base-impregnated filter measurements of total acidic chloride vapor.) With the currently improved precision of the remote spectroscopic instrumentation, profiles with associated uncertainties of perhaps less than 10% can be anticipated for the near future; thus, more sensitive tests of the seasonal variability of HCl could be made, provided sufficiently frequent observational opportunities are available.

21.2.8.3 Chlorine Nitrate ($ClONO_2$). The only specific detection claimed for $ClONO_2$ to date is that of Murcray et al. [1979] by IR absorption. The accuracy of the result is

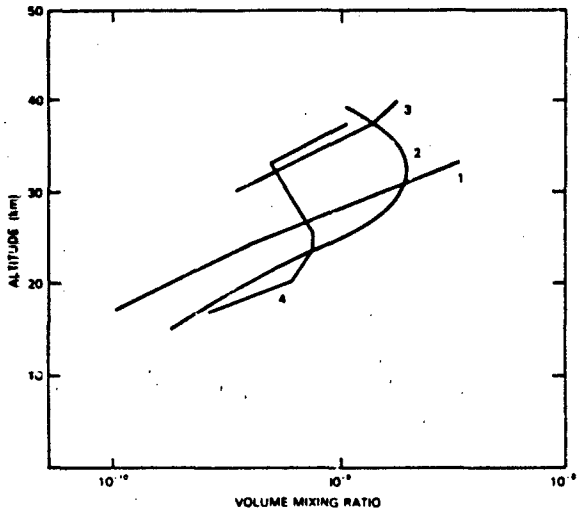


Figure 21-59. HCl measurements by (1) ground-based spectroscopy (June 79 profile), (2) pressure modulator radiometry, (3) far IR emission, and (4) *in situ* filter collection.

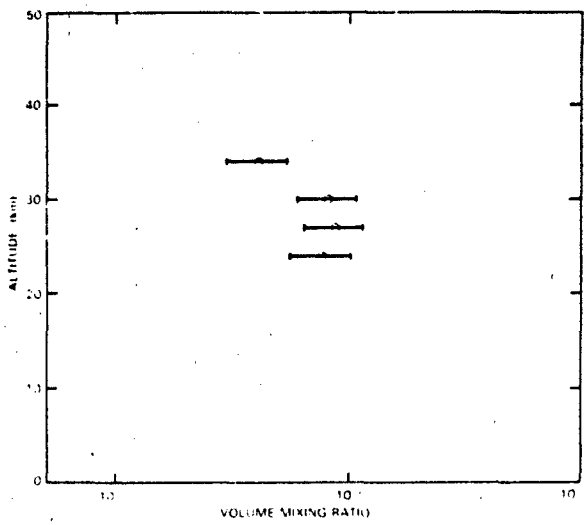


Figure 21-61. Distribution of $ClONO_2$ [Murcray et al., 1980].

affected both by modeling assumptions and by uncertainties in the intrinsic spectroscopic parameters involved in the analysis of the data (see below). The *in situ* sampling method, collection on base impregnated filters [Lazrus et al., 1977], is sensitive to all acidic chloride (see HCl) so that its results can only be used to provide an upper limit estimate for $ClONO_2$. Since the filter data are not compatible with the IR data for HCl, they cannot aid in evaluating the available remote sensing data for $ClONO_2$.

The published values from Murcray's observations, in October 1978 balloon flight (see Figure 21-61), differ from the preliminary values given in NASA RP 1049. The measurements were made by the limb absorption method, that is, by observing the sun through the atmosphere at sunrise or sunset from a stratospheric balloon platform. The strong infrared absorption by $ClONO_2$ at 1292 cm^{-1} was used by Murcray and his coworkers in their analysis; this band coincides in the stratospheric spectrum with strong absorptions by the natural gases, N_2O, CH_4, and H_2O. As the absorptions due to these constituents increase (i.e., at the lower tangents heights of observation), the superimposed $ClONO_2$ absorption is completely masked. However, even at the higher altitudes, where the $ClONO_2$ absorption might be discernible against the spectral background, the maximum expected effect does not exceed 5% or 6% depression of the HCl continuum. Thus the quantitative analysis is dependent both on the spectral model and on the quality of the experimental data. With these considerations in mind, the published data may at best indicate a possible specific identification of $ClONO_2$; the deduction of a profile of concentration with associated experimental errors of $\pm 25\%$ does not seem justified. An upper limit concentration of 10^{-9} by volume, between 25 and 35 km altitude is consistent with the observational data.

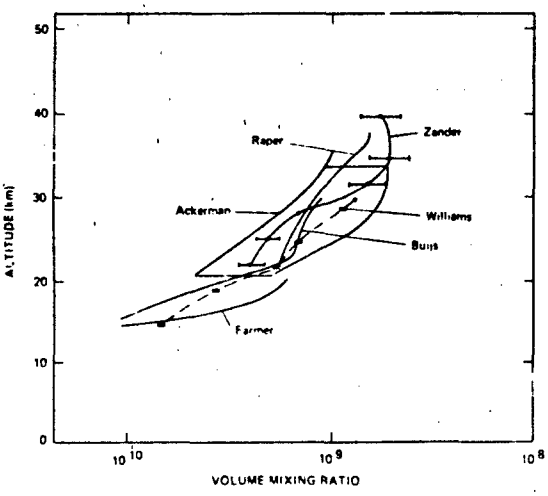


Figure 21-60. HCl distribution from balloon-borne IR absorption spectroscopy.

21.2.9 Other Halogens

21.2.9.1 Hydrogen Fluoride (HF). Stratospheric HF has been measured by several different groups using both remote sensing and *in situ* techniques. The measurements for the most part have been made at different locations and seasons and do not include a sub-set of observations similar to those for HCl from which a most probable profile can be derived. The measurements that have been made to date are summarized in Figure 21-62.

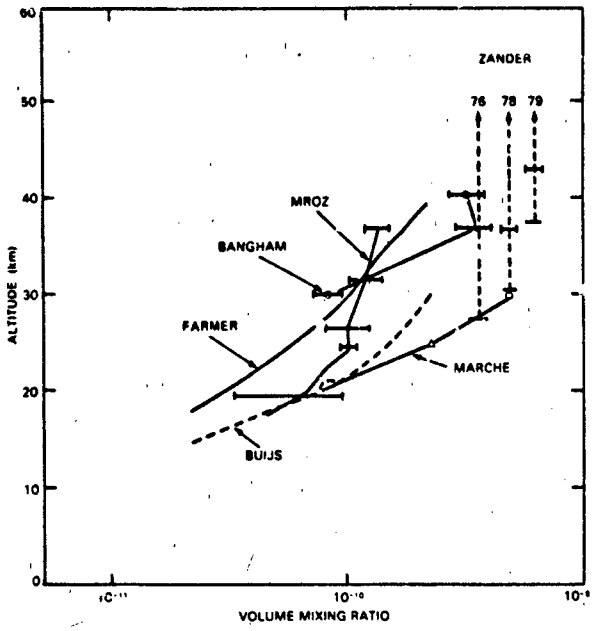


Figure 21-62. Stratospheric HF profile measurements.

21.3 MESOSPHERE

Within the mesosphere, the concentration of many of the minor species can vary by orders of magnitude during a diurnal cycle. The photochemical time constants are so short that extreme gradients are found when temporal variations are studied. However, it is still instructive to consider a case such as Figure 21-63a to indicate the approximate concentrations under a particular set of conditions. Here the results from a time dependent one-dimensional dynamic model are shown based on the work of Keneshea et al. [1979] for the case of midlatitude noontime conditions. There are additional minor species which are not shown in Figure 21-63a; however, these are of particular aeronomic interest. The time constants of some of the middle atmosphere minor species which are controlled by chemical lifetime and transport are shown in Figure 21-63b from a model by Allen et al. [1981]. Part of the discussion of the mesospheric composition will necessarily overlap into the region of lower thermosphere because of the important span between 90 and 120 km where the atmosphere departs from its mixed state to begin diffusive separation. Our understanding of this

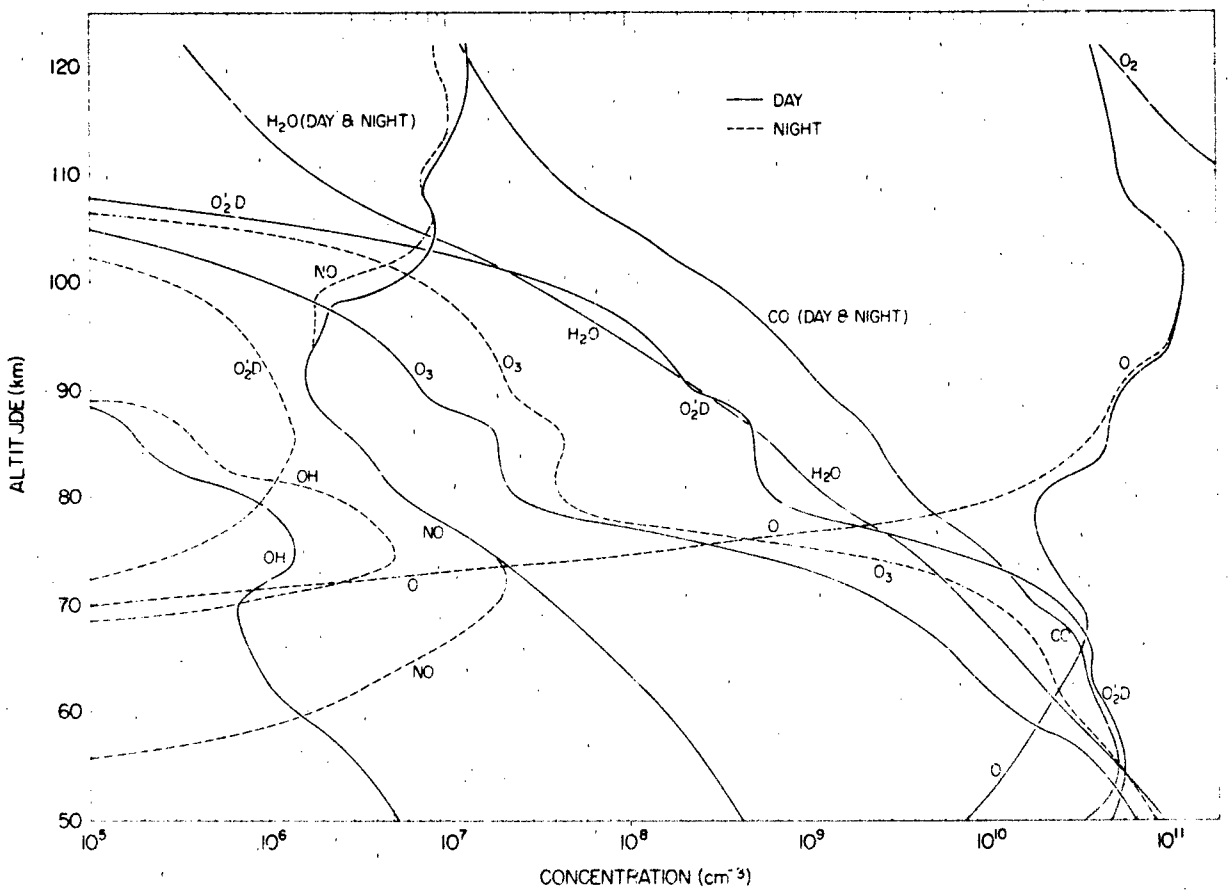


Figure 21-63a. Vertical profiles of neutral minor constituents for noontime conditions at 45°N latitude from time dependent model calculations [Keneshea et al., 1979] (Reprinted with permission from Pergamon Press Ltd. © 1979.)

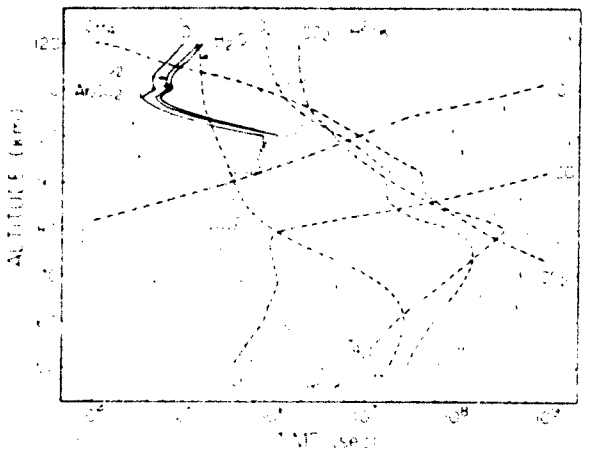


Figure 21-63b A comparison of the chemical lifetimes (dashed line) of middle atmosphere species and the time scales for transport due to eddy diffusion (dotted line) and molecular diffusion (solid line) [Allen et al., 1981].

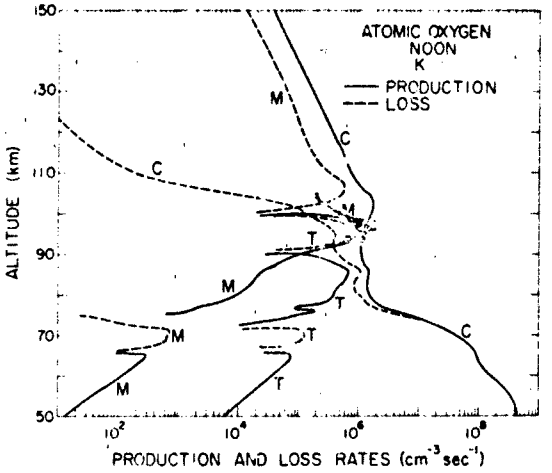


Figure 21-64a. Comparison of the production (———) and loss terms (– – – –) for atomic oxygen at noon. The curves are labeled to represent the contributions of chemistry, C, the amplitude of the molecular flux gradient, M, and the amplitude of the turbulent flux gradient, T [Keneshea et al., 1979]. (Reprinted with permission from Pergamon Press Ltd. © 1979.)

regional has progressed significantly during the last 10 years because of the development of more advanced instruments for rocket investigations [von Zahn, 1973]. For convenience, the discussion will be divided into the following catagories: the oxygen constituents (O, O_3), the carbon-oxygen constituents (CO_2, CO), the hydrogen-oxygen constituents (H_2O, OH, HO_2, H_2O_2), the nitrogen-oxygen constituents (NO, NO_2, N_2O), the excited state species ($O_2[^1\Delta_g]$, $N[^2D]$), the meteoric/dust constituents, and the inert gases (He, Ne, Ar, Kr, Xe). The acids, halogens and chlorine compounds and other trace species have been discussed in the section on the stratospheric species and these are less important in our current understanding of the mesosphere.

21.3.1 The Oxygen Constituents

The atomic oxygen, which results primarily from the dissociation of molecular oxygen in the Schumann-Runge bands, is an important minor species in the mesosphere and becomes the major atmospheric species in the thermosphere for altitudes between 160 and 700 km. The atomic oxygen in the mesosphere is strongly controlled by transport. Eddy diffusion mixes the oxygen atoms to lower altitudes where they are lost by three-body reactions to form ozone in the lower mesosphere.

Figures 21-64a and 21-64b show the results of the one-dimensional time dependent calculations for atomic oxygen at noon and at midnight. The curves show the production and loss associated with the chemistry, molecular transport and turbulent transport of the atomic oxygen. In Figure 21-65 the midlatitude diurnal variation of atomic oxygen is shown. At altitudes below 80 km, the very pronounced local time dependence is observed. The variation observed is due to the chemistry changes with the assumed constant diffusion profiles. The measurement of atomic oxygen in the mesosphere requires use of rather complicated techniques because of its highly reactive properties and the influence of the measuring instrument in the relatively high pressure of the mesosphere. Mass spectrometers in this region must be cryo-pumped to reduce the gas collisions within the analysing field. The techniques which have been used include mass spectrometers [Philbrick et al., 1973; Offermann and

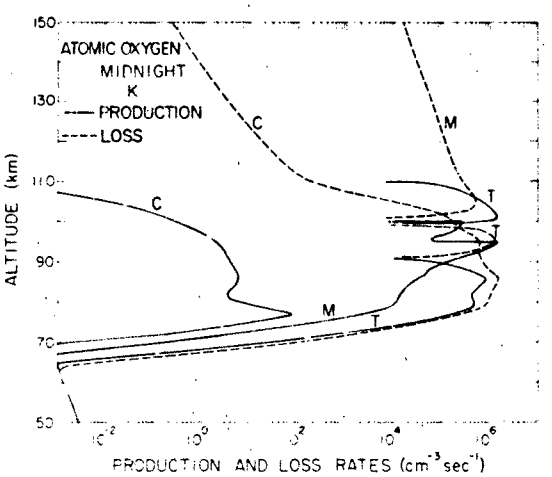


Figure 21-64b. Comparison of the production and loss terms in the chemical model for atomic oxygen at midnight. The labels are the same as in (a) [Keneshea et al., 1979]. (Reprinted with permission from Pergamon Press Ltd. © 1979.)

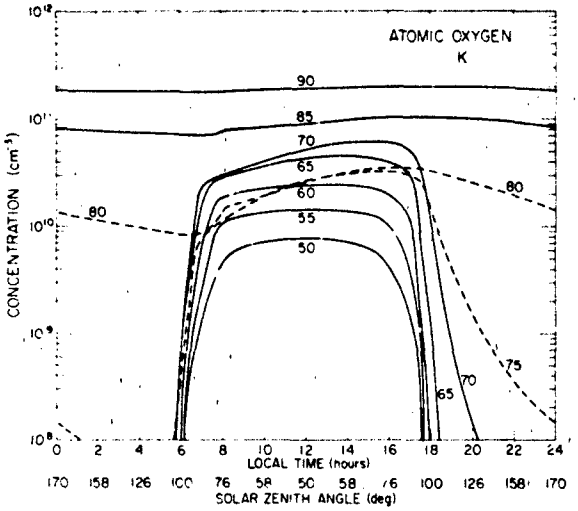


Figure 21-65. Diurnal variation of atomic oxygen between 50 and 90 km from model calculations [Keneshea et al., 1979] (Reprinted with permission from Pergamon Press Ltd. © 1979.).

Grossmann, 1973; Arnold and Krankowsky, 1977], photometers for 5577 Å emission [Dandekar, 1972; Offermann and Drescher, 1973], silver film oxidation [Henderson, 1971], resonance fluorescence [Dickinson et al., 1974; Howlett et al., 1980], NO chemiluminescence [Golomb and Good, 1972], atomic absorption [Dickinson et al., 1974], and from hydroxyl airglow [Good, 1976]. A study of the existing data by Offermann et al. [1981] has shown a strong correlation of the atomic oxygen column density between 80 and 120 km with the measured peak density. This correlation (see Figure 21-66) is quite good considering the large variability which is observed in the individual profiles shown in Figure 21-67. The oxygen profile is strongly affected by the transport properties at a given time. The fact that a large amplitude variation that may be expected due to eddy diffusion was shown in the study of Keneshea and Zimmerman [1970]. Offermann et al. [1981] also pointed out the anti-correlation of the argon and atomic oxygen behavior which is expected with changes in the eddy diffusion. Most of the measure-

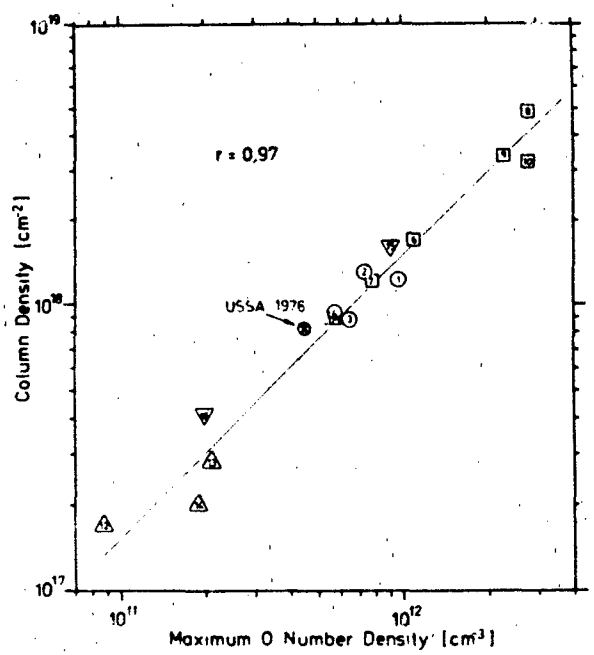


Figure 21-66. Atomic oxygen layer peak density vs O layer content. Layer content is represented by the column density up to 110 km. Different measuring techniques of 16 rocket flights are indicated by different symbols. Model values from USSA 1976 are given for comparison and the straight line is at 45°. Identification numbers in symbols are described in original text [Offermann et al., 1981] (Reprinted with permission from Pergamon Press Ltd. © 1981.)

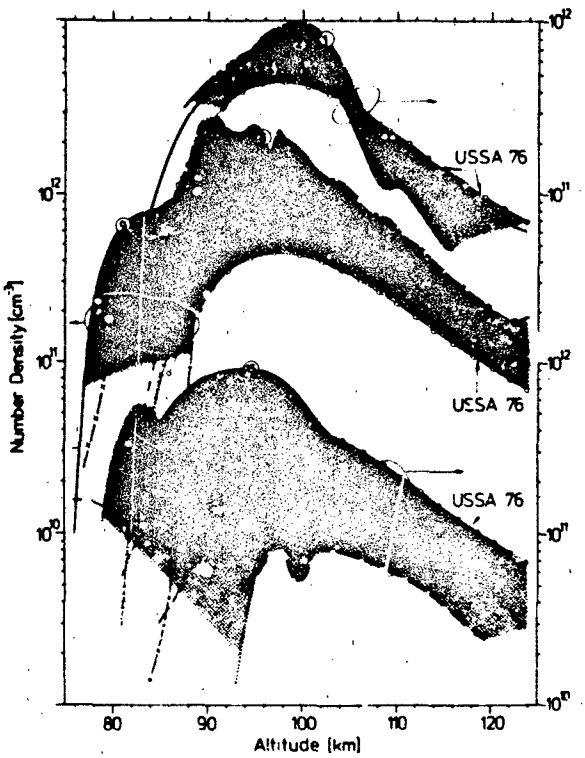


Figure 21-67. Atomic oxygen number density profiles as obtained from 16 *in situ* measurements. Results from 3 different measuring techniques are grouped together (shaded areas) and refer to 3 different density scales. USSA 1976 model profile is given for comparison with each of the 3 groups. Profile identification numbers are given in original text [Offermann et al., 1981] (Reprinted with permission from Pergamon Press Ltd. © 1981.)

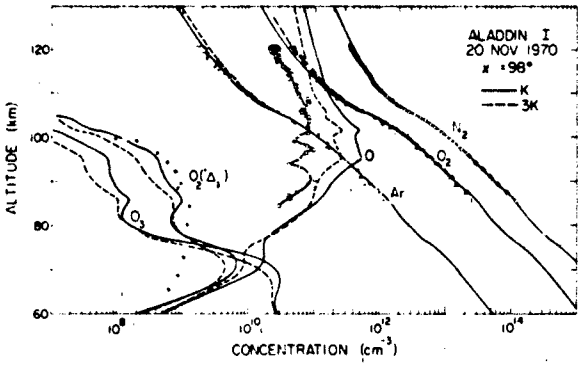


Figure 21-68. Comparison of computed species profiles for turbulent diffusivities K and 3 K (solid dashed curves) with measured profiles for O (squares), O_2 (triangle), N_2 (circles), Ar (crosses), and $O_2(^1\Delta_g)$ (dots) for the ALADDIN I program [Keneshea et al., 1979]. (Reprinted with permission from Pergamon Press Ltd. © 1979.)

ments of atomic oxygen show the irregular structure in the altitude region between 90 and 110 km due to the turbulent layers in that region. The calculations of atomic oxygen in a time-dependent transport model also show irregularities in this region where the eddy diffusion profile is irregular. An example of this structure is observed in the measurement and model calculations for the ALADDIN I experiment shown in Figure 21-68. The series of ALADDIN (*A*tmospheric *L*ayering *An* ' *D*ensity *D*istribution of *I*ons and *N*eutrals) experiments and models have provided some of the most important foundation for our current understanding of

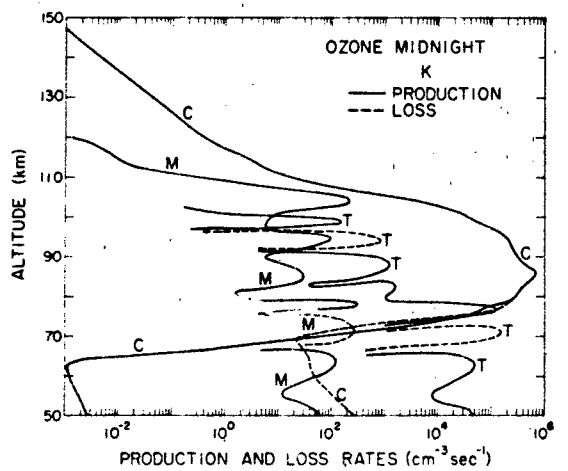


Figure 21-69. Model calculations for the ozone production and loss for midnight at midlatitudes. The label C represents chemistry effects, M represents molecular transport, and T represents turbulent transport [Keneshea et al., 1979]. (Reprinted with permission from Pergamon Press Ltd. © 1979.)

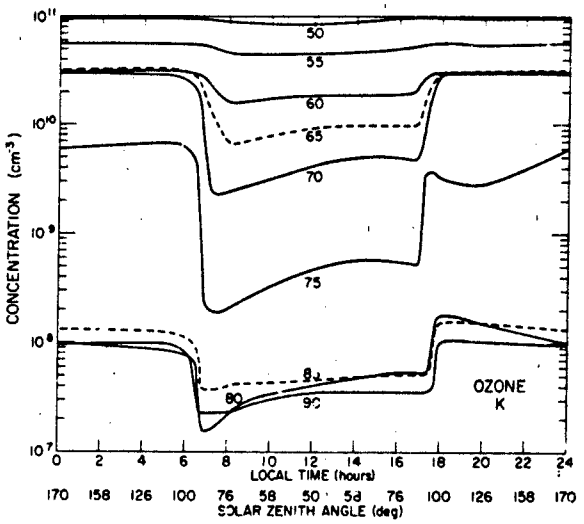


Figure 21-70. The model calculation of the diurnal variation of the O_3 [Keneshea et al., 1979]. (Reprinted with permission from Pergamon Press Ltd. © 1979.)

the middle atmosphere properties and therefore several of those results are used for illustration.

Model calculations for the ozone production and loss for midnight at midlatitudes is shown in Figure 21-69. The diurnal variation of the O_3 is shown in Figure 21-70. The measurements of ozone have been refined and both *in situ* and remote sensing techniques have provided accurate results in the stratosphere. However, there are relatively few measurements of the ozone in the mesosphere. A comparison of photometer measurements and model calculations is shown in Figure 21-71.

21.3.2 Carbon-Oxygen Constituents

The atmospheric carbon dioxide, CO_2, and its dissociative product, CO, are very important species for the role they contribute to the radiation balance and thermal structure of the middle atmosphere. The CO_2 molecules are also important in some of the chemical reactions for formation of the negative ion species of the D region. In Figure 21-72, the results of a calculation by Hays and Olivero [1970] are shown for the CO_2 dissociation and transport by eddy diffusion. The various curves represent two extremes for the recombination and a range of eddy diffusion values. The fact that the CO_2 remains near ground level mixing ratio values for altitudes up to the turbopause has been substantiated by measurements shown in Figure 21-73. Trinks and Fricke [1978] pointed out that in addition to solar dissociation, the CO_2 concentration at higher altitudes is also con-

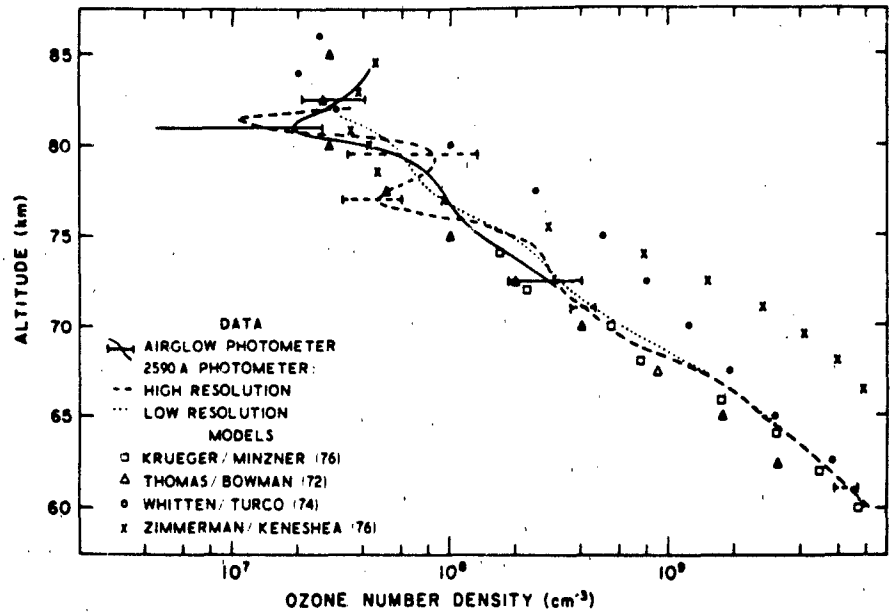


Figure 21-71. Comparison of photometer measurements and model calculations of ozone [Weeks et al., 1978].

trolled by loss from an ion molecule reaction with O^+. The calculations of the CO_2 and CO behavior in the mesosphere and thermosphere require solution of the time-dependent transport equations with appropriate considerations of the photodissociation, diffusion transport and loss by the ion molecule reaction. The CO_2 emission in the 15 μm and 4.3 μm bands has also been measured by rocket-borne IR spectrometers, Stair et al. [1975]. The measured IR signal results from a combination of the contributions of the CO_2 concentration and the temperature.

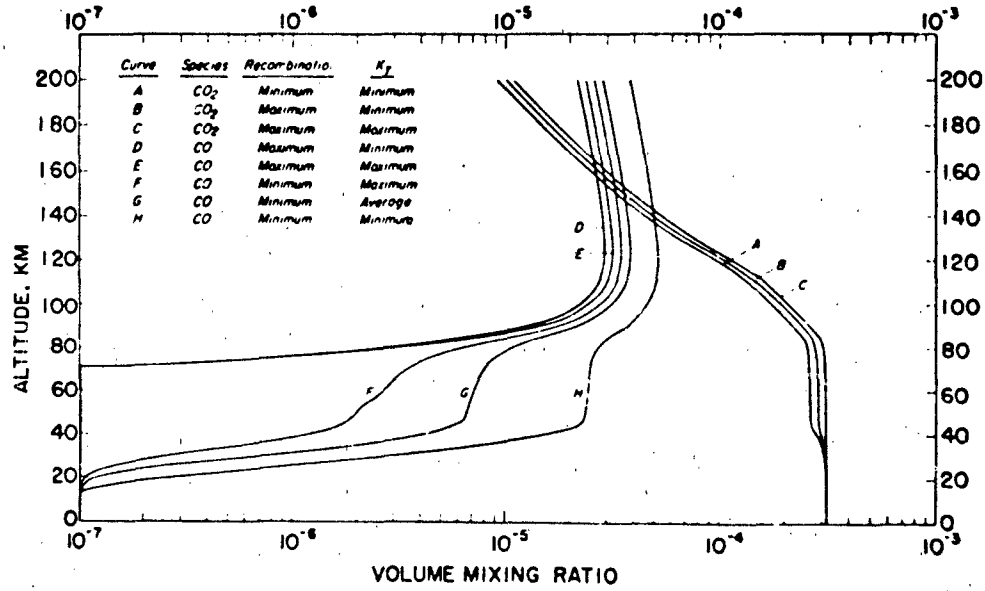


Figure 21-72. Calculated profiles of the volume mixing ratios of CO_2 and CO [Hays and Olivero, 1970]. (Reprinted with permission from Pergamon Press Ltd. © 1970.)

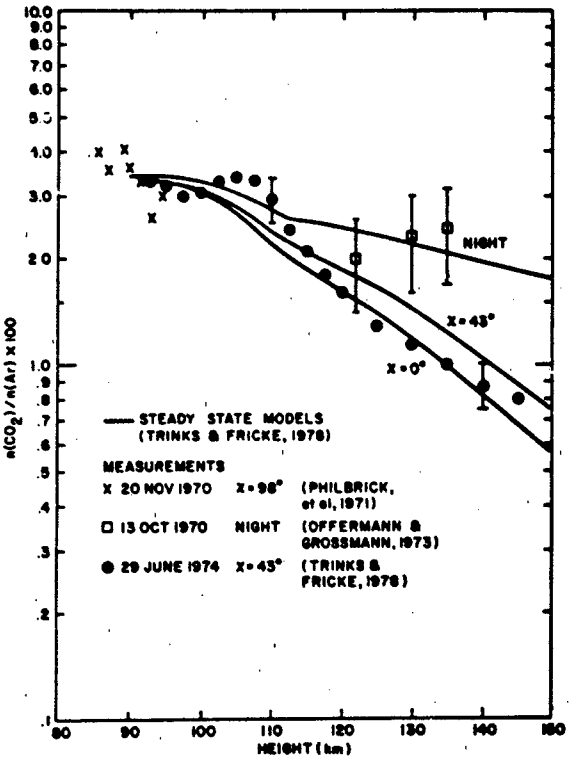


Figure 21-73. Calculated and measured $n(CO_2)/n(Ar)$ ratio for three zenith angles. Data are from the sources indicated. Day-night differences are expected due to solar dissociation of CO_2 and due to an ion chemistry reaction with O^+.

21.3.3 Hydrogen-Oxygen Constituents

The hydrogen constituents are important in many of the chemical processes in the middle atmosphere. Figure 21-74 provides an overview of the important photochemistry and transport involving the hydrogen constituents. The H_2O concentration leads to the formation of H and OH which in turn react with O, O_2, and O_3 as indicated in Figure 21-75. The oxygen chemistry in the mesosphere is strongly tied to the hydrogen concentration through reactions with OH and H_2O. In Figure 21-76 the differences between the steady state photochemical calculations and the time dependent calculations for O, H_2O, and H are shown for noon midlatitude conditions. This example, from the Keneshea et al. [1979] model, emphasizes the importance of the time-dependent calculations for the mesospheric processes. The H_2O in the mesosphere participates in another important process, the formation of the water cluster ions which are the dominant positive and negative ion species of the lower D region. The model calculations of the diurnal variations of OH and H are shown in Figures 21-77 and 21-78. The question of the mixing ratio of the H_2O which is appropriate to various mesosphere conditions is yet to be answered. One of the ways of inferring the possible concentrations of H_2O, as well as certain other minor species, is from the ion composition in the D region when the temperature is known and the laboratory measured rate constants are used, Ramseyer et al. [1983], Kopp and Philbrick [1983]. The H_2O reactions in the mesosphere are not as simple as would be inferred in Figure 21-75, but in fact include several of the other minor species as indicated in Figure 21-79.

21.3.4 The Nitrogen-oxygen Constituents

The nitric oxide concentration in the mesosphere is primarily responsible for the formation of the D-region. This results from solar Lyman- α radiation ionization at low altitudes within a window of the O_2 absorption. The NO production rate increases dramatically during periods of particle ionization. Figure 21-80 shows the nitric oxide production rate associated with some particle events. The model calculations of Brasseur [1984] for the verticle flux of nitric oxide in the mesosphere are shown in Figure 21-81. Figure 21-82 shows the meridional distribution of nitric oxide which is consistent with the vertical flux shown in Figure 21-81

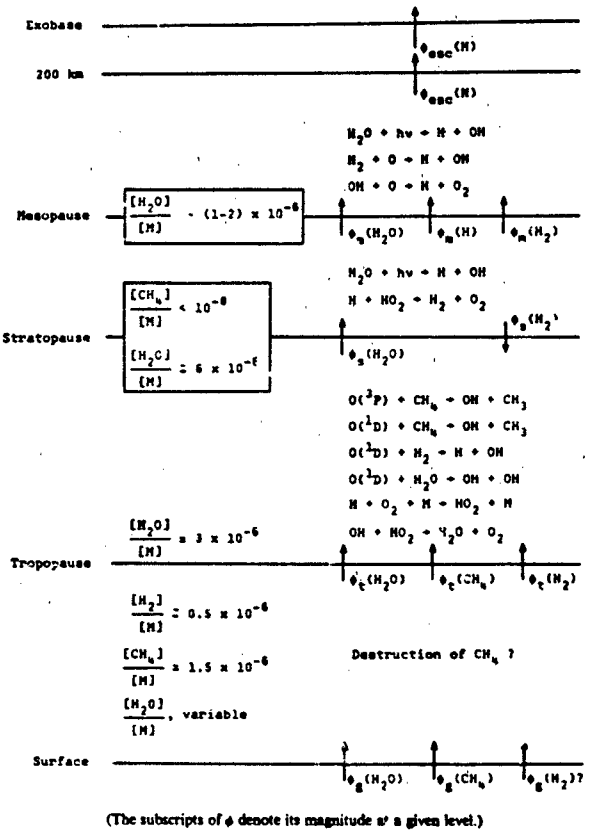


Figure 21-74. Photochemistry and transport of hydrogen constituents [after Strobel, 1972].

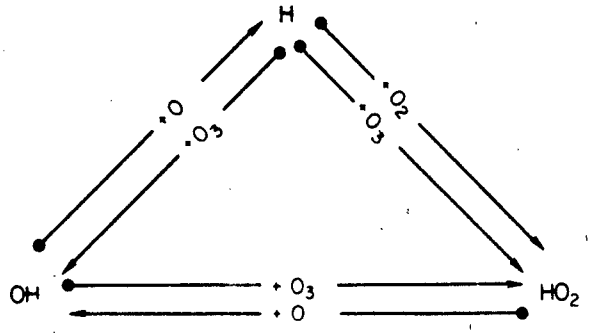


Figure 21-75. Principal reactions in which H, OH, and HO_2 are involved with O and O_3 [after Strobel, 1972].

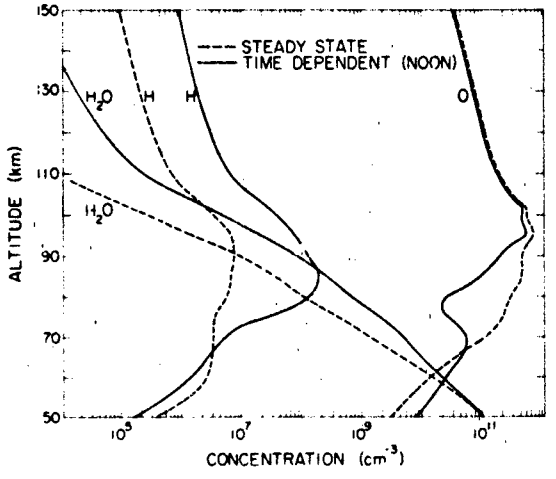


Figure 21-76. Hydroxyl radical variation from the time dependent model [Keneshea et al., 1979] (Reprinted with permission from Pergamon Press Ltd. © 1979.)

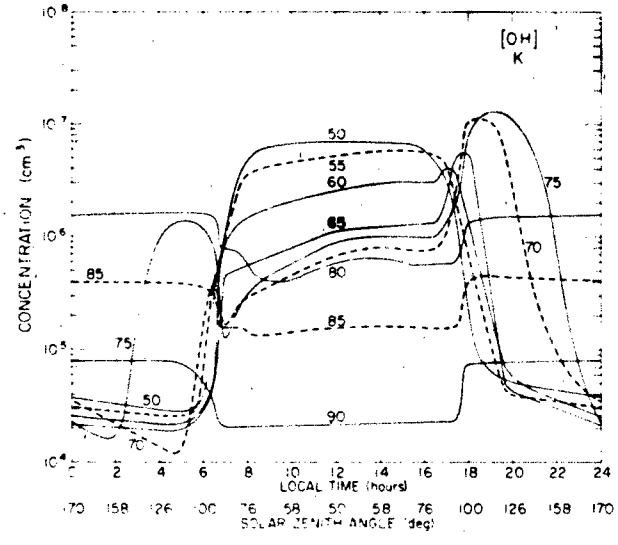


Figure 21-77. The steady-state profiles (dashed curves) and the noontime profiles (solid curves) taken from the last solution day of the time-dependent calculations for the species O, H, and H_2O [Keneshea et al., 1979] (Reprinted with permission from Pergamon Press Ltd. © 1979.)

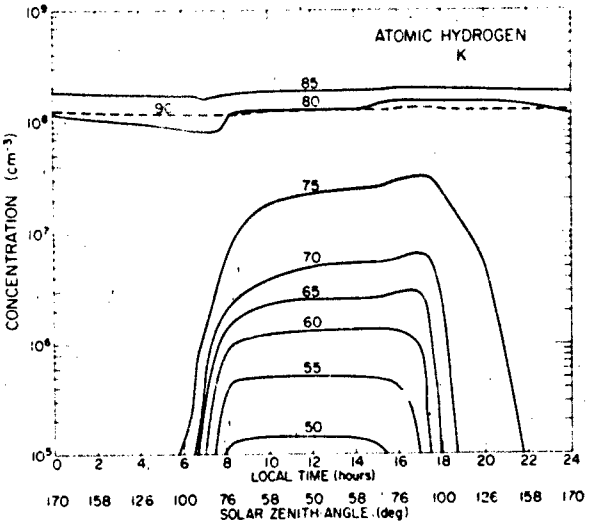


Figure 21-78. Atomic hydrogen variation from the time-dependent model [Keneshea et al., 1979]. (Reprinted with permission from Pergamon Press Ltd. © 1979.)

(based on Ebel [1980]). The meridional distribution of NO shown in Figure 21-83 is in better agreement with the measurements of NO profiles and results from a better choice of the eddy diffusion coefficient. The result emphasizes again the important role of the dynamics in the mesosphere. In Figure 21-84, the steady-state concentrations of ON (mean odd nitrogen - all nitrogen other than N_2), $N(^4S)$, and the fractional abundance of NO relative to odd nitrogen. The important reactions in the odd nitrogen constituents in the atmosphere are shown in Figure 21-85.

21.3.5 Meteoric and Dust Constituents

The resonance fluorescence dayglow emission of several of the meteoric debris species, such as sodium and lithium,

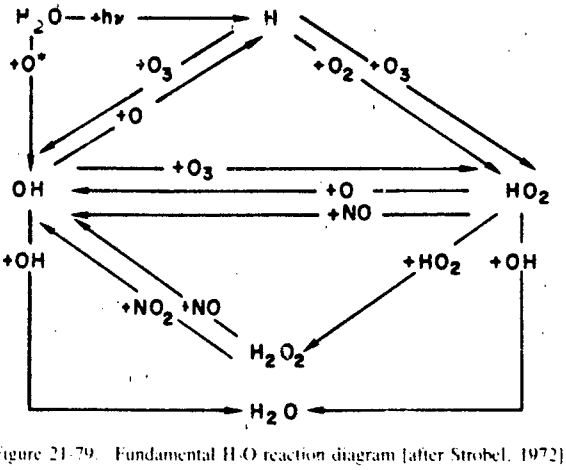


Figure 21-79. Fundamental HOx reaction diagram [after Strobel, 1972]

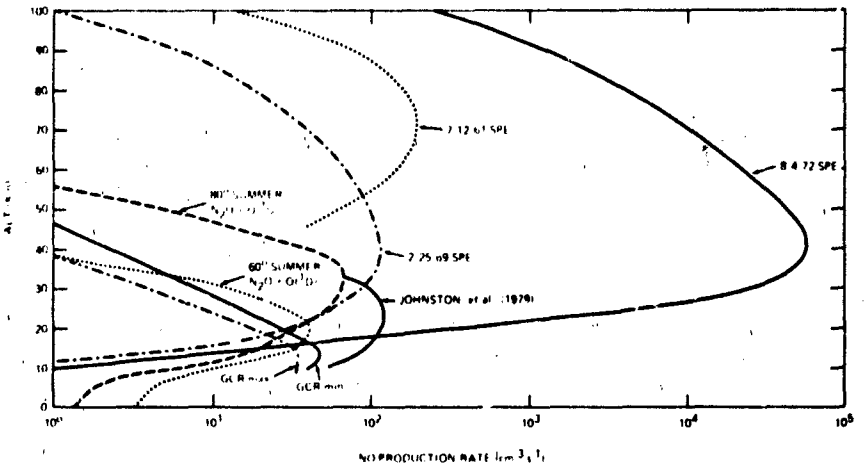


Figure 21-80. Instantaneous nitric oxide production rates for Galactic Cosmic Ray (GCR) (at solar maximum and minimum). Solar Proton Events (SPE) (July 12, 1961, February 25, 1969, August 4, 1972), and oxidation of nitrous oxide (60° summer, 80° summer, and Johnston et al. [1979] at 60° summer [Jackman et al., 1980]).

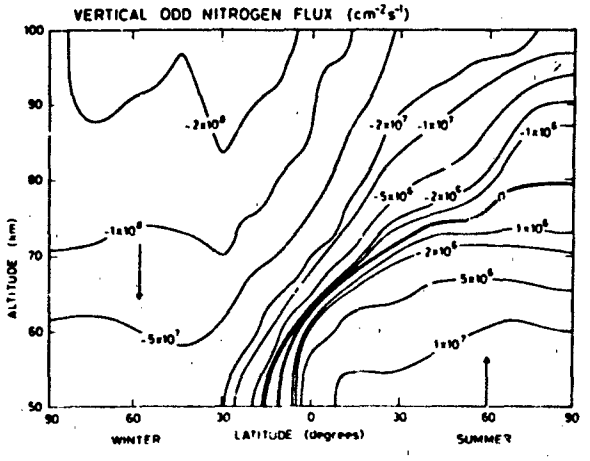


Figure 21-81. Meridional distribution of the vertical flux of nitric oxide based upon the exhange coefffficients suggested by Ebel, 1980 [Brasseur, 1984].

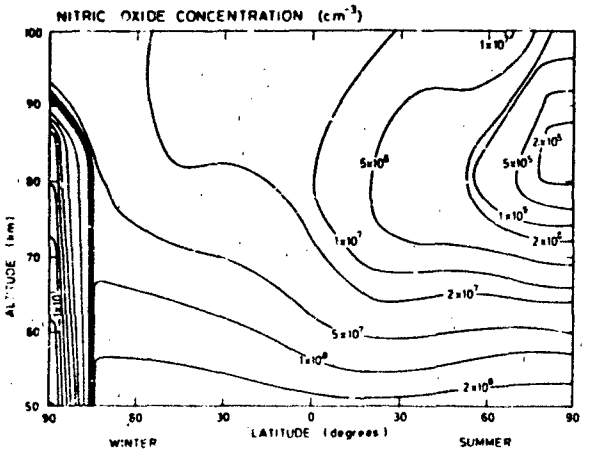


Figure 21-82. Meridional distribution of nitric oxide consistent with the flux values of Figure 21-85 [Brasseur, 1984].

has been monitored for many years. The ions of iron, magnesium, aluminum, calcium, etc. have been often measured and found to represent an almost continual input of the meteoric species to the upper atmosphere. Sodium and other species have been measured by LIDAR techniques by groups in France [Megie et al., 1978] and the U.S. [Richter et al., 1981]. The results of the sodium calculations from a model by Kirchoff et al. [1981] are shown in Figure 21-86. The meteoric and volcanic dust injected into the atmosphere is important in several areas. The dust can serve as a nucleation

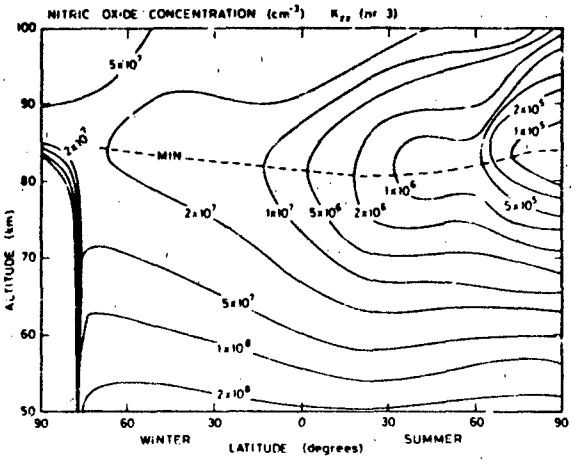


Figure 21-83. Improved model calculation for the nitric oxide meridional distribution by better choice of the dynamical properties in the calculation [Brasseur, 1984].

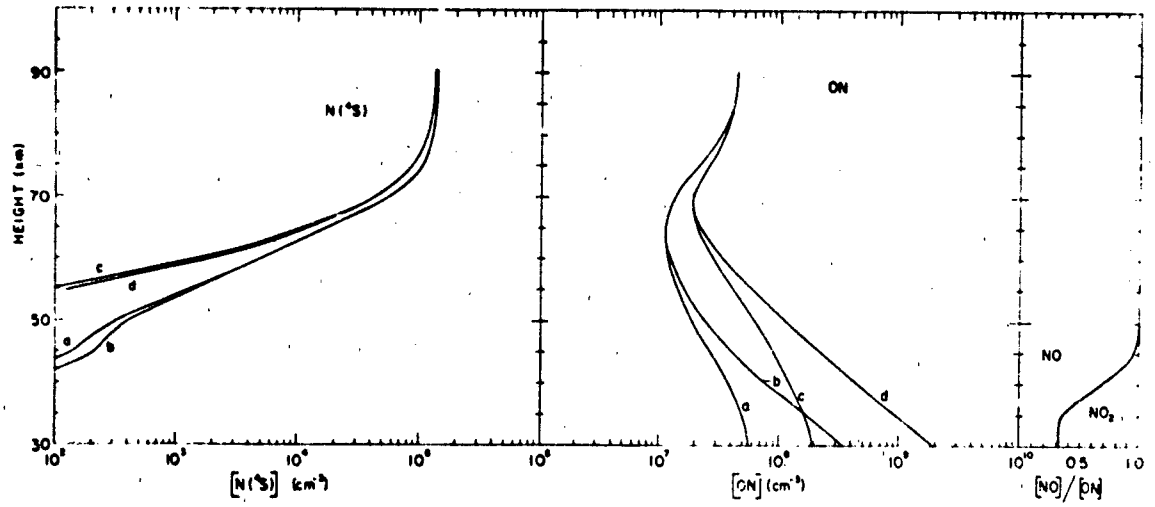


Figure 21-84. Density profiles for N(⁴S) and odd nitrogen and relative daytime concentrations of NO and NO_2 for various models. Curve *a*, large NO dissociation rate and $\partial[NO]/\partial z = 0$ at lower boundary; curve *b*, same as curve *a* but $\phi(ON) = 0$ at lower boundary; curve *c*, same as curve *a* but small NO dissociation rate; and curve *d*, small NO dissociation rate and $\phi(ON) = 0$ at lower boundary [after Strobel, 1971].

Level	Sources	Exchange Reactions	Sinks
Exobase			
THERMOSPHERE	$O^+(^4S) + \ldots \rightarrow NO^+ + N(^4S)$ $N_2^+ + O \rightarrow NO^+ + N(^4S \text{ or } ^2D)$ $N_2^+ + e \rightarrow 2N(^4S \text{ or } ^2D)$ $e(\text{fast}) + N_2 \rightarrow 2N(^4S \text{ or } ^2D)$ $N_2 + h\nu(\lambda\lambda 800\text{-}1000\text{Å}) \rightarrow 2N(^4S \text{ or } ^2D)$	$NO^+ + e \rightarrow N(^4S \text{ or } ^2D) + O$ $N(^4S) + O_2 \rightarrow NO + O$ $O_2^+ + N(^4S) \rightarrow NO^+ + O$ $N(^2D) + O_2 \rightarrow NO + O$ $N(^2D) + O \rightarrow N(^4S) + O$ $e + N(^2D) \rightarrow e + N(^4S)$ $O_2^+ + NO \rightarrow NO^+ + O_2$	$N(^4S) + NO \rightarrow N_2 + O$ $N(^2D) + NO \xrightarrow{?} N_2 + O$ $\downarrow \phi_1(NO) = 2.5 \times 10^8 cm^{-2} sec^{-1}$
Mesopause	$\downarrow \phi_1(NO) = 2.5 \times 10^8 cm^{-2} sec^{-1}$		
Mesosphere and Upper Stratosphere	$N_2 + h\nu(\lambda\lambda 1100\text{-}1250\text{Å}) \rightarrow 2N(^4S)$	$NO + h\nu(\lambda\lambda 1750\text{-}1910\text{Å}) \rightarrow N(^4S) + O$ $NO + h\nu(\lambda 1216\text{Å}) \rightarrow NO^+ + e$ $NO^+ + e \rightarrow N(^4S \text{ or } ^2D) + O$ $N(^4S) + O_3 \rightarrow NO + O_2$ $N(^4S) + OH \rightarrow NO + H$ $N(^4S) + O_2(^1\Delta) \rightarrow NO + O$ $N(^2D) + O_2 \rightarrow NO + O$ $O_3 + NO \rightarrow NO_2 + O_2$ $NO_2 + h\nu(\lambda < 3975\text{Å}) \rightarrow NO + O$ $+ h\nu(\lambda < 2750\text{Å}) \xrightarrow{?} N(^4S) + O_2$ $O + NO_2 \rightarrow NO + O_2$ $N(^4S) + NO_2 \rightarrow 2NO$	$N(^4S) + NO \rightarrow N_2 + O$ $N(^4S) + NO_2 \rightarrow N_2 + O_2$ $\rightarrow N_2 + 2O$ $\rightarrow N_2O + O$ $\downarrow \phi_2(NO + NO_2) < 10^7 cm^{-2} sec^{-1}$
Middle Stratosphere	$\downarrow \phi_2(NO + NO_2) < 10^7 cm^{-2} sec^{-1}$		
Lower Stratosphere and Troposphere	$O(^1D) + N_2O \rightarrow 2NO$ $N_2O + h\nu(\lambda < 2500\text{Å}) \xrightarrow{?} N(^4S) + NO$	$O_3 + NO \rightarrow NO_2 + O_2$ $NO, NO_2 \rightarrow HNO_2, HNO_3, NO_3, N_2O_5$	$N(^4S) + NO \rightarrow N_2 + O$ $N(^4S) + NO_2 \rightarrow$ see above removal of HNO_2 and HNO_3 by precipitation processes
	Air pollution		
Ground			$\downarrow \phi(HNO_2 + HNO_3) = ??$

Figure 21-85. Sources, sinks, and exchange reactions of odd nitrogen [after Strobel, 1971].

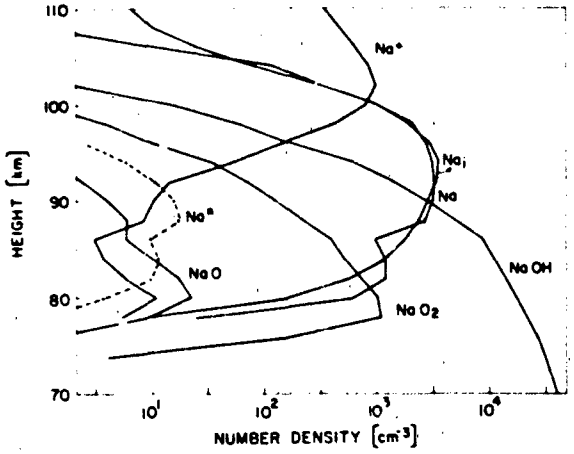


Figure 21-86. Calculated flux model sodium density profile. The production of excited sodium in units of cm^{-3} s^{-1} is designated by Na*, and Na, is the average measured profile shown for comparison [Kirchhoff et al., 1981].

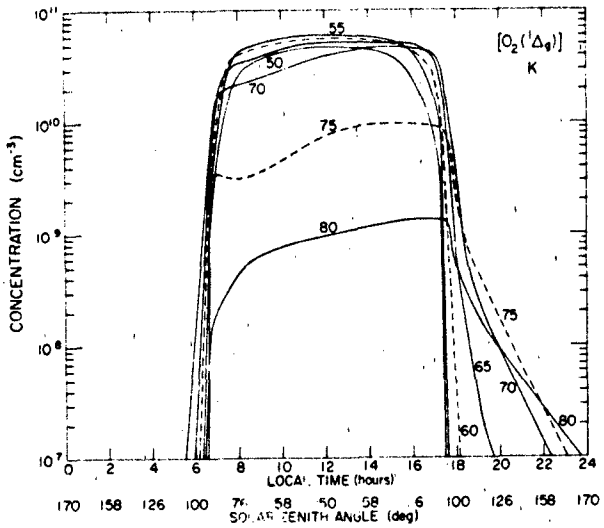


Figure 21-88. $O_2(^1\Delta_g)$ time dependence calculated from one-dimensional model [Keneshea et al., 1979]. (Reprinted with permission from Pergamon Press Ltd. © 1979.)

center for partical growth, possibly important in development of noctilucent clouds, or may be the source for species which can have significantly longer lifetimes, such as the meteoric ion lifetimes which are important in the formation of sporadic E layers.

21.3.6 The Inert Gases

The nobel gases provide the opportunity of studying the effects of transport without the complications of the chemical processes which also affect most species. The region near 100 km is of particular interest because of the turbopause or cessation of mixing of the species. This region is particularly important when attempting to model the thermosphere because of the effect on the distribution of the species. Figure 21-87 shows a summary of the published Ar/N_2 ratios in this region. The structure in the profiles indicates that the transition to molecular diffusion is not smooth, but is rather complicated by a region of alternating laminar and turbulent layers with rather strong peaks in the eddy diffusion. Argon is probably the most important of the inert species because it provides a good tracer for the transition from mixed to diffusive regions. It is also the "thermometer" in determining the temperature structure in the lower thermosphere

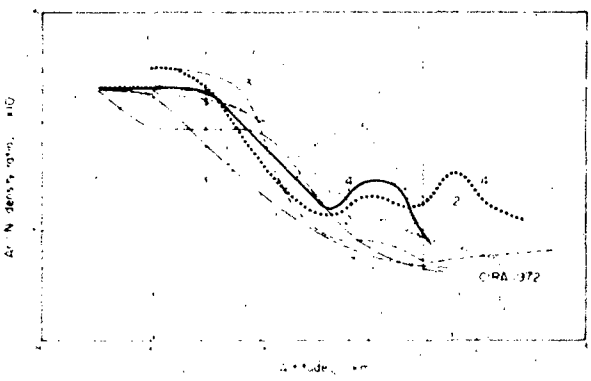


Figure 21-87. Number density ratios of argon and nitrogen vs altitude. CIRA 1972 Mean Reference Atmosphere profile is given for comparison. Profile identification numbers are given in original text [Offermann, 1981]. (Reprinted with permission from Pergamon Press Ltd. © 1981.)

21.3.7 Excited Species

Several species in excited states deserve special consideration because of the pronounced effect that they have in the mesospheric chemistry. Two of the most interesting are $O_2(^1\Delta_g)$ and $N(^2D)$. The $O_2(^1\Delta_g)$ provides an important source of ionization in the D region because of its higher energy state. In Figure 21-88 the model calculations of the temporal variation of the $O_2(^1\Delta_g)$ are shown. Figure 21-89 shows a comparison of several rocket flight measurements with model calculations of the $O_2(^1\Delta g)$. Measurements of the near-IR emission associated with this species are useful to infer some properties of the mesospheric chemistry because it is involved in the O and O_3 reaction schemes.

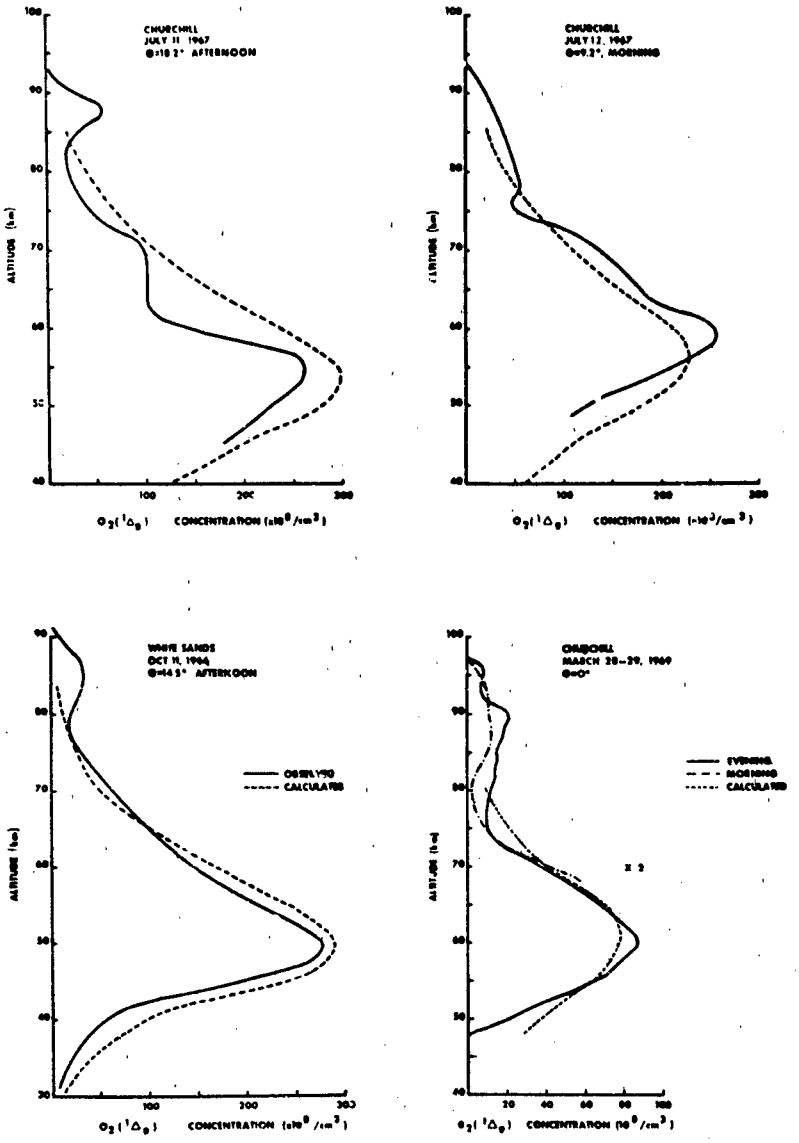


Figure 21-89. Comparison of calculated with observed altitude profiles of $[O_2(^1\Delta)]$ [after Evans and Llewellyn, 1970].

21.4 IONIZED CONSTITUENTS

21.4.1 Overview

The ionosphere is created mainly by the absorption of solar ultraviolet and x ray radiation in the earth's upper atmosphere. Hence, the initial ions are formed in some proportion to their respective neutral parent populations. The atmosphere is composed primarily of N_2 and O_2 below 100 km, the homosphere, where the atmospheric gases are generally well-mixed. Above this altitude, in the heterosphere, which virtually coincides with the thermosphere, atomic oxygen rapidly becomes a major constituent with increasing altitude and is the dominant species above about 170 km, depending upon thermospheric expansion due to solar activity. A fairly linear relationship exists, for example, between solar radiation monitored at 10.7 cm and thermospheric temperature [Nicolet, 1963].

The fact that the principal thermospheric species is atomic above a certain height has profound implications with respect to absolute electronic and ionic concentrations. The ionosphere is basically a weak, neutral plasma, that is, the total positive ion concentration equals the electron plus the

total negative ion concentration. (Negative ions are important only in the D region, below 90 km.) Ionization is depleted when oppositely charged particles recombine. However, the rate coefficient, $\sim 10^{-12}$ cm^3 s^{-1}, for (radiative) recombination between atomic positive ions and electrons, particularly $O^+ + e \rightarrow O + h\nu$ where $h\nu$ represents a photon, is five orders of magnitude slower than the (dissociative) recombination rates of common ionospheric molecular ions like O_2^+ and NO^+. These vastly different loss rates are the main reason that the maximum electron (or total positive ion) concentration of the F region, approximately 10^6 cm^{-3}, nominally occurs near 300 km, even though the maximum photoionization rate is near 150 km, the base of the F region. This region contains the overwhelming bulk of the total ionospheric charged particle content. This peak altitude is a function of latitude, solar activity, plus chemical and transport loss processes. Above the peak of the F layer, ions are lost mainly by transport to lower altitudes. Below the F layer maximum, ionization is depleted largely by ion-molecule reactions between O^+ ions and O_2 and N_2 which create O_2^+ and NO^+ ions, respectively. These diatomic ions recombine rapidly with electrons.

The E region, nominally 90 to 150 km, is comprised mainly of O_2^+ and NO^+ ions. A majority of the initial ions created by the absorption of solar x rays and EUV radiation in this region are N_2^+ ions, which react rapidly with O_2 and O at these altitudes to form O_2^+ and NO^+ ions. Some O_2^+ ions are formed directly from the solar ionization of O_2. Depending upon the relative electron and NO concentrations, a signficant fraction of O_2^+ ions may undergo charge transfer with NO to form NO^+ ions. The E region frequently is enhanced by precipitating kilovolt electrons at high latitudes, an energy (and hence altitude) regime common to auroral ionization. The transport of ionization generally is unimportant in the E region where the lifetime of molecular ions is rather short, roughly $3 \times 10^6/[e]$ s, where [e] is the electron concentration in cm^{-3}. An important exception occurs when atomic (meteoric metal) ions are present and the wind fields are favorable with respect to the creation of a sporadic E layer, a thin layer (2 km) of ionization that can attain ionic (electronic) concentrations exceeding 10^6 cm^{-3} in contrast to typical daytime midlatitude concentrations of 1 to 2×10^5 cm^{-3}.

The smallest portion of the ionosphere is the D region, roughly 50 to 90 km. It contains less than 0.1% of the total ionospheric charged particle content. This region acts as the upper boundary of the earth-ionospheric wave-guide for LF and VLF electromagnetic radiation. In addition, HF and VHF electromagnetic waves may suffer attenuation since the electron-neutral collision frequency is greatest in this lowest portion of the ionosphere because of the high neutral concentrations. This attenuation is minimal for the normal quiet D region which has a daytime maximum electron concentration of about 10^3 cm^{-3} near 80 km. However, solar flares may produce greatly enhanced x ray radiation below 10 Å which can increase electron concentrations by 10 fold or more, creating an SID (sudden ionospheric disturbance) lasting 10 to 20 minutes. At higher latitudes, precipitating protons (sometimes accompanied by energetic electrons) associated with a solar proton event (SPE) produce an enhanced D region, which in the SPE case, spreads out over the entire polar cap and may last several days. An SPE may be intense enough to blackout HF transmissions and, therefore, is known also as a polar cap absorption (PCA) event.

The D region is the most complex part of the ionosphere from a chemical standpoint. Neutral concentrations are sufficiently great to permit three-body processes to be important. Complex hydrated ions are observed, mainly hydronium ions, $H_3O^+ \cdot (H_2O)_n$, even though the initial ion formed in the quiet D region is apparently NO^+ which arises from the ionization of a minor gas, nitric oxide, by the strong solar radiation line, HLyα. For the disturbed D region, where the initial ions are formed roughly in proportion to their neutral concentrations, O_2^+ ions essentially are the only initial positive ions since N_2^+ ions very rapidly undergo charge transfer with O_2 at these altitudes. The effectiveness of three-body processes in the D region also allows for the formation of negative ions since the attachment rate of an electron to a molecule in a two-body collision is very slow. The negative species in the D region are mainly ionic below about 70 km by day and 80 km at night. Galactic particles maintain a constant ionization production rate of significance to the lower D region, < 65 km, and the stratosphere. A nominal midlatitude rate (cm^{-3} s^{-1}) is 10^{-17} times the total neutral concentration, [M], in cm^{-3}. The cosmic ionization rate varies with latitude, being a factor of 10 higher at the pole, and weakly, about a factor of two, with inverse solar activity [Swider, 1969]. Ionization of air (78% N_2, 21% O_2) by energetic particles leads to an ionic production rate relatively distributed [Swider, 1969] as 62 : 17 : 14 : 07 for $N_2^+ : O_2^+ : N^+ : O^+$ where the atomic ions originate from dissociative ionization.

At night, as already noted, galactic cosmic rays maintain a weak ionization in the low D region and stratosphere. Electrons have a feeble presence in the upper D region above about 80 km as a result of solar HLyα and HLyβ radiation which is transported to the nightside of the earth through scattering processes within the hydrogen geocorona. Ionization maintained by the absorption of this radiation is a maximum in the E region and prevents nighttime ionic concentrations from decreasing below about 10^3 cm^{-3}. Along the auroral oval, the nighttime (or daytime) E region can attain and even exceed normal midday, midlatitude, concentrations. The high latitude D region is also susceptible. Simple formulas are available to describe the total ion and electron profiles in the SPE-disturbed nighttime D region [Swider et al., 1971]. The electron concentration formula is

$$[e] = \left[\left(\frac{L(A)}{2\alpha_D}\right)^2 + \frac{q}{\alpha_D}\right]^{1/2} - \frac{L(A)}{2\alpha_D}$$

where $\alpha_D = 6 \times 10^{-7}$ cm^3 s^{-1}, q is the total ionization production rate and $L(A) = k_1 [O_2]^2 + k_2 [O_2] [N_2]$; the

bracketed terms refer to molecular concentrations in cm^{-3} and $k_1 = 1.4 \times 10^{-29} (300/T) \exp(-600/T)\ cm^6\ s^{-1}$ with $k_2 = 10^{-31}\ cm^6\ s^{-1}$ [Phelps, 1969]. Weak x ray and UV stellar sources of ionization [Gough, 1975] also may contribute to the ionization of the nighttime E and D regions. But, for MF and HF transmissions these nighttime regions are practically nonexistent except at high latitudes when disturbed and whenever sporadic E layers are prominent.

The F layer decays slowly at night since O^+ ions cannot effectively recombine with electrons except through transport to altitudes below about 250 km where there is sufficient O_2 and N_2 to form O_2^+ and NO^+, respectively, which (dissociatively) recombine rapidly with electrons. This action can be virtually stopped in its entirety by horizontal winds, and especially at the equator, electromagnetic drift mechanisms which can elevate the entire layer [Rishbeth, 1968]. Little scattered solar radiation is present in the nighttime F layer save for some He lines.

21.4.2 Positive Ion Distributions

Figure 21-90 is a composite picture [Narcisi, 1973] of the E and F regions during solar minimum when the F layer maximum has a relatively low height, about 250 km. The data were taken from Hoffman et al. [1969] for altitudes above 250 km, from Holmes et al. [1965] for altitudes 250 to 115 km, and from Narcisi [1968] for altitudes below 115 km. Atomic oxygen ions clearly are the major ions in the F region. Molecular ions NO^+, O_2^+ and N_2^+ decline rapidly with increasing altitude in the lower F region because of their short lifetimes in the presence of large quantities of electrons and their decreasing production rate with altitude, either being formed directly by the ionization of the neutral gas (N_2^+, O_2^+) or partly (O_2^+) or completely (NO^+) from ion-molecule interactions. The ions N_2^+ and N^+ are depleted rapidly below 200 km because of fast reactions with O_2 (N^+, N_2^+) and O (N_2^+). The N^+ ion originates from the dissociative ionization of N_2 [Nicolet and Swider, 1963]. The H_2O^+ ions are contaminents arising from the process $O^+ + H_2O \rightarrow H_2O^+ + O$. Water outgasses profusely from rockets and satellites although the latter "dry out" after several days. In the upper F region, He^+ and H^+ ions are observed since He and H are present at great altitudes in the thermosphere. A few doubly charged ions are created by solar radiation and are observed in the high F region where low neutral concentrations prevent their immediate disappearance by means of ion-molecule reactions.

Typical nighttime E and F regions are shown in Figure 21-91 [Narcisi, 1973]. Data above 200 km were taken from Hoffman (1967); data from 220 to 140 km are from Holmes et al. [1965]; and results below 140 km are from Narcisi [1971] with dashed curves from Giraud et al. [1971]. Meteoric ions are omitted. The depth of the valley near 150 km is quite varible as are many of the features in the figure including the peak concentrations of the E and F layers. The O^+ ion remains the major F region ion. However, hydrogen ions are now more prominent. This feature arises roughly because of the absence of solar radiation and because the reaction $H^+ + O \leftrightarrows O^+ + H$ is virtually thermoneutral. Figures 21-90 and 21-91 represent midlatitude conditions. The equatorial situation is not too different, but the high-latitude ionosphere does have two striking features (see discussion by Narcisi [1973]): (1) a persistent light ion trough in H^+ and He^+ with the former ion's concentration declining by 1 to 2 orders of magnitude near $\pm 60°$ geomagnetic latitude, and (2) a variable poleward peak of ionization. There are a number of other features such as a winter bulge in He^+.

The E region is the easiest portion of the ionosphere to

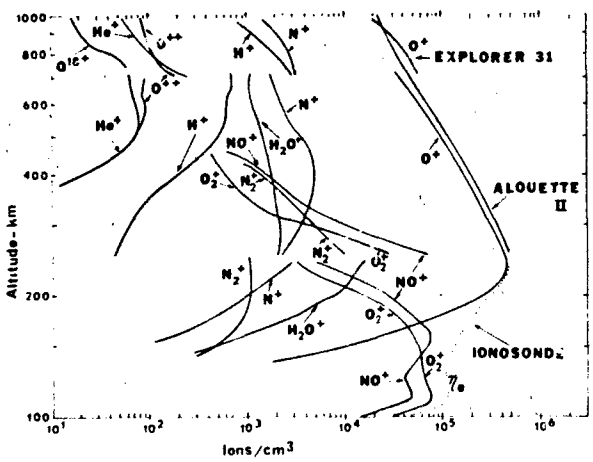


Figure 21-90. Ion composition of the midlatitude daytime ionosphere near solar minimum [Narcisi, 1973].

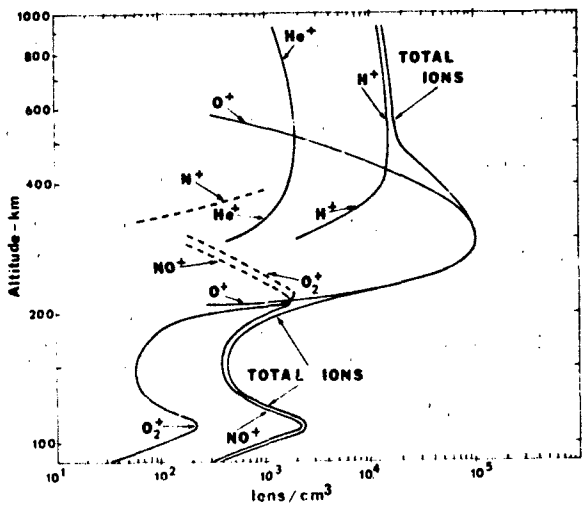


Figure 21-91. Ion composition of the midlatitude nighttime ionosphere [Narcisi, 1973].

assess since the transport of ionization is unimportant except for sporadic E layers and, perhaps at night, when charge concentrations are low (ion/electron lifetimes long). A model of the midlatitude E region [Keneshea et al., 1970] has been available for over a decade. In Figure 21-92, model electron concentrations are plotted for selected altitudes as a function of solar zenith angle, the angle between the sun and the local vertical coordinate. Since the calculations originally were performed to match twilight ionic composition data for Eglin AFB in Florida, the minimum solar zenith angle was about 22°. The accompanying figure (21-93) illustrates the ratio of the two major ions, NO^+ and O_2^+. The ratio is about unity for most altitudes in the daytime, but quite large at night. This behavior is dependent upon the quantity of nitric oxide present in the E region for the conversion of O_2^+ ions into NO^+ ions through charge exchange, $O_2^+ + NO \rightarrow NO^+ + O_2$. In the daytime, many of the O_2^+ ions recombine with electrons before charge transfer can occur. The much lower electron concentrations at night make the charge transfer process relatively more effective, and hence $[NO^+]/[O_2^+]$ greater. Strobel et al. [1974] calculated the nighttime E region using a more detailed picture of scattered HLyα and HLyβ as a function of solar position. However, the nighttime E region generally is quite variable because of electromagnetic drift, sporadic E layer formation, and, at the higher latitudes, electron precipitation, which apparently is a function of Kp, the geomagnetic planetary three-hour range index [Voss and Smith, 1980]. Figure 21-94 depicts a nighttime sporadic E layer having a peak electron concentration of 10^5 cm^{-3} near 90 km. The principal meteoric ions in the layer are Fe^+ (56 amu) and Mg^+ (24 amu) with lesser amounts of Ca^+ (40 amu) ions. The normal E region ions are O_2^+ (32 amu) and NO^+ (30 amu). The ion with mass 37 amu represents $H_3O^+ \cdot H_2O$, a D region ion. In general, intense sporadic E ($f_{E_s} > 5$ MHz) is most likely during daytime midlatitude summer [Smith and Matsushita, 1962].

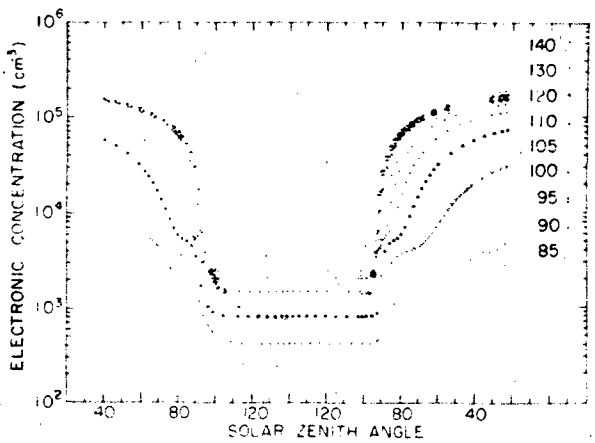


Figure 21-92. Diurnal model of quiet E region electron concentrations at altitudes from 85 to 140 km [Keneshea et al., 1970].

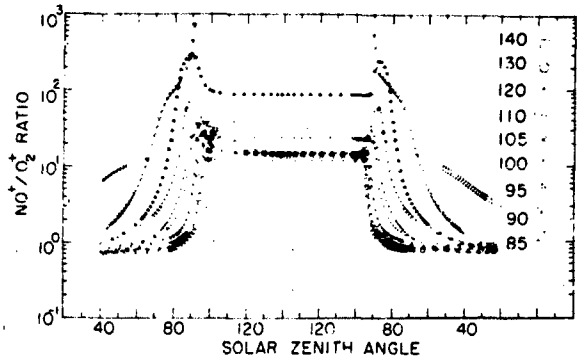


Figure 21-93. Diurnal model of quiet E region positive ion ratio $[NO^+/O_2^+]$ at altitudes from 85 to 140 km [Keneshea et al., 1970].

The importance of transport is evident in Figure 21-95 where an experimental sunset ion profile at a solar zenith angle of 99° and a theoretical profile excluding ion transport [Keneshea et al., 1970] is compared to a model including transport due to measured winds [Keneshea and MacLeod, 1970]. The disturbances apparent in Figures 21-94 and 21-95 should not detract from the fact that the Keneshea et al. [1970] model remains valid for the normal midlatitude E region since even recent measurements [Trost, 1979] are in good agreement with their model which is for moderate solar activity. Under auroral conditions, the $[NO^+]/[O_2^+]$ ratio tends to be higher than unity, the approximate midlatitide daytime value, as shown by a survey (Figure 21-96) of eight auroral flights by Swider and Narcisi [1977]. This higher ratio reflects a greater quantity of NO in the auroral or higher-latitude E region as compared with midlatitudes, particularly in the winter, when photodissociation of NO is minimal.

Narcisi and Bailey [1965] were the first to report that hydronium ions are the dominant positive ions of the D region. Figure 21-97 illustrates the difference between the E region where NO^+ (30 amu) and O_2^+ (32 amu) are dominant and the D region where $H_5O_2^+$ (37 amu), $H_7O_3^+$ (55 amu) and $H_9O_4^+$ (73 amu) are the principal ions. The ions labeled 24 amu and 56 amu represent meteoric Mg^+ and Fe^+ ions, respectively. Depending upon the temperature and $[H_2O]$, even heavier hydronium ions may be present in the D region like $H_{11}O_5^+$ and $H_{13}O_6^+$. Shock heating can lead to the observation of lighter species. Hydronium ions recombine with electrons at dissociative recombination rates 5 to 10 times greater [Biondi, 1973] than those for O_2^+ and NO^+ ions, respectively. Hence, the conversion of NO^+ and O_2^+ ions to hydronium ions provides for a lesser D region than would be present in the absence of this conversion. Under disturbed conditions, enhanced ion and electron concentrations tend to lower the crossover altitude between simple molecular ions ($NO^+ + O_2^+$) and hydronium ions, because dissociative recombination becomes more competitive with increased [e] as compared to ionic conversion

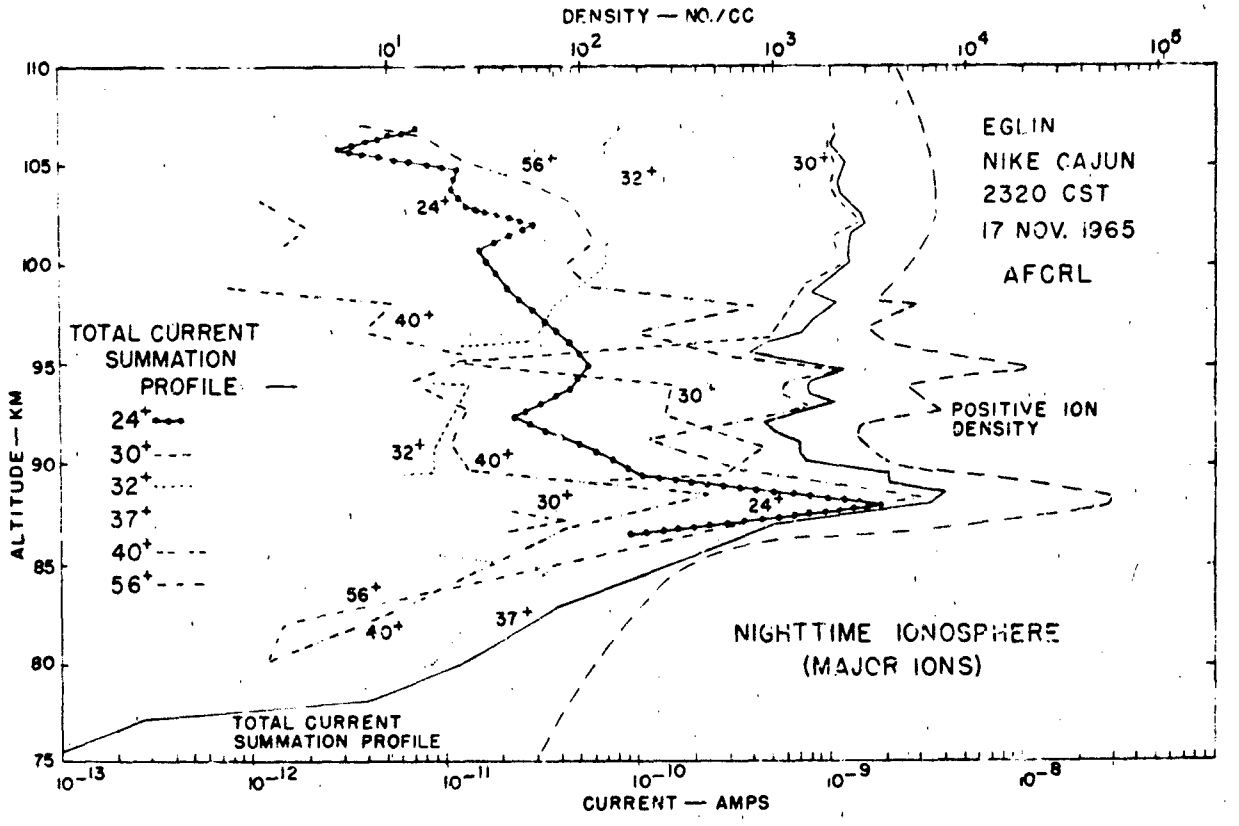


Figure 21-94. Ion composition of a nighttime sporadic E layer [Narcisi, 1969].

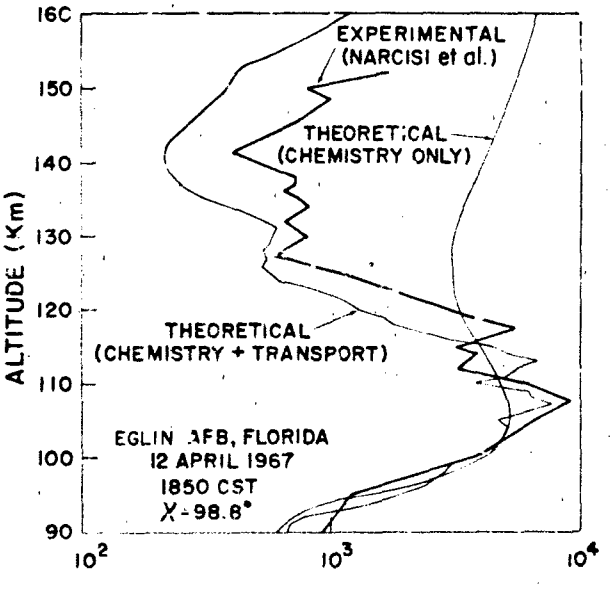


Figure 21-95. Total positive ion profiles in the sunset E region: Comparison of theory with experiment [Keneshea and MacLeod, 1970].

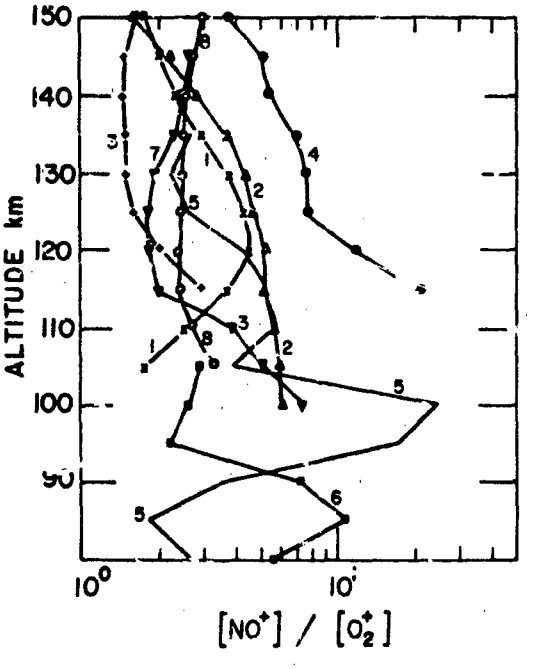


Figure 21-96. The positive ion ratios measured during eight known auroral E region rocket flights [Swider and Narcisi, 1977].

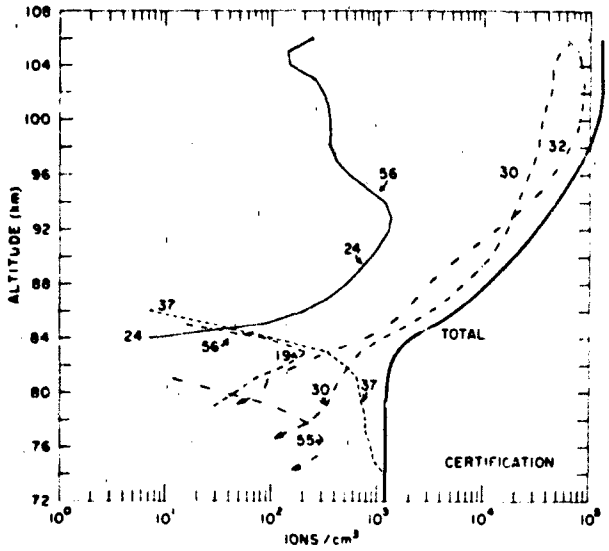


Figure 21-97. Positive ion distributions measured in the lower E and upper D regions at a 20° solar zenith angle [Narcisi et al., 1971].

processes which rely only upon the neutral concentrations. Figure 21-98 compares a theoretical model for the 3 November 1969 SPE (PCA) with the measured electron concentrations and is a good example of hydronium ion distributions in a disturbed D region. Similar behavior should occur for an SID. Below about 60 km, the hydronium ions in the figure appear to be in thermodynamic equilibrium and only a knowledge of the water vapor concentration (about 6 ppmv), the temperature and the thermodynamic parameters [Kebarle et al., 1967] are needed in order to derive the relative hydronium ion composition. Since electron concentrations are lower in the quiet D region, thermodynamic

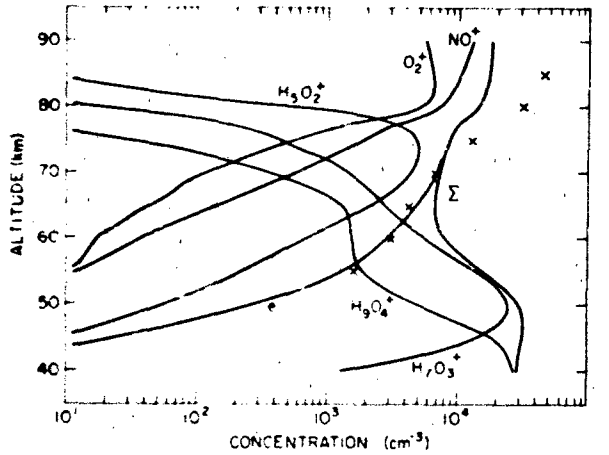


Figure 21-98. Theoretical model of the electron and major positive ion profiles in the SPE disturbed D region, as compared with the measured (crossed) electron concentrations [Swider et al., 1978].

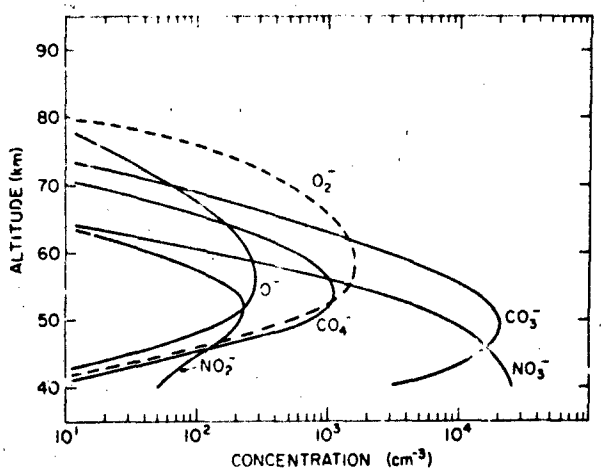


Figure 21-99. Theoretical negative ion profiles commensurate with Figure 21-100 [Swider et al., 1978].

equilibrium may be appropriate up to 70 km or even higher. However, in the atmosphere, processes involving CO_2 may aid the formation of heavier hydronium ions. Hence, the thermodynamic expression of Kebarle et al. [1967] should be used with caution.

The quiet D region positive ion composition differs little at night from day except that the concentrations are less at night. As might be expected, the crossover between simple ions (NO^+ + O_2^+) and hydronium ions rises—from about 82 km by day to 86 km at night [Narcisi, 1973]. The positive ion distributions of the daytime D region have been modelled extensively by Reid [1977].

21.4.3 Negative Ion Distributions

Our understanding of negative ions is poor. They occur only at D-region altitudes and below. Figure 21-99 is a representation of the negative ion composition computed in conjunction with the positive ions and electron distributions shown in the previous figure. The total negative ion populations calculated appear to be compatible with the calculated and observed electron and total positive ion distributions. However, this match is somewhat misleading since it has been assumed, based upon laboratory evidence (Smith et al., 1976), that all ion-ion recombinations (between positive and negative ions) have a rate coefficient of 6×10^{-8} $cm^3\ s^{-1}$. By definition then, mainly the total negative ion concentration is relevant and not their individual identities. Yet, the individual identities play some role because of their uniquely related ion-molecule interactions with the neutral atmosphere plus photodetachment and photodissociation processes. On the other hand, negative ion profiles computed at night for SPE conditions [Swider et al., 1978] resulted in distributions not unlike those calculated by Arnold and Krankowsky [1971] which were in good agreement

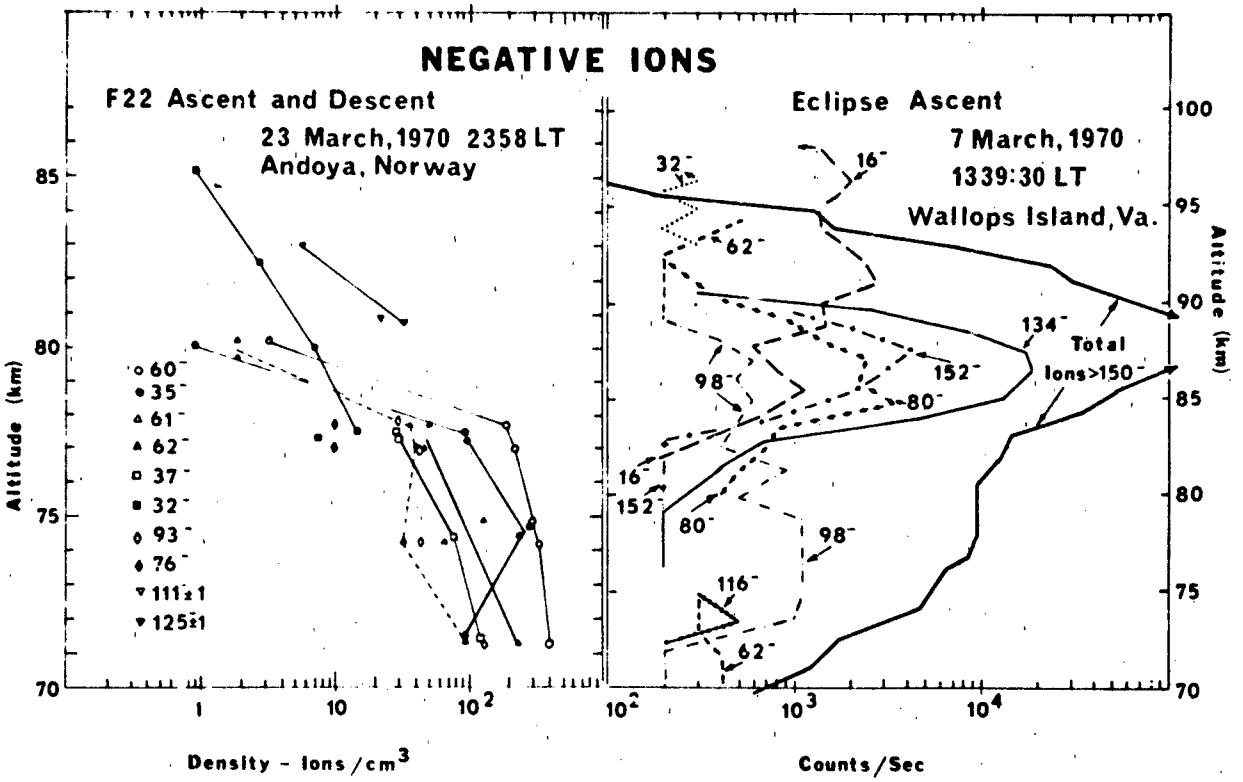


Figure 21-100. D region negative ion observations: Left panel from Arnold et al. [1971]; right panel from Narcisi et al. [1972]. Note the different altitude scales.

with the one observational result (nighttime) they obtained. However, a serious controversy arises over the fact that in fourteen flights to date [upleg and downleg, R.S. Narcisi, private communication, 1981], a layer of heavy negative ions was measured near 85 km. This is illustrated in Figure 21-100 where the Max Planck Institute at Heidelberg results are shown in conjunction with one AFGL result. The AFGL data indicate greater hydration, with mainly NO_3^- (62 amu), $NO_3^-H_2O$ (80 amu) plus higher hydrates than do the Heidelberg group's results. In addition, the AFGL data contain a layer of very heavy negative ions at 85–90 km whose origin cannot be explained by conventional gas-phase theory. During rocket flights into a disturbed and eclipsed D region on 26 February 1979, AFGL [Narcisi et al., 1983] again recorded a layer of heavy negative ions near 85 km as did a University of Bern group [Kopp et al., 1980] in a separate experiment. The Bern results suggest that the principal negative ions may be HCO_3^- and its hydrates. This conclusion is not incompatible with the AFGL results since such ion identifications are within the mass uncertainty of the AFGL instrument, about ±1 amu. Reported observations of Cl⁻ (35 amu and 37 amu) may arise from the interaction of ambient ions with chlorine contaminants [Narcisi, 1973].

CHAPTER 21

REFERENCES

Ackerman, M. and C. Muller, "Stratospheric Methane and Nitrogen Dioxide from infrared spectra," *Pure Appl. Geophys.*, **106–108:**1325–1335, 1973.

Ackerman, M., J.C. Fontanella, D. Frimout, A. Girard, N. Louisnard, and C. Muller, "Simultaneous Measurements of NO and NO_2 in the Stratosphere," *Planet. Space Sci.*, **23:**651–660, 1975.

Ackerman, M., D. Frimout, C. Muller, and D.J. Wuebbles, "Stratospheric Methane Measurements and Predictions," *Pure Appl. Geophys.* **117:**367–380, 1978/79.

Aiken, A.C., and E.J.R. Maier, "Balloon-Borne Photoionization Mass Spectrometer for Measurement of Stratospheric Gases," *Rev. Sci. Instrum.*, **49:**1034–1040, 1978.

Aiken, A.C., B. Woodgate, and H.J.P. Smith, "Atmospheric Ozone Determination by Solar Occulation Using the UV Spectrometer on the Solar Maximum Mission," *Appl. Opt.*, **21:**2421–2424, 1982.

Ainsworth, J.E. and J.R. Hagemeyer, "Measurement of Ozone by a Dasibi Ozone Monitor," in *The Stratcom VIII Effort*, edited by E.I. Reed, NASA Technical Paper 1640, 95–100, 1980.

Allen, M., Y.L. Yung, and J.W. Waters, "Vertical Transport and Photochemistry in the Terrestrial Mesosphere and Lower Thermosphere (50–120 km)," *J. Geophys. Res.*, **86:**3617–3627, 1981.

Allen, M., J.I. Lunine, and Y.L. Yung, "The Vertical Distribution of Ozone in the Mesosphere and Lower Thermosphere," *J. Geophys. Res.*, **89:**484–4872, 1984.

Anderson, G.P., C.A. Barth, F. Cayla, and J. London, "Satellite Observations of the Vertical Ozone Distribution in the Upper Stratosphere," *Ann. Geophys.*, **25:**341–343, 1969.

Anderson, J. G., Rocket-borne Ultraviolet Spectrometer Measurement of OH Resonance Fluorescence with a Diffusive Transport Model for Mesospheric Photochemistry," *J. Geophys. Res.*, **76:**4634–4652, 1971a.

Anderson, J.G., "Rocket Measurement of OH in the Mesosphere," *J. Geophys. Res.*, **76:**7820–7824, 1971b.

Anderson, J.G., "The Absolute Concentration of $O(^3P)$ in the Earth's Stratosphere," *Geophys. Res. Lett.*, **2:**231–238, 1975.

Anderson, J.G., "The Absolute Concentration of OH $(X^2\pi)$ in the Earth's Stratosphere," *Geophys. Res. Lett.*, **3:**165–168, 1976.

Anderson, J.G., "Free Radicals in the Earth's Stratosphere: A Review of Recent Results," *Proc. of the NATO Advanced Study Institute on Atmospheric Ozone*, edited by A.C. Aiken, FAA-EE-80-20, 233–251, 1980.

Anderson, J.G., J.J. Margitan, and D.H. Stedman, "Atomic Chlorine and the Chlorine Monoxide Radical in the Stratosphere: Three in situ Observations," *Science*, **198:**501–503, 1977.

Anderson, J.G., H.J. Grassl, R.E. Shetter, and J.J. Margitan, "Stratospheric Free Chlorine Measured by Balloon-Borne in situ Resonance Fluorescence," *J. Geophys. Res.*, **85:**2869–2887, 1980.

Anderson, J.G., H.J. Grassl, R.E. Shetter, and J.J. Margitan, "HO$_2$ in the Stratosphere: Three in situ Observations," *Geophys. Res. Lett.*, **8:**289–292, 1981.

Anderson, J.G., N.L. Hazen, B.E. McLaren, S.P. Rowe, C.M. Schiller, M.J. Schwab, L. Solomon, E.E. Thompson, and E.M. Winestock, "Free Radicals in the Stratosphere: A New Observational Technique," *Science*, **228:** 1309–1311, 1985.

Angell, J.K. and J. Korshover, "Global Variation in the Total Ozone and Layer-Mean Ozone: An Update Through 1981," *J. Climate Appl. Meteorol.*, **22:**1611–1627, 1983.

Arnold, F. and D. Krankowsky, "Negative Ions in the Lower Ionosphere: A Comparison of a Model Computation and a Mass-Spectrometric Measurement," *J. Atmos. Terr. Phys.*, **33:**1693–1702, 1971.

Arnold, F., J. Kissel, D. Krankowsky, H. Wieder, and J. Zahringer, "Negative Ions in the Lower Ionosphere: A Mass-Spectrometric Measurement," *J. Atmos. Terr. Phys.*, **33:**1169–1175, 1971.

Arnold, F., D. Krankowsky, K. H. Marien and W. Joos, "A Mass Spectrometer Probe for Composition and Structure Analysis of the Middle Atmosphere Plasma and Neutral Gas," *J. Geophys.*, **44:**125–138, 1977.

Arnold, F., R. Fabian, G. Hepschen, and W. Joos, "Stratospheric Trace Gas Analysis from Ions: H_2O and HNO_3," *Planet. Space Sci.*, **28:**681–685, 1980.

Barnett, J.J., J.T. Houghton, and J.A. Pyle, "The Temperature Dependence of the Ozone Concentration Near the Stratopause," *Quart. J. Roy. Meteorol. Soc.*, **101:**245–257, 1975.

Barth, C.A., D.W. Rusch, R.J. Thomas, G.H. Mount, R.J. Rottman, G.E. Thomas, R.W. Sanders, and G.M. Lawrence, "Solar Mesospheric Explorer: Scientific Objectives and Results," *Geophys. Res. Lett.*, **10:**237–240, 1983.

Biondi, M.A., "The Effects of Ion Complexity on Electron-Ion Recombination," *Comments At. Mol. Phys. Part D*, **4:**85–91, 1973.

Bischof, W., P. Fabian, and R. Borchers, "Decrease in CO_2 Mixing into Ratio Observed in the Stratosphere," *Nature*, **288:**347–348, 1980.

Blatherwick, R.D., A. Goldman, D.G. Murray, G.R. Cook, and J.W. VanAllen, "Simultaneous Mixing Ratio Profiles of Stratospheric NO and NO_2 as Derived from Balloon-Borne Infrared Solar Spectra," *Geophys. Res. Lett.*, **7:**471–473, 1980.

Bloomfield, P., G. Oehlert, M.L. Thompson, and S. Zeger, "A Frequency Domain Analysis of Trends in Dobson Total Ozone Records," *J. Geophys. Res.* **88:**8512–8522, 1983.

Bojkov, R.D., "Tropospheric Ozone, Its Changes and Possible Radiative Effects," WMO Special Environmental Report #16 to be published, World Meteorological Organization, Geneva, 1984.

Brasseur, G., *Physique et Chemie de l'Atmosphere Moyene*, Masson, Paris, 1982.

Brasseur, G., "Coupling Between the Thermosphere and the Stratosphere: The Role of Nitric Oxide," in *Middle Atmosphere Handbook* 10, SCOSTEP, University of Illinois, 1984.

Brasseur, G. and S. Solomon, *Aeronomy of the Middle Atmosphere*, Rydell, New York, 1984

Brewer, A.W., "Evidence for a World Circulation Provided by the Measurements of Helium and Water Vapour Distribution in the Stratosphere," *Quart. J. Roy. Meteorol. Soc.*, **75**:351–363, 1949.

Burkhardt, E.G., C.A. Lambert, and C.K.N. Patel, "Stratospheric Nitric Oxide: Measurements during Daytime and Sunset," *Science*, **188**:1111–1113, 1975.

Burnett, C.R., "Terrestrial OH Abundance Measurement by Spectroscopic Observation of Resonance Absorption of Sunlight," *Geophys. Res. Lett.*, **3**:319–322, 1976.

Burnett, C.R., "Spectroscopic Measurements of Atmospheric OH Abundance," *Bull. Am. Phys. Soc.*, **22**:539, 1977.

Burnett, C.R. and E.B. Burnett, "Spectroscopic Measurements of the Vertical Column Abundance of Hydroxyl (OH) in the Earth's Atmosphere," *J. Geophys. Res.*, **86**:5185–5202, 1981.

Bush, Y.A., A.L. Schmelekopf, F.C. Fehsenfeld, D.L. Albritton, J.R. McAfee, P.D. Goldan, and E.E. Ferguson, "Stratospheric Measurements of Methane atSeveral Latitudes," *Geophys. Res. Lett.*, **5**:1027–1029,1978.

Calbally, J.E. and C.R. Roy, "Destruction of Ozone at the Earth's Surface," *Quart. J. Roy. Meteorol. Soc.*, **106**:599–620, 1980.

Chaloner, C.P., J.R. Drummond, J.T. Houghton, R.F. Jarnot, and H.K. Roscoe, "Infrared Measurements of Stratospheric Composition I. The Balloon Instrument and Water Vapour Measurements," *Proc. Roy. Soc. London A*, **364**:145–159, 1978.

Chapman, S., "A Theory of Upper Atmosphere Ozone," *Mem. R. Met. Soc.*, **3**:103–125, 1930.

CIAP (U.S. Climatic Impact Assessment Program), "The Effects of Stratospheric Pollution by Aircraft," A.J. Grobecker, S.C. Coroniti, and R.H. Cannon, Jr., U.S. Dept. of Transportation, Washington, D.C., 1974.

CIAP "The Natural Stratosphere of 1974," Monograph 1 DOT-TST-74-51, Final Report, U.S. Dept. of Transportation, Washington, D.C., 1975.

Cicerone, R.J., "Minor Constituents in the Stratosphere and Mesosphere," *Rev. Geophys. Space Phys.*, **13**:900, 1975.

Clark, J.H.E. and T.G Rogers, "The Transport of Conservative Trace Gases by Planetary Waves," *J. Atmos. Sci.*, **35**:2232–2235, 1978.

Coffey, M.T., W.G. Mankin, and A. Goldman, "Simultaneous Spectroscopic Determination of the Latitudinal, Seasonal, and Diurnal Variability of Stratospheric N_2O, NO, NO_2, and HNO_3," *J. Geophys. Res.*, **86**:7331–7341, 1981.

COMESA, *The Report of the Committee on Meteorological Effects of Stratospheric Aircraft (1972–1975)*, Great Britain Meteorological Office, Bracknell, U.K., 1975.

COVOS, *Rapport Final Activities 1972–1976*, Presente par E.A. Brun, Societe Meteorologique de France, Boulogne, 1976.

CRC Handbook of Toxicology, Cleveland, Ohio, 1969.

Cronn, D. and E. Robinson, "Tropospheric and Lower Stratospheric Vertical Profiles of Ethane and Acetylene," *Geophys. Res. Lett.*, **6**:641–644, 1979.

Crosby, D.S., W.G. Planet, A.J. Miller, and R.M. Nagatani, "Evaluation and Comparison of Total Ozone Fields Derived from TOVS and SBUV," *Proceedings of the Quadrennial International Ozone Symposium*, 1980, Boulder, Colo, edited by J. London, International Ozone Commission, Vienna, Va., 161–167, 1980.

Crutzen, P.J., "Ozone Production Rates in an Oxygen, Hydrogen-Nitrogen Oxide Atmosphere," *J. Geophys. Res.*, **76**:7311–7327, 1971.

Crutzen, P.J., I.S.A. Isakesen, and G.C. Reid, "Solar Proton Events: Stratospheric Sources of Nitric Oxide," *Science*, **189**:457–458, 1975.

Cunnold, D.M., F.N. Alyea, and R.G. Prinn, "Preliminary Calculations Concerning the Maintenance of the Zonal Mean Ozone Distribution in the Northern Hemisphere," *Pure Appl. Geophys.*, **118**:329–354, 1980.

Dandekar, B. S., "Atomic Oxygen Concentration from the [OI] 5577 Å Line Emission at the Auroral Zone Latitude," *Plan. Space Sci.*, **20**:1781, 1972.

Danielsen, E.F. and R.S. Hipskind, "Stratospheric-Tropospheric Exchange at Polar Latitudes in Summer," *J. Geophys. Res.* **85**:393–400, 1980.

Danielsen, E. F., et al., "Observed Distribution of Radioactivity, Ozone, and Potential Vorticity Associated with Tropopause Folding," *J. Geophys. Res.*, **75**:2353–2361, 1970.

Deepak, A. (ed.), *Inversion Methods in Atmospheric Remote Sounding*, Academic Press, Nev York, 1977.

Dewan, E., "Turbulent Vertical Transport due to Thin Intermittent Mixing Layers in the Stratosphere and Other Stable Fluids," *Science*, **211**:1041–1042, 1981.

Dickinson, P.H.G., R.C. Bolden, and R.A. Young, "Measurements of Atomic Oxygen in the Lower Ionosphere using a Rocket-Borne Resonance Lamp," *Nature*, **252**:289–291, 1974.

Dobson, G.M.B., "Observations of the Amount of Ozone in the Earth's Atmosphere and Its Relation to Other Geophysical Conditions, Part IV," *Proc. Roy. Soc.* **A129**:411–433, 1930.

Dobson, G.M.B., *Exploring the Atmosphere*, Oxford University Press, London, 1963.

Dobson, G.M.B., A.W. Brewer, and B. Cwilong, "Meteorology of the Lower Stratosphere," *Proc. Roy. Soc. London A*, **185**:144–175, 1946.

DOE (Dept. of Environment, Great Britain), "Chlorofluorocarbons and Their Effect on Stratospheric Ozone," (Second Report), Pollution Paper No. 15, *HMSO*, London, 1979.

Drummond, J.R. and R.F. Jarnot, "Infrared Measurements of Stratospheric Composition II. Simultaneous NO and NO_2 Measurements," *Proc. Roy. Soc. London A*, **364**:237–254, 1978.

Drummond, J.W., J.M. Rosen, and D.J. Hoffman, "Balloon Borne Chemiluminesent Measurement of NO to 45 km," *Nature*, **265**:319–329, 1977.

Dutsch, H.U., "Vertical Ozone Distribution on a Global Scale," *Pure Appl. Geophys.*, **116**:511–529, 1978.

Dutsch, H.U. and C. Ling, "Six Years of Regular Ozone Soundings Over Switzerland," *Pure Appl. Geophys.* **106–108**:1151–1168, 1973.

CHAPTER 21

Ebel, A. "Eddy Diffusion Models for the Mesosphere and Lower Thermosphere," *J. Atmos. Terr. Phys.*, **42:** 617–628, 1980.

Ehhalt, D.H., "In Situ Measurements of Stratospheric Trace Constituents," *Rev. Geophys. Space Phys.*, **16:**217–224, 1978.

Ehhalt, D.H. and L.E. Heidt, "The Concentration of Molecular H_2 and CH_4 in the Stratosphere," *Pure Appl. Geophys.*, **106–108:**1352–1360, 1973a.

Ehhalt, D.H. and L.E. Heidt, "Vertical Profiles of CH_4 in the Troposphere and Stratosphere," *J. Geophys. Res.*, **78:**5265–5271, 1973b.

Ehhalt, D.H. and U. Schmidt, "Sources and Sinks of Atmospheric Methane," *Pure Appl. Geophys.*, **116:**452–464, 1978.

Ehhalt, D.H. and A. Tonnissen, "Hydrogen and Carbon Compounds in the Stratosphere," *Proc. of the NATO Advanced Study Institute on Atmospheric Ozone*, edited by A.C. Aiken, FAA-EE-80-20, 129–151, 1980.

Ehhalt, D.H., L.E. Heidt, R.H. Lueb, and N. Roper, "Vertical Profiles of CH_4, H_2, CO, N_2O, and CO_2 in the Stratosphere," *Proc. III CIAP Conf.*, Boston, edited by A.J. Broderick and T.M. Hard, DOT-TSC-OST-74-15, U.S. Dept. of Transportation, 153–159, 1974.

Ehhalt, D.H., L.E. Heidt, R.H. Lueb, and E.A. Martell, "Concentrations of CH_4, CO, CO_2, H_2, H_2O, and N_2O in the Upper Stratosphere," *J. Atmos. Sci.*, **32:**163–169, 1975a.

Ehhalt, D.H., L.E. Heidt, R.H. Lueb, and W. Pollock, "The Vertical Distribution of Trace Gases in the Stratosphere," *Pure Appl. Geophys.*, **113:**389–402, 1975b.

Ellsaesser, H.W., J.E. Harries, D. Kley, and R. Penndorf, "Stratospheric H_2O," *Planet. Space Sci.*, **28:**827–835, 1980.

Evans, W.F.J. and E.J. Llewellyn, "Molecular Oxygen Emissions in the Airglow," *Ann. Geophys.* **26:**167–178, 1970.

Evans, W.F.J., H. Fast, J.B. Kerr, C.T. McElroy, R.S. O'Brien, D.I. Wardle, J.C. McConnell, and B.A. Ridley, "Stratospheric Constituent Measurements from Project Stratoprobe," *Proc. WMO Symposium on the Geophysical Aspects and Consequences of Change in the Composition of the Stratosphere*, World Meteorological Organization, **511:**55–60, Geneva, 1978.

FAA (U.S. Federal Aviation Administration), *Second Biennial Report Prepared in Accordance with the Ozone Protection Provision, Section 153(g) of the Clean Air Act Amendments of 1977*, Federal Aviation Administration, U.S. Dept. of Transportation, Washington, D.C., 1979.

Fabian, P., R. Borchers, K.H. Weiler, U. Schmidt, A. Volz, D.H. Ehhalt, W. Seiler, and F. Muller, "Simultaneously Measured Vertical Profiles of H_2, CH_4, CO, N_2O, $CFCl_3$, and CF_2Cl_2 in the Mid-Latitude Stratosphere and Troposphere," *J. Geophys. Res.* **84:**3149–3154, 1979.

Fabian, P., R. Borchers, G. Flentje, W.A. Matthews, W. Seiler, H. Giehl, K. Bunse, F. Muller, U. Schmidt, A. Volz, A. Khedim, and F.J. Johnen, "The Vertical Distribution of Stable Trace Gases at Mid-Latitudes," *J. Geophys. Res.*, **86:**5179–5184, 1981.

Farmer, C.B., O.F. Raper, B.D. Robbins, R.A. Toth, and C. Muller, "Simultaneous Spectroscopic Measurements of Stratospheric Species: O_3, CH_4, CO, CO_2, N_2O, H_2O, HCl, and Hf at Northern and Southern Mid-Latitudes," *J. Geophys. Res.*, **85:**1621–1632, 1980.

Farmer, C.B., O.F. Raper, C. Muller, B.D. Robbins, and R.A. Toth, "Simultaneous Spectroscopic Measurements of Stratospheric Species: O_3, CH_4, CO, CO_2, N_2O, H_2O, HCl and HF at Northern and Southern Mid-Latitudes," *J. Geophys. Res.*, **85:**1621–1632, 1980.

Fels, S.B., J.D. Mahlman, M.D. Schwarzkopf, and R.W. Sinclair, "Stratospheric Sensitivity to Perturbations in Ozone and Carbon Dioxide: Radiative and Dynamical Response," *J. Atmos. Sci.*, **37:**2265–2297, 1980.

Fischer, H., F. Fergg, and D. Rabus, "Radiometric Measurements of Stratospheric H_2O, HNO_3, and NO_2 Profiles," *Proc. Intl. Radiation Symposium*, Fort Collins, August 1980.

Fishman, J., "Ozone in the Troposphere," in *Ozone in the Free Atmosphere*, edited by R.C. Whitten and S. Prasad, Von Nostrand Reinhold, New York, 1985.

Fishman, J., S. Solomon, and P.J. Crutzen, "Observational and Theoretical Evidence in Support of a Significant in situ Photochemical Source of Tropospheric Ozone," *Tellus*, **31:**432–446, 1979.

Fontanella, J.C., A. Girard, L. Gramont, and N. Louisnard, "Vertical Distribution of NO, NO_2, and HNO_3 as Derived from Stratospheric Absorption Infrared Spectra," *Appl. Opt.*, **14:**825–839, 1975.

Garcia, R.R. and D.L. Hartmann, "The Role of Planetary Waves in the Maintenance of the Zonally Averaged Ozone Distribution of the Upper Stratosphere," *J. Atmos. Sci.*, **37:**2248–2264, 1980.

Garcia, R.R. and S. Solomon, "A Numerical Model of the Zonally-Averaged Dynamical and Chemical Structure of the Middle Atmosphere *J. Geophys. Res.*, **88:**1379–1400, 1983.

Garcia, R.R., S. Solomon, R.G. Roble, and D.W. Rusch, "Numerical Study of the Response of the Middle Atmosphere to the 11-Year Solar Cycle," *Planet. Space Sci.*, in press, 1984.

Gidel, L.T., P.J. Crutzen, and J. Fishman, "A Two-Dimensional Photochemical Model of the Atmosphere," *J. Geophys. Res.*, **88:**6622–6640, 1983.

Gille, J.C., G.P. Anderson, W.J. Kohri, and P.L. Bailey, "Observations of the Interaction of Ozone and Dynamics," *Proceedings of the Quadrennial International Ozone Symposium*, 1980, Boulder, Colo, edited by J. London, International Ozone Commission, Vienna Va., 1007–1011, 1980a.

Gille, J.C., P.L. Bailey, and J.M. Russell III, "Temperature and Composition Measurements from the LRIR and LIMS Experiments on Nimbus 6 and 7," *Phil. Trans. Roy Soc. London*, **A296:**205–218, 1980b.

Gille, J.C., C.M. Smythe, and D.F. Heath, "Observed Ozone Response to Variations in Solar UV," *Science*, accepted for publication in 1984.

Giraud, A., G. Scialom, A. Pokhounkov, S. Poloskov, and

G. Tulinov, "Daytime and Nighttime Ion Composition Between 100 and 400 km Under Solar Maximum Winter Conditions," *Space Res.*, **11**:1057–1062, 1971.

Girard, A., J. Besson, R. Giraudet, and L. Gramont, "Correlated Seasonal and Climate Variations of Trace Constituents in the Stratosphere," *Pure Appl. Geophys.*, **117**:381–394, 1978/79.

Goldan, P.D., F.G. Fernald, W.J. Williams, and D.G. Murcray, "Vertical Distribution of NO_2 in the Stratosphere as Determined from Balloon Measurements of Solar Spectra in the 4500 A Region," *Geophys. Res. Lett.*, **5**:257–260, 1978.

Goldan, P.D., W.C. Kuster, D.L. Albritton, and A.L. Schmeltokopf, "Stratospheric $CFCl_3$, CF_2Cl_2, and N_2O Height Profile Measurements at Several Latitudes," *J. Geophys. Res.*, **85**:413–423, 1980.

Goldan, P.D., W.C. Kuster, A.L. Schmeltekopf, F.C. Fehsenfeld, and D.L. Albritton, "Correction of Atmospheric N_2O Mixing-Ratio Data," *J. Geophys. Res.*, **85**:5385–5386, 1981.

Golomb, J.D. and R.E. Good, "Atomic Oxygen Profiles over Churchill and Hawaii from Chemical Releases," *Space Res.*, **12**:675–683, Akademie-Verlag, Berlin, 1972.

Good, R.E., "Determination of Atomic Oxygen Density from Rocket-Borne Measurements of Hydroxyl Airglow," *Planet. Space Sci.*, **24**:389, 1976.

Goody, R.M., *Atmospheric Radiation I: Theoretical Basis*, Oxford University Press, London, 1964.

Gordley, L.L., and J.M. Russell III, "Rapid Inversion of Limb Radiance Data Using an Emissivity Growth Approximation," *Appl. Opt.*, 20:807–813, 1981.

Gough, M.P., "On the Ultraviolet Source of the Nighttime E-Region," *Planet Space Sci.*, **23**:1236, 1975.

Harries, J.E., "Spectroscopic Observations of Middle Atmosphere Composition," in *The Middle Atmosphere as Observed from Balloons, Rockets and Satellites*, Cambridge University Press, Cambridge, 161–173, 1980.

Harries, J.E., D.G. Moss, N.R.W. Swann, G.F. Neill, and P. Gildwarg, "Simultaneous Measurements of H_2O, NO_2, and HNO_3 in the Daytime Stratosphere from 15 to 35 km," *Nature*, **259**:300–302, 1976.

Harrison, H., J.E. Johnson, and J.D. Cline, "Light Hydrocarbon in the Atmospheric Boundary Layer Over the North Pacific," Unpublished Manuscript, Dept. of Atmos. Sci., Univ. of Washington, Seattle, 1979.

Hartmann, D.L., Comments on "Stratospheric Long Waves: Comparison of Thermal Structure in the Northern and Southern Hemispheres," *J. Atmos. Sci.*, **34**:434–435, 1977.

Hartmann, D.L., "A Note Concerning the Effect of Varying Extinction on Radiative-Photochemical Relaxation," *J. Atmos. Sci.*, **35**:1125–1130, 1978.

Hays, P.B. and J.J. Olivero, "Carbon Dioxide and Monoxide above the Troposphere," *Planet. Space Sci.*, **18**:1729, 1970.

Hays, P.B. and R.G. Roble, "Observations of Mesospheric Ozone at Low Latitudes," *Planet. Space Sci.*, **11**:273, 1973.

Heaps, W.S., T.J. McGee, R.D. Hudson, and L.O. Caudill, "Stratospheric Ozone and Hydroxyl Radical Measurements by Balloon-borne Lidar," *Appl. Optics* **21**:2265–2274, 1982.

Heath, D.F., L.L. Mateer, and A.J. Krueger, "The Nimbus-4 BUV Atmospheric Ozone Experiment: Two Years Operation," *Pure Appl. Geophys.*, **106–108**,1238–1253, 1973.

Heath, D.F., A.J. Krueger, and P.J. Crutzen, "Solar Proton Event: Influence on Stratospheric Ozone," *Science*, **197**:886–889, 1977.

Heidt, L.E. and D.H. Ehhalt, "Correlations of CH_4 Concentrations Measured Prior to 1974," *Geophys. Res. Lett.*, **7**:1023, 1980.

Heidt, L.E., R. Lueb, W. Pollock, and D.H. Ehhalt, "Stratospheric Profiles of CCl_3F and CCl_2F_2," *Geophys. Res. Lett.*, **2**:445–447, 1975.

Heidt, L.E., J.P. Krasnec, R.A. Lueb, W.H. Pollock, B.E. Henry, and P.J. Crutzen, "Latitudinal Distributions of CO and CH_4 Over the Pacific," *J. Geophys. Res.*, **85**:7329–7336, 1980.

Henderson, W.R., "D-Region Atomic Oxygen Measurement," *J. Geophys. Res.*, **76**:3166, 1971.

Herman, J.R., "The Response of Stratospheric Constituents to a Solar Eclipse, Sunrise, and Sunset," *J. Geophys. Res.*, **84**:3701–3710, 1979.

Hesstvedt, E., "On the Water Vapor Content of the High Atmosphere," *Geofys. Publ. Oslo*, **25**:1, 1964.

Hilsenrath, E., "Rocket Observations of the Vertical Distribution of Ozone in the Polar Night and During a Mid-Winter Stratospheric Warming," *Geophys. Res. Lett.*, **7**:581–584, 1980.

Hilsenrath, E. and B. Schlesinger, "The Seasonal and Interannual Variability of Total Ozone as Revealed by the BUV Nimbus-4 Experiment," *Fourth NASA Weather and Climate Program Science Review*, NASA Conference Publication 2976, edited by E.R. Kreins, 277–286, 1979.

Hilsenrath, E., and B. Schlesinger, "Total Ozone Seasonal and Interannual Variations Derived from the 7 Year Nimbus-4 BUV Data Set," *J. Geophys. Res.*, **86**:12087–12096, 1981.

Hoffman, J.H., "A Mass Spectrometric Determination of the Composition of the Nighttime Topside Ionosphere," *J. Geophys. Res.*, **72**:1883–1888, 1967.

Hoffman, J.H., C.Y. Johnson, J.C. Holmes, and J.M. Young, "Daytime Mid-Latitude Ion Composition Measurements," *J. Geophys. Res.*, **74**:6281–6290, 1969.

Holmes, J.C., C.Y. Johnson, and J.M. Young, "Ionospheric Chemistry," *Space Res.*, **5**:756–766, 1965.

Holton, J.R., "The Dynamic Meteorology of the Stratosphere and Mesosphere," *Meteorol. Monographs*, **15**:1980.

Horvath, J.J. and C.J. Mason, "Nitric Oxide Mixing Ratios Near the Stratopause Measured by a Rocket-Borne Chemiluminescent Detector," *Geophys. Res. Lett.*, **5**:1023–1026, 1978.

Howlett, L.C., K.D. Baker, L.R. Megill, A.W. Shaw, and W.R. Pendleton, "Measurement of a Structured Profile of Atomic Oxygen in the Mesosphere and Lower Thermosphere," *J. Geophys. Res.*, **85**:1291–1298, 1980.

CHAPTER 21

Hsu, C.P.F., "A Numerical Study of the Role of Wave-Wave Interactions During Sudden Stratospheric Warmings," *J. Atmos. sci.*, **38**:189–214, 1981.

Hudson, R.D., "Minor Constituents in the Stratosphere and Mesosphere," *Rev. Geophys. Space Phys.*, **17**:467, 1979.

Hunt, B.G., "Photochemistry of Ozone in a Moist Atmosphere," *J. Geophys. Res.*, **71**:1385–1398, 1966.

Inn, E.C.Y., J.F. Vedder, B.J. Tyson, and D. O'Hara, "COS in the Stratosphere," *Geophys. Res. Lett.*, **6**:191–193, 1979.

Jackman, G.H., J.E. Frederick, and R.S. Stolarski, "Production of Odd Nitrogen in the Stratosphere and Mesosphere: An Intercomparison of Source Strengths," *J. Geophys. Res.*, **85**:7495, 7575, 1980.

Johnson, C.Y. "Ionospheric Composition and Density from 90 to 1200 km at Solar Minimum," *J. Geophys. Res.*, **71**:330–332, 1966.

Johnston, H.S., O. Serang, and J. Podolske, "Instantaneous Global Nitrous Oxide Photochemical Rates," *J. Geophys. Res.*, **84**:5077–5082, 1979.

Kebarle, P., S.K. Searles, A. Zolla, J. Scarborough, and M. Arshadi, "The Solvation of the Hydrogen Ion by Water Molecules in the Gas Phase. Heats and Entropies of Solvation of Individual Reactions: $H^+(H_2O)_{n-1} + H_2O \rightarrow H^+(H_2O)_n$," *J. Amer. Chem. Soc.*, **89**:6393–6399, 1967.

Keneshea, T.J. and M.A. MacLeod, "Wind Induced Modification of E Region Ionization Profiles," *J. Atmos. Sci.*, **27**:981–984, 1970.

Keneshea, T.J. and S.P. Zimmerman, "The Effect of Mixing upon Atomic and Molecular Oxygen in the 70–170 km Region of the Atmosphere," *J. Atmos. Sci.*, **27**:831–840, 1970.

Keneshea, T.J., R.S. Narcisi, and W. Swider, "Diurnal Model of the E-Region," *J. Geophys. Res.*, **75**:845–854, 1970.

Keneshea, T.J., S.P. Zimmerman, and C.R. Philbrick, "A Dynamic Model of the Mesosphere and Lower Thermosphere," *Planet. Space Sci.*, **27**:385–401, 1979.

Kerr, J.B. and C.T. McElroy, "Measurement of Stratospheric Nitrogen Dioxide from the AES Stratospheric Balloon Program," *Atmosphere*, **14**:166–171, 1976.

Kirchhoff, V.W.J.H., B.R. Clemesha, and D.M. Simonich, "The Atmospheric Neutral Sodium Layer: Recent Modeling Compared to Measurements," *J. Geophys. Res.*, **86**:6892–6898, 1981.

Klenk, K.F., P.K. Bhartia, E. Hilsenrath, and A.J. Fleig, "Standard Ozone Profiles from Balloon and Satellite Data Sets," *J. Climate Appl. Meteorol.*, **22**:2012–2022, 1983.

Kley, D. and M. McFarland, "Chemiluminescence Detector for NO and NO_2," *Atmos. Technol.*, **12**:63–68, 1980.

Kley, D., E.J. Stone, W.R. Henderson, J.W. Drummond, W.J. Harrop, A.L. Schmeltekopf, T.L. Thompson, and R.H. Winkler, "In Situ Measurements of the Mixing Ratio of Water Vapor in the Stratosphere," *J. Atmos. Sci.*, **36**:2513–2524, 1979.

Kley, D., J.W. Drummond, and A.L. Schmeltekopf, "On the Structure and Microstructure of Stratospheric Water Vapor," in *Atmospheric Water Vapor*, edited by A. Deepak, T.D. Wilkerson, and L.H. Ruhnke, Academic Press, New York, 314–327, 1980.

Kley, J.W., Drummond, M. McFarland, and S.C. Liu, "Tropospheric Profiles of NO," *J. Geophys. Res.*, **86**:3153–3161, 1981.

Komhyr, W.D. and T.B. Harris, "Development of an ECC Ozonesonde," NOAA Technical Report ERL-ARCL18, Boulder, Colo., 1971.

Kopp, E. and C.R. Philbrick, "Wind and Temperature Induced Effects on Mesospheric Ion Composition and Inferred Minor Constituents," Sixth ESA Symposium on European Rocket and Balloon Programs and Related Research (Interlaken, 1983), ESA Report SP-183, European Space Agency, Paris, 1983.

Kopp, E., L. Andre, P. Eberhardt, and U. Herrmann, "Negative Ion Composition of the D-Region during the 1979 Solar Eclipse," *Eos*, **61**:311, 1980.

Krueger, A.J., "The Mean Ozone Distribution from Several Series of Rocket Soundings to 52 km at Latitudes from 58°S to 64°N," *Pure Appl. Geophys.*, **106–108**: 1272–1280, 1973.

Krueger, A.J., "Nimbus 7 Total Ozone Mapping Spectrometer (TOMS) Data during Gap, France Ozone Intercomparison of June 1981," *Planet. Space Sci.*, **31**:773–777, 1983.

Krueger, A.J., and R.A. Minzner, "A Mid-Latitude Model for the 1976 U.S. Standard Atmosphere," *J. Geophys. Res.*, **81**:4477–4481, 1976.

Krueger, A.J., B. Guenther, A.J. Fleig, D.F. Heath, E. Hilsenrath, R. McPeters, and C. Prabakhara, "Satellite Ozone Measurements," *Philos. Trans. Roy. Soc. London*, **A296**:991–204, 1980.

Lazrus, A.L. and B.W. Gandrud, "Distribution of Stratospheric Nitric Acid Vapor," *J. Atmos. Sci.*, **31**:1102–1108, 1974.

Lazrus, A.L., B.W. Gandrud, J. Greenberg, J. Bonelli, E. Mroz, and W.A. Sedlacek, "Midlatitude Seasonal Measurements of Stratospheric Acidic Chlorine Vapor," *Geophys. Res. Lett.*, **4**:587–589, 1977.

Lean, J.L., "Observation of the Diurnal Variation of Atmospheric Ozone," *J. Geophys. Res.*, **87**:4973–4980, 1982.

Lippens, C., and C. Muller, "Atmospheric Nitric Acid and Chlorofluoromethane from Interferometric Spectra Obtained at the Observations du Pic du Midi," *Aeronomica Acta* **A277** Inst. d'Aeronomie Spatiale Belgique, Brussels, 1980.

Liu, S.C., D. Kley, M. McFarland, J.D. Mahlman, and H. Levy II, "On the Origin of Tropospheric Ozone," *J. Geophys. Res.*, **85**:7546–7552, 1980.

Loewenstein, M. and H. Savage, "Latitudinal Measurements of NO and O_3 in the Lower Stratosphere from 5° to 82° North," *Geophys. Res. Lett.*, **2**:448–450, 1975.

Loewenstein, M., H.F. Savage, and R.C. Whitten, "Seasonal Variation of NO and O_3 at Altitudes of 18.3 and 21.3 km," *J. Atmos. Sci.*, **32**:2185–2190, 1975.

Loewenstein, M., H.F. Savage, and J.B. Borucki, "Geographical Variations of NO and O_3 in the Lower Stratosphere," *International Conference on Problems Related to the Stratosphere*, Utah State University, Logan,

JPL Publication 77-12 Pasedena, Jet Propulsion Lab, Calif. Inst. of Tech. 230–233, 1977.

Loewenstein, M., W.J. Starr, and D.G. Murcray, "Stratospheric NO and HNO_3 Observations in the Northern Hemisphere for Three Seasons," *Geophys. Res. Lett.*, **5**:531–534, 1978a.

Loewenstein, M., W.J. Borucki, H.F. Savage, J.G. Borucki, and R.C. Whitten, "Geographical Variations of NO and O_3 in the Lower Stratosphere," *J. Geophys. Res.*, **83**:1874–1882, 1978b.

Logan, J.A., M. Prather, S. Wofsy, and M.B. McElroy, "Atmospheric Chemistry: Response to Human Influence," *Phil. Trans. Roy. Soc. London, A*, **290**:187–234, 1978.

Logan, J.A., M.J. Prather, S.C. Wofsy, and M.B. McElroy, "Tropospheric Chemistry: A Global Perspective," *J. Geophys. Res.*, **86**:7210–7254, 1981.

London, J. and C.A. Reber, "Solar Activity and Total Atmospheric Ozone," *Geophys. Res. Lett.*, **6**:869–872, 1979.

London, J., R. Bojkov, S. Oltmans, and J. Kelley, "Atlas of the Global Distribution of Total Ozone July 1957–July 1967," NCAR Technical Report NCAR/TN/**113**, National Center for Atmospheric Research, Boulder, Colo., 1976.

Lordi, N.J., A. Kasahara, and S.K. Kao, "Numerical Simulation of Sudden Warmings with a Primitive Equation Spectral Model," *J. Atmos. Sci.*, **37**:3746–2767, 1980.

Louisnard, N., G. Fergant, A. Girard, L. Gramont, O. Lado-Bradowsky, J. Laurent, S. LeBoiteux, and M.P. Lemitre, "Infrared Absorption Spectroscopy Applied to Stratospheric Profiles of Minor Constituents,", *J. Geophys. Res.*, **28**:5365–5377, 1983.

Louisnard, N., G. Fergant, and A. Girard, "Simultaneous Measurements of Methane and Water Vapor Vertical Profiles in the Stratosphere," *Proceedings of the Quadrennial International Ozone Symposium*, 1980, Boulder, Colo, edited by J. London, International Ozone Commission, Vienna, Va, 797–802, 1980a.

Louisnard, N., A. Girard, and G. Eichen, "Mesures du Profil Vertical de Concentration de la Vapeur d'eau Stratospherique," *Comptes Rendus Acad. Sci. Paris, Ser. B*, **290**:385–388, 1980b.

Lovill, J.E., et al., "Total Ozone Retrieval from Satellite Multichannel Filter Radiometer Measurements," *UCRL-52773*, Lawrence Livermore Laboratory, University of California Livermore, 1978.

Maeda, K., "Semiannual Oscillation of Stratospheric Ozone," *Geophys. Res. Lett.*, **11**:583–586, 1984.

Maier, E.J., A.C. Aiken, and J.E. Ainsworth, "Stratospheric Nitric Oxide and Ozone Measurements Using Photoionization Mass Spectrometer and UV Absorption," *Geophys. Res. Lett.*, **5**:37–40, 1978.

Mason, C.J. and J.J. Horvath, "The Direct Measurement of Nitric Oxide Concentration in the Upper Atmosphere by a Rocket-Borne Chemiluminescent Detector," *Geophys. Res. Lett.*, **3**:391–394, 1976.

Mastenbrook, H.J., "Water Vapor Distribution in the Stratosphere and High Troposphere," *J. Atmos. Sci.*, **25**:299–311, 1968.

Mastenbrook, H.J. and R.E. Daniels, "Measurements of Stratospheric Water Vapor Using a Frost Point Hygrometer," in *Atmospheric Water Vapor*, edited by A. Deepak, T.W. Wilkerson, and L.H. Ruhnke, Academic Press, New York, 329–342, 1980.

Mateer, C.L., "On the Information Content of Umkehr Observations," *J. Atmos. Sci.*, **22**:370–381, 1965.

Mateer, C.L., J.J. DeLuisi, and C.C. Porco, "The Short Umkehr Method, Part I: Standard Ozone Profiles for use in the Estimation of Ozone Profiles by the Inversion of Short Umkehr Observations," NOAA Technical Memorandum ERL ARL-86, 1980.

Mauersberger, K., "Measurement of Heavy Ozone in the Stratosphere," *Geophys. Res. Lett.*, **8**:935–937, 1981.

Mauersberger, L., and R. Finstad, "Carbon Dioxide Measurements in the Stratosphere," *Geophys. Res. Lett.*, **7**:873–876, 1980.

Mauersberger, K., R. Finstad, S. Anderson, and D. Robbins, "A comparison of Ozone Measurements," *Geophys. Res. Lett.*, **8**:361–364, 1981.

Maugh, T.H. II, "What is the Risk from Chlorofluorocarbons?" *Science*, **223**:1051–1052, 1984.

McPeters, R.D., D.F. Heath, and P.K. Bhartia, "Average Ozone Profiles for 1979 from the NIMBUS 7 SBUV Instrument," *J. Geophys. Res.*, **89**:5199–5214, 1984.

Megie, G., J.Y. Allain, M.L. Chanin, and J.E. Blamont, "Vertical Profile of Stratospheric Ozone by Lidar Sounding from the Ground," *Nature*, **270**:329–331, 1977.

Megie, G., F. Bos, J.E. Blamont, and M.L. Chanin, "Simultaneous Nighttime Lidar Measurement of Atmospheric Sodium and Potassium," *Planet. Space Sci.*, **26**:27, 1978.

Mentall, J.E., J.P. Herman, and B. Zak, "Ozone Profile and Solar Fluxes from a Grating Spectrometer," *The Stratcom VIII Effort*, NASA Technical Paper 1640, edited by E. Reed, 117–123, 1980.

Mihelcic, D., D.H. Ehhalt, G.F. Kulessa, J. Klomfass, M. Trainer, U. Schmidt, and H. Rohrs, "Measurements of Free Radicals in the Atmosphere by Matrix Isolation and Electron Paramagnetic Resonance," *Pure Appl. Geophys.*, **116**:530–536, 1978.

Miller, A.J., R.M. Nagatani, J.D. Laver, and B. Korty, "Utilization of 100 mb Midlatitude Height Fields as an Indicator of Sampling Effects in Total Ozone Variations," *Mon. Wea. Rev.*, **107**:782–787, 1979.

Miller, A.J., R.M. Nagatani, T.G. Rogers, A.J. Fleig, and D.F. Heath, "Total Ozone Variations 1970–74 Using Backscattered Ultraviolet (BUV) and Ground-Based Observations," *J. Appl. Meteorol.*, **21**:621–630, 1982.

Miller, E., B.A. Emery, and R.G. Roble, "OSO-8 Lower Mesospheric Ozone Number Density Profiles," in *Proceedings of the Quadrennial International Ozone Symposium*, 1980, Boulder, Colo., edited by J. London, International Ozone Commission, Vienna, Va., **1**:572–575, 1981.

Molina, M.J. and F.S. Rowland, "Stratospheric Sink for Chlorofluoromethanes: Chlorine Atomic-catalysed Destruction of Ozone," *Nature*, **249**:810, 1974.

Murcray, D.G., A. Goldman, W.J. Williams, F.H. Mur-

cray, J.N. Brooks, J. Van Allen, R.N. Stocker, J.J. Kosters, D.B. Barker, and D.E. Snider, "Recent Results of Stratospheric Trace Gas Measurements from Balloon-Borne Spectrometers," Proc. Third CIAP Conf., edited by A.J. Broderick and T.M. Hard, DOT-TSC-OST-74-15, U.S. Dept. of Transportation, Washington, D.C., 184–192, 1974.

Murcray, D.G., D.B. Barker, J.N. Brooks, A. Goldman, and W.J. Williams, "Seasonal and Latitudinal Variation of the Stratospheric Concentration of HNO_3," *Geophys. Res. Lett.*, **2**:223–225, 1975.

Murcray, D.G., A. Goldman, F.H. Murcray, F.J. Murcray, and W.J. Williams, "Stratospheric Distribution of $ClONO_2$," *Geophys. Res. Lett.*, **6**:856–859, 1979.

Murcray, D.G., W.J. Williams, D.B. Barker, A. Goldman, C. Bradford, and G. Cook, "Measurements of Constituents of Interest in the Chemistry of the Ozone Layer Using IR Techniques," *Proceedings WMO Symposium on the Geophysical Aspects and Consequences of Change in the Composition of the Stratosphere*," WMO NO. 511, 61–68, Toronto, Canada, June 1978.

Murgatroyd, R.J., "Recent Progress in Studies of the Stratosphere," *Q. J. Roy. Meteorol. Soc.*, **108**:271–312, 1982.

Nagatani, R.M. and A.J. Miller, "Stratospheric Ozone Changes During the First Year of SBUV Observations," *J. Geophys. Res.*, **89**:519–5198, 1984.

Narcisi, R.S., "Processes Associated with Metal-Ion Layers in the E Region of the Ionosphere," *Space Res.*, **8**:360–369, 1968.

Narcisi, R.S., "Discussion, in Meteorological and Chemical Factors in D-Region Aeronomy - Record of the Third Aeronomy Conference," *Aeronomy Rep. 34*, University of Illinois, 284, 1969.

Narcisi, R.S., "Composition Studies of the Lower Ionosphere," in *Physics of the Upper Atmosphere*, edited by F. Verniani, Editrice Compositori, Bologna, 1971.

Narcisi, R.S., "Mass Spectrometer Measurements in the Ionosphere," 171–183 in *Physics and Chemistry of the Upper Atmosphere*, edited by B.M. McCormac, D. Reidel, Dordrecht, Holland, 1973.

Narcisi, R.S. and A.D Bailey, "Mass Spectrometric Measurements of Positive Ions at Altitudes from 64 to 112 Kilometers," *J. Geophys. Res.*, **70**:3687–3700, 1965.

Narcisi, R.S., A.D. Bailey, L.E. Wlodyka, and C.R. Philbrick, "Ion Composition Measurements in the Lower Ionosphere during the November 1966 and March 1970 Solar Eclipse," *J. Atmos. Terr. Phys.*, **34**:647–668, 1972.

Narcisi, R.S., A. Bailey, G. Frederico, and L. Wlodyka, "Positive and Negative Ion Composition Measurements in the D- and E-Regions During the 26 Fegruary 1979 Solar Eclipse," *J. Atmos. Terr. Phys.*, **45**:461–478, 1983.

National Academy of Sciences, *Environmental Impact of Stratospheric Flight*, National Academy Press, Washington, D.C., 1975.

National Academy of Sciences, *Stratospheric Ozone Depletion by Halocarbons: Chemistry and Transport*, National Academy Press, Washington, D.C., 1979.

National Academy of Sciences, *Causes and Effects of Stratospheric Ozone Reduction: An Update*, National Academy Press, Washington, D.C., 1982

Nat'onal Academy of Sciences, *Causes and Effects of Stratospheric Ozone: Update, 1983*, National Academy Press, Washington, D.C., 1984.

NASA (U.S. National Aeronautics and Space Administration), *Chlorofluoromethanes and the Stratosphere*, Reference Publication 1010, edited by R.D. Hudson, 1977.

NASA (U.S. National Aeronautics and Space Administration), *The Stratosphere: Present and Future*, Reference Publication 1049, edited by R.D. Hudson and E.I. Reed, 1979.

NASA (U.S. National Aeronautics and Space Administration), *Present State of Knowledge of the Upper Atmosphere: An Assessment Report*, NASA/GSFC, Greenbelt, Md., 1984.

NASA/JPL, "Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling," Eval. Number 5, JPL Publ. 82-54, Jet Propulsion Laboratory, Calif. Inst. of Tech, Pasadena, Calif., 1982.

Naudet, J.P., R. Rigaud, and D. Huguenin, "Stratospheric NO_2 at Night from Stellar Spectra in the 440 nm Region," *Geophys. Res. Lett.*, **7**:701–703, 1980.

Naudet, J.P., D. Huguenin, P. Rigaud, and D. Cariolle, "Stratospheric Observations of NO_2 and its Experimental and Theoretical Distribution Between 20 and 40 km," *Planet. Space Sci.*, **29**:707–712, 1981.

Nicolet, M., "Solar Radio Flux and the Temperature of the Upper Atmosphere," *J. Geophys. Res.*, **68**:6121–6144, 1963.

Nicolet, M., "Stratospheric Ozone: An Introduction to Its Study," *Rev. Geophys. Space Phys.*, **13**:593, 1975.

Nicolet, M. and W. Swider, "Ionospheric Conditions," *Planet. Space Sci.*, **11**:1459–1482, 1963.

Niple, E., W.G. Mankin, A. Goldman, D.G. Murcray, and F.J. Murcray, "Stratospheric NO_2 and H_2O Mixing Ratio Profiles from High Resolution Infrared Solar Spectra Using Nonlinear Least Squares," *Geophys. Res. Lett.*, **7**:489–492, 1980.

Noxon, J.F., "Atomic Nitrogen Fixation by Lightning," *Geophys. Res. Lett.*, **3**:463–465, 1976.

Noxon, J.F., "Stratospheric NO_2 in the Antarctic Winter," *Geophys. Res. Lett.*, **5**:1021–1022, 1978.

Noxon, J.F., "Stratospheric NO_2, Global Behavior," *J. Geophys. Res.*, **84**:5067–5076, 1979.

Noxon, J.F., "Correction," *J. Geophys. Res.*, **85**:4560–4561, 1980.

Noxon, J.F., E.C. Whipple, Jr., and R.S. Hyde, "Stratospheric NO_2, I. Observational Method and Behavior at Mid-Latitude," *J. Geophys. Res.*, **84**:5047–5065, 1979.

O'Brien, R.S. and W.F.J. Evans, "Rocket Measurements of the Distribution of Water Vapor in the Stratosphere at High Latitudes," *J. Geophys. Res.*, **86**:12,101–12,107, 1981.

Offermann, D. and A. Drescher, "Atomic Oxygen Densities in the Lower Thermosphere as Derived from in situ 5577 A Night Airglow and Mass Spectrometer Measurements," *J. Geophys. Res.*, **78**:6690, 1973.

Offermann, D. and K.U. Grossmann, "Thermospheric Density and Composition as Determined by a Mass Spectrometer with Cryo Ion Source," *J. Geophys. Res.*, **78**:8296, 1973.

Offermann, D. and U. von Zahn, "Atomic Oxygen and Carbon Dioxide in the Lower Thermosphere," *J. Geophys. Res.*, **76**:2520, 1971.

Offermann, D., V. Friedrich, P. Ross, and U. von Zahn, "Neutral Gas Composition Measurements Between 80 and 120 km," *Planet. Space Sci.*, **29**:747, 1981.

Ogawa, T., K. Shibasaki, and K. Suzuki, "Balloon Observation of the Stratospheric NO_2 Profile by Visible Absorption Spectroscopy," *J. Meteorol. Soc. Japan*, **59**:410–416, 1981.

Parent, R.A., "A Review of Ozone Toxicology Studies, Air Quality Meteorology and Atmospheric Ozone," SSTM STP 653, edited by A.L. Morris and R.C. Barnes, American Society for Testing and Materials, Philadelphia, 575-396, 1978.

Parrish, A., R. deZafra, and P. Solomon, "Groundbased MM-Wave Emission Spectroscopy for the Detection and Monitoring of Stratospheric Ozone," in *Proceedings of the Quadrennial International Ozone Symposium*, edited by J. London, IAMAP, NCAR, Boulder, Colo., 122–130, 1981.

Patel, C.K., N.E.G. Burkhardt, and C.A. Lambert, "Spectroscopic Measurements of Stratospheric Nitric Oxide and Water Vapor," *Science*, **184**:1173–1176, 1974.

Pelon, J. and G. Megie, "Ozone Monitoring in the Troposphere and Lower Stratosphere: Evaluation and Operation of a Ground-Based Lidar Station," *J. Geophys. Res.*, **87**:4947–4955, 1982.

Penkett, S.A., K.A. Brice, R.G. Derwent, and A.E.J. Eggleton, "Measurement of CCl_3 and CCl_4 at Harwell Over the Period January 1975–November 1977," *Atmos. Environ.*, **13**:1011–1019, 1979.

Penner, J.E., "Trend Prediction of O_3: An Analysis of Model Uncertainty with Comparison to Detection Thresholds," *Atmos. Environ.*, **16**:1109–1115, 1982.

Persson, R., "World Forest Resources," *Res. Notes 17*, Dept. of Forestry Surv., Royal Coll. of Forestry, Stockholm, 1974.

Peters, R.D., C.H. Jackman, and E. G. Stassinopoulous, "Observations of Ozone Depletion Associated with Solar Proton Events," *J. Geophys. Res.*, **86**:12071–12081, 1981.

Phelps, A.V., "Laboratory Studies of Electron Attachment and Detachment Processes of Aeronomic Interest," *Can. J. Chem.*, **47**: 1783–1793, 1969.

Philbrick, C.R., G.A. Faucher, and R.A. Wlodyka, "Neutral Composition Measurements of the Mesosphere and Lower Thermosphere," AFCRL-71-0602, AD739169, 1971.

Philbrick, C.R., G.A. Faucher, and E. Trzcinski, "Rocket Measurements of Mesospheric and Lower Thermospheric Composition," *Space Res.*, **13**:255–260, Akademie Verlag, Berlin, 1973.

Pollock, W., L.E. Heidt, R. Lueb, and D.H. Ehhalt, "Measurement of Stratospheric Water Vapor by Cryogenic Collection," *J. Geophys. Res.*, **85**:5555–5568, 1980.

Pommereau, J.P. and A. Hauchecorne, "Observations Spectroscopiques Depuis le Sol du Dioxyde d'Azote Atmospherique," *Comptes Rendus, Acad. Sci. Paris, Ser. B.*, **288**:135–138, 1979.

Prabhakara, C., B.J. Conrath, R.A. Hanel, and E.J. Williamson, "Remote Sensing of Atmospheric Ozone Using the 9.6 Micron Band," *J. Atmos. Sci.*, **27**:689–697, 1970.

Profitt, M.H. and R.J. McLaughlin, "Fast-Response Dual-Beam UV-Absorption Ozone Parameter Suitable for Use on Stratospheric Balloons," *Rev. Sci. Instrum.* **54**:1983.

Radford, H.E., M.M. Litvak, C.A. Gottlieb, E.W. Gottlieb, S.K. Rosethal, and A.E. Lilley, "Mesospheric Water Vapor Measured from Ground-Based Microwave Observations," *J. Geophys. Res.*, **82**:472–478, 1977.

Ramseyer, H., P. Eberhardt, and E. Kopp, "Silicon Chemistry and Interred Water Vapor in the Lower Ionosphere," Sixth ESA Symposium on European Rocket and Balloon Programs and Related Research (Interlaken, 1983), ESA Report SP-183, European Space Agency, Paris, 1983.

Reed, R.J. and K.E. German, "A Contribution to the Problem of Stratospheric Diffusion by Large-Scale Mixing," *Mon. Weather Rev.*, **93**:313–321, 1965.

Regener, V.H., "On a Sensitive Method for the Recording of Atmospheric Ozone," *J. Geophys. Res.*, **65**:3975, 1960.

Reid, G.C., "The Production of Water-Cluster Positive Ions in the Quiet Daytime D Region," *Planet. Space Sci.*, **25**:275–290, 1977.

Reinsel, G.C., "Analysis of Total Ozone Data for the Detection of Recent Trends and the Effects of Nuclear Testing During the 1960's," *Geophys. Res. Lett.*, **8**:1227–1230, 1981.

Remsberg, E.E. and L.L. Gordley, "Analysis of Differential Absorption Lidar from the Space Shuttle," *Appl. Opt.*, **17**:624, 1978.

Remsberg, E.E., J.M. Russel III, J.C. Gille, L.L. Gordley, P.L. Bailey, W.G. Planet, and J.E. Harries, "The Validation of NIMBUS 7 LIMS Measurements of Ozone," *J. Geophys. Res.*, **89**:5161–5178, 1984.

Richter, E.S., J.R. Rowlett, C.S. Gardner, and C.F. Sechrist, "Lidar Observation of the Mesospheric Sodium Layer Over Urbana, Illinois," *J. Atmos. Terr. Phys.* **43**:327–337, 1981.

Ridley, B.A. and D.R. Hastie, "Stratospheric Odd-Nitrogen: NO Measurements at 51°N in Summer," *J. Geophys. Res.*, **86**:3162–3166, 1981.

Ridley, B.A. and L.C. Howlett, "An Instrument for Nitric Oxide Measurements in the Stratosphere," *Rev. Sci. Instr.*, **45**:742–746, 1974.

Ridley, B.A. and H.I. Schiff, "Stratospheric Odd-Nitrogen: Nitric Oxide Measurements at 32°N in Autumn," *J. Geophys. Res.*, **86**:3167–3172, 1981.

Ridley, B.A., M. McFarland, J.T. Bruin, H.I. Schiff, and J.C. McConnell, "Sunrise Measurements of Stratospheric Nitric Oxide," *Can. J. Phys.*, **55**:212–221, 1977.

Rigaud, P., J.P. Naudet, and D. Huguenin, "Etude de la Repartition Verticale de NO_2 Stratospherique Durant la Nuit," *Comptes Rendus Academie des Sciences, Paris, Ser. B*, **284**:331–334, 1977.

Rishbeth, H., "On Explaining the Behavior of the Ionospheric F Region", *Rev. Geophys.*, **6**:33–71, 1968.

Robbins, D.E., "NASA-JSC Ozone Observations for Validation of NIMBUS-7 LIMS Data," NASA Tech Memo 58227, 1980.

Robbins, D.E. and J.G. Carnes, "Variations in the Upper Stratosphere's Ozone Profile," *WMO Symposium on the Geophysical Aspects and Consequences of Changes in the Composition of the Stratosphere*, Toronto, Canada, June, 1978, 511, World Meteorological Organization, Geneva, WMO Publ. No. 131–137, 1978.

Robinson, G.D., "The Transport of Minor Atmospheric Constituents Between Troposphere and Stratosphere," *Quart. J. Roy. Meteorol. Soc.*, **106:**227–253, 1980.

Rodgers, C.D., "Retrieval of Atmospheric Temperature and Composition from Remote Measurements of Thermal Radiation," *Rev. Geophys. Space Phys.*, **14:**609–624, 1976

Roe, J.M., "A Climatology of a Newly-Defined Tropopause Using Simultaneous Ozone-Temperature Profiles: Final Report," AFGL TR-81-0190, ADA106399 1981.

Rogers, J.W., A.T. Stair, Jr., T.C. Degges, C.L. Wyatt, and D.J. Baker, Rocketborne Measurement of Mesospheric H_2O in the Auroral Zone," *Geophys. Res. Lett.*, **4:**366–368, 1977.

Rowland, F.S., and M.J. Molina, "Chlorofluoromethanes in the Environment," *Rev. Geophys. Space Phys.*, **13:**1, 1975.

Roscoe, H.K., J.R. Drummond, and R.F. Jarnot, "Infrared Measurements of Stratospheric Composition III. The Daytime Changes of NO and NO_2," *Proc. Roy. Soc. London, A*, **375:**507–528, 1981.

Roy, C.R., I.E. Galbally, and B.A. Ridley, "Measurements of Nitric Oxide in the Stratosphere of the Southern Hemisphere," *Quart. J. Roy. Meteorol. Soc.*, **106:**887–894, 1980.

Rudolph, J., D.H. Ehhalt, and G. Gravenhorst, "Recent Measurements of Light Hydrocarbons in Remote Areas," *Proceedings of the First European Symposium on Physico Chemical Behavior of Atmospheric Pollutants*, Ispra, Italy Oct 16–18, 1979, D. Reidel, Dordrecht, Holland 41–51, 1979.

Rudolph, J., D.H. Ehhalt, and A. Tonnissen, "Vertical Profiles of Ethane and Propane in the Stratosphere," *J. Geophys. Res.*, **86:**7267–7272, 1981.

Rusch, D.W., G.H. Mount, C.A. Barth, G.J. Rottman, R.J. Thomas, G.E. Thomas, R.W. Sanders, G.M. Lawrence, and R.W. Eckman, "Ozone Densities in the Lower Mesosphere Measured by Limb Scanning Ultraviolet Spectrometer," *Geophys. Res. Lett.*, **10:**241–244, 1983.

SCEP, "Man's Impact on the Global Environment," *Report of the Study of Critical Environmental Problems*, MIT Press, Cambridge, Mass., 1970.

Shapiro, M.A., A.J. Krueger, and P.J. Kennedy, "Nowcasting the Position and Intensity of Jet Streams Using a Satellite-Borne Total Ozone Mapping Spectrometer," in *Nowcasting*, edited by K.A. Browning, Academic Press, New York, 1982.

Shapiro, M.A., E.R. Reiter, R.D. Cadle, and W.A. Sedlacek, "Vertical Mass and Trace Constituent Transport in the Vicinity of Jet Streams," *Arch. Met. Geoph. Biokl. B*, **28:**193–206, 1980.

Singh, H.B. and P.L. Hanst, "Peroxyacetyl Nitrate (PAN) in the Unpolluted Atmosphere: An Important Reservoir for Nitrogen Oxides," *Geophys. Res. Lett.*, **4:**453–456, 1981.

Singh, H.B., L.J. Salas, H. Shigeishi, and E. Scribner, "Atmospheric Halocarbons and Sulfur Hexafluoride: Global Distributions, Sources, and Sinks," *Science*, **203:**899–903, 1979.

Smith, E.K. Jr., and S. Matsushita (eds.), *Ionospheric Sporadic E*, Pergamon Press, New York, 1962.

Smith, D., N.G. Adams, and M.J. Church, "Mutual Neutralization Rates of Ionospherically Important Ions," *Planet. Space Sci.*, **24:**697–703, 1976.

Soloman, S. "Minor Constituents in the Stratosphere and Mesosphere," *Rev. Geophys. Space Phys.*, **21:**276–283, 1983.

Solomon, S., P.J. Crutzen, and R.G. Roble, "Photochemical Coupling Between the Thermosphere and Lower Atmosphere," *J. Geophys. Res.*, **87:**7206, 1982.

Solomon, S., G.C. Reid, D.W. Rusch, and R.J. Thomas, "Mesospheric ozone depletion during the solar proton event of July 13, 1982: Part II," *Geophys. Res. Lett.*, **10:** , 1983.

Stair, A.T. Jr., J.C. Ulwick, K.D. Baker, and D.J. Baker, *Atmospheres of Earth and Planets*, edited by B.M. McCormac, D. Reidel, Dordrecht, Holland, 1975.

Stair, A.T. Jr., R.D. Sharma, R.M. Nadile, D.J. Baker, and W.F. Grieder, "Observations of Limb Radiance with Cryogenic Spectral Infrared Rocket Experiment (SPIRE)," *J. Geophys Res.*, 1984.

Stolarski, R.S. and R.J. Cicerone, "Stratospheric Chlorine: A Possible Sink for Ozone," *Can J. Chem.*, **52:** 1610–1615, 1974.

Strobel, D.F., "Nitric Oxide in the Upper Atmosphere," *J. Geophys. Res.*, **76:**2442–2452, 1971.

Strobel, D.F., "Minor Neutral Constituents in the Mesosphere and Lower Thermosphere," *Radio Sci.*, **7:**1–21, 1972.

Strobel, D.F., T.R. Young, R.R. Meier, T.P. Coffey, and A.W. Ali, "The Nighttime Ionosphere: E Region and Lower F Region," *J. Geophys. Res.*, **79:**3171–3178, 1974.

Swider, W., "Ionization Rates Due to the Attenuation of 1-100 A Nonflare Solar X Rays in the Terrestrial Atmosphere," *Rev. Geophys.*, **7:**573–599, 1969.

Swider, W., and R.S. Narcisi, "Auroral E-Region: Ion Composition and Nitric Oxide," *Planet. Space Sci.*, **25:**103–116, 1977.

Swider, W., T.J. Keneshea, and C.I. Foley, "An SPE-Disturbed D-Region Model," *Planet. Space Sci.*, **26:**883–892, 1978.

Swider, W., R.S. Narcisi, T.J. Keneshea, and J.C. Ulwick, "Electron Loss During a Nighttime PCA Event," *J. Geophys. Res.*, **76:**4691–4694, 1971.

Sze, N.D., M.K.W. Ko, R. Specht, and M. Livshits, "Modeling of Chemical Processes in the Troposphere and Stratosphere: Final Report," AFGL TR-80-0251, ADA092704, 1980.

Thomas, L., and M.R. Bowman, "The Diurnal Variations of Hydrogen and Oxygen Constituents in the Mesosphere and Lower Thermosphere," *J. Atmos. Terr. Phys.* **34**:1843, 1972.

Thomas, R.J., C.A. Barth, G.J. Rottman, D.W. Rusch, G.H. Mount, G.M. Lawrence, R.W. Sanders, G.E. Thomas, and L.E. Clemens, "Ozone Density Distribution in the Mesosphere (50–90 km) Measured by the SME Limb Scanning Near Infrared Spectrometer," *Geophys. Res. Lett.*, **10**:249–252, 1983.

Thompson, A.M., "The Effects of Clouds on Photolysis Rates and Ozone Formation in the Unpolluted Troposphere," *J. Geophys. Res.*, **89**:134–1349, 1984

Thrush, B.A., "Chemistry of the Stratosphere," *Philos. Trans. Roy. Soc. London*, **296**, 149–160, 1980.

Trinks, H. and K.H. Fricke, "Carbon Dioxide Concentrations in the Lower Thermosphere," *J. Geophys. Res.*, **84**: 3883, 1978.

Trost, T.F., "Electron Concentrations in the E and Upper D Region at Arecibo," *J. Geophys. Res.*, **84**: 2736–2742, 1979.

Twomey, S., B. Herman, and R. Rabinoff, "An Extension to the Chahine Method of Inverting the Radiative Transfer Equation," *J. Atmos. Sci.*, **34**:1085, 1977.

Tyson, B.J., J.F. Vedder, J.C. Arvesen, and R.B. Brewer, "Stratospheric Measurements of CF_2Cl_2 and N_2O," *Geophys. Res. Lett.*, **5**:369–372, 1978.

UNEP, *Report of the 3rd Session*, United Nations Environmental Programme, Coordinating Committee on the Ozone Layer, United Nations, New York, 1979.

UNEP, *Report on the 4th Session*; United Nations Environmental Programme, Coordinating Committee on the Ozone Layer, New York, 1980.

Vedder, J.F., B.J. Tyson, R.B. Brewer, C.A. Boitnott, and E.C.Y. Inn, "Lower Stratosphere Measurements of Variation with Latitude of CF_2Cl_2, $CFCl_3$, CCl_4, and N_2O Profiles in the Northern Hemisphere," *Geophys. Res. Lett.*, **5**:33–36, 1978.

Vedder, J.F., E.C.Y. Inn, B.J. Tyson, C.A. Boitnott, and D. O'Hara, "Measurements of CF_2Cl_2, $CFCl_3$, and N_2O in the Lower Stratosphere between 2°S and 73°N Latitude," *J. Geopays. Res.*, **86**:7363–7368, 1981.

Volz, A., U. Schmidt, J. Rudolph, D.H. Ehhalt, F.J. Johnson, and A. Khedim, "Vertical Profiles of Trace Gases at Mid-Latitudes," Jul-Report No. 1742, Kernforschungsanlage, Julich, Federal Republic of Germany, 1981.

von Zahn, U., "Neutral Air Density and Composition," *Physics and Chemistry of the Upper Atmosphere*, D. Reidel, Dordrecht, Holland, 1973.

Voss, H.D., and L.G. Smith, "Global Zones of Energetic Particle Precipitation," *J. Atmos. Terr. Phys.*, **42**:227–239, 1980.

Wang, P., M.P. McCormick, and W.P. Chu, "A Study on the Planetary Wave Transport of Ozone During the Late February 1979 Stratospheric Warming Using the SAGE Ozone Observation and Meteorological Information," *J. Atmos. Sci.*, **40**:2419–2431, 1983.

Waters, J.W., J.J. Gustincic, P.N. Swanson, and A.R. Kerr, "Measurements of Upper Atmospheric H_2O Emission at 183 GHz," *Atmospheric Water Vapor*, edited by A. Deepak, T.D. Wilderson, and L.H. Ruhnke, Academic Press, New York, 229–440, 1980.

Waters, J.W., J.C. Hardy, R.F. Jarnot, and H.M. Pickett, "Chlorine Monoxide Radical, Ozone, and Hydrogen Peroxide: Stratospheric Measurements by Microwave Limb Sounding," *Science*, **214**:61–64, 1981.

Weeks, L.H. R.E. Good, J.S. Randhawa and H. Trinks, "Ozone Measurements in the Stratosphere, Mesosphere, and Lower Thermosphere During Aladdin 74," *J. Geophys. Res.*, **83**:978–982, 1978.

Weinrab, M.P., W.A. Morgan, I-L. Chang, L.D. Johnson, F.A. Bridges, and A.C. Neverdoffer, "High Altitude Balloon Test of Satellite Solar Occultation Instrument for Monitoring Stratospheric O_3, H_2O, and HNO_3," *J. Atmos. Oceanic Tech.*, **1**:87–100, 1984.

Weinstock, E.M., M.J. Phillips, and J.G. Anderson, "*In situ* Observations of ClO in the Stratosphere: A review of Recent Results," *J. Geophys. Res.*, **86**:7273–7278, 1981.

Whitten, R.C. and S. Prasad (eds), *Ozone in the Free Atmosphere*, Von Nostrand Reinhold, New York, 1985.

Wilcox, R., "Total Ozone Trend Significance from Space and Time Variability of Daily Dobson Values," *J. Appl. Meteorol.*, **17**:1569–1591, 1978.

Wilcox, R., G. Nastrom, and A. Belmont, "Periodic Variations of Total Ozone and Its Vertical Distribution," *J. Appl. Meteorol.*, **16**:290–298, 1977.

WMO (World Meteorological Organization), *Assessment of Performance Characteristics of Various Ozone Observing Systems*, Report of the meeting of Experts, Boulder, July 1980, WMO Global Ozone Research and Monitoring Project Report No. 9, 1981.

WMO (World Meteorological Organization), *The Stratosphere 1981, Theory and Measurements*, WMO Global Ozone Research and Monitoring Project Report No. 11, NASA/GSFC, Greenbelt, Md., 1982.

Wofsy, S.C. and M.B. McElroy, "HO_x, NO_x and ClO_x: Their Role in Atmospheric Photochemistry," *Can. J. Chem.*, **52**:1582–1592, 1974.

Wuebbles, D.J., F.M. Luther, and J.E. Penner, "Effect of Coupled Anthropogenic Perturbations on Stratospheric Ozone," *J. Geophys. Res.*, **88**:1444–1456, 1983.

Zander, R., H. Leclercq, and L.D. Kaplan, "Concentration of Carbon Monoxide in the Upper Stratosphere," *Geophys. Res. Lett.*, **8**:365–368, 1981a.

Zimmerman, P.R., R.B. Chatfield, J. Fishman, P.J. Crutzen, P.L. Hanst, "Estimates of the Production of CO and H_2 from the Oxidation of Hydrocarbon Emissions from Vegetation," *Geophys. Res. Lett.*, **5**:679–682, 1978.

Chapter 22

ATMOSPHERIC EMISSION AND ABSORPTION OF ULTRAVIOLET RADIATION

R. E. Huffman

22.1 VUV/UV BACKGROUND ATMOSPHERIC RADIANCE

The radiance of the earth's atmosphere as seen from space is shown in Figure 22-1. The day and night radiance levels are shown for the earth-center or nadir viewing direction with the observer at an altitude of about 500 km or more. The units are Rayleigh per Å with a Rayleigh equal to 10^6 photon/(s cm^2 column 4 π steradian). The values in Figure 22-1 are based on a number of rocket and satellite observations. Only prominent spectral features are shown.

The day radiance values are shown in the upper part of Figure 22-1. The upper curve labeled day maximum gives observed radiance values for overhead sun (solar zenith angle of zero). The radiance is less for larger solar zenith angles. At wavelengths longer than about 2000 Å in the figure, solar scatter is the principal component. Representative references are Elliot et al. [1967], Barth and Mackey [1969], Huffman et al. [1980], and Rawcliffe and Elliot [1966]. Two broad minima are present, centered at about 2500 Å due to ozone and at about 1500 Å due to oxygen. In each case, the atmospheric constituent absorbs back-

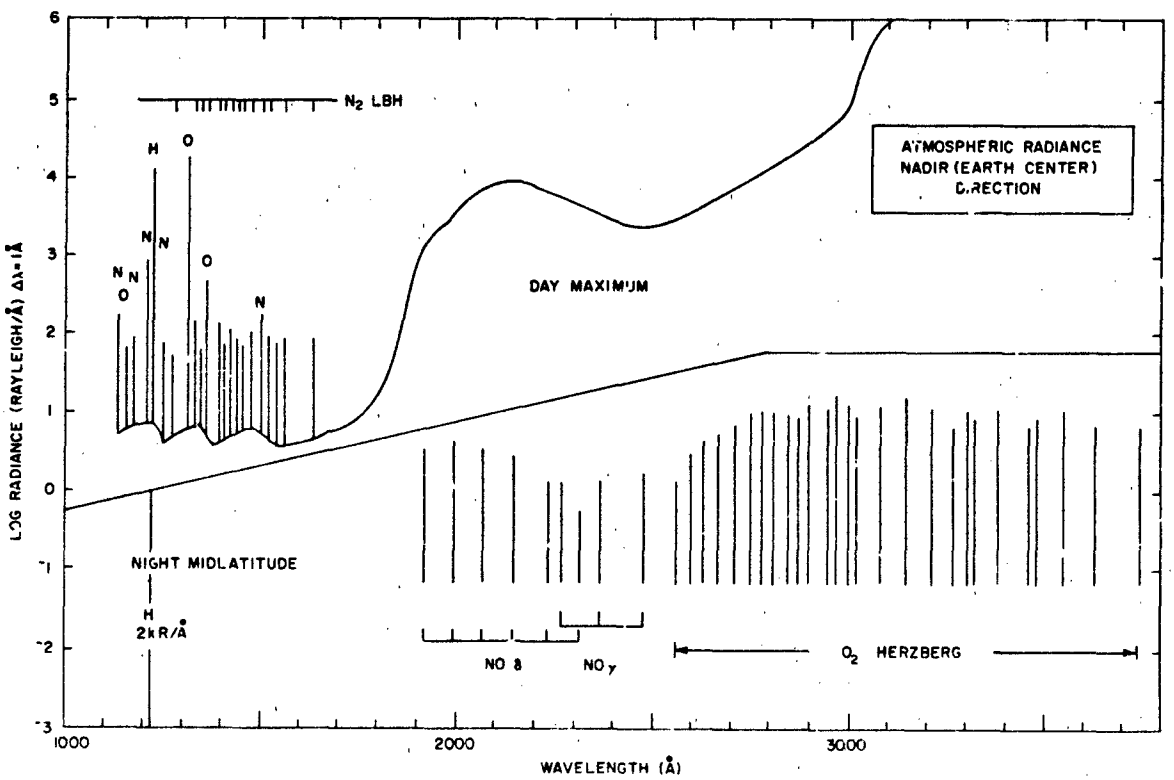


Figure 22-1. Atmospheric radiance in the ultraviolet.

cattered emission leading to a minimum in the background Heath, 1973].

The emission at wavelengths shorter than 2000 Å in the igure is due primarily to the dayglow. The Lyman-Birge-lopfield (LBH) bands of molecular nitrogen are the most rominent band system [Meier et al., 1980; Huffman et al., 980; Prinz and Meier, 1971; Takacs and Feldman, 1977]. trong multiplets of atomic oxygen occur at 1304 and 1356 Å, and there are several lines of atomic nitrogen, with the trongest at 1493 and 1200 Å. Hydrogen Lyman-alpha 1216Å) emission due to solar resonance scattering from he geocorona has been extensively measured for both day nd night situations [Carruthers et al., 1976; Meier and Mange, 1970, 1973; Thomas, 1963].

The night midlatitude radiance values in the lower part f Figure 22-1 are representative of observations found be-ween the tropical ultraviolet airglow belts and the auroral egion. The nitric oxide delta and gamma bands and the xygen Herzberg band emission are due to chemilumines-ence. References to these bands are as follows: Cohen-abban and Vuillemin [1973]; Huffman et al. [1980]; and eed and Chandra [1975]. The minimum value in a band entered at 1550 Å was measured by Huffman et al. [1980].

There are several localized sources of ultraviolet emis-ion from the atmosphere, as summarized in Table 22-1. he tropical ultraviolet airglow, due to radiative recombi-ation of oxygen ions, occurs mainly in the pre-midnight ector [Barth and Schaffner, 1970; Carruthers and Page, 976; and Hicks and Chubb, 1970]. The twilight spectrum s seen near the terminator when only upper altitudes are lluminated by solar ultraviolet, resulting in a low back-round from solar scatter [Gerard et al., 1970; Barth et al., 973]. The auroral region as seen from space in the ultra-iolet produces many emission features, as measured by eiting and Feldman [1978]; Chubb and Hicks [197]; Peek 1970]; and Huffman et al. [1980].

The extreme ultraviolet wavelength region from about 100 Å to the soft x-ray regior near 100 Å has not been as ell characterized as VUV/UV wavelengths. Emission is rimarily oxygen and nitrogen lines. Rocket observations available to date are by Christensen [1976], Gentieu et al. [1979], and Feldman et al. [1981].

22.2 ABSORPTION AND IONIZATION CROSS SECTIONS OF MAJOR GASES

A general introduction to the role of ultraviolet absorption and ionization in aeronomy is given by Watanabe [1958], although most of the cross-section values have been obtained at improved resolution in later work. The major gases discussed here are O_2, O_3, N_2, and O. Cross sections for use in aeronomy are available from several sources. Huffman [1969]; Hudson [1971]; Kirby et al. [1979]. These compilations generally concentrate on solar emission features.

A detailed review including the major atmospheric gases is available by Berkowitz [1979]. Complete bibliographies of photon cross-section references can be obtained through the Low Energy Atomic Collision Cross-Section Information Center, Joint Institute for Laboratory Astrophysics, University of Colorado, Boulder, Colorado 80309.

The transmission through a gas is given by $\exp[-\sigma N]$, where σ is the total absorption cross section in cm^2 and N is the column density in cm^{-2}. The photoionization cross section is the portion of σ that results in ionized products. Specific cross sections for producing specific ionized or neutral products can be defined in a similar manner. The cross section must be measured at a resolution considerably less than any structure in the spectrum. This difficulty occurs for narrow lines, and in these regions the simple exponential transmission calculation is not accurate.

Cross sections are given in Figure 22-2 for molecular oxygen (O_2) and ozone (O_3) and in Figure 22-3 for molecular nitrogen (N_2) and atomic oxygen (O). These figures give the location in wavelength of the photon absorption of most importance to aeronomic problems. In regions of complicated structure such as molecular bands or atomic lines, representative values are shown. In these regions, the detailed measurements must be utilized, as found through the

Table 22-1. Ultraviolet emission from localized sources

Source	Constituent	Nadir Intensity	Remarks
Aurora	O,I Lines N_2 Bands (NO?)	up to 10 kR (variable)	Arcs and diffuse aurora seen day and night in auroral oval
Tropical UV Airglow	O 1304,1356	up to 1 kR (variable)	Night, two bands ± 10° from dip equator.
Twilight	NO (Gamma bands) Mg	up to 1 kR up to 100R	Near terminator only

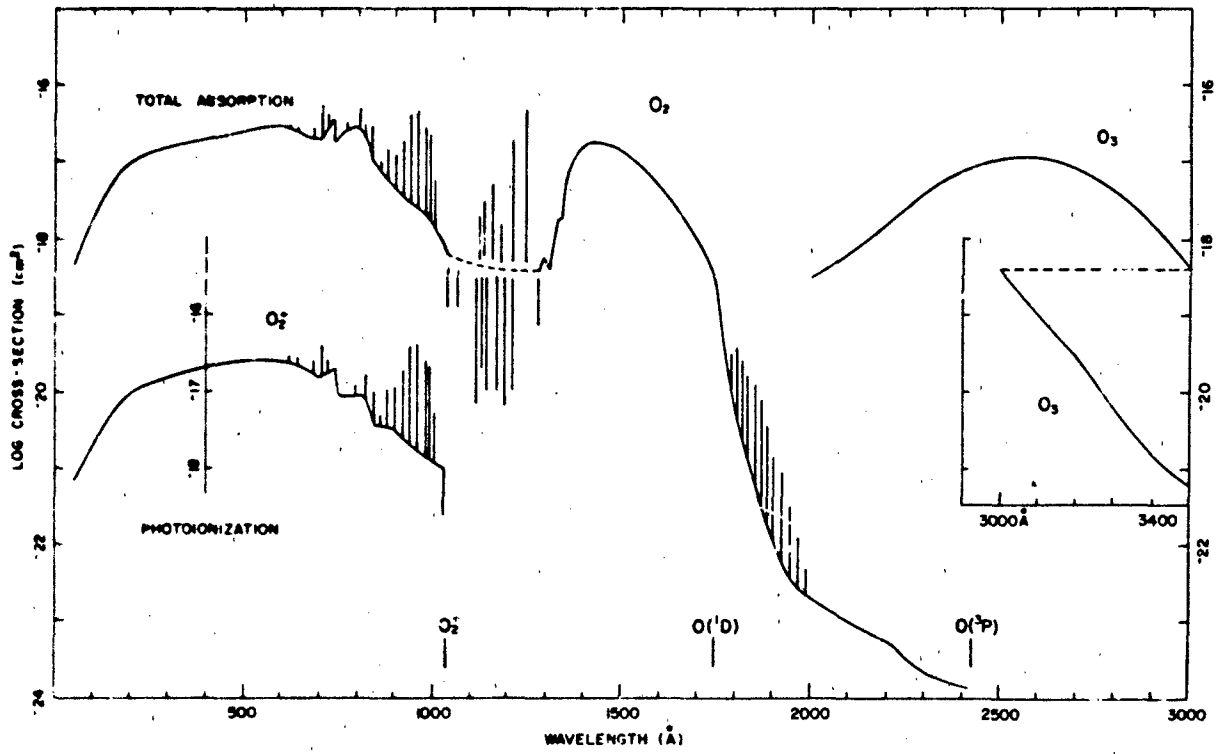


Figure 22-2 Molecular oxygen (O_2) and ozone (O_3) photon cross sections. Detailed structure and all molecular bands are not shown. See text for further discussion

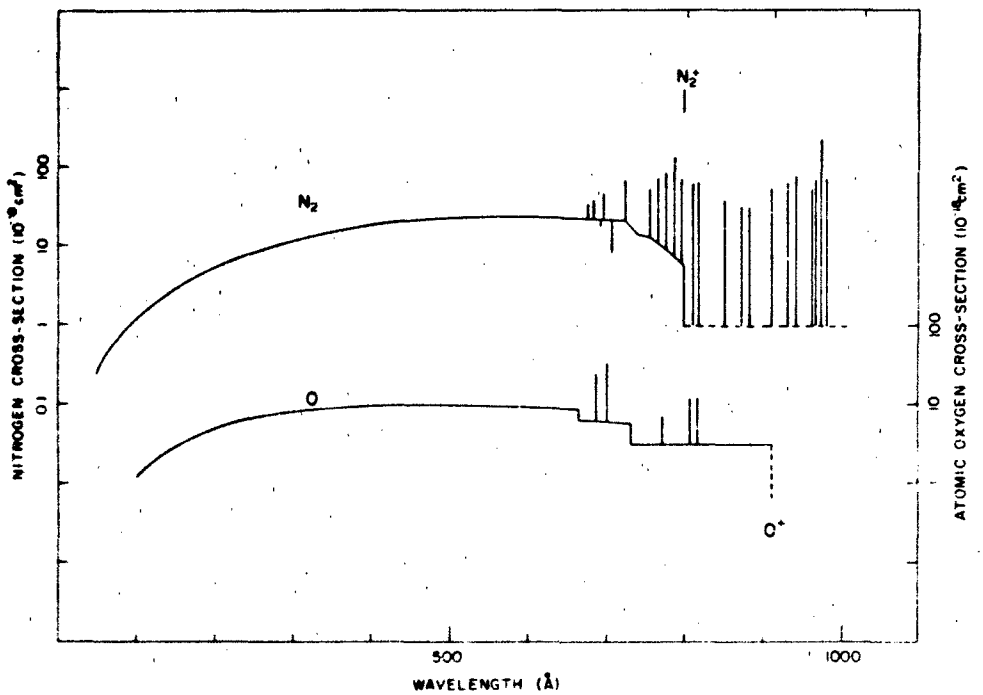


Figure 22-3 Molecular nitrogen (N_2) and atomic oxygen (O) photon cross sections. Detailed structure and all molecular bands and atomic lines are not shown. See text for further discussion

references. The curves given are based on measurements from many sources. Reference to measurements will be necessary for detailed cross sections, especially where discrete structure is indicated.

Molecular oxygen (O_2) absorption occurs weakly in the Herzberg continuum between about 2400 and 2050 Å. The Schumann-Runge bands are located between about 1750 and 2050 Å. These bands are composed of many narrow rotational lines. Transmission models have been developed [Blake, 1979]. The Schumann-Runge continuum occurs between about 1250 and 1750 Å. The region between 1027 and 1250 Å has a number of "windows," or regions of low absorption, as well as strong absorption bands. Nitric oxide is photoionized by the intense solar hydrogen Lyman-alpha line (1216 Å) at one of these windows with an ionization cross section of 2.0×10^{-18} cm^2.

The molecular oxygen ionization threshold occurs at 1027 Å; however, strong molecular bands continue to be observed to about 600 Å. The photoionization cross section is also shown as it is appreciably different from the total cross section to about 500 Å. The bands may show autoionization structure. Toward the shorter wavelengths of the soft x-ray region, the cross section decreases.

The principal ozone (O_3) absorption is centered near 2500 Å. Weaker absorption continues to about 3500 Å toward longer wavelengths. Bands are not shown in the figure.

Molecular nitrogen (N_2) begins to absorb strongly at about 1000 Å. Only a few of the numerous bands are shown. A continuum underlying these bands must be weak and an upper limit is shown. The ionization threshold is at 796 Å. A separate photoionization curve is not shown as it is approximately the same as the total absorption curve. Autoionized molecular bands are observed from about 650 to 796 Å.

Atomic oxygen (O) is photoionized at wavelengths shorter than 911 Å. At longer wavelengths absorption is confined to line series converging to the ionization threshold. At wavelengths shorter than 911 Å, there are autoionized series converging to more energetic ionization thresholds. A few of the more intense members are shown.

22.3 RATE OF PHOTODISSOCIATION IN THE ATMOSPHERE

Photodissociation resulting in neutral products occurs in the atmosphere due to the absorption of solar radiation by molecules. The photon absorption process resulting in the formation of ions and electrons is called photoionization. It is covered in Chapter 21 of this handbook. In general, photodissociation into neutral products requires less energy than photoionization. It occurs at longer wavelengths and lower altitudes. At many wavelengths and altitudes, both photodissociation into neutral products and photoionization are occurring. The internal and kinetic energy states of the neutral products are very important in determining reaction mechanisms in the atmosphere and where information is available it is included.

This section is an introduction to the most important photodissociation processes and their role in the atmosphere. It is based on Turco and Huffman [1979] and Turco [1975].

The wavelengths of solar radiation of interest in photodissociation are primarily from the strong hydrogen Lyman-alpha line at 1216 Å through the ultraviolet to about 7500 Å. The solar flux available for photodissociation in the ultraviolet at wavelengths less than about 3000 Å is controlled by absorption of molecular oxygen and ozone. This absorption acts as a shutter to decrease the solar flux as a function of altitude and determine the rate of photodissociation of species present at smaller concentrations.

The solar atmospherically attenuated irradiance as a function of solar zenith angle is shown in Figures 22-4, 22-5 and 22-6. Absorption between 3300 and 2000 Å is largely by ozone. Between 2000 Å and 1216 Å, molecular oxygen is the important absorbing molecule. Three representative solar zenith angles of 75°, 60°, and 30° are given in the figures. The oxygen vertical profile is the mean reference profile of Champion [1972]. The ozone profile used is a midlatitude empirical model due to Krueger and Minzer [1976]. The solar flux at the top of the atmosphere used in Figures 22-4, 22-5, and 22-6 is the same as given in Chapter 2 by Heroux and Hinteregger (Figures 2-2, 2-3).

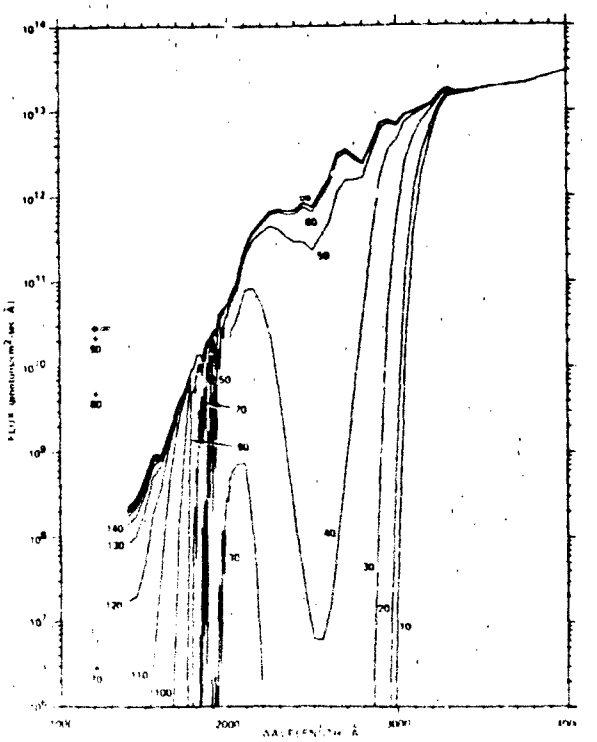


Figure 22-4. Atmospherically attenuated solar irradiance for solar zenith angle of 75° for various altitudes (km).

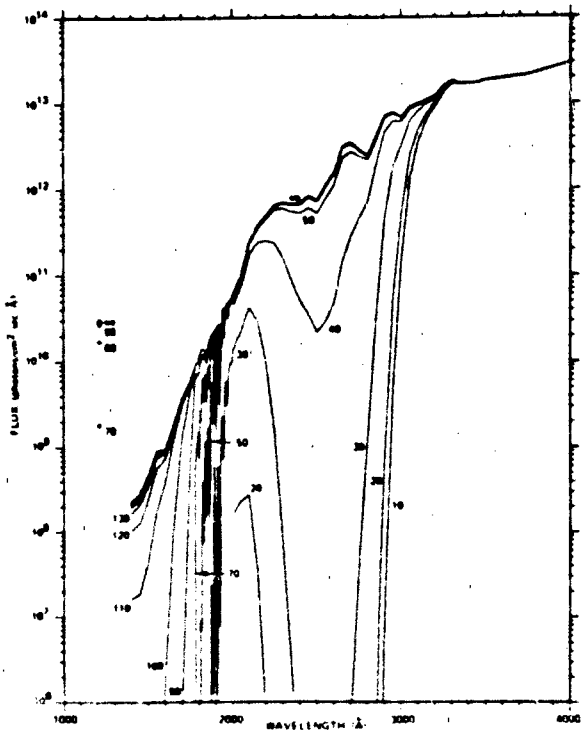


Figure 22-5 Atmospherically attenuated solar irradiance for solar zenith angle of 60° for various altitudes (km).

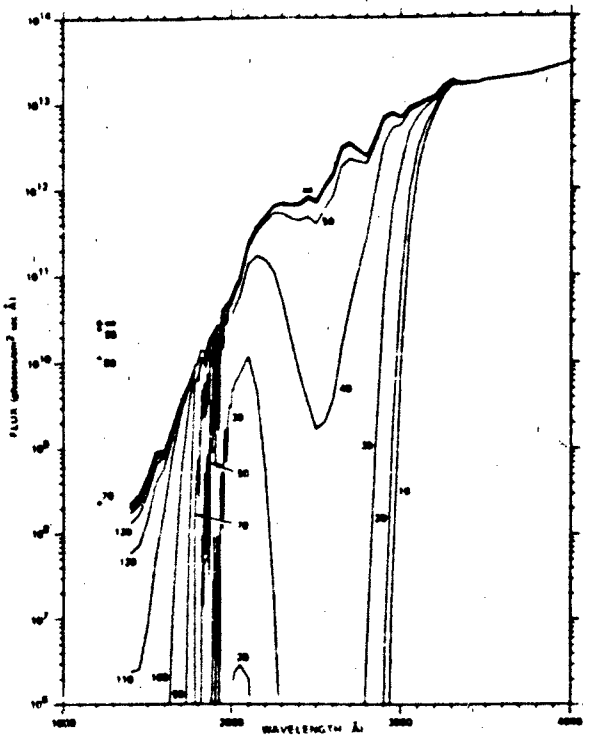


Figure 22-6 Atmospherically attenuated solar irradiance for solar zenith angle of 30° for various altitudes (km).

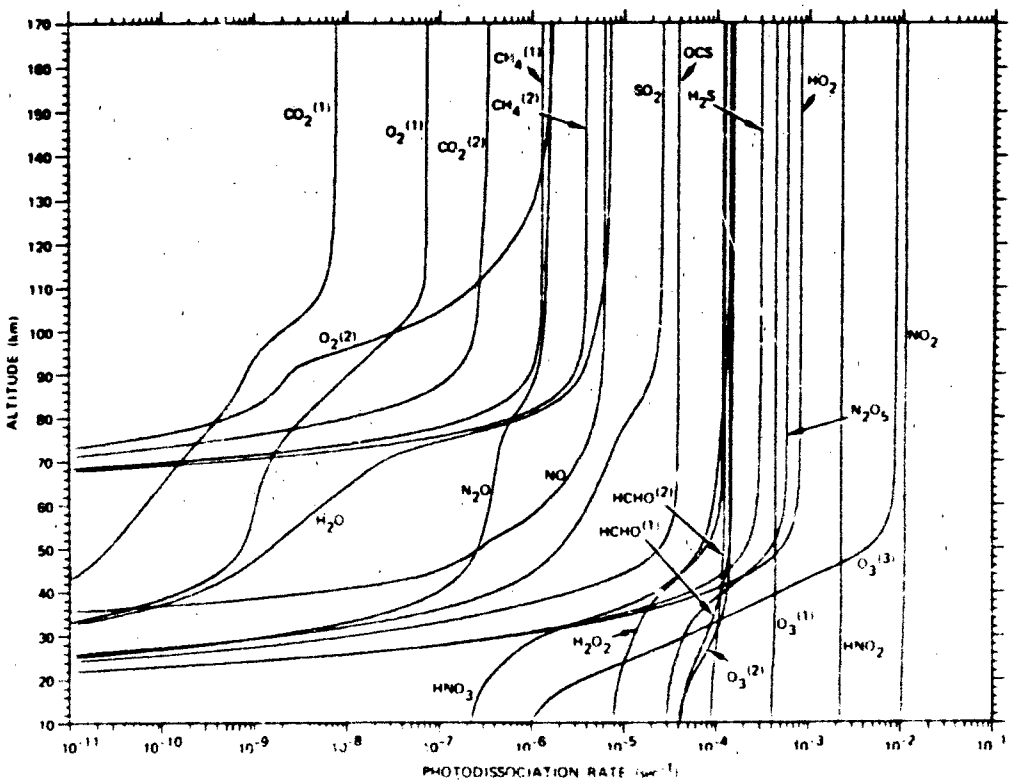


Figure 22-7 Photodissociation rates for solar zenith angle of 75°.

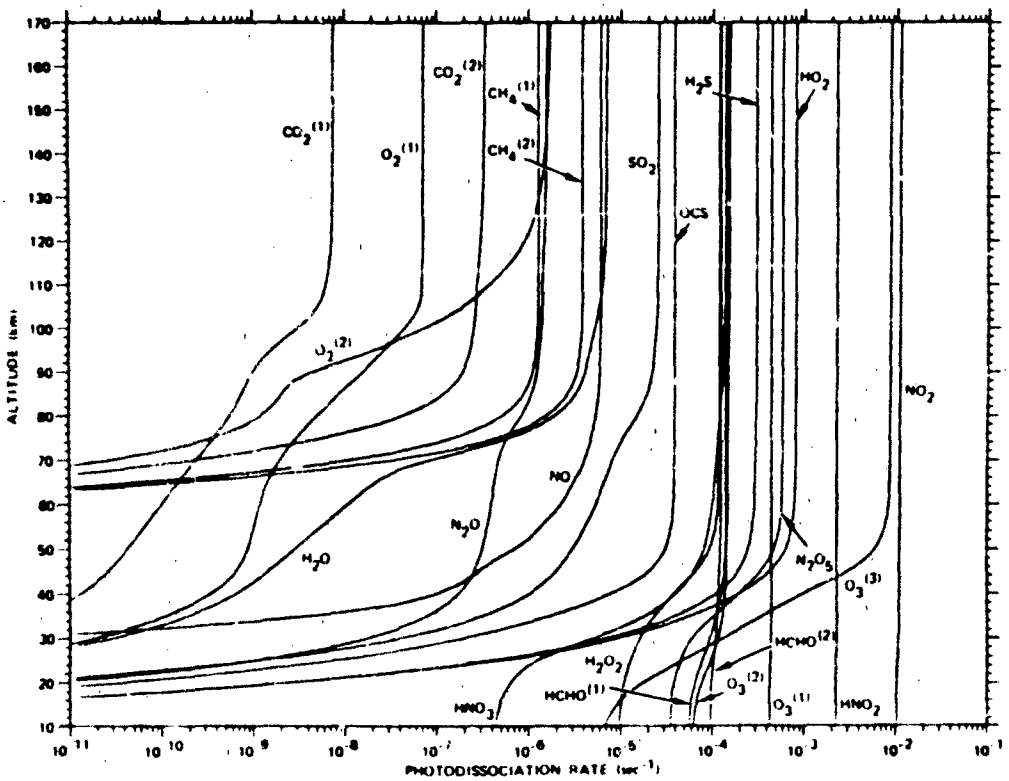


Figure 22-8 Photodissociation rates for solar zenith angle of 60°.

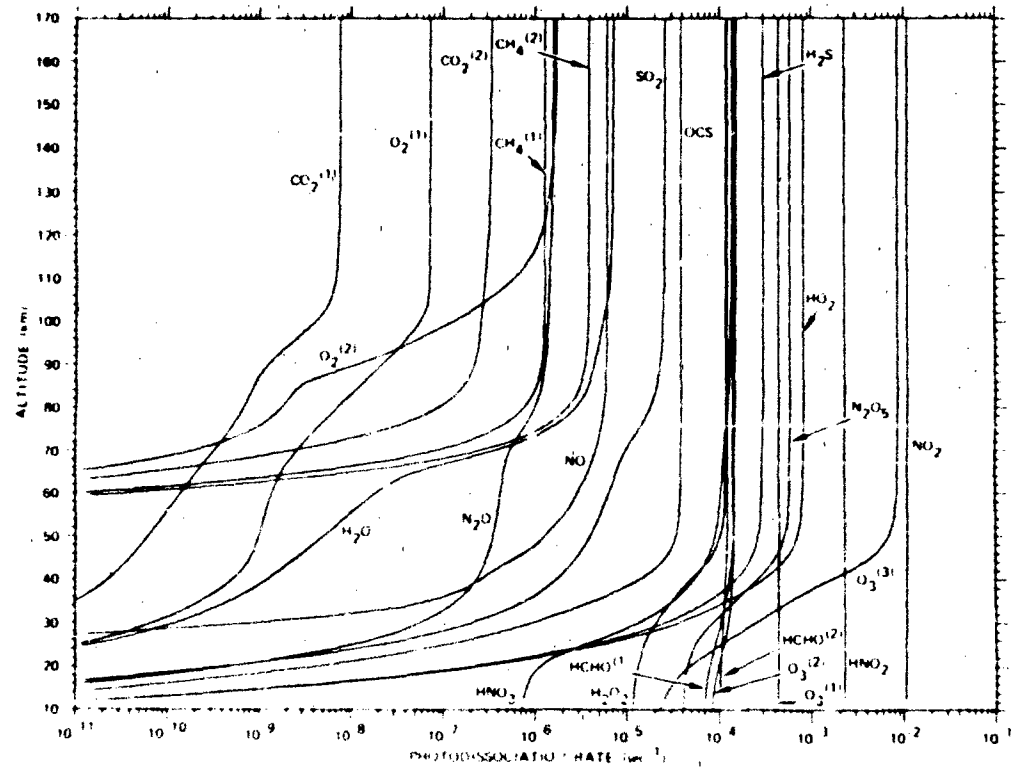


Figure 22-9 Photodissociation rates for solar zenith angle of 80.

Table 22-2 Molecular photodissociation processes. Product is ground state unless indicated otherwise

Designation in Figures 22-7, 8 & 9	Photodissociation Process
$O_2(1)$	$O_2 \rightarrow O + O$
$O_2(2)$	$O_2 \rightarrow O + O(^1D)$
$O_3(1)$	$O_3 \rightarrow O + O_2$
$O_3(2)$	$O_3 \rightarrow O + O_2(^1\Delta_g)$
NO	$NO \rightarrow N + O$
NO_2	$NO_2 \rightarrow NO + O$
N_2O	$N_2O \rightarrow N_2 + O(^1D)$
N_2O_5	$N_2O_5 \rightarrow 2NO_2 + O$
HNO_2	$HNO_2 \rightarrow OH + NO$
HNO_3	$HNO_3 \rightarrow OH + NO_2$
HO_2	$HO_2 \rightarrow OH + O$
H_2O	$H_2O \rightarrow OH + H$
H_2O_2	$H_2O_2 \rightarrow OH + OH$
$CO_2(1)$	$CO_2 \rightarrow CO + O$
$CO_2(2)$	$CO_2 \rightarrow CO + O(^1D)$
$CH_4(1)$	$CH_4 \rightarrow CH_3 + H$
$CH_4(2)$	$CH_4 \rightarrow CH_2 + H_2$
HCHO(1)	$HCHO \rightarrow CHO + H$
HCHO(2)	$HCHO \rightarrow CO + H_2$
SO_2	$SO_2 \rightarrow SO + O$
OCS	$OCS \rightarrow CO + S(^1D)$
H_2S	$H_2S \rightarrow HS + H$

The number of photodissociation events occurring per unit volume per second at a given point in the atmosphere is equal to the local number density multiplied by the photodissociation rate per molecule per second. The latter factor is the integral of the absorption cross section times the solar flux at a given altitude over all significant wavelengths. If known, the cross sections for specific products can also be used to give production rates for specific energy states.

Photodissociation rates are shown in Figures 22-7, 22-8 and 22-9 for the three solar zenith angles of 75°, 60°, and 30°. A total of twenty-three photodissociation processes are shown for seventeen absorbing molecules. The molecules included are O_2, O_3, NO, NO_2, N_2O, N_2O_5, HO_2, H_2O_2, HNO_2, HNO_3, CO_2, CH_4, HCHO, SO_2, H_2S, and OCS. The absorption cross sections for O_2 and O_3 are given in Figure 22-2. References to the remaining cross sections are in Turco and Huffman [1979] and Turco [1975].

The total photodissociation rate per unit volume as a function of altitude can be obtained from Figures 22-7, 22-8, and 22-9 and the number density of the molecule of interest. These densities will vary depending on the problem under consideration. Some of the molecules are included because they are important in the atmospheric chemistry associated with nitrogen oxides, water vapor, and/or other gases released in the atmosphere at greater than naturally occurring levels.

Table 22-2 gives the specific photodissociation reactions that apply to the curves in Figures 22-7, 22-8, and 22-9. In many cases, the excited electronic state of the product is known. The internal energy in the products can be very important in determining subsequent reactions.

CHAPTER 22

REFERENCES

Barth, C.A. and E.F. Mackey, "OGO-4 Ultraviolet Airglow Spectrometer," *IEEE Transactions on Geoscience Electronics*, **GE-7** (2): 114–119, 1969.

Barth, C.A. and S. Schaffner, "OGO-4 Spectrometer Measurements of the Tropical Ultraviolet Airglow," *J. Geophys. Res.*, **75:** 4299–4306, 1970.

Barth, C.A., D.W. Rusch, and A.I. Stewart, "The UV Nitric Oxide Experiment for Atmosphere Explorer," *Radio Sci.*, **8:** 377, 1973.

Beiting, E.J. and P.D. Feldman, "A Search for Nitric Oxide Gamma Band Emission in an Aurora," *Geophys. Res. Lett.*, **5:** 51, 1978.

Berkowitz, J., *Photoabsorption, Photoionization and Photoelectron Spectroscopy*, Academic Press, New York, 1979.

Blake, A.J., "An Atmospheric Absorption Model for the Schumann-Runge Bands of Oxygen," *J. Geophys. Res.*, **84:** 3272, 1979.

Carruthers, G.R. and T. Page, "Apollo 16 Far Ultraviolet Imagery of the Polar Auroras, Tropical Airglow Belts, and General Airglow," *J. Geophys. Res.*, **81:** 483–496, 1976a.

Carruthers, G.E., T. Page, and R.R. Meier, "Apollo 16 Lyman-alpha Imagery of the Hydrogen Geocorona," *J. Geophys. Res.*, **81:** 1164–1672, 1976b.

Champion, K.S.W., "The Natural Atmosphere: Atmospheric Structure," *Defense Nuclear Agency Reaction Rate Handbook*, DNA 1948H, edited by M.H. Borner and T. Bauer, Chapter 2, 1972.

Christensen, A.B., "A Rocket Measurement of the Extreme Ultraviolet Dayglow," *Geophys. Res. Lett.*, **3:** 221, 1976.

Chubb, T.A. and G.T. Hicks, "Observations of the Aurora in the Far Ultraviolet from OGO-4," *J. Geophys. Res.*, **75:** 1290–1311, 1970.

Cohen-Sabban, J. and A. Vuillemin, "Ultraviolet Nightglow Spectrum from 1900 A and 3400 A," *Astrophys. Space Sci.*, **24:** 127–132, 1973.

Elliot, D.D., M.A. Clark and R.D. Hudson, "Latitude Distribution of the Daytime Ozone Profile above the Peak," Aerospace Report No. TR-0158 (3260-10)-2, 1967.

Feldman, P.D., D.E. Anderson, R.R. Meier, and E.P. Gentieu, "The Ultraviolet Dayglow 4. The Spectrum and Excitation of Singly Ionized Oxygen," *J. Geophys. Res.*, **86:** 3583, 1981.

Gentieu, E.P., P.D. Feldman, and R.R. Meier, "Spectroscopy of the Extreme Ultraviolet Dayglow at 6.5 A Resolution: Atomic and Ionic Emissions Between 530 and 1240 A," *Geophys. Res. Lett.*, **6:** 325, 1979.

Gerard, J.-C., D.W. Rusch, P.B. Hays, and C.L. Fesen, "The Morphology of Equatorial Mg^+ Ion Distribution Deduced from 2800 Å Airglow Observations," *J. Geophys. Res.*, **84:** 5249, 1979.

Heath, D.F., C.L. Mateer, and A.J. Krueger, "The Nimbus-4 Backscatter Ultraviolet (BUV) Atmospheric Ozone Experiment – Two Years Operation," *Pure Appl. Geophys.*, **106–108:** 1238–1253, 1973.

Hicks, G.T. and T.A. Chubb, "Equatorial Auroral Airglow in the Far Ultraviolet," *J. Geophys. Res.*, **75:** 6233–6248, 1970.

Hudson, R.D., "Critical Review of Ultraviolet Photoabsorption Cross-Sections for Molecules of Astrophysical and Aeronomic Interest," *Rev. Geophys. Space Phys.*, **9:** 305, 1971.

Huffman, R.E., "Absorption Cross-Sections of Atmospheric Gases for use in Aeronomy," *Canad. J. Chem.*, **47:** 1823, 1969.

Huffman, R.E., F.J. LeBlanc, J.C. Larrabee, and D.E. Paulsen, "Satellite Vacuum Ultraviolet Airglow and Auroral Observations," *J. Geophys. Res.*, **85:** 2201–15, 1980.

Kirby, K., E.R. Constantinides, S. Babeu, M. Oppenheimer, and G.A. Victor, "Photoionization and Photoabsorption Cross Sections of He, O, N_2 and O_2 for Aeronomic Calculations," *At. Data Nucl. Tables*, **23:** 63, 1979.

Krueger, A.J. and R.A. Minzner, "A Mid-Latitude Ozone Model for the 1976 U.S. Standard Atmosphere," *J. Geophys. Res.*, **81:** 4477, 1976.

Meier, R.R. and P. Mange, "Geocoronal Hydrogen: An Analysis of the Lyman Alpha Airglow Observed from OGO-4," *Planet. Space Sci.*, **18:** 803, 1970.

Meier, R.R. and P. Mange, "Spatial and Temporal Variations of the Lyman Alpha Airglow and Related Atomic Hydrogen Distributions," *Planet. Space Sci.*, **21:** 309–327, 1973.

Meier, R.R., D.J. Strickland, P.D. Feldman, and E.P. Gentieu, "The Ultraviolet Dayglow: I. Far UV Emissions of N and N_2," *J. Geophys. Res.*, **85:** 2177, 1980.

Peek, H.M., "Vacuum Ultraviolet Emission from Auroras," *J. Geophys. Res.*, **75:** 6209–6217, 1970.

Prinz, D.K. and R.R. Meier, "OGO-4 Observations of the Lyman-Birge-Hopfield Emission in the Day Airglow," *J. Geophys. Res.*, **76:** 6146–6158, 1971.

Rawcliffe, R.D. and D.D. Elliott, "Latitude Distributions of Ozone at High Altitudes Deduced from Satellite Measurement of the Earth's Radiance at 2840 Å," *J. Geophys. Res.*, **71:** 5077–5089, 1966.

Reed, E.I. and S. Chandra, "The Global Characteristics of Atmospheric Emissions in the Lower Thermosphere and Their Aeronomic Implications," *J. Geophys. Res.*, **80:** 3053–3062, 1975.

Takacs, P.Z. and P.D. Feldman, "Far Ultraviolet Atomic and Molecular Nitrogen Emission in the Dayglow," *J. Geophys. Res.*, **82:** 5011–5023, 1977.

Thomas, G.E., "Lyman-alpha Scattering in the Earth's Hydrogen Geocorona," *J. Geophys. Res.*, **68:** 2639, 1963.

Turco, R.P., "Photodissociation Rates in the Atmosphere Below 100 km," *Geophys. Surveys*, **2:** 153–192, 1975.

Turco, R.P. and R.E. Huffman, "Solar Photodissociation in the Atmosphere," *Defense Nuclear Agency Reaction Rate Handbook*, DNA 1948H, edited by M.H. Borner and T. Bauer, Chapter 13B, 1979.

Watanabe, K., "Ultraviolet Absorption Processes in the Upper Atmosphere," *Adv. Geophys.*, **5:** 154, 1958.

Chapter 23

GEOKINETICS

Sections 23.1, 23.2, and 23.3 J.C. Johnston and J.C. Battis
Section 23.4 J.J. Ciper and G.H. Cabaniss
Section 23.5 J.A. Shearer and J.M. Novak Capt. USAF

The surface of the earth is in constant motion in response to natural (gravitational, rotational, hydrological, meteorological, tectonic) and manmade (detonations, surface loading, vibrating machinery) forces. Some of these motions are of great interest to earth scientists and engineers, but many compose the background noise spectrum from which the signals of interest must be extracted.

The nature of the motions range from transient vibrations caused by earthquakes through continuous long period oscillations. There are also non-periodic rotations of large blocks of the outer shell of the earth. Due to the small amplitude or low frequency of these motions, most go unnoticed in the common experience. The most catastrophic of these motions are the transient waves generated by large earthquakes or explosions. In this chapter the general characteristics, magnitude, and measurement of crustal motions will be discussed.

23.1 STRUCTURE OF THE EARTH

The mechanical and chemical properties of the earth, such as density, rigidity and compressibility, are strong functions of depth and therefore pressure [Anderson and Hart, 1976]. Four major discontinuities in these properties have been identified and on this basis the interior of the earth has been subdivided into four regions that on a gross scale are laterally homogeneous (Figure 23-1). From the surface downward these regions are as follows:

1. Crust—This region ranges in thickness from 5 km beneath oceans to 35 km beneath continents, to a maximum of 60 km beneath the Himalaya Mountains. The crust, which is composed of silicates, exhibits radial and lateral homogeneity to a much greater extent than any of the other three zones. The lower boundary of the crust is defined by a sharp seismic velocity discontinuity, known as the Mohorovicic discontinuity, commonly called the "Moho." Typical upper crustal rock density is 2.5 g/cm^3. The variation in depth to the Moho is thought to be an expression of isostatic equilibrium since the crustal material is of lower density than that of the mantle. In other words, the so called "crustal roots" beneath continents or mountain ranges compensate for the additional material in such a way that the pressure from a vertical column of mass at some depth anywhere around the planet will be equal (Figure 23-2). This process is not instantaneous and it can be demonstrated that some areas of the world have not reached isostatic equilibrium.
2. Mantle—Starting at the base of the crust, the mantle extends to a depth of 2885 km. The upper

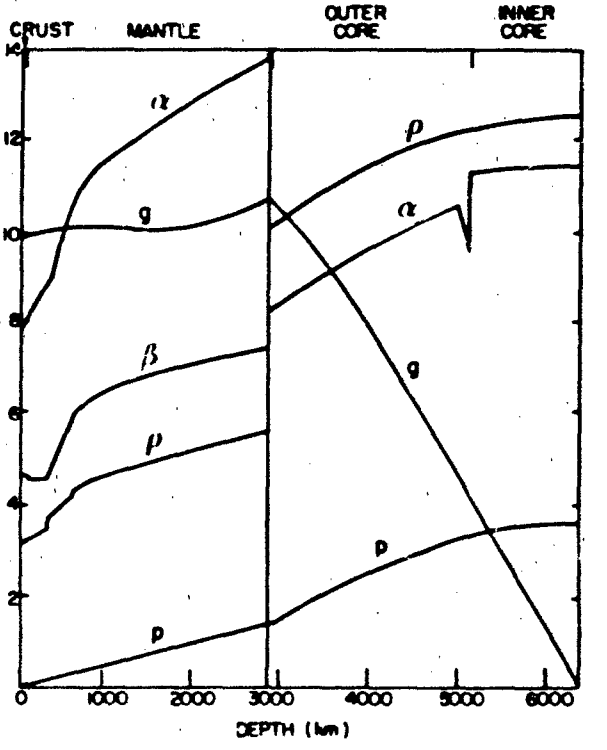


Figure 23-1 Variation of physical properties with depth in the earth (P, pressure in 10^{11} dyn/cm^2; ρ, density in gm/cm^3; α, β, P and S wave velocities in km/s; g, acceleration of gravity in 10^{3}cm/s^2) [Anderson and Hart, 1976]

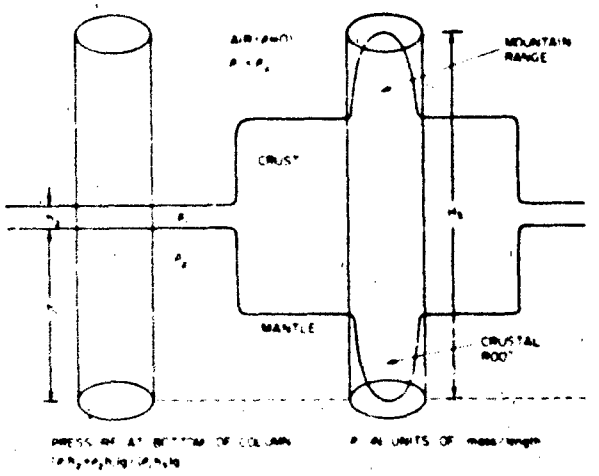


Figure 23.2 Isostacy principle—crustal roots under continents and mountain ranges permit a spherical shell of equal pressure to exist at some depth beneath the surface of the earth

part of the mantle is probably composed mainly of peridotites (olivines and pyroxines). The term "pyrolite" has been applied to the composition of the upper mantle [Clark and Ringwood, 1964]. The composition of the lower mantle is uncertain. It may consist of oxides and silicates of magnesium and iron with small amounts of aluminum and calcium [Verhoogen et al. 1970]. Typical mantle densities range from 3 to 5 g/cm^3.

3. Outer core—This region ranges from a depth of 2885 to 5155 km. It is believed to be composed of liquid iron and nickel and therefore cannot support shear forces. The density of the outer core ranges from approximately 9 g/cm^3 to 12 g/cm^3.
4. Inner core—Beginning at a depth of 5155 km, the inner core extends to the center (6371 km depth) of the earth. Like the outer core, it is believed to consist of iron and nickel; however, as a result of the extreme pressure it is solid. The density of the inner core material is approximately 11 g/cm^3.

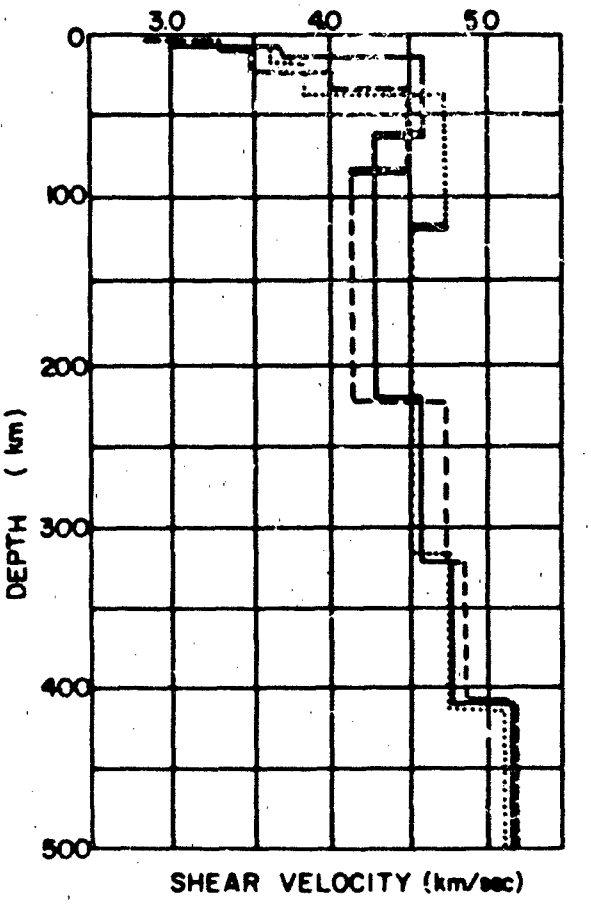
In geophysical literature there is another radial division of the earth into the lithosphere and the asthenosphere. This is primarily a division based on strength, and the boundary does not correspond to any of the ones previously discussed. The next section, which is an overview of the theory of plate tectonics, is concerned with this division.

23.2 PLATE TECTONICS

The development of the theory of plate tectonics represents a revolution in the understanding of the evolution of the surface of the earth. Because the theory is referred to in several of the following sections, a brief survey of the subject will be presented here. More comprehensive discussions can be found in geological and geophysical texts published since 1970 (for example, Garland [1971] and Wyllie [1971]). There are also several collections of scientific papers (such as Pird and Isacks [1972] and Cox [1973]) that illustrate the chronological development of the theory.

Plate tectonics is primarily concerned with the kinematics of the first several hundred kilometers below the surface of the earth. This region is divided into two zones: the lithosphere, from the Greek word meaning "rock", and the asthenosphere, from the Greek word meaning "weak" or lacking strength. The lithosphere exhibits finite strength and is capable of brittle fracture, while the asthenosphere can be viewed as a basement of weak material having low mechanical strength. The lithosphere includes the crust, but also extends into the upper mantle. The boundary between these two regions is associated with the top of a seismic low velocity zone (LVZ) which typically occurs at approximately 100 km depth. See Dorman [1969] and Figure

Figure 23.3 Crust and upper mantle velocity distributions (solid line—Pacific Ocean, dash line—Alpine, dotted line—Canadian Shield). Thick horizontal bars indicate top of low velocity zone [Dorman, 1969]



23-3. The depth of this interface varies with tectonic province, that is, whether it is beneath oceans, continents, or mountains. It should be noted that the lithosphere-asthenosphere boundary does not correspond to the classical crust-mantle velocity discontinuity, the Mohо.

Major fracture systems occur in the lithosphere, dividing the outer shell of the earth into a number of sections or plates (Figure 23-4). Twelve major plates and numerous minor plates capable of motion along the surface of the earth have been identified. The interior of the plates can be considered essentially rigid and the effect of the interactions of the moving plates confined primarily to the plate margins. Plates can be composed of either oceanic floor, continental material, or both. The major distinction between the theory of plate tectonics and the old concept of continental drift is that continents are not separate mobile entities, but are part of the plates. For example, as seen in Figure 23-4, the North American plate includes a large portion of the Atlantic Ocean as well as the continent of North America.

Three basic types of plate boundaries have been identified. The first type is an extensional margin where new oceanic lithosphere is created. This is referred to as sea-floor spreading and occurs when two adjacent plates are moving away from each other. Here basaltic magma from the asthenosphere is intruded at the margin (Figure 23-5a). Normal faulting, an expression of extension with strike (direction) parallel to the margin axis, is found here. In these areas, the lithosphere is very thin, elevated, and has high heat flow. Examples of this type of boundary are the East Pacific Rise and the Mid-Atlantic Ridge. Since the area of the surface of the earth must be conserved, it is obvious that there must be a mechanism for destroying lithosphere. This type of margin is called a consuming plate margin or a subduction zone. Subduction occurs when oceanic lithosphere converges with either oceanic or continental lithosphere (Figure 23-5b). The oceanic plate will be overridden by the other plate and sink because it is colder, older, and thus denser than its surroundings. At the point of contact, the subducting plate will down-warp the overriding plate forming a trench. As the subducting plate enters the asthenosphere it is heated until it melts. A chain of volcanoes is formed at the rear of the trench system.

Typical examples of this type of plate boundary are the Japanese island arc system and the Aleutian Islands. Subduction zones exhibit marked gravity anomalies because formation of a trench creates a mass deficiency. These zones are often the sites of the world's largest (fault lengths 100-800 km) and deepest (to depths of 700 km) earthquakes. This zone of earthquakes, not to be confused with the actual slab, is known as the Benioff zone. Various geometries of subducted lithosphere and subduction rates have been identified. Generally, we expect that the faster the subduction rate, the deeper the plate can descend before melting; however, this is also dependent on the time constant for assimilation into the athenosphere. The dip angle of the slabs varies and in some cases fragments of plate seem to have broken off (Figure 23-6). The type of faulting varies with the position on the slab. Because the plate has some flexural

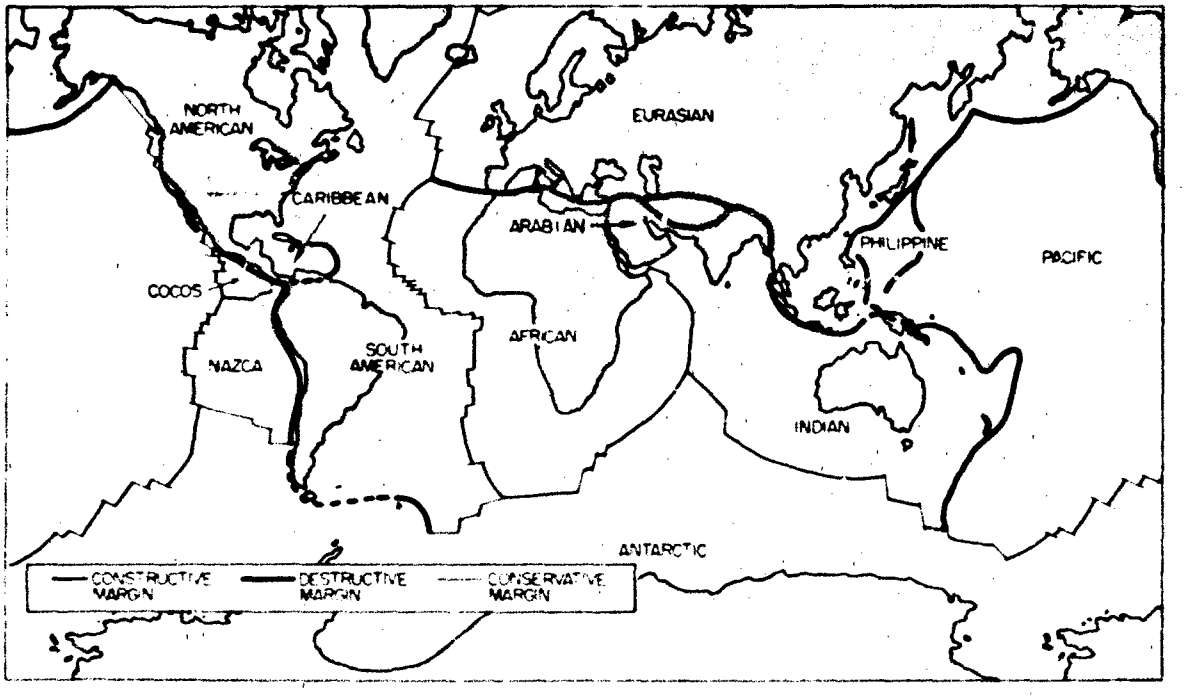


Figure 23-4. Major plate boundaries of the world.

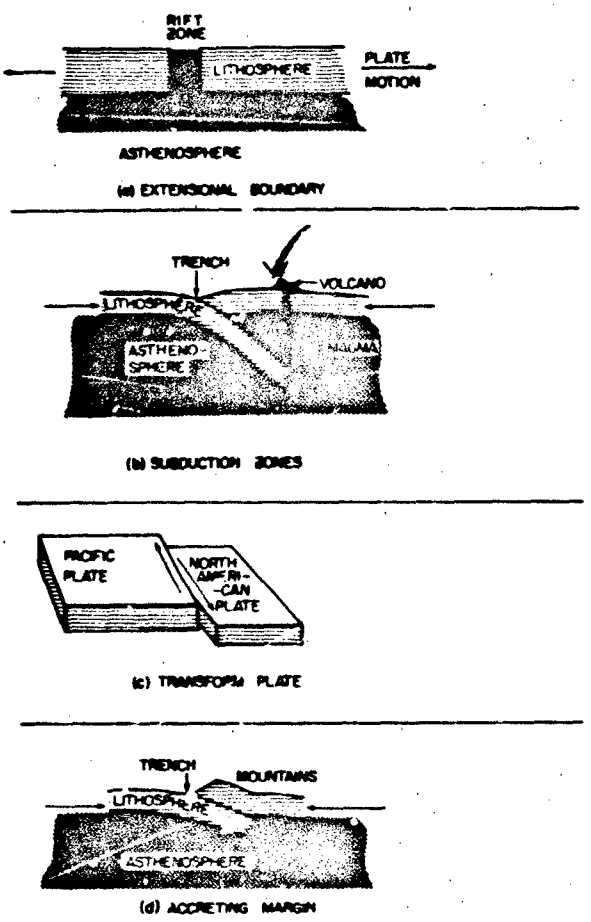


Figure 23-5. Plate boundary types.

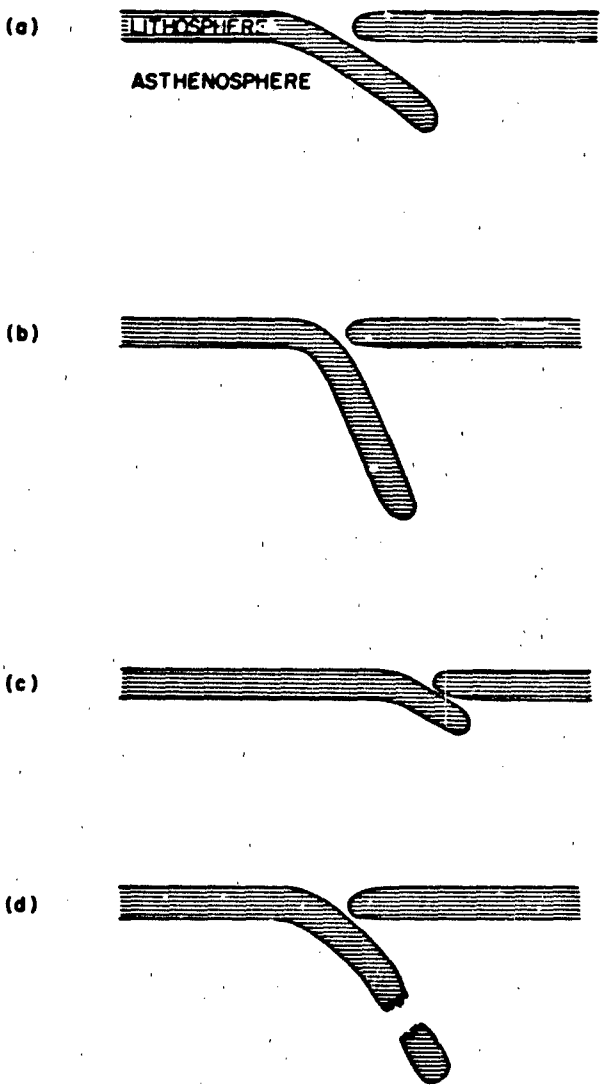


Figure 23-6. Various geometries of subducting slabs
(a) shallow angle
(b) steep angle
(c) short slab
(d) broken slab

rigidity, the downgoing slab actually warps upward about the radius of curvature. Tensional faulting is observed in this region. Compressional thrust faulting is found further down the subducted slab. In some areas of the world, parts of the downgoing slab appear to be breaking off and are in tension (axis of least compressive stress is parallel to the slab and dipping down). This could be caused by gravitational sinking [Isacks et al., 1968]. In some cases of subduction of oceanic lithosphere beneath continental lithosphere, slivers of ocean crust have been scraped off onto the continent. These fragments are called "ophiolites" [Gass, 1982].

While the combination of the spreading mid-ocean ridge systems with the subduction zones creates a recycling scheme for the surface of the earth, a third type of boundary neither creates nor destroys lithosphere. If two plates do not converge but slide past each other, the plate margin is called a transform or strike-slip fault. See Figure 23-5c and Wilson [1965]. The San Andreas Fault in California is the best known example of a transform fault. In this case, the only resistance to the relative motion results from irregularities in the plate boundaries and friction between the two plates. In this case, the Pacific plate is rotating in a counterclockwise direction relative to the North American Plate.

It should be noted that the recycling scheme mentioned above refers only to oceanic lithosphere. Continental lithosphere, which is buoyant and thick, generally does not subduct. As a consequence, the oldest oceanic lithosphere (about 200 million years) is younger than most continental lithosphere. When the collision involves continental lithosphere of both plates, the buoyant forces of the overridden plate produce uplift of the overriding plate resulting in mountain building forces (Figure 23-5d). The Tibetan Plateau, site of the convergence of the India and Eurasian

plates, is an example of this type of plate margin. The plate approximation becomes less useful for continent to continent collisions and the plastic properties of the lithosphere must be considered. Since continental lithosphere cannot be destroyed, the oldest [illegible] in the world are found on the stable, inactive continental shields such as Canada and Africa. Island arcs, which are not part of the continental lithosphere, are usually "scraped off" in slivers by the overriding plate when the plate subducts. This is thought to be a mechanism for creation of the continents.

23.2.1 Driving Mechanism

The driving mechanism for plate motion is unknown; however, several hypotheses have been proposed. Initially, thermal convection cells in the mantle resulting from dissipation of heat in the deep mantle from radioactive decay were believed to drive the system by exerting viscous drag on the bottom of the lithospheric plates. Although this mechanism has not been discounted, other explanations have been offered. Other hypothetical mechanisms include gravitational pulling by subducting plates, pushing by injection of new crust at mid-ocean ridges (considered unlikely on the basis of simple energy arguments), and excitation by the Chandler wobble or change in the length of day (rotation rate). In the latter, it is unclear which is the cause and which is the effect.

The geometric boundaries of the plates are not static since the processes occurring at any margin may change with time and new plate margins may be created by the subdivision of existing plates. Approximately 200 million years ago, the continental land masses were united in one paleocontinent known as Pangaea. See Figure 23-7 and Dietz and Holden [1970]. Gradually, this single continent has been subdivided by rifting and the creation of new oceans including the Atlantic and Indian Oceans. In addition, the rifting of continents has been initiated at several points but

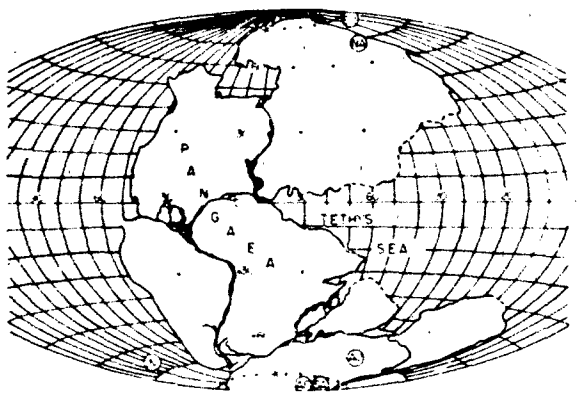


Figure 23-7 Reconstruction of Pangaea from present continents (approximately 200 million years before present) [Dietz and Holden, 1970]

seems to have terminated prior to significant separation. The Mississippi and St. Lawrence River Valleys may be regions of failed rifting. This theory postulates that a plate begins to break up because of upwelling mantle plumes which are convective cells of molten material that reach the surface. It can be shown that the least energy configuration for fracture from a point source is a triple junction with radially symmetric fracture arms at 120 degrees. If several of these spreading arms from other plumes intersect, the coalescing arms become a spreading ridge. The third faulted arm becomes relatively dormant. A name sometimes applied to these "failed arms" is alachogen.

23.2.2 Plate Motions

Relative plate motions can be estimated by the analysis of spreading rates at extensional margins, displacement across transform faults, and earthquake slip at subduction zones. Spreading rates at sea floor spreading ridges are generally measured by paleomagnetic anomalies. When magma is intruded at these ridges, it is heated past the Curie temperature, the temperature at which material gives up its remanent magnetization and realigns itself with the ambient magnetic field. When the magma cools it "freezes" in the ambient magnetic field. This makes estimates of spreading rates possible because of two facts. First, the polarity of the earth's magnetic field reverses itself on a seemingly random basis at periods ranging from thousands to tens of thousands of years. As the plates move away from the ridge, they carry distinctive magnetic signatures on the sea floor. The cause of the reversal is not known. Second, each magnetic signature can be dated by other means, such as radioactive isotope rock dating and fossil remains.

Using a simplified plate structure, Minster and Jordan [1978] estimated the relative plate motions by using an averaging interval of less than three million years. Because any motion on a sphere can be depicted as a rotation, plate motions are given as relative rotation vectors. Each vector gives the latitude and longitude of the radial axis of rotation and the angular velocity. Their calculated relative rotation vectors for abutting plate pairs are given in Table 23-1. It must be remembered that these poles are mathematical and have no physical expressions.

In addition to the relative movement of the plates, the entire lithosphere is apparently free to slip around on the surface of the earth like a shell. Mantle plumes, which are theorized to be areas of hot, upwelling, mantle material, appear to provide a stationary reference frame for the estimation of these absolute motions [Morgan, 1972]. Mantle plumes have as surface manifestations localized volcanic activity such as that found in Iceland and the Hawaiian Islands. As a plate moves over a plume, the active center of volcanism appears to move in a direction opposite to that of the plate motion leaving a trail of extinct volcanism on the plate, for example the Emperor Seamount Chain off

Table 23.1 Locations of poles of rotation for relative plate motions [Minster and Jordan, 1978].

	Relative Rotation Vector*						Error Ellipse†			Importance Distribution			
Plate Pair	θ °N	σ_θ deg	φ °E	σ_ϕ deg	ω deg m.y.	σ_ω deg m.y.	ζ_{max} deg	σ_{max} deg	σ_{min} deg	RA	TF	SV	Total
NOAM-PCFC	48.77	1.10	−73.91	1.94	0.852	0.025	S71°E	1.30	1.08	0.405	0.398	0.694	1.497
COCO-PCFC	38.72	0.89	−107.39	1.01	2.208	0.070	S37°E	1.00	0.63	0.977	0.272	0.009	1.258
NAZC-PCFC	56.64	1.89	−87.88	1.81	1.539	0.029	N09°E	1.91	0.96	0.849	0.341	0.038	1.228
EURA-PCFC	60.64	1.04	−78.92	3.04	0.977	0.027	S78°E	1.51	1.02	0	0	0	0
INDI-PCFC	60.71	0.77	−5.79	1.83	1.246	0.023	S82°E	0.90	0.76	0	0	0.246	0.246
ANTA-PCFC	64.67	0.90	−80.23	2.32	0.964	0.014	N52°E	1.11	0.75	1.200	0.811	0.039	2.050
COCO-NOAM	29.50	1.06	−121.28	2.07	1.489	0.070	S75°E	1.84	0.99	0	0	0.165	0.165
AFRC-NOAM	80.43	1.57	56.36	35.29	0.258	0.019	N86°E	5.68	1.51	0.851	0.246	0.091	1.188
EURA-NOAM	65.85	6.17	133.44	5.06	0.231	0.015	S14°E	6.36	1.39	1.055	0.626	0.366	2.047
NOAM-CARB	−33.83	9.19	−70.48	2.76	0.219	0.052	S13°E	9.42	0.97	0.952	1.741	0.253	2.946
COCO-CARB	23.60	1.48	−115.55	2.26	1.543	0.084	S63°E	2.24	1.21	0	0	0.111	0.111
NAZC-CARB	47.30	5.37	−97.57	4.57	0.711	0.056	S19°E	5.59	0	0	0	0	
COCO-NAZC	5.63	1.40	−124.40	2.61	0.972	0.065	N89°E	2.60	1.40	1.829	0.732	0.076	2.637
NOAM-SOAM	25.57	7.12	−53.82	6.22	0.167	0.029	S14°E	7.22	5.49	0	0	0	0
CARB-SOAM	73.51	11.75	60.84	48.86	0.202	0.038	S52°E	16.84	6.84	0	0	0	0
NAZC-SOAM	59.08	3.76	−94.75	3.73	0.835	0.034	S05°E	3.77	1.90	0	0	0.464	0.464
AFRC-SOAM	66.56	2.83	−37.29	2.65	0.356	0.010	S08°E	2.85	0.98	1.201	1.108	0.072	2.381
ANTA-SOAM	87.69	1.30	75.20	79.29	0.302	0.018	N84°E	3.22	1.26	0.167	0.608	0.283	1.058
INDI-AFRC	17.27	0.97	46.02	1.06	0.644	0.014	S47°E	1.24	0.66	0.843	1.098	0	1.941
ARAB-AFRC	30.82	3.44	6.43	11.48	0.260	0.047	S79°E	10.02	2.93	1.989	0.934	0.077	3.000
AFRC-EURA	25.23	4.25	−21.19	0.98	0.104	0.036	S01°E	4.25	0.89	0	0.783	1.167	1.950
INDI-EURA	19.71	1.40	38.46	2.66	0.698	0.024	S65°E	2.72	0.90	0	0	0	0
ARAB-EURA	29.82	2.53	−1.64	9.57	0.357	0.054	S85°E	8.33	2.45	0	0	0	0
INDI-ARAB	7.08	2.15	63.86	2.30	0.469	0.066	S51°E	2.51	1.89	0	0	0	0
NAZC-ANTA	43.21	4.50	−95.02	3.28	0.605	0.039	S01°E	4.50	2.39	0.246	0.058	0.222	0.526
AFRC-ANTA	9.46	3.77	−41.70	3.55	0.149	0.009	S42°E	4.93	1.45	0.697	1.243	0.195	2.135
INDI-ANTA	18.67	1.16	32.74	1.41	0.673	0.011	S62°E	1.39	1.10	1.012	0.135	0.025	1.172
Total										14.273	11.134	4.593	30.000

*First plate named moves counterclockwise with respect to the second. Uncertainties are the standard deviations of marginal distributions.
†One-sigma error ellipses are specified by the azimuth ζ_{max} of the major axis, lengths of the axes are geocentric angles.

Hawaii. By dating the volcanic rocks at various places along the island chain, the plate motion relative to this so-called "hot spot frame of reference" can be determined. The absolute rotation vectors, (with respect to the rotation axis of the earth) as determined by Minster and Jordan [1978] assuming stationary mantle plumes, are shown in Table 23-2.

23.3 SHORT PERIOD MOTIONS SEISMOLOGY

Seismology is commonly defined as the study of elastic waves that propagate through the earth, the study of their source, and the analysis of the waves to determine the physical properties and processes of the earth. Although seismology is classically concerned with the study of earthquakes, it has expanded to include the study of any source of elastic waves ranging from nuclear explosions to cultural activities. The common characteristics of these sources are the generation of transient signals with a concentration of energy at relatively high frequencies as compared to other earth motions such as solid earth tides.

The most significant class of elastic wave sources, in terms of radiated energy and frequency of occurrence, is earthquakes. An earthquake is defined as the sudden release of strain energy by the rupturing of rocks in which the strain energy has gradually accumulated. Although additional mechanisms such as instantaneous volumetric change accompanying a chemical phase transition, ice fractures (cryoseisms), and rock bursts have been suggested for limited classes of earthquakes, almost all earthquakes can be considered rupture type events. Fracture begins at one point, spreading over the fault plane at a rupture velocity of 2 to 3 km/s. As the rupture front reaches each point, the abutting surfaces break and slide against each other. The velocity of the particle movement on either side of the fault plane is typically on the order of 1 to 2 m/s. Note that two processes are at work: the propagation of the rupture over the fault plane and also relative motion of the sides of the fault. In some instances, fault rupture has stopped on one branch of a fault system and restarted on an adjacent preexisting fault. This behavior, known as "incoherent" rupture, can generate complexities on the seismic recordings known as starting and stopping phases. The June 1966 Parkfield event [Lindh and Boore, 1981] is an example.

Table 23-2. Locations of poles of rotation for absolute plate motions [Minster and Jordan, 1978].

	Absolute Rotation Vector						Error Ellipse		
Plate	θ °N	σ_θ deg	ϕ °E	σ_ϕ deg	ω deg/m.y.	σ_ω deg/m.y.	ζ_{max} deg	σ_{max} deg	σ_{min} deg
AFRC	18.76	33.93	−21.76	42.20	0.139	0.055	S73°E	40.40	33.24
ANTA	21.85	91.81	75.55	63.20	0.054	0.091	N12°E	93.01	56.12
ARAB	27.29	12.40	−3.94	18.22	0.388	0.067	S76°E	16.38	12.11
CARB	−42.80	39.20	66.75	40.98	0.129	0.104	N30°E	43.21	23.90
COCC	21.89	3.08	−115.71	2.81	1.422	0.119	S32°E	3.35	2.25
EURA	0.70	124.35	−23.19	146.67	0.038	0.057	S67°E	151.10	118.90
INDI	19.23	6.96	35.64	6.57	0.716	0.076	S25°E	7.16	5.97
NAZC	47.99	9.36	−93.81	8.14	0.585	0.097	S02°E	9.37	5.43
NOAM	−58.31	16.21	−40.67	39.62	0.247	0.080	S57°E	23.12	12.14
PCFC	−61.66	5.11	97.19	7.71	0.967	0.085	S16°E	5.23	3.50
SOAM	−82.28	19.27	75.67	85.88	0.285	0.084	N03°E	19.28	11.38

Earthquake location is described by the following convention: the actual location, possibly at some depth below the surface, is called the hypocenter or the focus; the projection of the focus onto the surface of the earth is called the earthquake epicenter.

Seismologists measure earth motion as a function of time with seismometers or accelerometers. See Aki and Richards [1980] for a discussion of these instruments. A typical seismometer is a velocity sensor consisting of a damped pendulum free to move in a magnetic coil.

23.3.1 Seismic Waves

For any given earthquake source, the characteristics of the mechanism of faulting (for example, focal mechanism and rupture velocity) will determine the energy distribution with respect to frequency and the radiation pattern of that energy. However, earth motions produced by any seismic source are limited to only four categories of waves: two are body waves and two are surface waves [Bath, 1973]. Body waves propagate through the earth while surface waves propagate along the surface of the earth. The first type of body wave is the compressional or dilatational wave, designated as the P-wave, ("primary" or "pressure") and the second is the S-wave ("shear" or "secondary"). The P-wave is analogous to an acoustic pressure wave with particle motion parallel to the direction of propagation as indicated in Figure 23-8. For the S-wave, particle motion is perpendicular to the direction of motion. The velocities of these waves are given by

$$V_p = \alpha = \sqrt{\frac{\lambda + 2\mu}{\rho}} = \sqrt{\frac{\kappa + (4\mu/3)}{\rho}}$$

[Bullen, 1965]

$$V_s = \beta = \sqrt{\frac{\mu}{\rho}}$$

where V_p is the P-wave velocity, V_s is the S-wave velocity, ρ is the material density, κ is the bulk modulus, μ is the shear modulus, and λ is the Lame parameter. V_p is always greater than V_s. For liquids, the value of μ is zero implying that S-waves can not propagate through liquids. Typical velocity ranges for earth materials are given in Table 23-3 [Crowford, 1974].

A typical simplifying assumption of seismology is that the variation of seismic properties within the earth can be approximated by horizontal layers or radial shells of uniform physical properties. For flat layers, the path of body waves through the medium can be approximated by optical ray theory or Snell's law: sin i/V = P, where i is the angle of

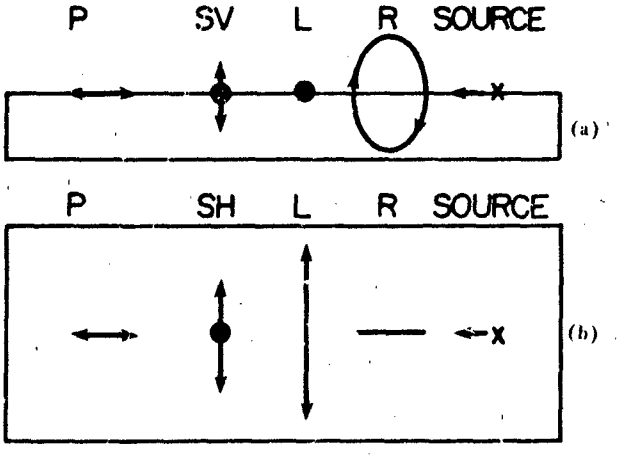


Figure 23-8 Propagational modes for waves traveling along a strike (as seen from the side (a) and above (b)

Table 23-3 Typical seismic velocities.

Material	Seismic Velocity ft/s	Seismic Velocity m/s
Loose and dry soils	600– 3 300	180–1 000
Clay and wet soils	2 500– 6 300	760–1 900
Coarse and compact soils	3 000– 8 500	910–2 600
Sandstone and cemented soils	3 000–14 000	910–4 300
Shale and marl	6 000–17 500	1 800–5 300
Limestone—chalk	7 000–21 000	2 100–6 400
Metamorphic rocks	10 000–21 000	3 000–6 400
Volcanic rocks	10 000–22 000	3 000–6 700
Sound plutonic rocks	13 000–25 000	4 000–7 600
Jointed granite	8 000–15 000	2 400–4 600
Weathered rocks	2 000–10 000	600–3 100

ncidence of a ray at the interface. V is the appropriate velocity for P or S waves and p, the ray parameter, is a constant along the ray path. At each interface, an impinging P or S wave can generate both reflected and transmitted P and S waves. Thus, the potential paths over which body wave energy can travel from the source to the point of observation grow quickly with an increasing number of layers. Over many paths, however, insufficient energy is transmitted to define a distinct arrival at the observation point.

There is a standard method for identifying body waves as they pass through the earth. See Richter [1958] and Ben-Menahem and Singh [1981] for a complete discussion.

Briefly the convention is

P	P-wave
S	S-wave
PP	P-wave reflected once by the surface
SS	S-wave reflected once from the surface
PPP...,SSS...	Multiply reflected Ps or Ss
SP	S wave reflected by surface changed into P wave.
PSPS...	Reflected wave which changed on reflection into P or S
K	Wave is transmitted through outer core (no shear wave in liquid outer core). Example: PKP
I	Wave passes through inner core
c	Wave reflects off outer core
i	Wave reflects off inner core
p,s...pP,sS...	Depth phase p and s waves travel up to the surface from deep focus earthquakes. Example: pP,sS

hese symbols are combined to uniquely describe the wave ype and path. Several of these waves are shown in Figure 3-9. The last phases listed, the depth phases, are important or determining the depth of an earthquake. Seismic waves avel from a focus in all directions; when an event occurs, ome of the rays travel up to the surface, then reflect downard into the familiar P or S patterns. Since the velocity of ropagation is known, the arrival time delay on a seismo-

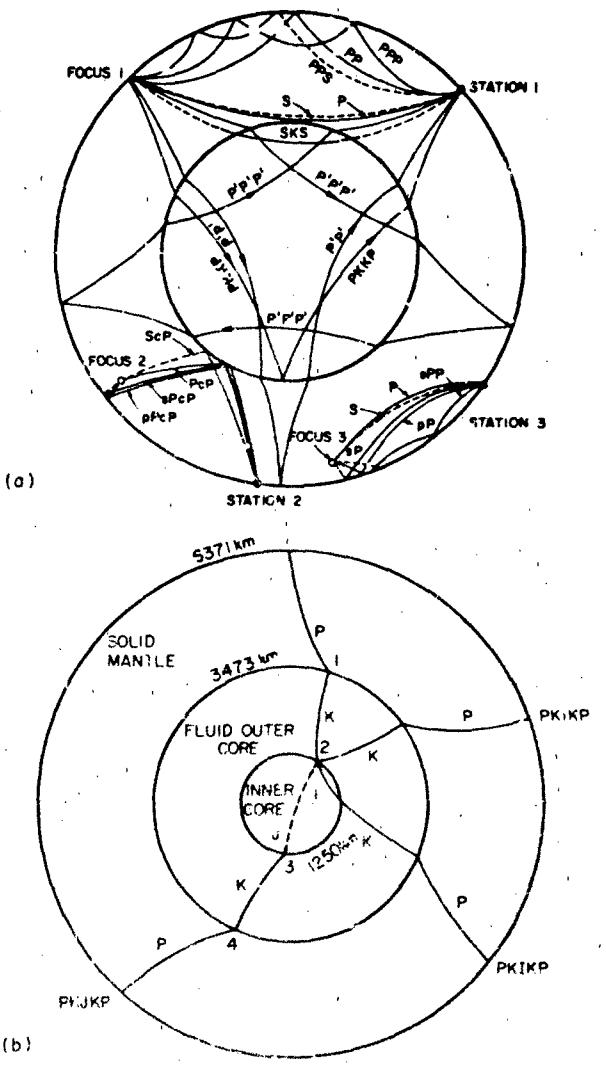


Figure 23-9 Types of body waves, modified from Ben-Menahem and Singh [1981]

gram between a pP and P for example, will be a function of the depth.

In addition to reflection and transmission, seismic body waves can undergo refraction and diffraction. Because of diffraction about the core, for example, a certain distance range on the surface of the earth from the source will experience a "shadow zone" in which direct waves cannot be detected (Figure 23-10).

Surface waves observed on the free surface of the earth are either Love (L) or Rayleigh (R) waves. The particle motion of Love waves is perpendicular to the propagation direction and in the plane of the surface, while Rayleigh waves have retrograde elliptical particle motion in the vertical plane of propagation (Figure 23-8). The particle motion decreases with depth as a function of frequency, with longer periods sampling greater depths. Thus, the velocity of surface waves is a function of the period of the wave and of the velocity structure of the earth (dispersion). In general, Love wave velocities are higher than Rayleigh wave velocities and the velocity of long period waves are faster than short period surface waves. In Figure 23-11 typical Rayleigh and Love wave phase velocities (the velocities at which distinct frequency waves travel) are shown for several geologic structures [Brune, 1969].

A typical seismogram recording will show the following characteristics. The first signal to arrive will be P-waves (including direct P, pP, PP, etc.) which, in a general layered medium, will arrive at the receiver over different paths. This packet will be followed by S-waves, also arriving along different paths. Next, the Love wave will appear with longer periods arriving first. The frequency of Love waves becomes higher and the later Love waves overlap the long-period Rayleigh wave arrivals. Short-period Rayleigh waves arrive last. Many arrivals in the seismogram can be identified with distinct paths through the earth; hence the nomenclature for these distinct signals has been developed. The travel times of these arrivals are well known as functions of event depth and distance from the observation point. By fitting the arrival times of these phases at a number of stations to the known travel time curves for each phase, the epicenter, depth, and origin time of a seismic event can be determined.

Earthquakes or large explosions radiate surface waves in all directions with amplitudes that depend on the radiation pattern [Ben-Menahem and Singh, 1981]. Most commonly observed are surface waves that travel the shortest arc of a great circle between source and station. Large events possess sufficient energy so that surface waves traveling the long arc are also recorded. Surface waves from the largest earthquakes, for example the 1960 Chilean shock, can be detected after multiple trips around the globe. Important information on earthquake source parameters and the deep interior of the earth can be gleaned by studying the velocity and amplitudes of these waves. See, for example, Kanamori [1970] and Kanamori and Cipar [1974].

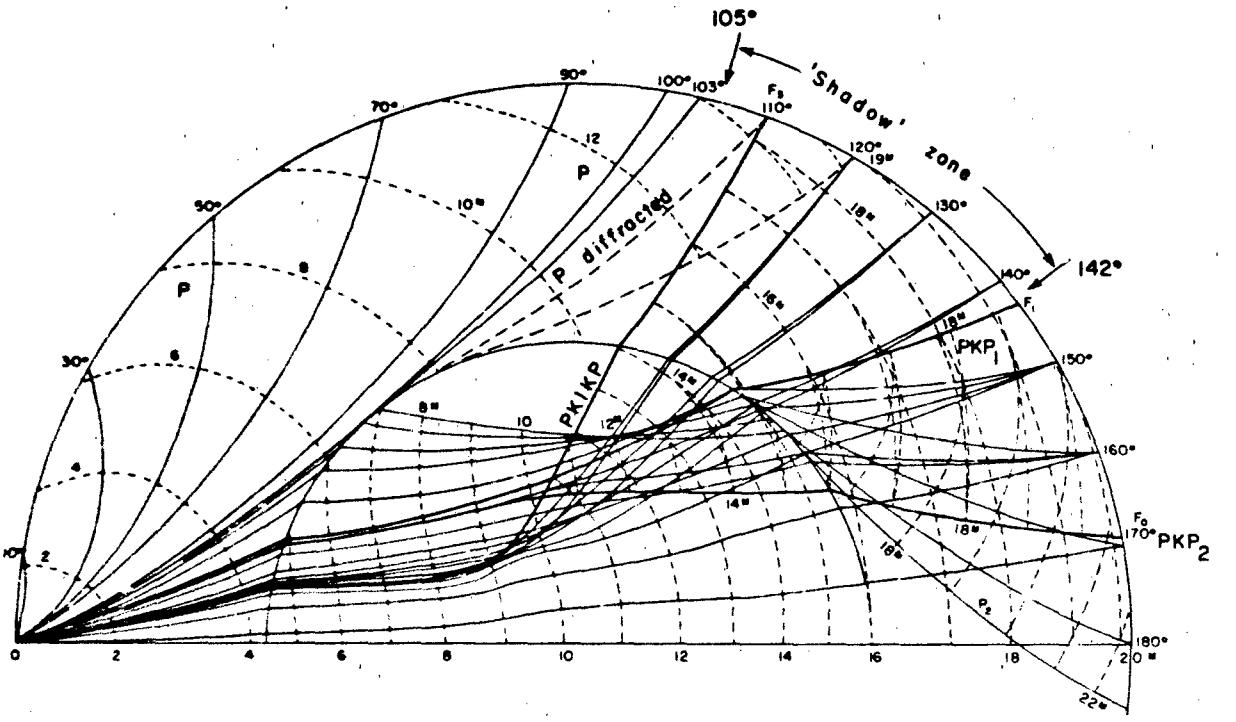


Figure 23-10 P wave, diffracted by the core-mantle interface, creates a "shadow zone" at 105° to 142° distance, modified from Ben-Menahem and Singh [1981].

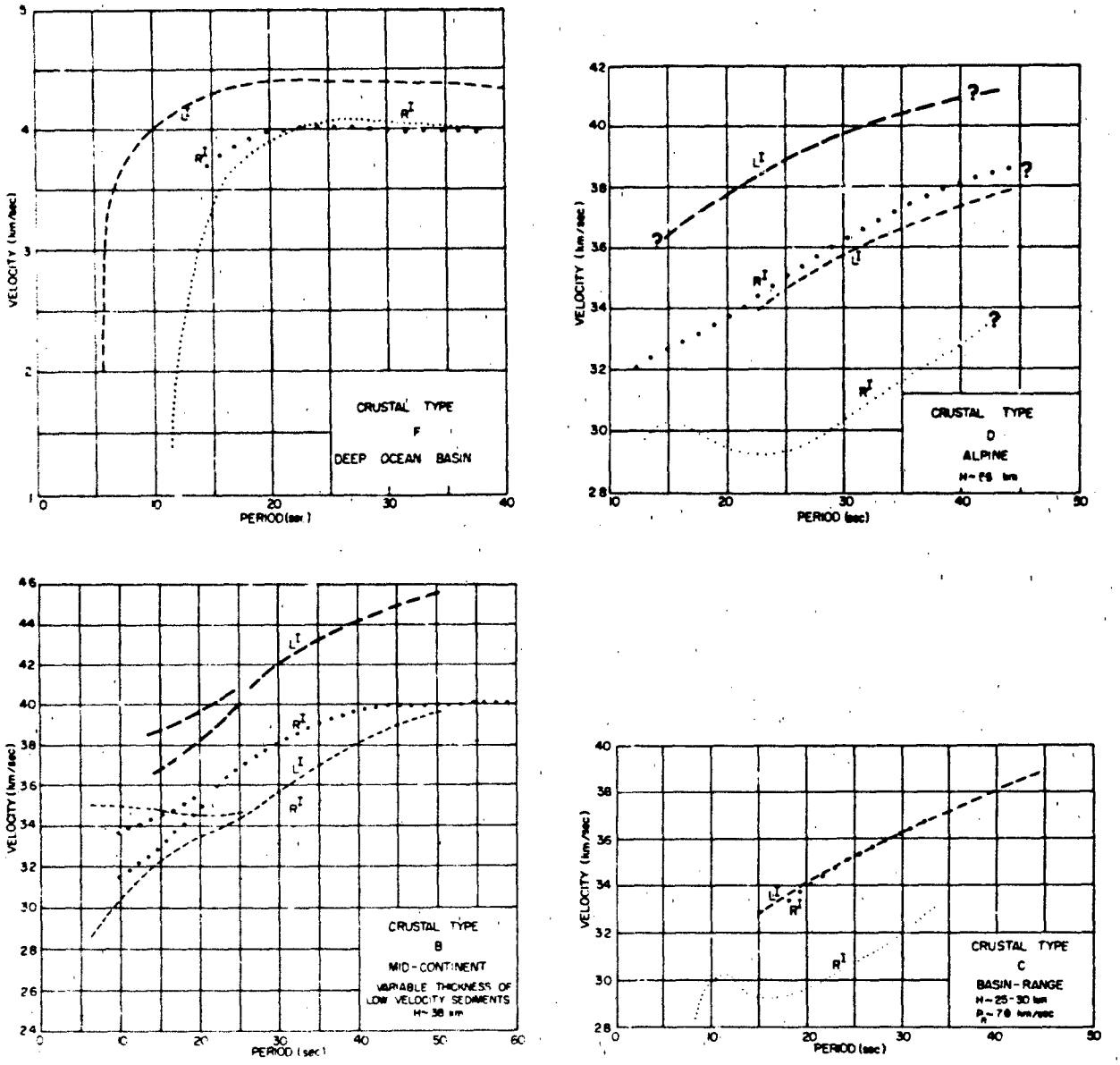


Figure 23-11 Rayleigh (R) and Love (L) wave phase velocity (bold dots and dashes) and group velocity (small dots and dashes) curves for various crustal structures [Brune, 1969]

Surface waves have been described above as traveling over the surface of the earth. Alternatively, the motion of the earth following a giant earthquake can be pictured as superimposed fundamental and higher mode oscillations, analogous to those produced by the ringing of a bell. These free oscillations, in effect standing waves, can be detected up to several days following a large earthquake. A line spectrum of a long-period horizontal seismogram exhibits both torsional and spheroidal oscillations, which are the free oscillation representations of Love and Rayleigh waves, respectively. Such line spectra exhibit line broadening due to earth rotation, damping, and non-symmetrical earth structure. Details of spectral amplitudes are diagnostic of the source processes of the giant earthquakes which generate these motions. The lowest order free oscillations sample the entire earth (the lowest spheroidal oscillation has a period of 54 minutes) and thus are diagnostic of the structure of the deep interior.

No discussion of wave propagation is complete without mention of attenuation effects. Basically, there are two pro-

cesses at work. The first, geometric spreading, is an inverse square function of the distance from the source, analogous to the falloff of gravity as $1/r^2$ where r is the distance. This occurs because the area at distance r subtended by a steradian (solid angle) is proportional to r^2. The second type, called intrinsic attentuation or internal friction, is a function of the distance from the source and the wavelength of the wave, that is, the number of cycles traversed over the distance. Some of the causes of anelastic damping are changes in internal vibrational energy in the molecules, coulomb force (friction) effects on the particles as a result of the physical interaction, viscous damping, and mechanical damping during the acceleration of the particles. The energy is dissipated in the form of heat. The form of the expressions in the equations for the anelastic damping factor, Q, varies and the reader should refer to one of the standard texts, for example, Aki and Richards [1980], for details. Generally, a high value of Q connotes low attenuation and vice versa.

23.3.2 Earthquakes

23.3.2.1 Spatial Distribution. On a worldwide scale, the distribution of earthquakes is not random. The vast majority of earthquakes occur along plate boundaries (Figure 23-12) where relative motions of the plates produce large stress concentrations. The depths of earthquakes range between the surface of the earth and about 700 km. Earthquakes occurring away from consuming plate boundaries (at transform boundaries and at extensional margins) are generally shallow with depths less than 30 km. At subduction zones the whole range of depths can be seen in a distinct pattern. At the trench side of the margin, the depths are shallow, increasing away from the trench and under the overriding plate. The focal depth increases with the depth of penetration of the subducting plate into the asthenosphere.

In California, where almost all major and minor fault zones are well mapped, most earthquakes can be associated with pre-existing fault zones. This is readily explained by the fact that a pre-existing fault represents a zone of weakness in the crust where stresses accumulate and lower stress concentrations are required to cause rupturing. In other regions away from major plate boundaries, such as New England or the Central Mississippi River Valley, the patterns of earthquake occurrence are less clear. In these regions the lack of surficial expression of faults due to the thickness of soil and alluvial coverage over bedrock and the lower temporal rate of occurrence of earthquakes tends to produce more diffuse patterns of seismic activity.

23.3.2.2 Measurement. The size of an earthquake can be described by its epicentral intensity, magnitude, and seismic moment. In modern seismological practice, magnitude is the most commonly reported measurement.

Intensity of shaking is a measure of the damage effects produced by an earthquake at a specific location. The maximum damage usually occurs in the epicentral region and this intensity is assigned to an event for purposes of com-

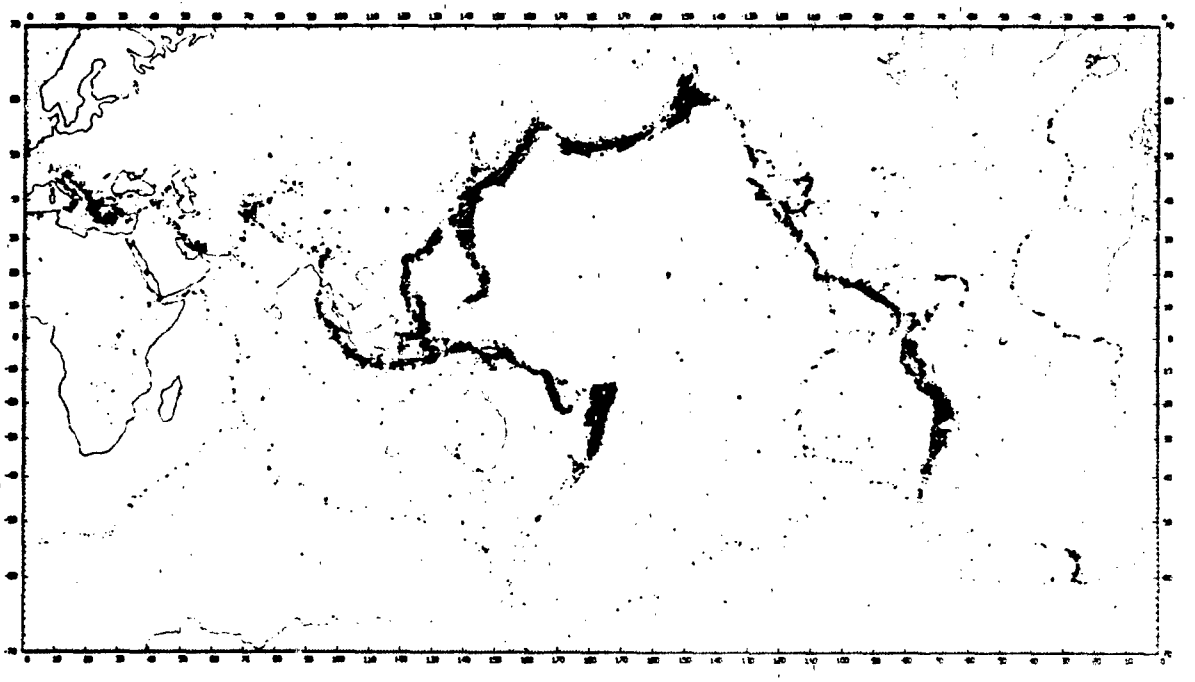


Figure 23-12. Worldwide distribution of epicenters for 1961 through 1967 [Isacks, 1968]

parison. Several intensity scales have been developed; the most commonly used in the United States is the Modified Mercalli Intensity Scale (MMI). Table 23-4 describes the intensities, typically shown as Roman numerals. The MMI ranges from I to XII, with XII being the maximum intensity that can be assigned to an event. While intensity scales have been superseded by more quantitative measures of earthquake size based on instrumental recordings, they are essential in comparing the size of historical earthquakes with modern events. There are few instrumentally recorded events prior to about 1925, but accounts of damage from earthquakes have been published by journals in the past with sufficient detail to assign an intensity and an approximate epicenter. Magnitude-versus-intensity relationships can be used to construct an earthquake history of an area. Using reports of intensity at different distances from the same event, attenuation of ground motion with distance can be examined. This information is particularly valuable in areas such as Cape Ann, Mass., Charlestown, S.C., and New Madrid, Mo. that have experienced large earthquakes in the past but currently exhibit low seismicity. A drawback of this system is that level of shaking, level of panic, and so forth are subjective. Local damage levels are functions of the size of an earthquake and of several variable factors such as construction quality in the area, geological foundation characteristics, population density, and the scientific

Table 23-4 Modified Mercalli intensity scale of 1931

I. Not felt except by a very few under specially favorable circumstances. (I Rossi-Forel scale.)

II. Felt only by a few persons at rest, especially on upper floors of buildings. Delicately suspended objects may swing. (I to II Rossi-Forel scale.)

III. Felt quite noticeably indoors, especially on upper floors of buildings, but many people do not recognize it as an earthquake. Standing motorcars may rock slightly. Vibration like passing of truck. Duration estimates. (III Rossi-Forel scale.)

IV. During the day felt indoors by many, outdoors by few. At night some awakened. Dishes, windows, doors disturbed; walls make creaking sound. Sensation like heavy truck striking building. Standing motorcars rocked noticeably. (IV to V Rossi-Forel scale.)

V. Felt by nearly everyone, many awakened. Some dishes, windows, etc., broken; a few instances of cracked plaster; unstable objects overturned. Disturbances of trees, poles, and other tall objects sometimes noticed. Pendulum clocks may stop. (V to VI Rossi-Forel scale.)

VI. Felt by all, many frightened and run outdoors. Some heavy furniture moved; a few instances of fallen plaster or damaged chimneys. Damage slight. (VI to VII Rossi-Forel scale.)

VII. Everybody runs outdoors. Damage negligible in buildings of good design and construction; slight to moderate in well-built ordinary structures; considerable in poorly built or badly designed structures; some chimneys broken. Noticed by persons driving motorcars. (VIII Rossi-Forel scale.)

VIII. Damage slight in specially designed structures; considerable in ordinary substantial buildings with partial collapse; great in poorly built structures. Panel walls thrown out of frame structures. Fall of chimneys, factory stacks, columns, monuments, walls. Heavy furniture overturned. Sand and mud ejected in small amounts. Changes in well water. Persons driving motorcars disturbed. (VIII + to IX − Rossi-Forel scale.)

IX. Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb; great in substantial buildings, with partial collapse. Buildings shifted off foundations. Ground cracked conspicuously. Underground pipes broken. (IX + Rossi Forel scale.)

X. Some well-built wooden structures destroyed; most masonry and frame structures destroyed with foundations; ground badly cracked. Rails bent. Landslides considerble from riverbanks and steep slopes. Shifted sand and mud. Water splashed (slopped) over banks. (X Rossi-Forel scale.)

XI. Few, if any, (masonry) structures remain standing. Bridges destroyed. Broad fissues in ground. Underground pipelines completely out of service. Earth slumps and land slips in soft ground. Rails bent greatly.

XII. Damage total. Waves seen on ground surfaces. Lines of sight and level distorted. Objects thrown upward into air.

integrity of the observer. If population is sparse in the epicentral area, reports of the maximum damage may be missing, thus underestimating the size of the event or mislocating the epicenter.

Because of these problems, data used to compile relationships of intensity to quantitative measures of earthquake size such as peak acceleration or magnitude are derived from scattered data sets. A potentially more informative intensity parameter, earthquake felt area, is often computed. It is defined as the area enclosed by a particular "isoseismal" line. These lines are contours of intensity data drawn around an epicenter. Felt area is usually computed to the transition between intensity III to IV or II to III.

The magnitude of an earthquake is calculated from an instrumental measurement of the amplitude of a specified wave packet, such as body waves or surface waves. The general form is $M_x = (\log[A/T] \text{ or } \log A) + C_p + C_c + C_r$, where M is the magnitude, x is the subscript that identifies the type of magnitude, A is amplitude, T is the period of the waves (equal to approximately 1 s for body waves, 20 s for surface waves), C_p is the correction factor whereby stations at all distances theoretically calculate the same magnitude for an event (a function of the depth of the event and the epicentral distance to the event from the station), C_c is a correction factor for local conditions and C_r encompasses regional correction factors allowing for effects of recording site, wave path, and focal mechanism [Willmore, 1979]. As can be seen from the above representation, earthquake magnitude is a logarithmic scale so that a change of one unit of magnitude roughly represents a change of a factor of 10 in actual ground motion.

The concept of magnitude was developed by Richter in the 1930's for California earthquakes and has been expanded for other areas and different wave packets by Richter and many others. Among the most common magnitudes used today are local magnitude (M_L), the original scale proposed by Richter (M), teleseismic body wave magnitude (m_b), and surface wave magnitude (M_S).

Due to many factors, the magnitude determined for each scale usually will not be identical for a given earthquake. The shape of the seismic wave spectrum is generally flat up to some "corner frequency" after which the amplitude decays. The level of the flat portion is a function of the size of an earthquake: generally the larger the earthquake, the larger the amplitude of the spectrum at a given frequency. This is why m_b and M_s magnitudes, which are measures of the amplitudes at specific frequencies, are used to measure size. However, the corner frequency moves to lower frequencies and longer periods as the size of the earthquake increases. Thus the relationships between magnitude scales, besides being almost completely empirical, are often nonlinear. Mathematically, there is no upper limit to the earthquake magnitude scale; however, because of this shift in corner frequency, the size of an earthquake can get increasingly larger with a smaller and smaller increase in M_s as discussed by Chinnery [1975]. This effect is called saturation of the magnitude scale. The highest magnitude ever evaluated is 8.9M_s for an event near Japan on 2 March 1933. This event approached the theoretical maximum magnitude based on the strengths of rocks. On the lower end of the scale, the sensitivity of the instruments and the distance from the event are the only limiting values with negative magnitudes being determined for some events detected at very close range. Approximate relationships between these scales have been determined by several authors [for example, Nuttli, 1979; Bath, 1973]. Examples of such empirical relationships are $m_b = 1.7 + 0.8M_L - 0.01\ M_L^2$ and $m_b = 0.56M_s + 2.9$.

Local magnitude is strictly defined only for California, or places with very similar attenuation characteristics, while the body wave and surface wave magnitudes are applied on a worldwide basis. A common mistake found in the literature, especially in engineering studies, is to combine all magnitudes or to simply assume they are all the same. Such a practice, as indicated by the relationships above, can completely invalidate the study. Even though some reports and agencies may simply quote "Magnitude" or "Richter Magnitude", care should be taken to understand how the magnitude was calculated. Richter magnitude often means "M_L" up to M = 6.0 and "M_S" above M = 6.0. Improper discrimination of magnitude types causes major problems in communication and comparison of data, for example the estimation of maximum credible earthquakes in specific regions.

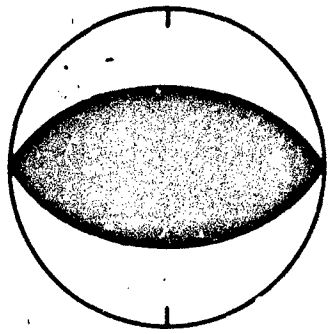
Relationships between magnitude and intensity have also been developed and these functions have strong regional dependencies [Brazee, 1976]. For the Western United States, the relation between maximum intensity, I_m, and body wave magnitude is $m_b = 2.886 + 0.365I_m$, while for the Central United States the relation is $m_b = 2.607 + 0.341\ I_m$. The importance of this regional variation and the difficulties of using intensity scales for determination of earthquake size are demonstrated by comparing the 1906 San Francisco earthquake and the 1811-1812 New Madrid earthquakes [Nuttli, 1973]. The San Francisco earthquake probably had a higher magnitude than any of the New Madrid earthquakes. However, due to regional differences in seismic wave attenuation, the San Francisco earthquake was felt over an area less than one-tenth of that for the largest of the New Madrid series of events. This is the result of lower seismic wave attenuation in the Central and Eastern United States than in California. Attenuation will be discussed further in the section on seismic hazard (Section 23.3.2.5).

The seismic moment (M_o) of an earthquake is defined as the product of the rigidity (μ), the average offset along the fault (D) and the area of the fault (S) [Kanamori, 1977]. It is usually calculated in units of dyne-cm. When these parameters are not directly observed, spectral analysis and speculation must be used. This measurement of earthquake size is closer to describing the physical dimensions of a rupture then other kinds of measures (I_o, MMI, M). Relative to earthquake magnitude, which is systematically calculated

by various international agencies such as the International Seismological Centre in Edinburg, only a handful of moments, mostly for larger events, have been determined. A new type of magnitude based on the strain energy drop (the moment) has been computed for some events. It was developed by Kanamori [1977] and is defined by log $W_o = 1.5M_w + 11.8$ where W_o is the minimum strain energy drop and M_w is the moment magnitude. This magnitude scale circumvents the problem of saturation of M_s that was discussed earlier in this section. The great Chilean earthquake had an M_w of 9.5 but an M_s of only 8.3.

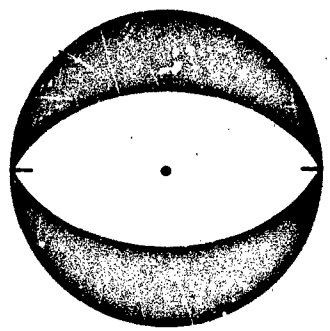
There are numerous magnitude scales developed for use with events too small to be recorded teleseismically (definitions of this term vary; here 1000 km is considered the cutoff distance). Measurement of such events presents a problem because the recordings are heavily influenced by the path through the crust.

23.3.2.3 Focal Mechanisms and Fault Plane Solutions. Another common tool used to understand the nature of earthquakes is the fault plane solution. This is determined quantitatively by a special plot of the polarity of the first motions of seismic waves recorded at all stations where the first motion is readable. Figure 23-13 shows examples of the three main kinds of fault motion: thrust, normal and strike slip. Thrust faulting occurs when two fault blocks are pushed together and the overhanging wall rises over the foot block wall. Normal faulting occurs when blocks are pulled apart (tension) and one block drops down. Strike-slip faulting occurs when two fault blocks slide past each other. The strike of the fault (the angle the fault plane makes with north) and the dip (the angle that the fault plane makes with the surface) are easily read off these plots. There is, however, an ambiguity over which focal plane is the fault plane. This must be resolved either through surficial expressions of the fault or knowledge of the tectonic regime. Figure 23-14a shows how a fault plane solution is constructed for thrust or normal faulting. Basically, it represents a projection of a hemisphere, usually a lower hemisphere, centered about the focus of the earthquake onto a plane. The angles of incidence of the seismic waves leaving the scene of an earthquake to recording stations have been calculated for various distances. Compressional or dilatational waves reaching a station are coded on a plot (solid circles for compressional, open circles for dilatational). Focal planes are drawn dividing compressional and dilatational regions, with the stipulation that the planes are mutually perpendicular. Figure 23-14b shows the rationale for the characteristics of the strike-slip fault plane solution. These faults are often described as "right lateral" or "left lateral". This means that if the ground on the opposite side of a fault from an observer moves to the right, the fault is said to be right lateral; if it moves to the left, left lateral (Figure 23-15). Faults can also exhibit combinations of the motions described in this section.



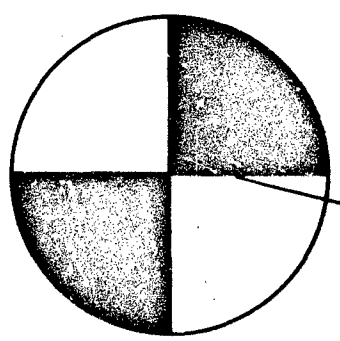
STRIKE=90° DIP=45°

(a) PURE THRUST FAULTING



STRIKE=90° DIP=45°

(b) PURE NORMAL FAULTING



STRIKE 90° DIP=90°

(c) PURE STRIKE-SLIP FAULTING

Figure 23-13. Examples of fault plane solutions for different types of earthquake mechanisms

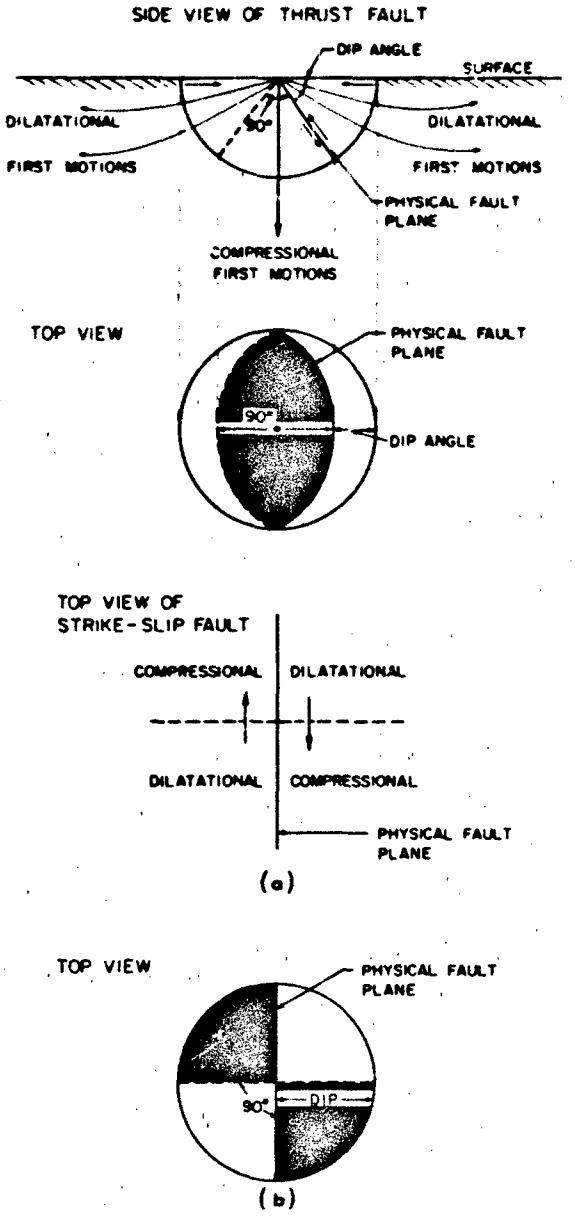


Figure 23-14. How to construct a fault-plane solution.
(a) Side and top view of thrust fault plane solution
(b) Top view of strike-slip fault and top view of fault plane solution

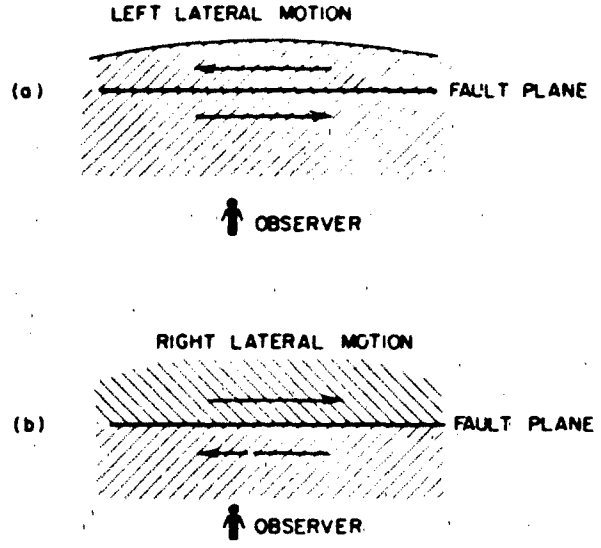


Figure 23-15. Fault motion convention
(a) Observer sees land on opposite side of fault line move to his left.
(b) Observer sees land on opposite side of fault line move to his right.

23.3.2.4 Temporal Distribution. The rate of earthquake occurrence has been found to be described approximately by the function $\log N = a - bM$, where N is the number of earthquakes of magnitude M per unit time. For the entire earth from the period 1918 to 1964 the constants of this equation have been found to be $a = 8.73$ and $b = 1.15$, where N is the number of events per year based on earthquakes in the range of 6.0 to 8.9M [Bath, 1973]. Recurrence curves of this form are useful for areas as small as $10^4 km^2$. The slope, or b-value, of the curve is normally found to lie between 0.5 and 1.5. It is not clear whether the variations are caused by scatter or are indicators of the seismic processes in a region. The value of "a" is highly variable, however. Obviously regions of high seismic activity, such as southern California, have higher "a" values than less seismically active regions, such as New England.

23.3.2.5 Hazard Evaluation. Strong ground motions from earthquakes constitute a great threat to facilities near the causative fault. Earthquakes capable of producing strong ground motions have been recorded in almost every state of the United States (Figure 23-16). Two approaches have been developed by seismologists to estimate potential strong ground motion levels at particular sites. The deterministic method attempts to predict the maximum ground motion possible while the probabilistic approach predicts the likelihood that a given ground motion will occur over some specified period of time. The two methods need not be mutually independent. The most physically realistic model would ideally incorporate elements common to both.

The deterministic approach to seismic hazard evaluation requires knowledge of faults in the vicinity (radius of 100 km or more) of the facility. In general, faults that have shown any activity during the Quaternary (past two million years) are considered capable of sustaining seismic activity. An estimate is made of the maximum length of each fault that could rupture in one earthquake and this value is used

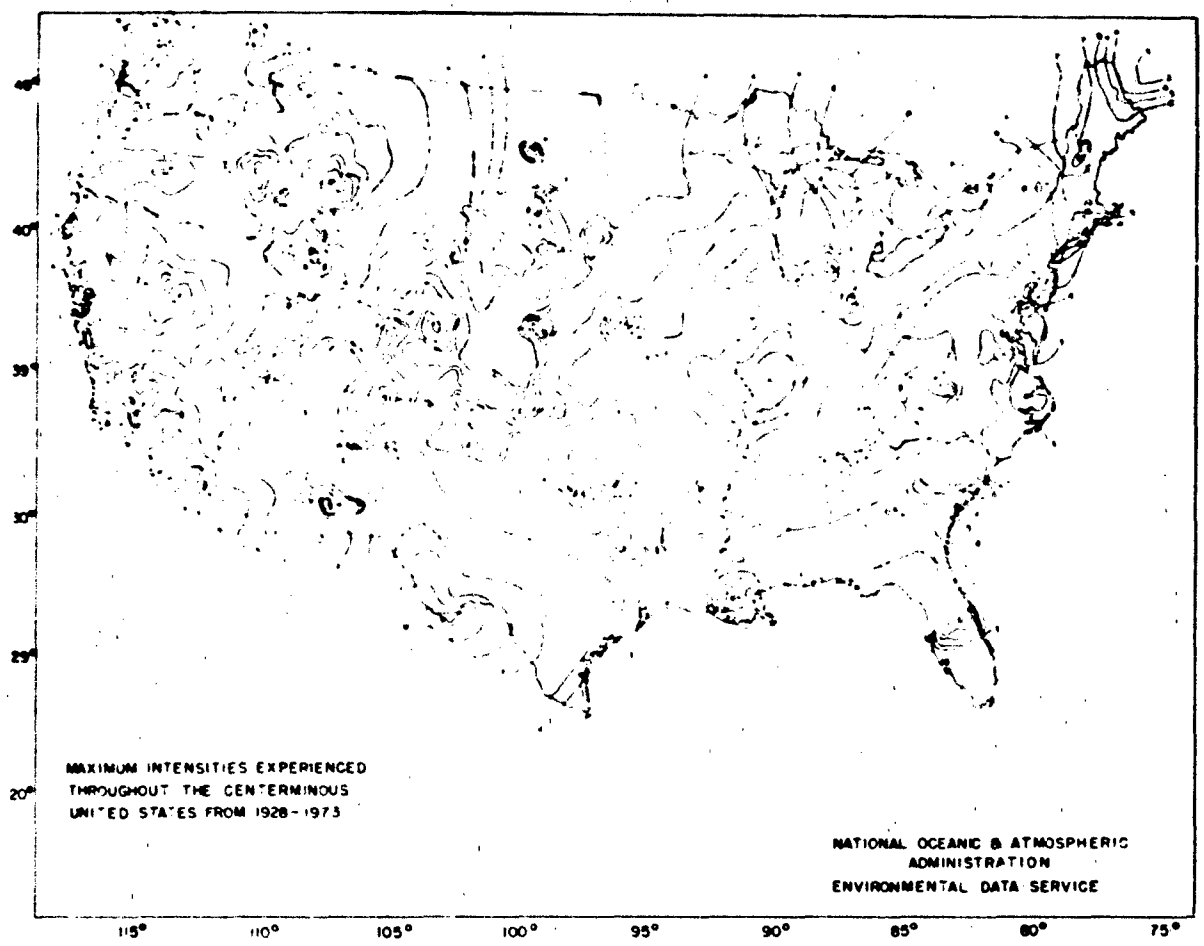


Figure 23-16 Maximum earthquake intensities throughout the United States from 1928 to 1973 [Brazee, 1976]

to evaluate a maximum credible earthquake, that is, the largest earthquake with a reasonable chance of occurring on the given fault. The evaluation is done using any of several empirical equations relating magnitude to maximum fault rupture length [Slemmons, 1977]. One equation of this form is $M_L = 5.4 + 1.4 \log L$, $\sigma = 0.26$, where L is the length of the fault in kilometers [Greensfelder, 1974]. This phase of the deterministic approach, estimating the maximum possible earthquake, is the most difficult. Often the length of the fault is unknown and estimates of magnitude of past events from geological examination of fault offsets do not always yield unique values. When the estimation is made, strong ground motions at a site can then be evaluated using empirical relations between magnitude, distance, acceleration, velocity, or displacement. A compilation of various relationships can be found in McGuire [1976]. Caution must be used however in selecting a set of equations since they are not all valid worldwide; some were developed for specific areas, such as California. One set of equations of this type is

$$\ln a_s = 6.16 + 0.645M_L - 1.3(R+25)$$

$$\ln v_s = 1.63 + 0.921M_L - 1.2(R+25)$$

$$\ln d_s = 0.393 + 0.99M_L - 0.88(R+25),$$

where a_s, v_s and d_s are the site acceleration, velocity, and displacement, M_L is the event magnitude, and R is the distance from the fault to the site in kilometers. Note that these equations are valid only for rock sites in the western United States. A modified acceleration function for the central United States is given by Battis [1981]: $\ln a_s = 3.155 + 1.240m_b - 1.244 \ln(R + 25)$. Further modification of these equations is required to compensate for local site conditions, but these methods are beyond the scope of this review.

One problem with a purely deterministic approach is that the return period of the maximum event may be so large that it is unreasonable logistically and financially to design a relatively short lived facility to a very high value. For

example, if a magnitude 7.0 event had a return period of 100 000 years, the annual rate of occurrence would be 1×10^{-5}. If the life of the structure in question is 40 years, then the probability of exceeding a magnitude 7 during the lifetime of the structure would be approximately 4×10^{-4}. At this point the probabilistic approach should be utilized.

Another problem is that these relationships for acceleration do not predict the level of ground motion over the entire frequency range. Standardized spectra have been developed (Nuclear Regulatory Guide 1.60) which can be anchored at a peak acceleration value. As yet, modification of the shape of the spectra for the type of soil at the site and for the magnitude of the event is not a routine process [Johnston et al., 1980].

The method of using seismic hazard estimates to design structures has gained acceptance in recent years. A computer program by McGuire [1976] that will perform the necessary calculations has been made available to the public. Since the concepts and assumptions that make up the calculation must be understood in order to assess the appropriate use and the limitations of the method, a brief overview is presented. Once the site of interest has been selected, a catalogue of earthquakes in terms of date, location, and size (magnitude or intensity) is examined. If the catalogue is not homogeneous, then conversions from one magnitude scale to another (for example m_b to M_s) or from intensity to magnitude must be made. Care must be taken to select the proper conversion relationship (Section 23.3.3). A range of scenarios are tried, using source zones of various sizes, to model the seismicity (Figure 23-17a).

Once source zones are defined, the enclosed seismicity is tabulated. Earthquakes are counted according to magnitude, or any other measure of size and time interval. The seismologist decides on the appropriate time intervals using judgements of the completeness of the catalogue for various years for the different magnitude ranges. A recurrence curve is derived for each source zone (Figure 23-17b). These curves represent the modeled cumulative number of earthquakes that will occur in a given source zone at any magnitude (per year if the calculations are for annual risk). Next, all the modeled seismicity that occurs annually in this zone is distributed evenly throughout the zone according to location. Even if the natural occurring seismicity is clustered in a small area within the zone, it will be spread out uniformly. This is an important detail because of the distance calculation (Figure 23-17c). Therefore, the geometry of the zones plays an important part. The next step requires an attenuation relationship that relates magnitude to acceleration, velocity, or displacement and distance to the site (Figure 23-17d). Distances are then computed to the seismicity which is now distributed evenly in the zone. When all this has been tabulated, the result is Figure 23-17e, the annual risk of exceeding the ground motion parameter used in the figure.

This type of analysis is probably most useful for com-

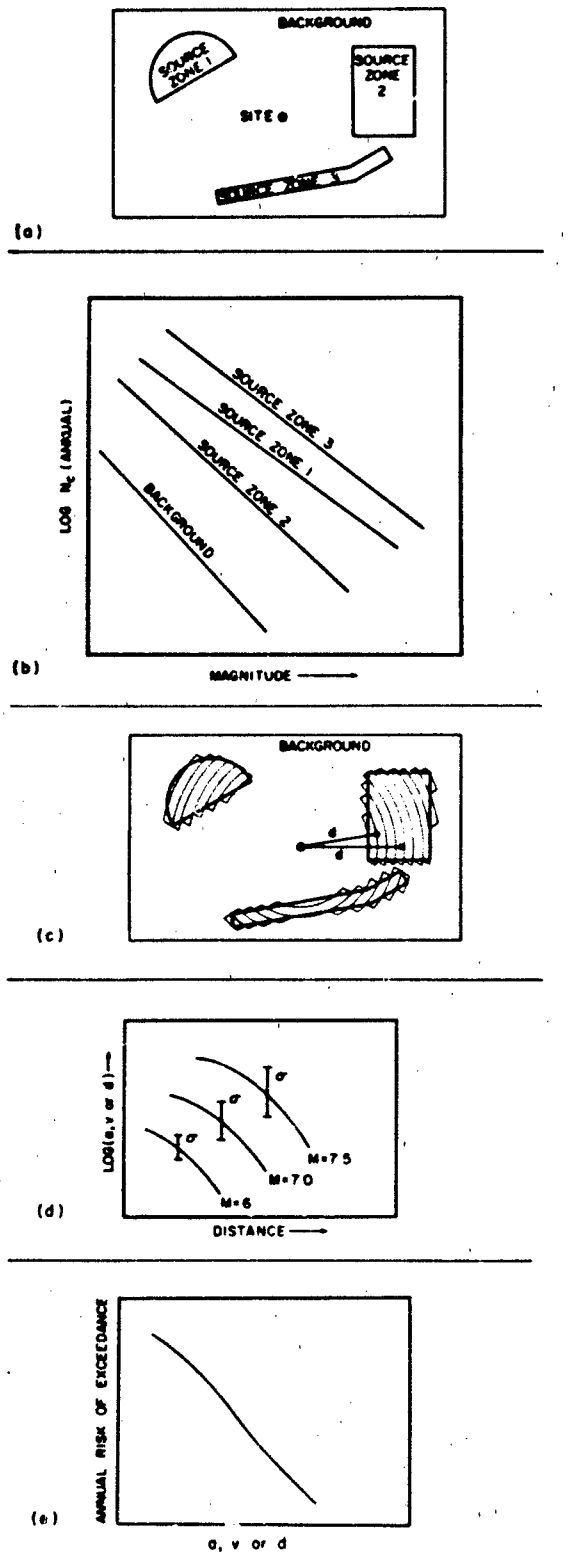


Figure 23-17. Performance of seismic hazard calculation

parison purposes: performing this calculation for two distinct sites that have a fairly accurate earthquake catalog will yield a good estimate of the relative seismic hazard of the two regions. Since there are some elements of the deterministic approach incorporated into the hazard, such as choosing the source zone boundaries and an event upper magnitude cutoff, this method gives more information about the actual hazard to a facility during its useful lifetime (typically 30 to 50 years). It also allows designers or regulating agencies to quantitatively incorporate conservatism into the design or to set guidelines. Often a 10 000 year return period is considered to be an acceptable hazard. The actual probability of damage to a stucture using seismic hazard and engineering methods is called the seismic risk.

Another limitation of the hazard method is that it does not include the duration of the ground motion contributing to the annual risk. Seismic waves from a magnitude 7.0 earthquake at 100 km distance may, for example, have higher amplitude ground motion and longer time duration at periods greater than approximately 0.8 s than will a magnitude 5.5 which is located within 10 km of the site. Since structures respond to periodic signals more readily than to a single spike of motion and have specific natural frequencies, these complications should be considered.

The maximum credible earthquake may have a much longer return period than the useful life of most structures. Probabilistic methods are generally considered to provide more realistic estimates of the ground motions and may be more useful in the case of life-line structures or structures such as nuclear reactor facilities whose failure would produce a significant threat.

Several methods have been devised to provide regionalized seismic hazard maps for the United States. The results of two of these studies are shown in Figures 23-18, 23-19, and 23-20 [Algermissen, 1969; Donovan et al., 1977]. Figure 23-18 is the result of a deterministic approach while Figures 23-19 and 23-20 result from probabilistic methods.

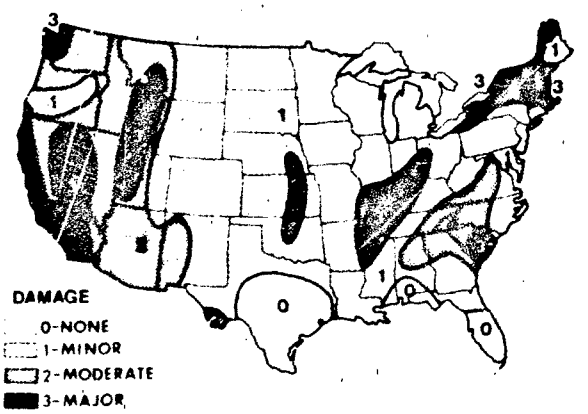


Figure 23-18 Seismic hazard zonation of the United States based on deterministic methods [Algermissen, 1969].

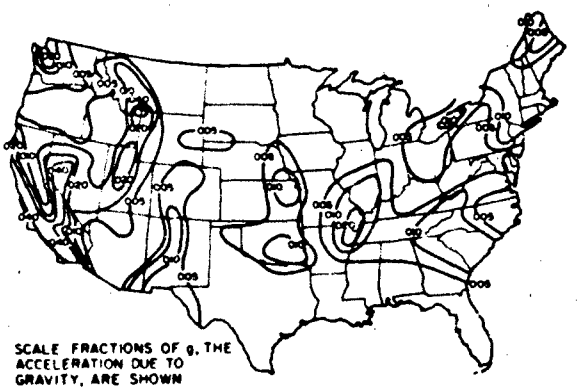


Figure 23-19 Contours of peak accelerations due to earthquakes at the 90% confidence level for any 50-yr period [Donovan et al., 1977]. Reprinted with permission from *Technology Review*, © 1977.

23.3.2.6 Premonitory Phenomena. In recent years earthquake prediction has become one of the major areas of study in seismology. The basic premise of prediction is that as stresses accumulate in the crust, physical properties of the rock will change and these changes are detectable. Compression of rock along the future earthquake rupture zone is expected to produce changes in porosity, electrical resistivity, seismic velocities, water table levels, and the radon content of ground water. Tilt and elevation changes may also be indicators of future significant events. Each of these phenomena are detectable by various means. For short term prediction, abnormal animal behavior has been associated with seismic events. In addition, the spatial and temporal distribution of low magnitude earthquakes is thought to be altered by stress accumulation and provides an indication of an imminent large earthquake. To date, significant advances have been made in earthquake prediction with at least one significant major prediction (4 Feb 1975, Liaoning

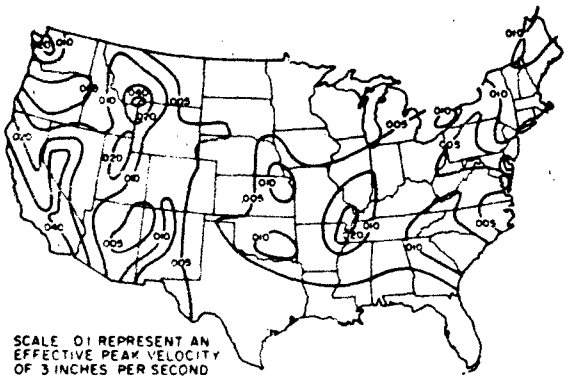


Figure 23-20 Contours of maximum peak velocity due to earthquakes at the 90% confidence level for any 50-yr period [Donovan et al., 1977]. Reprinted with permission from *Technology Review*, © 1977.

Province, China, 7.3 M_L) possibly saving thousands of lives. However, prediction is more complex than it first appeared to be. Many of the phenomena which are believed to precede an earthquake have been detected without the occurrence of an earthquake. Methods used to predict an earthquake at one place in the world do not always work at another time or location. At the present time, earthquake prediction is not an exact science and will require more time and research to achieve routine practical application.

23.3.3 Other Sources of Seismic Radiation

23.3.3.1 Seismic Noise. A seismometer placed anywhere on the surface of the earth will record virtually continuous earth motions of variable intensity. This seismic background noise is known as microseisms. Typically, the generation of microseisms is the result of cultural, meteorologic, and oceanographic conditions and is unrelated to tectonic processes responsible for earthquakes. Worldwide, the documented sources of microseisms appear to be surf or other near-shore action along steep coast lines and cyclonic activity over deep water [Bath, 1973]. More localized sources include traffic, heavy machinery, rivers, and wind-induced motion in buildings or trees. At any given site, seismic noise can be classified into three frequency dependent branches.

1. Microseisms having periods of less than 2 s are generally caused by local sources. Maximum amplitudes usually occur near 1 Hz and are rapidly attenuated away from the source [Ossing and Gray, 1978]. The localized nature of generation makes the amplitudes of these microseisms highly site dependent. These sources can normally be readily identified as cultural in origin.

2. Between 2 and 10 s lie the storm microseisms. In this band, the occurrence of high level activity has been strongly correlated with deep low pressure areas and frontal passages impinging on continental margins. The nature of the storm microseismic source suggests two facts: first, the intensity of storm microseisms are significantly higher during the winter, and second, coastal sites are affected to a much greater degree than sites in the continental interiors. The mode of propagation approaches the Rayleigh surface wave mode traveling at a velocity of 2.7 km/s [Iyer, 1964]. Amplitude attenuation is low for propagation through uniform geologic structures but is greatly increased at discontinuities such as continental margins or mountain boundaries [Ossing and Gray, 1978].

3. The long period microseisms, greater than 10 s, peak at approximately 20 to 25 s. This peak is associated with microseisms generated by the coupling of the ocean floor with cyclonic induced ocean waves over deep water. Attempts to track tropical cyclones using these microseisms have met with varying success [Bullen, 1965]. A pronounced minimum in the noise spectrum is also noted for periods between 30 and 40 s. This minimum appears to be global in character and exhibits stable spatial and temporal characteristics.

Brune and Oliver [1959] have derived minimum, average and maximum displacement amplitude curves for nonlocal seismic noise levels in the period range of 0.1 to 31.5 s. These curves, converted to amplitude spectral density, are shown in Figure 23-21 and represent limiting values of the microseism spectrum. The actual spectrum calculated for any given site over a short time interval would be expected to lie between the minimum and maximum curves but they might have a significantly different shape than the plotted curves. For example, the depicted 8-s storm microseism peak could vary between a 4 and 10 s period.

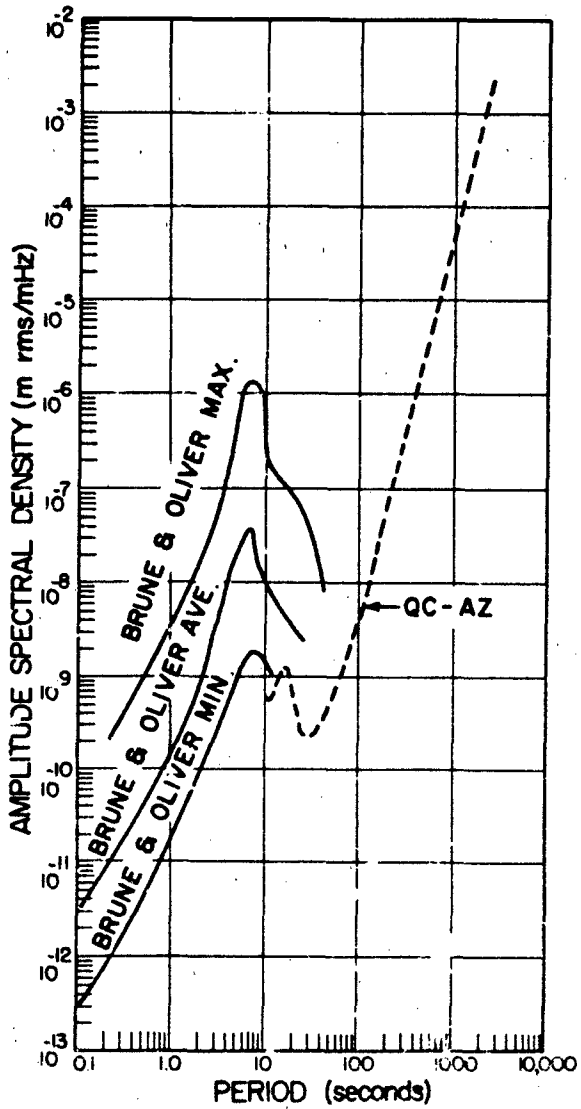


Figure 23-21 Amplitude spectral density of seismic noise from 10 to 2560 s period [compiled from Brune & Oliver, 1959 and Fix, 1972].

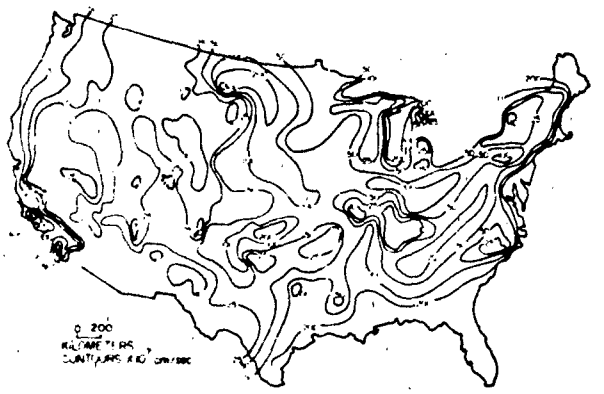


Figure 23-22. Contour map of seismic noise levels in the United States in terms of particle velocity [Ossing and Gray, 1978].

Also plotted in Figure 23-21 is the seismic noise amplitude spectral density for data recorded at a quiet site (Queen Creek, Arizona) for periods between 10 and 2560 s [Fix, 1972]. Particular care was taken in the installation of this station to minimize all locally induced noises. It is thought that this spectrum is a good representation of the minimum noise levels in this period range.

The seismic noise level observed at any site is also a function of geographic location. The proximity to coastlines and population, the type of vegetation and local geologic structure will all influence noise levels. In Figures 23-22 and 23-23, contour maps of average seismic background noise levels for displacements and particle velocities near 1 Hz in the continental United States are shown [Ossing and Gray, 1978].

23.3.3.2 Other Sources of Seismic Radiation. While earthquakes (and volcanoes to some extent) are the most prominent natural source of destructive seismic radiation, many other sources are of interest for different reasons. Some of these sources are nuclear and chemical explosions, air coupled seismic waves generated by aircraft acoustic energy (especially low-flying aircraft such as the cruise missile) or atmospheric explosions, vibrating machinery, rockbursts (collapsing mines), cryoseisms (ice fracturing), moving vehicles (trucks or landing airplanes), and people. Following are some examples of how these sources affect life on earth.

Large chemical or nuclear explosions generate significant motions which must be considered in many military construction projects. Away from the source region where cratering and other source particular effects must be considered, these motions can often be treated like earthquake generated motion. Differences in the characteristics of explosions and earthquakes, however, also allow the seismologist to provide information on the detection and dis-

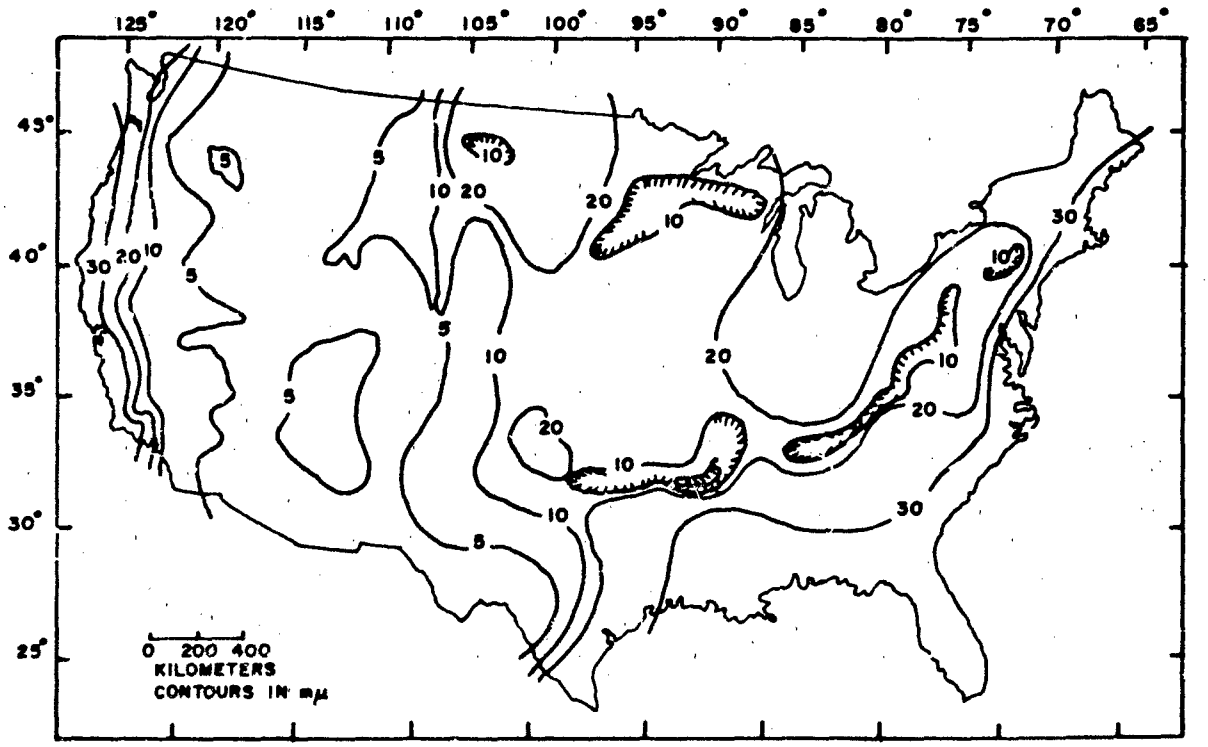


Figure 23-23. Contour map of average background noise displacements in the United States [Ossing and Gray, 1978].

crimination of nuclear explosions for the purpose of test ban treaty verification. In the siting of motion sensitive instruments, local disturbances caused by any source should be considered, especially where large concrete structures are constructed. Vibrations caused by even small quarry blasts during the critical curing period can prevent the proper cohesion from developing

The fact that moving vehicles and personnel produce seismic signals has been used to develop security systems and methods of remote battlefield sensing. Pressure waves generated by military aircraft can, under certain conditions, couple with the ground to produce large amplitude seismic motions which could affect ground facilities or be used to track the aircraft.

The study of seismic sources in the high frequency regime has almost unlimited practical applications in geologic exploration, nuclear detection, and earthquake prediction.

23.4 LONG PERIOD AND SECULAR EARTH MOTIONS

This section briefly treats measurement of crustal motion and puts some bounds on the magnitude of such deformations and the accuracy of the observations. Some of the most important techniques used are from the field of geodesy, which is reviewed in Chapter 24. In the first subsection, details of the motions caused by tectonic processes, which were reviewed in Section 23.2 will be examined; in the second, the earth tides; and in the third, other more localized types of motions, which often form the noise background and hinder the performance of instruments and systems. Specifically excluded from consideration are various geological processes, such as soil creep, landslides, glaciological activity, and volcanic activity, all of which can exhibit substantial motions.

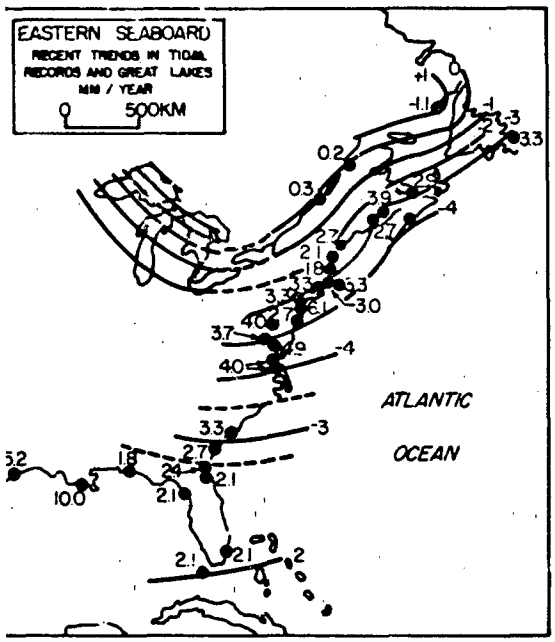
23.4.1 Tectonic Motions

We know from geological and paleomagnetic studies that portions of the earth have moved, at least in a relative sense, thousands of kilometers horizontally and tens of kilometers vertically (the top of Mt. Everest is marine limestone). We also have found from such studies that these movement velocities are not constant but are discontinuous, episodic, or even cyclical. One of the main challenges in solid earth geophysics is to directly measure such motions, relate them to the past rates determined geologically or paleomagnetically, and develop empirical or physical models to predict future motions. These goals are not without practical application, for the most dramatic form of tectonic motion is the sudden release of accumulated strain during major earthquakes with the concurrent generation of potentially destructive seismic waves. Volcanoes, which are another manifestation of tectonic processes, are capable of producing large, if localized, motion.

Recent tectonic theory has given the broad picture of relative motions between rigid plates, but the details are considerably blurred for several reasons. First, the plate boundaries are not simple discontinuities but are generally zones up to 100 km or more in width where strain can accumulate. A second problem is that plate tectonics does not explain crustal motions in plate interiors. Another problem of a different nature is the proclivity of many investigators to discover "micro-plates". Then, of course, there are large-scale secular motions that have no direct relationship to plate tectonics, for example, the mid-continental downwarps associated with deposition in sedimentary basins.

23.4.1.1 Horizontal Motions. Almost all earthquakes that cause observable surface displacements are associated with faults or fault zones. Coseismic fault displacements range from negligible for small earthquakes, to several centimeters for magnitude 5, to several meters over hundreds of kilometers for magnitude 9 shocks. The maximum total displacement for the great Alaskan earthquake of 1964 was over 25 m. The ratio of displacement to fault length ranges from 10^{-4} to 10^{-6}. Well-defined, natural or artificial (planned or fortuitous) linear or planar features (for example, fences, alignment arrays, stream channels, or shorelines) crossing the fault can be used to determine the offset. Repeated geodetic surveys (Chapter 24) are also made to determine both earthquake displacements and accumulation of strain. The U.S. Geological Survey runs an active deformation monitoring program in California, using laser geodimeters to measure the distances among benchmarks in several geodetic networks [Savage et al., 1981]. The formal errors (standard deviations) in these surveys are between 10^{-6} and 10^{-7} for lines between 20 and 50 km long. Most of the error is attributable to variations of the index of refraction along the line of sight. The strain accumulation rate in the San Andreas fault zone is about 0.3×10^{-6}/yr, or where the motion is creep on the faults themselves with little or no strain buildup, about 3-5cm/yr. Thus the expected line length changes are of the order of the present day survey accuracy and several years must elapse between surveys before reliable results can be obtained. Since several crustal plates converge in Japan, the deformation rate is high and the motion pattern complex [Mogi, 1981]. Similar programs are in progress in Turkey, India, New Zealand, the USSR, Central America, Alaska, Canada, and many others. See Simpson and Richards [1981] for examples.

23.4.1.2 Vertical Displacements. Secular vertical crustal motions rarely exceed 1 mm/yr. Some exceptions are Japan (7 mm/yr in some regions), where tectonic activity is high, the Hudson Bay region (10 mm/yr), which is still responding to the removal of the continental ice sheet 10 000 years ago, and several other, usually local, areas.



igure 23-24. Recent trends in tidal gages [Anderson, et al., 1978].

e Figure 23-24 and Walcott [1972]. Thrust and normal ulting during earthquakes generally produce relative ver- al displacements of less than 5 m, except for the very gest tremors. Classically, vertical movements have been easured by repeated spirit leveling (Chapter 24), a simple rveying procedure that has been more or less routine for e past 100 years in which errors accumulate as some nction of line length. An examination of leveling surveys Castle et al. [1976] revealed that a broad area of crustal warp of up to 25 cm had occurred between 1960 and 74 in southern California. This aseismic uplift was soon ristened the "Palmdale Bulge" for a small town 50 km rth of Los Angeles near the zone of maximum uplift. dditional surveys over the past decade showed that not ly was the bulge much larger than previously expected, tending over most of the Mojave Desert region, but also at there had been a partial collapse since 1974. Theore- ians and modelers were puzzled because of the lack of ismicity and the difficulty of fitting the bulge into the own horizontal movements of the region. At this point, ckson and Lee [1979] questioned the very existence of e bulge, attributing it to systematic rod calibration and nospheric refraction errors that can accumulate over large evation changes, such as those occurring between Los ngeles and Palmdale. After a heated debate, tests by Na- onal Geodetic Survey (NGS) [Strange, 1981], and refu- ion by the U.S Geological Survey [Stein, 1981], a good rtion of the Palmdale Bulge has disappeared. Further- ore, many other surveys that suggested recent vertical ovements were found to be contaminated by these topo- graphically correlated errors [Reilinger and Brown, 1981]. This can be seen in a paper by Brown [1978] where surveys along the east coast of the U.S. suggest substantial local movement, but the tide gauges along the coast do not. Incidentally, new survey procedures published by the NGS should eliminate these errors by tightening rod calibration standards and reducing sight lengths.

23.4.1.3 Measurement Technique . Tide gauges have long been used to estimate vertical movements (Figure 23-24). However, depending on the record length there are some problems which include long-period tides, climatic and seasonal effects (pressure and temperature), meteorological effects (wind), river discharge variations, local site stability, instrumentation, and eustatic (world-wide) sea level changes [Anderson, 1978]. It appears that sea level is probably rising at a rate of less than 1 mm/yr, although it apparently rose about 6 m in the last 18 000 years. The release of water stored as ice in the great continental ice sheets is accommodated to a great extent by the deflection of the crust under the redistributed load [Walcott, 1972].

Another method often used to infer elevation changes is repeated gravimeter surveys. After corrections for earth tides (Section 23.4.2) and atmospheric pressure variations, temporal changes of gravity can occur from either (1) displacement of the observation point along the free-air gravity gradient (-3.086 μgal/cm, $\pm 5\%$ typically), or (2) variations in the subsurface density field (for example, 0.43σ μgal/cm for an extended sheet of material with density contrast σ). These two factors generally interact, constructively or destructively, depending on the process operating. Usually this is not known, so a model must be assumed. For example, a section of material subjected to horizontal elastic compression would bulge upward, thus reducing the gravity in accord with (1) above; however, the increase in density, effect (2), would cancel this almost exactly, resulting in no net change of gravity. Whitcomb [1976] presents a number of models and the corresponding gravity and elevation changes. Measurement of both could strongly constrain such models.

Aside from the fact that some tectonic phenomena produce elevation but little or no gravity change, detection requires careful measurement and removal of spurious effects as well as the variations caused by the solid earth and ocean tides (Section 23.4.2). Lambert and Liard [1981] recently assessed their accuracy of measurement in a strongly interconnected gravity network established with La Coste and Romberg "D" gravimeters, the best instrument currently available. They estimated that with very careful work, the conservative long-term accuracy of a single station after least-squares adjustment of the network is about 4 μgal, which corresponds to a free-air elevation change of about 1 cm. Everndon [1981] found a relative accuracy (RMS) of about 20 μgal for stations along profiles in southern California. A major contributor to changes in gravity is the

movement of groundwater. For example, water filling the intergranular voids of a 1 m thick sandstone with 10% porosity would cause a gravity increase of 4.3 μgal. Lambert and Beaumont [1977] found seasonal gravity changes as high as 12 μgal due to fluctuations in ground water level. In locales where the geohydrology is simple and wells are plentiful it is perhaps possible to correct for such effects, but in complex regions such as those with many aquifers or with strong lateral heterogeneity, ground water changes strongly limit this technique. Variations in the elevation of the water surface in reservoirs, rivers, lakes, and seas also cause gravity variations at nearby stations. These can usually be easily calculated if the phenomena are adequately monitored.

Despite the problems and ambiguities associated with gravity surveys, there are two distinct advantages over leveling. First, it is much easier to make spot measurements over a wide area. Second, gravity survey errors propagate as a function of time rather than with distance, so given the proper logistical support, better precision at distant points could be achieved.

Repeated Very Long Baseline Interferometry (VLBI) measurements using satellites as sources, apparently have the precision (~ ± 10 cm over 5000 km lines) to detect relative plate motions. With expected improvements, the formal error for repeated observations should drop to about ± 5 cm or better for the horizontal components. The precision of the vertical component is typically twice that because of the observing geometry. Table 23-5, derived from Minster and Jordan [1978] shows expected plate velocities among the existing VLBI radio observatories. Thus, for example, to detect the 2 cm/yr separation between Europe and North America, caused by sea floor spreading at the Mid-Atlantic Ridge at the 95% confidence level, would require several years between two measurements. In practice, the precision can be improved by numerous repeated observations. The major limitation to VLBI measurements is the presence of water vapor in the atmosphere. At a 20° elevation angle the correction is 30 cm. With only surface measurement of water vapor the uncertainty would be about 5 cm. Radiometers are being developed to make such corrections.

Satellite laser ranging from a number of ground stations to an orbiting vehicle equipped with retroreflectors has proved sufficiently precise to measure crustal displacements of several centimeters over 10^2 to 10^3 km. This technique was used for the San Andreas Fault Experiment (SAFE) to determine the relative velocity between the North American and Pacific plates [Smith et al., 1979]. The average velocity, determined from measurements taken in 1972, 1974, 1975, and 1979, was 9 ± 3 cm/yr, which is considerably greater than the Minster and Jordan [1978] value of 5-6 cm/yr and even greater than the observed creep or strain (equivalent to 2-3 cm/yr) in the fault zone. This could be due to the fact that the plate model has a time resolution of no better than 10 000 years and that the present motion is greater

Table 23-5. Interplate velocities (cm/yr) for proposed VLBI baselines

From:	Onsala (Sweden)	Effelsberg (W. Germany)	Kashima (Japan)	Brazil	Hawaii	Kwajalein	Arecibo (Puerto Rico)
To:			NORTH AMERICAN PLATE				
Haystack, Mass.	+1.7	+1.9	−0.4	−0.2	+0.8	+0.5	+0.4
NRAO, (W. Va.)	+1.7	+1.9	−0.5	−0.4	+1.5	+1.2	+0.4
Ft. Davis (Tex.)	+1.5	+1.6	−0.8	−0.7	+3.1	+2.8	+1.7
OVRO (Calif.)	+1.4	+1.5	−0.9	−0.6	+1.6	+2.3	+1.5
Alaska	+1.0	+1.1	−0.7	−0.2	−5.2	−2.3	N/A
Algonquin (Canada)	+1.7	+1.9	−0.5	−0.3	+0.4	+0.4	+0.6
To:			PACIFIC PLATE				
Hawaii	−2.7	−2.0	−8.7	+3.5	—	—	N/A
Kwajalein	−4.0	−3.6	−9.4	+2.4	—	—	N/A
To:			EURASIAN PLATE				
Onsala	—	—	—	+1.1	−2.7	−4.0	N/A
Effelsberg	—	—	—	+1.3	−2.0	−3.6	N/A
Kashima	—	—	—	−0.1	−8.7	−9.4	N/A
To:			SOUTH AMERICAN PLATE				
Brazil	+1.1	+1.3	−0.1	—	+3.5	+2.4	+1.3
To:			CARIBBEAN PLATE				
Arecibo	0.6	0.6	N/A	−1.3	N/A	N/A	—

than that average, or that slip between the plates is distributed over a much wider area than the San Andreas fault zone, possibly including the Basic-Range province of Nevada and Utah.

A third space-based system that has been proposed is a space-based laser ranging system [Smith, 1978]. In this concept, the laser would be placed in a circular orbit at 1000 km altitude at a 50° inclination to the equator. The spacecraft would be programmed to range to large numbers of retroreflectors placed on monuments in areas of interest. Simulations show that intersite distances of 20 to 200 km can be repeated to better than ±1 cm. The space-based laser ranging system has an advantage over VLBI techniques in that atmospheric corrections are small at optical wavelengths.

Continuous or high-frequency measurements of crustal deformation are made with tiltmeters, strainmeters, and multiwavelength ranging devices. These are reviewed thoroughly elsewhere [Committees on Geodesy and Seismology, 1981] so will be mentioned only briefly. Long (750 m) laser strainmeters (interferometers) installed on the ground surface have been operated successfully for a number of years in southern California [Berger and Lovberg, 1970]. Pendulum-type tiltmeters have been used to infer long-period crustal motions, but their short baselengths make them susceptable to very small reference point instabilities from which even long-base line instruments are not entirely free. Deep borehole tiltmeters have shown more stability than those installed at shallow depths because of greater isolation from near-surface effects, but agreement among closely-spaced instruments has not been good. Long fluid tiltmeters, from which tilt is inferred from changes in the height of a liquid a known distance apart, promise greater stability by averaging out short wavelength spatial noise [Wyatt et al., 1982]. Laser surveying instruments operating at two wavelengths can correct for variations in atmospheric refraction because of dispersion effects. Operated with a third (microwave) frequency to correct for water vapor content, accuracies of a few parts in 10^9 are possible. One of these has been operated on a continual basis, ranging to a number of targets, near Hollister, California [Huggett et al., 1977].

Rapid advances are being made on both measuring and understanding earth motions. Although problems discussed in Section 23.3.3 are a limitation, long-base instruments and space techniques show great promise in producing increasingly accurate observations of crustal deformation [Committees on Geodesy and Seismology, 1981].

23.4.2 Earth Tides

Historically, it was thought that a study of the solid earth's response to the tidal effects of the sun and moon would shed important light on global mechanical properties and the rate of tidal energy dissipation. Over the past ten years a number of perturbations, which cannot be adequately modeled, have rendered this goal futile. Much better models can be constructed from higher-frequency seismic data, and earth tidal effect over a wide range of these is only 1% -2% at most. Distinguishing among these is almost certainly not within the province of tidal measurements.

However, there have been rapid advances in theory, instrumentation, and techniques, many occurring within the past ten years that make the subject a fruitful field for research. In addition, there have been and still are practical reasons for tidal studies. A list, a very brief commentary, and references, follows. For many of these, the relative accuracy must be of the order of 1%.

1. Earth tide measurement can be used to help constrain ocean tidal models using shore-based gravity and tilt measurements [Kuo, 1978; Baker, 1978].
2. Baker [1978] has shown that the response to tidal loading is primarily due to the spatially averaged Young's modulus rather than density and Poisson's ratio, which are the primary parameters in seismic wave propagation. Thus tidal measurements can provide important supplementary information in elucidating crust and upper mantle structure.
3. The tidal admittances can be predicted initially from theory, or at worst, an observational period of several months (interior continental or coastal site, respectively) quite accurately. Changes in the tidal admittance could signal changes of physical properties at depth. Such observations could be useful in studies of earthquake mechanisms and prediction [Beaumont and Berger, 1978].
4. Because of the earth's liquid outer core, there is a near-diurnal (1.00217 cpd) free wobble [Wahr, 1979] that affects the tides. Very accurate measurements of the tidal response might reveal the details of the resonance and elucidate the nature of the hydrodynamic coupling between core and mantle.
5. The tidal response can be used to map gross spatial variations in earth structure, such as the postulated partially molten rock underlying the Yellowstone region of Wyoming [Harrison, 1978]. As this response is dependent on rock parameters other than those governing the propagation of seismic waves, as in (2) above, valuable information is obtained.
6. New geodetic techniques such as Very Long Base Interferometry (VLBI) and Lunar or Satellite Laser Ranging (LLR or SLR) are approaching a precision of several centimeters (Chapter 24). The ranges, particularly the vertical component, must be corrected for tidal deformations of the sites, which can be of the order of tens of centimeters.
7. Earth tide measurements might be useful in observing possible nonlinearities in the deformational behavior of rock. These effects are much more detectable in tidal than in seismic measurements because the expected signal is much better known.

8. Advanced inertial components (gyros and accelerometers) are becoming sufficiently sensitive that the tides are detectable. They can thus provide low level, precisely-defined calibration signals.
9. Similarly, the tidal signal can act as a useful monitoring device of the response of instruments measuring long-period or secular deformations. The tidal amplitudes are of the same order as the annual tilt rate in tectonically active regions, although the periods are considerably different.
10. Tidal corrections have long been made to gravimeter surveys made for oil prospecting where the signals are of the order of 10–100 μgal (1 μgal = 1×10^{-6} cm/s^2). Normally only the solid earth tide, uncorrected for ocean loading, is applied although loading should be considered in high-precision surveys near coasts.
11. Similarly, earth tide corrections must be made to absolute gravity measurements to remove the time dependence. As the precision of these measurements is approaching 10 μgal, ocean loading terms must be included.
12. Finally, it should be mentioned that the tides affect satellite orbits, the rotational period of the earth, polar motion, and geodetic coordinates.

23.4.2.1 Solid Earth Tides. Numerous authors have derived expressions for the tidal potential on the earth caused by an exterior mass such as the sun or moon. Because of the nonlinear complexities of the orbits of the sun, moon, and earth, the tides display energy at the sum and difference frequencies corresponding to the periods

τ_1 = 27.3 days (lunar declination)
τ_2 = 365.24 days (solar declination)
τ_3 = 8.85 years (lunar perigee)
τ_4 = 18.61 years (lunar node)
τ_5 = 20940 years (perihelion)
τ_a = 1 solar day (24 solar hours)
τ_c = 1 lunar day (24.84 solar hours)

$$\text{that is, } \sigma = S\begin{pmatrix}\omega o\\ \omega c\end{pmatrix} + \sum_{k=1}^{5} M_k\,\omega_k$$

$$S = 0,1,2,3$$

$$\omega_k = 2\pi/\tau_k$$

$$\phi = \text{longitude}$$

$$\Theta = \text{phase angle}$$

in harmonic representation,

$$\sum C_i \cos(\sigma_i t + S_{ij} + \Theta_i)$$

for the long period (S = 0), the diurnal (S = 1), the semidiurnal (S = 2), or the terdiurnal (S = 3) tide-generating potential (at time t, longitude ϕ, and phase Θ with respect to Greenwich). The tidal force thus has energy in frequency bands about 0.3 cycles per day (cpd) wide, centered about 0, 1, 2, and 3 cpd with the frequency separation between components in each band corresponding to τ_1 and further splitting corresponding to longer and longer periods, as shown in Figure 23-25 [Agnew, 1979]. The number of harmonics required to adequately represent the tide ranges from 13 for a period of 1 to 6 months to 500 for decades [Cartwright and Edden, 1973]. Refer to Schureman [1958] for an exhaustive derivation of tidal harmonics.

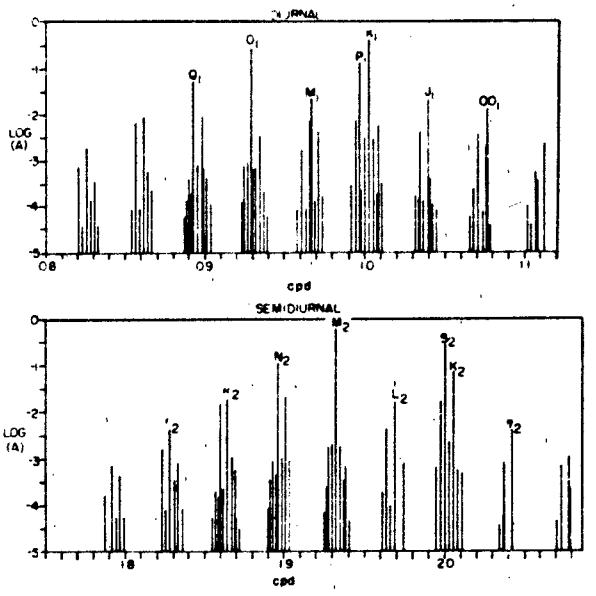


Figure 23-25. Tidal harmonics [Agnew, 1979].

The tidal potential itself cannot be observed, only changes in the vertical and horizontal components of the gravitational acceleration force g, which is the derivative of the potential. These components are the gravity tides $g = \bar{g} \cdot \bar{z}$ and tilt tides $h = |\bar{g} \times \bar{z}|/|g|$, where $\bar{z}$ is an earth-fixed unit vector parallel to mean $\bar{g}$. Occasionally tidal strain, the variation in distance between two nearby points, is measured while changes in the height of a liquid surface (the ocean tides) have been measured for thousands of years. Only in the past decade have tidal displacements of the earth's solid surface been measured with techniques such as satellite or lunar laser ranging and VLBI.

Although tides are almost always analyzed in the frequency domain, it is much more convenient to generate the theoretical tides as functions of time, the position of the disturbing body, and the position of the observer. The theoretical vertical g_m and horizontal h_m components of the

tide producing force of the moon on a rigid earth to the fourth order of lunar parallax to better than 0.04% are

$$g_m = \frac{GMr}{d^3}3\cos^2\theta + \frac{3}{2}\frac{GMr}{d^4}(5\cos^3\theta - 3\cos\theta)$$

$$h_m = \frac{3}{2}\frac{GMr}{d^3}\sin 2\theta + \frac{3GMr}{2\,d^4}(5\cos^2\theta - 1)\sin\theta,$$

where

- G = Newtonian gravitational constant
- M = mass of the moon
- d = distance between the centers of the earth and moon
- θ = zenith distance (angle) to the moon at the observation point

The vertical component is commonly expressed in microgal (10^{-8}cm/s^2) with reversal of sign to correspond to gravity observations, that is, a maximum tidal force produces a gravity minimum. The horizontal force is usually divided by the local value of gravity and multiplied by 1000, yielding the deflection of the vertical, or tilt, in nanoradians (10^{-9} radians), which may be converted to milliarcseconds (ms) by multiplying by 0.2063.

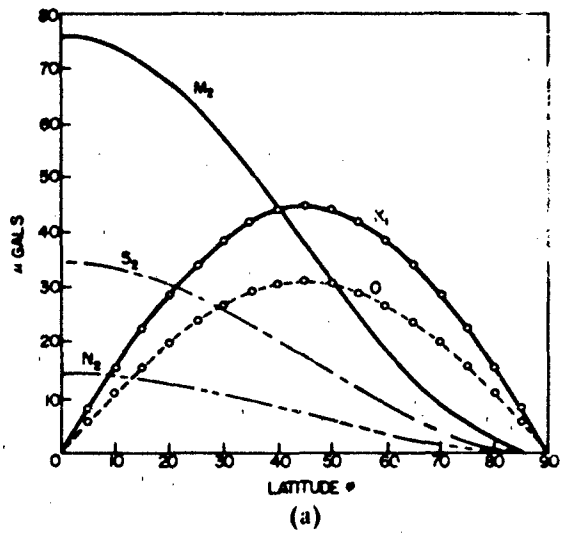
For the sun, the first term in each of the two expressions given above is adequate to the 0.002% level. The solar induced tides are 46% of the lunar. Numerous computer programs are available to calculate tidal time series [Cabaniss and Eckhardt, 1973].

On a rigid earth the maximum equilibrium tidal ranges are

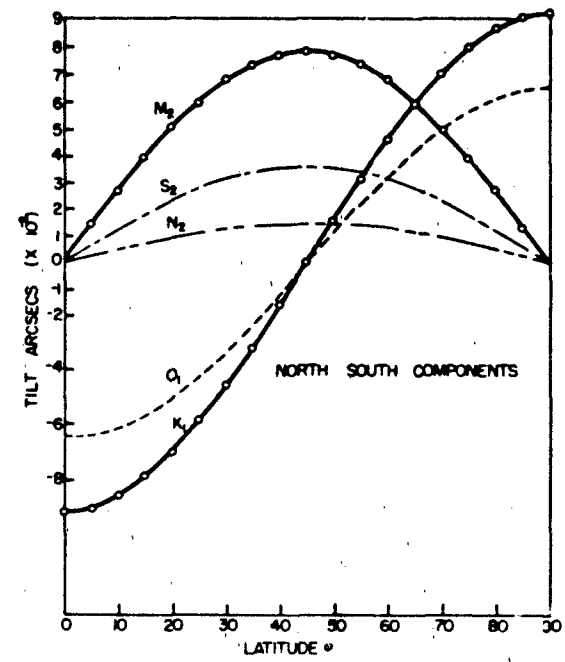
	Gravity	Tilt	Geoid (or fluid earth)
moon	164.5 μgal	34.6 ms	53.4 cm
sun	75.8	15.5	24.6
Total	240.3	51.1	78.0

The theoretical amplitudes of gravity and tilt of the main tidal components as a function of latitude are given as Figure 23-26.

The earth is neither rigid or fluid but essentially elastic at tidal periods. Because the period of the longest free oscillation (54 min) of the earth is more than 10 times shorter than the tidal frequencies, the earth has more than adequate time to adjust to the tidal forces and its response is nearly independent of frequency. Thus the earth tides are of the equilibrium type: that is, the relative amplitudes and phases of the constituents should correspond to the theoretical ratios. Such behavior of a solid, spherical, oceanless earth can be described in terms of three characteristic numbers, h, k, and l. Historically, the reason for studying earth tides was to determine these so-called Love numbers, but they



(a)

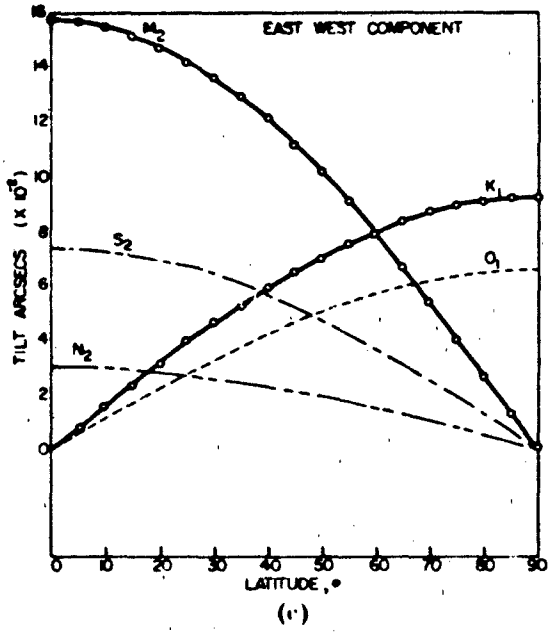


(b)

Figure 23-26. Amplitudes of the main diurnal and semi-diurnal earth tidal components versus latitude: (a) gravity, (b) north-south tilt, (c) east-west tilt.

can be derived much more accurately from seismic measurement because the tides are so perturbed by various indirect effects which will be discussed later. They are defined as follows:

- h represents the ratio between the height of the tide on the elastic earth to the theoretical equilibrium height;
- k is the ratio of the additional potential due to the deformation of the elastic earth to the tidal potential; and



(c)

Figure 23-26 (*Continued*)

l is the ratio of the horizontal displacement of the elastic earth to the theoretical tide.

See Melchior [1966] for futher discussion.

From seismic data, h = 0.603, k = 0.298, and l = 0.084 with a maximum variation of 1%–2% for reasonable seismic models. The ellipticity of the earth introduces latitudinally-dependent terms, but the effects are small [Wahr, 1979]. Thus the solid earth tides can be considered known on the surface of an oceanless spherically symmetrical earth. Here a tidal gravity meter would measure the equilibrium gravity tide $g_m + g$, multiplied by the constant $\delta = 1 + h - (3/2)k = 1.160$. Similarly, a tiltmeter would sense $(h_m + h)/g$ multiplied by $\gamma = 1 + k - h = 0.70$.

The gravity tides are increased by over 15% and the tilt tides diminished by about 30% because of elasticity. Surface strains depend on both h and l; the expressions are complex and will not be given here. See Agnew [1979] for details.

The one exception to the frequency independence of the earth to tidal forces is near 1 cpd, where the resonance caused by core-mantle coupling becomes important. Wahr [1979] finds from theoretical considerations that this effect should not exceed 20%. A number of investigators, including Levine [1978], have detected the phenomena but have not been able to elucidate the details of the resonance.

23.4.2.2 Ocean Loading Effects. The usual cause of the discrepancy between tidal theory and observation is the result of ocean tides, which can easily account for 10% of the total earth tide in gravity, 20% in strain, and 90% in tilt. The ocean tidal effect acts in three ways: (1) by the direct attraction of the water, (2) by the physical bending of the crust beneath the weight of the water, and (3) by the distortion of the equipotential from the first two effects. Near seacoasts, the region of the ocean closest to the observation point causes the greatest perturbation and local ocean tidal models reasonably match the observations. In the centers of continents, however, all of the oceans are significant contributors (several percent) to the total signal. This is particularly true for the gravity tide, because the vertical component of the mass attraction of water decreases only as r^{-1} while the effect on strain and tilt falls off as r^{-2}, where r is distance.

Although the existence of ocean loading effects has been known for years, it was treated simplistically or even ignored until the 1960s for two reasons: (1) lack of methodology to calculate the effect loading and (2) poor open-ocean tidal models. Rapid advances have been made in both areas over the past decade, both requiring high speed computers.

The first attempts to correct for ocean loading used the solution for the displacements of an elastic, homogeneous, isotopic half-space subjected to a normal surface point load, which is called the Boussinesq problem. See Farrell [1972] or Lewkowicz and Cabaniss [1980]. When the attraction of the water mass and effects of the deformed potential are taken into account, the half-space model is often adequate for tilt corrections because most of the ocean loading contribution occurs within 1000 km of the observation point. Because the earth is spherical and radially heterogenous, more realistic solutions are required. Takeuchi [1951], Longman [1962], and Kuo [1969], for example, made important contributions.

The complete problem was solved by Farrell [1972], who computed Green's functions for vertical displacement, gravity, tilt, strain, and horizontal displacements for a unit load on several realistic standard earth models. Using these functions $G(\Delta)$ for any of the above quantities, a mass M an angular distance Δ away would produce a change $MG(\Delta)$ in that quantity. Given an ocean tide model, one can convolve the tidal heights, with Farrell's tabulated Green's function, to compute the ocean loading. In practice, the oceans are compartmentalized and given the average tidal (complex) amplitude within each. Far away from the computation point the grid spacing can be coarse, but for points near coasts, the discretization must be comparatively fine [Goad, 1980].

The finite element method [Zienkiewicz, 1967] has been used by Beaumont and Lambert [1972] to deal with more complex structures, such as those with lateral variations, than those considered by Farrell. Such models are especially useful for computing tilts from nearby loads. This technique is quite adequate to an angular distance of about 30° but deteriorates beyond that because of the neglect of self-gravitation terms and the rigid bottom boundary required at the base of the model.

Concurrently, major advances have been made in the development of ocean tide models, a much more formidable

task than that for the earth tides. This is primarily because the natural frequencies of many of the ocean basins, gulfs, etc., are near the frequencies of tidal forces, so the dynamic resonant terms are important. Furthermore, nonlinearity plays a much greater role, particularly in shallow seas, because of friction and the nonlinearity of the equations for waves in shallow water, thus producing nonlinear tides called overtides by analogy with overtones.

Ocean tidal modelers integrate the Laplace Tidal Equations (LTE) to various boundary conditions, such as coastal tidal observations. Problems include dealing with realistic bathymetry, coastlines, and energy dissipation. In addition, only in the past 10 years have the self-attraction, loading, and solid earth tide term, all significant, been taken into account. Parke [1978] gives a review of the progress through 1978.

More recently Schwiderski [1979, 1980] has published a series of models for many of the major tidal constituents (M_2, O_1, K_1, N_2). These are available in geographical grid form on magnetic tape. In 1981, the Permanent Commission on Earth Tides, formed under the International Association of Geodesy of the International Union of Geodesy and Geophysics, adopted the Schweiderski models as the standard for ocean loading corrections.

Two problems still exist with ocean tide models. The first is that tide gauges are used for defining the boundary conditions. Almost invariably they are located in anomalous sites, that is, sheltered harbors, where tides are distorted for a number of reasons. Open ocean measurements using pressure transducers on the ocean floor are needed, but these are sparse and the number only slowly increasing [Cartwright, 1982]. Another problem is that world ocean models lack sufficient detail and accuracy near the coasts for earth tide measurements taken there. In some locales, for example the North Sea and southern California, local models are available which can be integrated with the coarser grids. Future advances will probably be made with satellite altimetry (Chapter 24), which has become sufficiently precise (fractions of a meter) to directly detect ocean height changes. The major problem to be overcome is orbital errors, an active area of research in the satellite geodesy community.

23.4.2.3 Cavity, Topographic, and Geological Effects. To avoid the overwhelming noise generally caused by temperature changes, rainfall, and soil moisture variations, tidal strain and tilt measurements have traditionally been made in underground mines, tunnels, or special cavities excavated for specific geophysical purposes. These openings, often in mountainous terrain of complex geology, are almost always excavated by explosions. Furthermore, short-base (less than 1 m) tiltmeters have often been placed in special "niches," carved with hand tools or jackhammers, in tunnel walls [Melchior, 1966].

The spatial variations of tidal tilt parameters measured in Europe through the early 1970s showed a disturbing lack of smoothness far beyond that attributable to crustal structure, inadequate ocean loading corrections, or instrumental unknowns. At some sites, significant changes occurred among instruments separated by meters or less [Lennon and Baker, 1973]. In 1973 King and Bilham pointed out that the cavities in which most instruments had been operated were subject to distortion by tidal strains. Tiltmeters installed, for example, in a circular tunnel at A and B in Figure 23-27 will

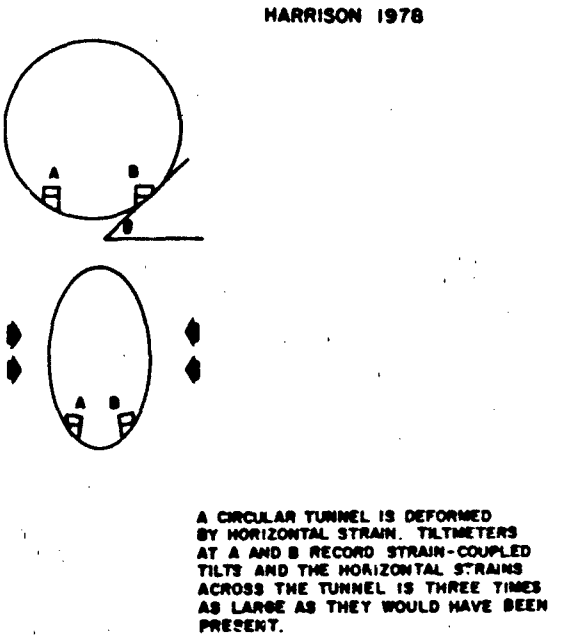


Figure 23-27. Strain-induced tilt in a circular tunnel [Harrison, 1978]

record strain-induced tilts equivalent to twice the horizontal strain. Similar effects are generated by topographic and geological irregularities. Since most theories dealing with the earth tides or seismic wave generation and propagation consider large-scale deformations on the free surface of comparatively simple earth models, it is difficult to relate these to the measurements made in locally complex regions.

Harrison [1976] has treated analytically the effects of ellipsoidal and cylindrical cavities and has made finite element calculations of more complex cavities, topography and geology. For example, the floor of a long tunnel shows no strain-tilt effect along the axis, nor does the side of a vertical borehole. Cracks and narrow cavities, however, induce large strains and tilts. Topographic distortion can also be substantial, with strain effects of several hundred percent. Berger and Beaumont [1976] have corrected the results from a number of earth strain observatories for cavity, topographic, and geological (that is, material inhomogeneity) effects using finite elements. These corrections generally improved the agreement between theory and observation, primarily because strain measurements are usually made over long ($\sim$30m) baselines along tunnels, where the corrections are normally small and cracks and inhomogeneities relatively

insignificant. For measurements made with short-base tiltmeters, however, the case is much more serious, not because of the complexity of the known cavity, which can be modeled sometimes expensively, but because of the unknown fracture pattern in the vicinity of the tiltmeters. Thus all earth tide measurements made with these instruments were essentially useless, except those made in vertical boreholes and possibly those designed to monitor temporal variations in the tidal parameters.

These cavity, topographic, and geological effects are frequency independent, including both slow tectonic movements and sudden earthquake offsets, as well as tides. Gravity measurements, however, are vitually unperturbed. To be free of such perturbations, tilts must be measured with long base instruments in shallow trenches excavated in flat terrain or borehole tiltmeters referenced to the side of a vertical hole [Wyatt et al., 1982].

23.4.3 Other Motions

In this section the causes, magnitude, and methods of measurement of other types of earth motions are discussed. Also addressed are those spurious motions that are often measured by an investigator who is attempting to record tectonic or tidal deformation. These include the effects of surface loading, groundwater and soil moisture, and temperature fluctuations, as well as those from the pumping of oil and water from wells. While probably not as geophysically interesting as the broader scale phenomena, these deformations are often larger and their causes more subtle, and they are often more likely to affect the testing and operation of sensitive systems, such as gyrocompasses.

23.4.3.1 Motion Background. In Section 23.3.3.1, the noise spectrum for short-period motions was discussed in terms of acceleration, velocity, and displacement. The large site-dependent amplitude variations cited are a result of both the type and strength of noise source and the proximity and characteristics of the measurement site; that is, the amplitude is much higher at a point on unconsolidated sediments near a freeway than on a pier on bedrock at an isolated location. Also, the short-period noise spectrum is characterized by a number of strong peaks and pronounced minima.

In contrast is the background strain spectrum derived from measurements at one site in Colorado and two sites in California [Berger and Levine, 1974]. They showed (Figure 23-28) that the background rises at an almost constant slope of f^{-2} over 10 decades (10^{-2} to 10^{-8} Hz) of frequency corresponding to periods ranging from 10 ms to about 1 yr, and that the power levels for two different types and lengths (30 m versus 750 m) of instruments at very different sites, Colorado mine versus California desert surface, are almost identical. Of course certain portions of the spectrum are not time or space invariant: large earthquakes excite the normal modes and the microseisms show pronounced peaks at 8

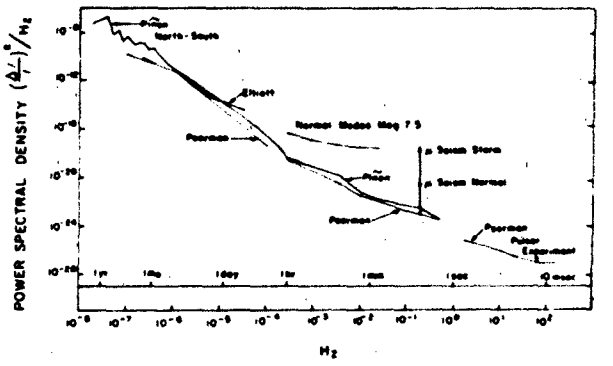


Figure 23-28 Background strain spectrum [Berger and Levine, 1974]

and 16 s periods. The earthtides were removed prior to computing the spectrum. Wyatt [1982] presents similar data on motions of reference monuments. These authors surmise that these observations are probably representative of the strain at any reasonably quiet site in the world or are an upper limit. The tilt spectrum should be essentially identical because the amplitudes of the rotational components, including tilt of the strain tensor, should be about the same as those of the linear ones.

23.4.3.2 Surface Loading. Ocean loading has a significant effect on earth tides. Their removal is particularly difficult because both occupy the same frequency space. Other variable or moving loads (trucks, trains, people, water in lakes or rivers, snow) can produce measurable deformations of observation sites. Their contributions can sometimes be calculated and removed, but almost always order of magnitude estimates can be made.

Farrell [1972] presents the method for computing Green's functions (point load response) for realistic earth models for tilt, gravity, and strain. Far simpler are the expressions for displacements at any depth in an elastic, homogeneous half-space, also included in Farrell's paper. These, as well as formulas for tilt of vertical and horizontal line elements, are given by Lewkowicz and Cabaniss [1980]. Near the load, the horizontal tilt can exceed the vertical by as much as a factor of three, the exact ratio being a fraction of Poisson's ratio. A number of authors, including Kuo [1969], have derived approximations of the response of horizontally layered elastic models to vertical point loads. The finite element method is useful for more complex models and loads. Formulas for distributed loads, such as discs, hemispheres, and hemi-ellipsoids, are given in a number of texts including Farrell [1972] and Herbst [1979]. Usually, effects of distributed loads can be very adequately and easily approximated by subdividing the surface into suitably sized cells, computing the point load contribution of each, and summing the results.

Numerous cases of atmospheric pressure changes causing significant (several times the tidal amplitude) tilts and

strains appeared until the early 1970s in the literature [Melchior, 1966]. Some authors postulated the coherent movement of crustal blocks hundreds of kilometers wide and partially decoupled from adjoining blocks. Herbst [1979] argues convincingly that such observations must result from local or instrumental phenomena. He modeled an unusually intense cyclonic low of 50 mb and found that the maximum vertical displacement for any reasonable load distribution was 16 mm at the center of the loaded region, with a tilt of 45 nrads at the center and edge, far less than the typical tidal tilt of 100 nrads. He was able to show that the apparent correlation with air pressure changes was actually a result of the concurrent rainfall draining into the fractured rock near his borehole instruments, a phenomenon that will be treated in greater detail later. This is not to say, of course, that air or water pressure changes acting normal to sloping terrain cannot produce significant tilts. This is especially true in clays and other low-permeability materials where diffusion of pore fluids becomes important [Zschau, 1976]. Direct pressure effects on instruments act by either distorting compliant sealed cases or buoying the mass in unsealed cases. The variations of mass attraction of the atmosphere must be considered in very precise earth tide gravimetry.

Vehicles and fluid storage tanks can produce significant tilts and tilt rates at nearby test facilities built on unconsolidated soils. The load produced by a vehicle can be modeled as a single point source at distances greater than three times the vehicle diameter, and as the superposition of four point loads, the wheels, at lesser distances. Lewkowicz and Cabaniss [1980] examined the tilts produced by a 1500 kg vehicle. They found the maximum vertical tilt at a depth of 1.3 m was 1.25 μrads at a horizontal distance of 1.3 m (nearest axle), which rapidly diminished to 0.2 μrads at 6 m. The maximum, unmeasured horizontal tilt was calculated to be a factor of three or larger. Water tanks or similar loads can be treated as disc loads nearby and point loads at distances greater than five times their radius.

23.4.3.3 Thermoelastic Effects. As materials are heated, they almost always expand, thus producing stresses and, usually, strains and tilts. These are almost always observable on instruments installed in buildings and on or near the earth's surface. Numerous authors give formulas, tables, and curves for calculating heat flow from various types of source distributions through media of some complexity. Finite element methods are also particularly suited to making such computations. In geological materials, problems almost always arise because the relevant material parameters are unknown; they vary both spatially and temporally, and the movement of groundwater and soil moisture greatly complicates the issue. Thus, most models used in geophysics are comparatively simple.

Harrison and Herbst [1977] discussed thermoelastic strains and tilts from the geophysical instrumentation standpoint. The most important components are the diurnal and annual temperature waves, which penetrate as damped progressive waves in semi-infinite halfspaces. The skin depth at which the amplitude decays to 1/e is about 1 m and 20 m in typical earth materials for these two periods. In this model the rock or soil can expand only vertically because it is confined laterally by equally heated materials, and horizontal compressive stresses develop; therefore, there is only vertical strain, no tilts or horizontal strains. However, instruments installed even in such an ideal material could be affected by asymmetry of their mounts or inherent temperature sensitivity. In the real world with finite surface topography and variations of thermal properties, sideways expansion can take place, thus generating tilts and horizontal strains. Harrison and Herbst [1977] derive an expression for tilts and strains on uniformly sloping topography with a constant thickness thermal boundary layer. For a temperature wave of 283 K on a slope of 6°, the strain is 2×10^{-6} and the tilt is 20 μrads, both falling off at greater depths as does temperature.

For more complex topography, finite element calculations have proved quite satisfactory. Herbst [1979] was able to explain annual tilts of 0.625 and 0.125 μrads observed at depths of 15 and 30 m, respectively, with a 2-D, three-layer, finite element model of this site, which had an annual temperature wave of 281 K amplitude at a depth of 1 m. This technique should be useful in assessing the temperature sensitivity of various structures.

In buildings and other facilities, especially those exposed to solar insolation, the tilts can be very large. As one side of a building expands, it bends, flexing the floors and transmitting stresses to its foundation. These can affect test piers completely decoupled from the building floors and walls unless such piers are some distance from the foundation. Figure 23-29 [Tsutsumi, 1970] shows the tilt on a pier in an insulated building standing on exposed bedrock on a low, forested hill and with no temperature control. The annual variation is over 100 arcseconds (about 485 μrads). Even in buildings with some temperature control, thermoelastic effects are difficult to avoid because it is almost impossible to change temperatures uniformly throughout an entire room. Figure 23-30 shows the tilts and air temperature (clearly diurnal) measured near the center of a massive concrete pier. First, note that the phasing between temperature

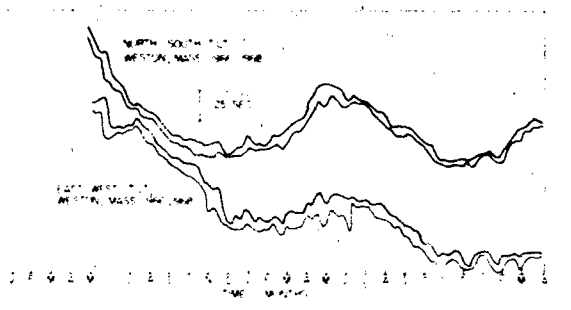


Figure 23-29. Weston vault tilt [Tsutsumi, 1970]

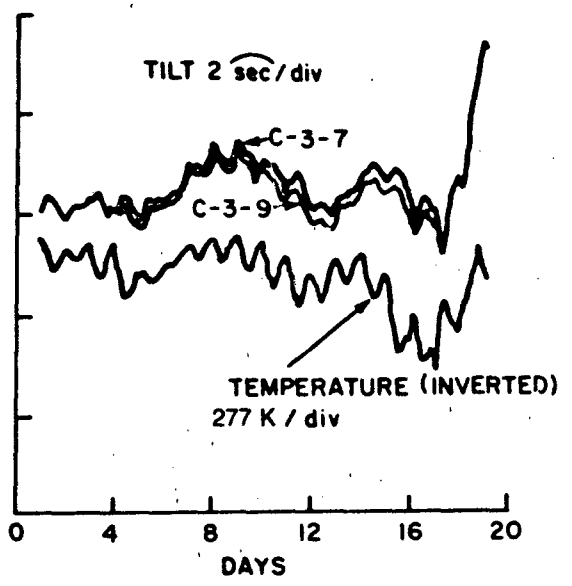


Figure 23-30. Haskell pier tilts and temperature.

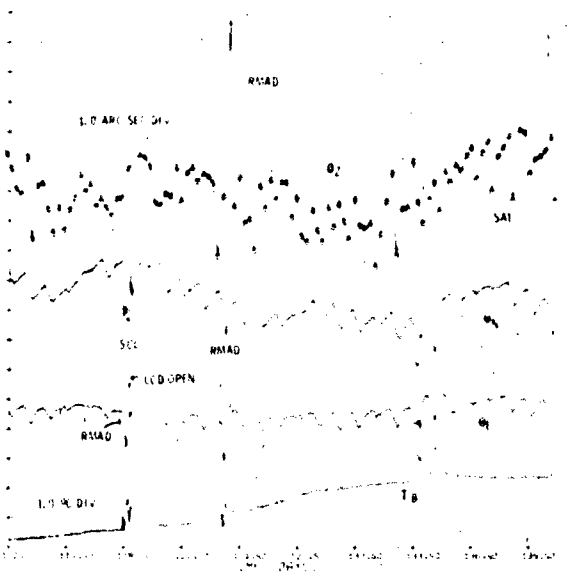


Figure 23-32. Silo rotations and temperature.

and tilt is not constant. This is because the pier is affected by thermoelastically-induced tilts, transmitted through the foundation and bedrock to the pier, caused by the outside air temperature and insolation variations that are out of phase with the room temperature. Second, note the disagreement between two tiltmeters set on the same plate several centimeters apart. The differences between two other pairs of adjacent instruments are shown on an expanded scale in Figure 23-31. Blk 2 and Blk 1 are in the same housing and the discrepancies, probably due to small misalignments, do not increase with temperature; however, the C3 and C4 obviously have different temperature coefficients, an example of the direct temperature effect. A clear and dramatic example of thermoelastic rotations is presented in Figure 23-32, which shows rotations about the vertical axis ϕ_z and two horizontal axes, Θ_{South} and Θ_{East}. Also shown is the temperature (T_B) measured at a depth of 10 m beneath the ground surface in a missile silo. Aside from various activities in the silo, the most striking feature is the strong diurnal character of the rotations and complete lack of daily fluctuation in air temperature. The tilts even have the shape of the typical diurnal outside temperature curve, a rapid change of temperature until noon and a decreasingly slow return during the afternoon and night. The troughs are within one hour of local noon and the measured flexure is in an ENE-WSW direction. The effects of heating the exposed concrete are obviously transmitted thermoelastically far down the structure. Such large motions would never be predicted from half-space calculations.

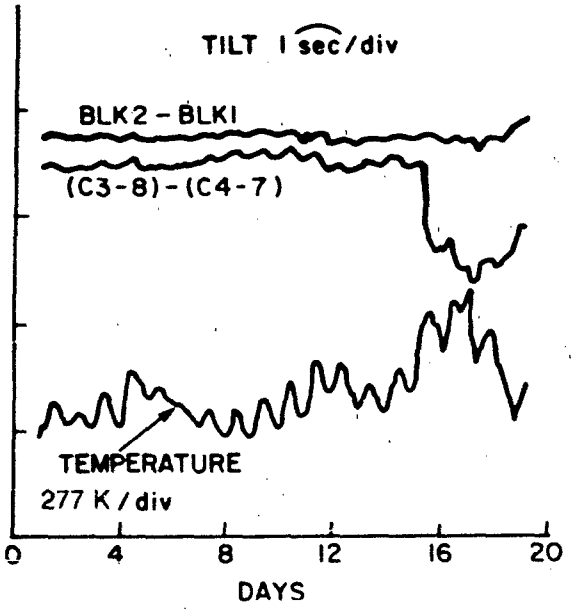


Figure 23-31. Haskell difference tilts and temperature.

23.4.3.4 Soil Moisture and Groundwater. Soil moisture and the movement of groundwater through rock fractures, soil, and other unconsolidated sediments can cause large ground motions, and these are the most difficult to model and predict because of the complex mechanical behavior of soil materials.

The most obvious phenomenon is the lowering of the ground surface in regions where oil or water is being extracted from poorly consolidated sediments. Dramatic examples of this occur along the Texas-Louisiana Gulf Coast, Las Vegas, and the San Joaquin Valley of California, where

the elevation has changed 8.5 m from 1926 to 1972 due to groundwater pumping. Subsidence rates can be extremely high: 13 mm/yr in Louisiana and as great as 24 cm/yr in California [Committees on Geodesy and Seismology, 1981]. In Arizona, between Phoenix and Tucson, subsidence caused by groundwater pumping has produced a fault over 13 km long and with offsets up to a meter [Holzner et al., 1979]. The mechanism is the removal of the supporting pore pressure between the soil particles, particularly the clay fractions. The process can be halted and partially reversed by recharging the affected strata. In Long Beach, California, serious subsidence caused by oil pumping was arrested and is no longer a threat despite continued exploitation because water is injected at the same rate as petroleum is extracted. Damage appears in the form of ruptured well cases, coastal flooding, shifted foundations, and improper flow in surface canals and aqueducts. Subsidence patterns are mapped by releveling and by measurements on well casings. GPS and VLBI surveys (Chapter 24) should be applicable here.

The usually unpredictable effects of the subsurface movement of water can seriously hamper geophysical measurements. Herbst [1979] was able to show that surface water from rain and melting snow moving into vertically-fractured rock caused large, erratic tilts of borehole instruments. Apparently the pressure head differences in adjacent fractures caused sufficient bending movement in the rock between them. The noise levels created were so great they significantly degraded the tidal results, even from an instrument as deep as 30 m, and made it impossible to measure tectonic motion, if any. Attempts were made to model this mechanism using linear systems techniques to predict the effects and correct the tilt measurements given the amount of precipitation and temperature, but these were unsuccessful because of the nonlinear, time-varying nature of the phenomenon and other variables.

One of the most successful and stable sets of long-term strain measurements is from the Pinon Flat Observatory in southern California [Berger and Levine, 1974], but these results too are contaminated by tilts and horizontal motions of the large granite piers at each end of the three 750 m laser strain-meters. Wyatt [1982] has made an extensive study of six years of Pinon Flat data and concludes the random motion in the partially weathered bedrock, although obviously related to temperature and precipitation, is unpredictable. It seems that small-scale inhomogeneities and the subtle variations in the stress field caused by soil moisture and temperature variations are too time-variable to model. These effects are also obvious in the Pinon Flat shallow tiltmeter data [Figure 23-33, from Wyatt et al., 1982]. Beta and Delta are Kinemetric instruments (sensor in the base of a 1.1 m stainless steel tube) sand-packed into holes 4.5 m deep. ADL is an A. D. Little instrument at 26 m, and LFT is a 535 m long fluid tiltmeter. The precipitation effects on Beta and Delta are obvious, as are the thermal ones in their spectra [Figure 23-34, from Wyatt et al., 1982]. Here the meteorological inputs have raised the noise levels of the

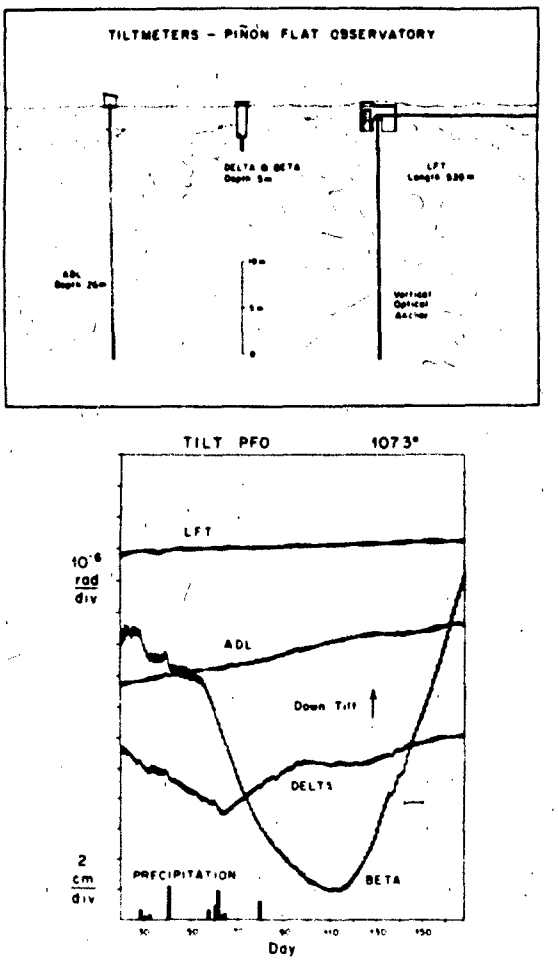


Figure 23-33. Pinon flat tilts [Wyatt et al., 1982].

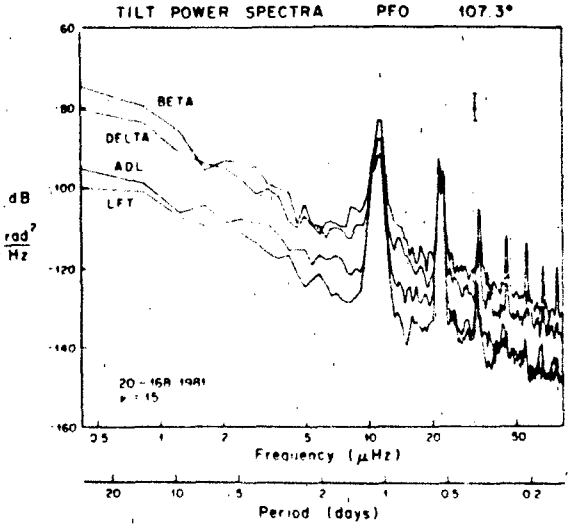


Figure 23-34. Pinon flat tilt spectrum [Wyatt et al., 1982].

shallow instruments about 20 dB above those of the other two and rendered them useless for the measurement of long-term tectonic movements. The unpredictable monument motions also have serious implications for strain measurements. Wyatt [1982] points out that a 7 km long instrument would be required to measure the approximately 1×10^{-7} yearly strain-rate observed in California to within 10%. His solution has been to reference the surface monuments to the deeper, presumably more stable, bedrock with an "optical anchor," which is essentially an interferometer with two paths inclined at 45° to the surface down to a depth of 15 m. Even with lines of 20-50 km, the stability of monuments is important: 0.5 mm being 0.1×10^{-7} on a 50 km line, movement easily caused by thermal or soil creep phenomena.

Another example of the problems associated with continually monitoring secular crustal movements is shown in Figure 23-35 [Cabaniss, 1978]. Data from three borehole tiltmeters installed 3 m from the bottom of cased holes 100 m apart and 100 m into bedrock beneath 20 m of overburden have been rotated into North and East components. Relevels have not been removed from the first eight months of data from #1 or two months from #2. The rapid drift starting at Day 740 seen on #2 North actually occurred almost

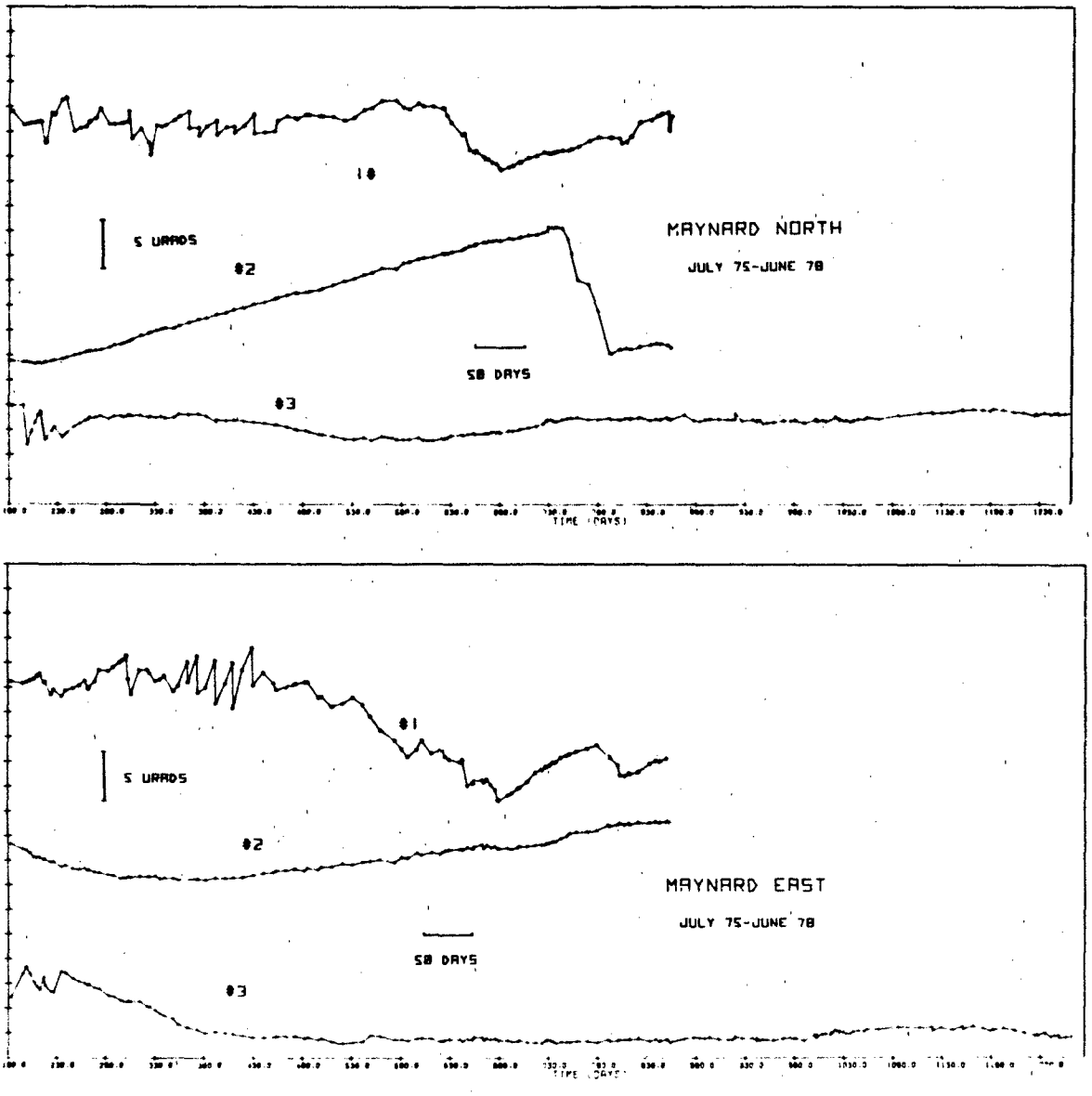


Figure 23-35. Maynard tilts

equally on both channels but in opposite directions. It lasted for 45 days and was possibly caused by corrosion of the mild steel casing. Instrument #3 displayed remarkable stability over the three year period, the net change being 0.3 μrad, with a maximum excursion of 2.0 μrads. The #3 drift rate is about 0.1 μrad/year down to the SW. Leveling [Brown, 1978] and tide gauge data (Figure 23-24) suggest a rate of 0.01 μrad/year down to the North.

23.4.3.5 Conclusions. A host of extraneous motions make the measurement of tectonic and other geophysically interesting crustal deformations difficult, particularly for long time periods. These same motions are usually the ones that cause structural damage and corrupt test results and operation of sensitive systems in near-surface facilities. In most cases, the usual geophysical deformations are of little or no consequence except, of course, for earthquakes and volcanic activity.

The measurement of long-term crustal deformation is highly dependent on reference point stability. At the present time only repeated annual surface-based geodetic measurements over 20 to 50 km lines and continuous long base tiltmeter measurements appear to have the necessary precision to monitor the strain-rates observed in California. Space-based techniques have the potential for extending this precision over longer distances and vastly increasing the frequency of measurement.

23.5 ROTATIONAL MOTIONS

Rotational motions are typically classified by geophysicists as either azimuthal rotations or tilt. Azimuthal rotations are those motions that occur in a horizontal plane about a vertical axis, while tilts occur about axes in the horizontal plane. Both the relative motion of a reference device over time and the absolute motion with respect to a permanent reference, such as an astronomical or earth fixed coordinate system, are of interest to geophysicists. However, since recent research has emphasized improving azimuth measurement accuracy, the discussion in the following sections will focus primarily on azimuthal rotations.

Knowlege of true azimuth or azimuthal rotations has applications for pointing, tracking, navigation, and evaluation of geophysical phenomena. Historically, except for surveying applications, the Air Force interest in making precise azimuth measurements has been limited to (1) establishing or verifying the azimuth of a reference mirror in a test laboratory, (2) establishing a reference for a tracking device such as a radar antenna, or (3) initializing and/or verifying the performance of an inertial navigation system.

Until recently, the accuracy required for these functions has been on the order of 5 arcseconds of uncertainty. Since short period motions of azimuth references are typically less than or equal to this measurement uncertainty, azimuth was normally considered to be a quasistatic parameter and was measured intermittently (every few months or longer). A time history, referred to as a "data base," was frequently used to describe the long period or step motions of an azimuth reference.

Little improvement was made to either instrumentation or measurement techniques for several decades until the need for high accuracy missile navigation systems forced the R&D community to explore new technology for azimuth measurement. These systems require measurement accuracy of less than 1 arcsecond over much shorter measurement periods. Thus, the smaller geophysically and culturally induced motions of the target point as well as those of the measuring device are now significant, and classical measurement techniques are nearly obsolete.

Recent research has advanced both instrumentation and measurement techniques. This section addresses the most common of these advancements, including their improvements as well as their limitations. The discussion highlights both instrumentation errors and the geokinetic influence on instrument performance.

23.5.1 Definition of Azimuth

Prior to discussing the various measurement techniques, however, it is first important to define azimuth to ensure consistency in performing and reporting azimuth measurements. Generally, the azimuth of a point is defined in terms of either an astronomical or geodetic coordinate system [Department of Defense Glossary of Mapping, Charting, and Geodetic Terms, 1973].

The astronomic azimuth between two points P and Q is illustrated in Figure 23-36 [Mueller, 1969]. It is defined as the angle "α" formed by the intersection of the observer's astronomical meridian with the plane containing the observed point and the true normal (vertical) of the observer, measured in the local horizontal plane clockwise from true North.

The geodetic azimuth is the angle between the geodetic

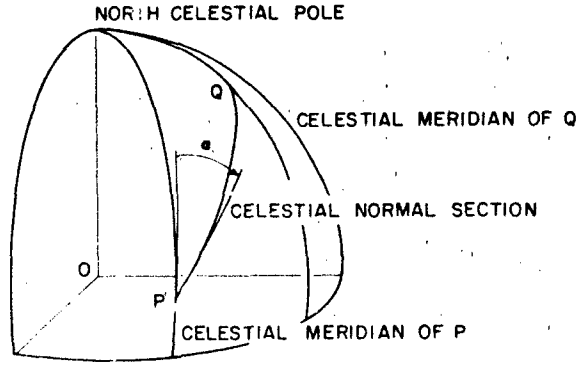


Figure 23-36. Azimuth on the celestial sphere [Mueller, 1969]

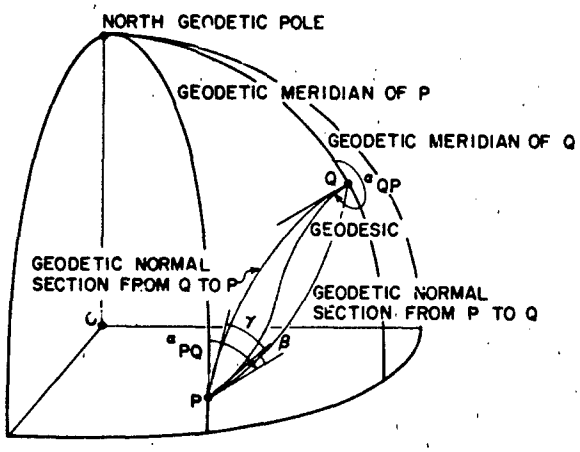


Figure 23-37. Azimuth on the ellipsoid [Mueller, 1969].

meridian and the tangent to the geodesic line at the observer, measured in the plane perpendicular to the ellipsoidal normal of the observer, clockwise from South. This is illustrated in Figure 23-37.

It is sufficient to say that the difference between the two definitions is the figure of reference. Astronomical azimuth refers to the astronomic rotation axis of the earth while geodetic azimuth refers to an ellipsoidal model of the earth and its respective mathematically defined pole. For our discussion, we will refer only to astronomic azimuth.

23.5.2 Conventional Azimuth Determination

Normally, we want to know the azimuth of a line of sight normal to the mean surface of some target or test fixture which is located at some angle with respect to true North. Thus, the problem is divided into two parts: (1) determining the azimuth of the line between the observer and a known reference such as a star at a specific epoch of time, and (2) measuring the horizontal angle between the star and the target. Star observations are made using theodolites. The circumpolar star Polaris is preferably observed since its near-zero azimuth minimizes the effect of errors in knowledge of the station latitude on the azimuth determination. The celestial coordinates of Polaris required to compute its azimuth are obtained from a star catalogue.

Figure 23-38 illustrates how this technique may be used in a test facility to establish an azimuth reference with a fixed mirrored cube. The master cube is used as a mean azimuth reference after repeated observations of the star and the cube from the theodolite station. The azimuth of the master cube can be calculated simply by adding the angle between the cube and the star to the azimuth of the star. The pentaprism is used solely to observe the surface of the master cube from a geometric setup which does not allow direct access from the theodolite station.

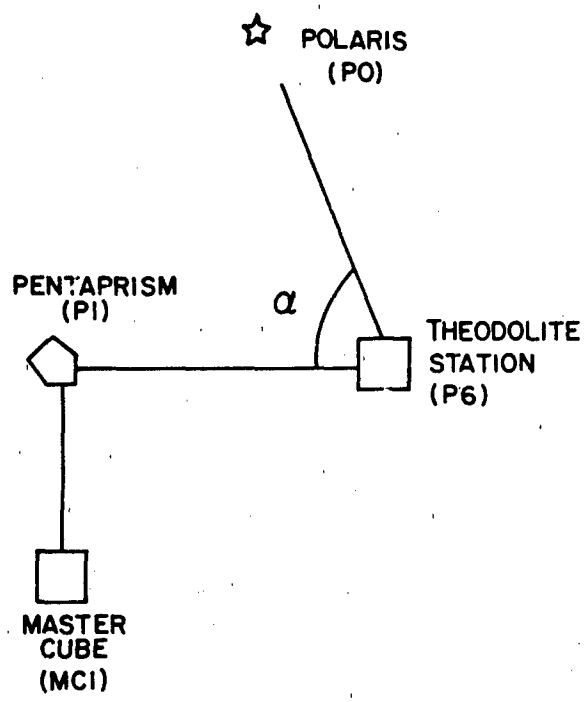


Figure 23-38. Schematic of typical astronomic azimuth observations in an inertial test facility.

Conventional azimuth measurement techniques are severely limited by a number of factors as illustrated in Table 23-6 [Roberts, 1980]. The largest single error source is the "personal equation" of the observer, which can contribute several arcseconds of bias and uncertainty to the azimuth estimate.

In addition, conventional observations are limited to clear nights and, therefore, inherently prohibit observation of motions of the desired target within periods of less than 24 hours. These motions are typically 1 arcsecond and can be as large as 5 arcseconds.

Few improvements to conventional measurement technolgy have been made in recent years. However, one significant improvement may result from research at the University of Maryland sponsored by the Defense Mapping Agency under the supervision of Air Force Geophysics Laboratory. This research has led to the development of a prototype array eyepiece system for a Wild T-4 theodolite known as a Charge Coupled Device (CCD). The prototype CCD automatically "observes" and records the transit of a star across a photo-diode array which has known coordinates with respect to the T-4, thereby removing observer errors. The path of the star across the array is then used to determine an estimate of the reference azimuth that is independent of observer bias. This azimuth is then transferred in the conventional fashion to the desired target [Currie and Salvermoser, 1981].

ble 23-6. Azimuth or horizontal angle error sources [Roberts, 1980].

Error Source	Instrument	Magnitude (sec)	Ref	Notes
Elastic deformation of eveling screws	T4 DKM3	0.05 0.15	8	
Clamping error		0.1 to 0.2	1	Cancels by averaging for equal torques
Thermal flexing of elescope by touching ne side	DKM3	0.16	9	May tend to cancel by reversal
Mechanical flexing of elescope	T4	0.7	10	
Trunnion axis roll rror	T4 DKM3	0.6 0.1	8	
Trunnion axis wobble rror	T3 T4 DKM3	−1 to +0.5 0.4 0.1	11 8 8	
Level vials	DKM3	0.5	1	Tends to cancel over many sets of observations
Atmospheric efraction		0.1	10	
Observer biases	T3	±2.84	12	
Timing error		0.06 to 0.15	1	
etaprism residual rror		±0.25	1	
Elliptic horizontal ircle distortion of 0^{-4} inches		2		Cancels by circle rotation
ccentric graduations n horizontal circle y 10^{-4} inches		0.4		Cancels by circle rotation
orro prism can't ngle		±0.14	1	
Misalignment of orizontal circle and lescope vertical axes 10^{-4} inches		0.4		Does not cancel by circle rotation
lovement of long ne mark by 1 cm in nm		1		
Blunder. (Acceptance one 5 s blunder in sets)		0.3		Possible with wide rejection tolerance
odel assumptions		±0.45		This report

Table 23-6 (*Continued*)

Error Source	Instrument	Magnitude (sec)	Ref	Notes
Star Models		0.1		
Random circle graduation errors		~1.5 to 1	13	
Optical window		±0.125	2	
Non-flatness of cube or mirror faces		±0.06	2	
Thermal circle index drift		0.3	10	

23.5.3 Inertial Azimuth Determination

The heart of any inertial azimuth system is the gyroscope. When it is operated as a gyrocompass like the Azimuth Laying Set (ALS) in Figure 23-39, the gyro senses a component of the earth's rotation rate which is proportional to azimuth.

The single degree of freedom (SDOF) floated gyro used in the ALS may be described as a sealed cylindrical canister within another canister. The inner canister called the float, immersed in a fluid at neutral buoyancy, contains a rotating wheel which generates angular momentum. The SDOF gyro has three principal axes which are defined by the wheel and float orientation and the gimbals on which the gyro is mounted as illustrated in Figure 23-40a. The spin axis (SA) and output axis (OA) are physically defined by the wheel axis and the longitudinal axis of the float, respectively. The input axis (IA) is defined as mutually perpendicular to SA and OA. In general, any gyroscopic element (wheel and axle) has an infinite number of pairs of input and output axes at right angles to each other. The number of gimbals, in our case two, defines the degrees of freedom and therefore the sensing axes.

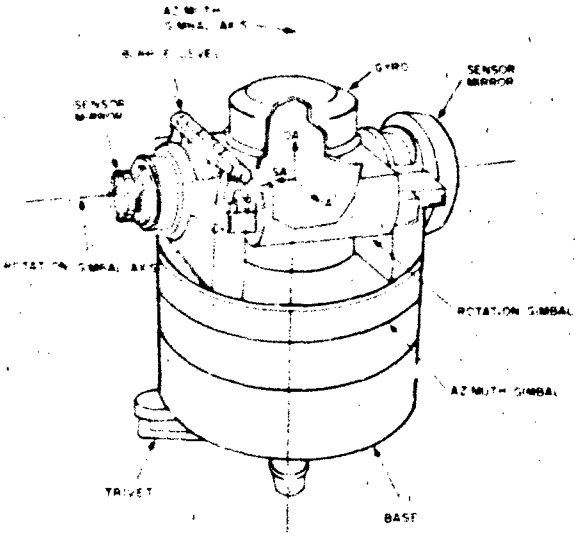


Figure 23-39. Azimuth laying set gyrocompass

During operation, the gyro wheel is rotated rapidly about the spin axis which generates angular momentum (H). When an angular rate is introduced about the input axis (ω_{IA}), a precessional torque (Mo) is generated about the output axis which is proportional to the vector cross product of H and ω_{IA}. This torque can then be expressed mathematically as

$$Mo = H \times \omega_{IA}$$

23.5.3.1 Gyrocompassing. The gyro is capable of sensing any rotation about its input axis. If the gyro is maintained in a fixed orientation on the earth, it senses a component of the earth's rotation rate that is proportional to the angle between the input axis and the earth's rotation axis.

To measure azimuth with a sensor like the ALS, the gyro is oriented so that the input axis is nearly east or west in a local horizontal plane where the sensitivity to small azimuth change (sensed earth rate) is maximized. The earth rate sensed by the gyro is then $\omega_{IA} = \omega_{EH} \sin \alpha$, where ω_{EH} is the horizontal component of the earth's rotation rate [$\omega_{EH} = \omega_E \cos(LAT)$], and α is the desired azimuth of the input axis defined from astronomic East. Note that the effect of latitude on gyro output is compensated for by applying a scale factor to the gyro output signal. By combining the equations for torque and IA rate, we can solve for azimuth of the IA in terms of the gyro's torque output.

$$\sin \alpha = \frac{Mo}{HW_{EH}} + \varepsilon$$

where ε represents the sum of other error torques (to be discussed later).

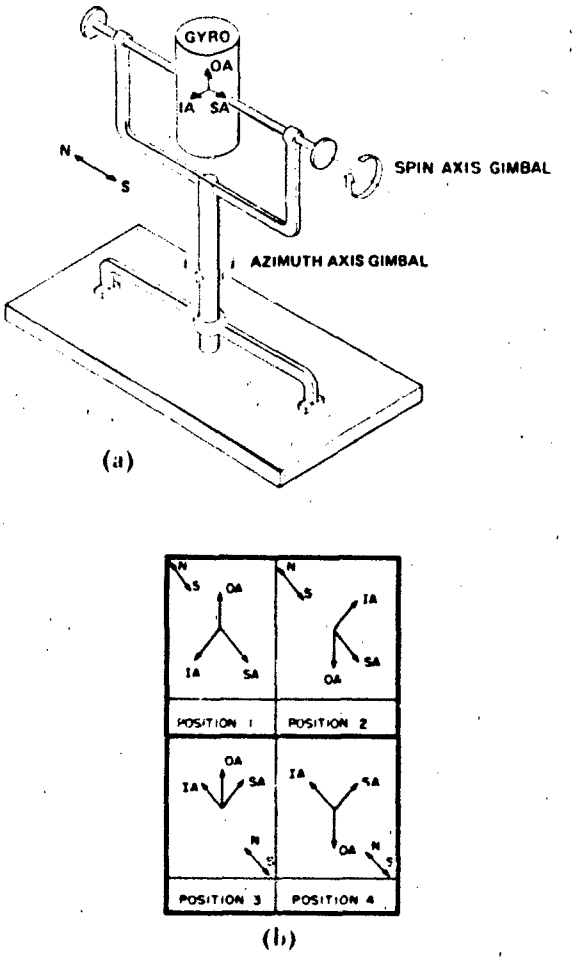


Figure 23-40 (a) Functional representation of gyro gimbal mounting (b) Functional representation of gyro axes during four position gyrocompassing

In this orientation, the ALS sensor mirrors are aligned nearly North-South because of their perpendicularity with the gyro IA. The combination of the gyro torque output and theodolite or other angle measurements between the sensor mirror and the target then provide the desired target azimuth.

Unfortunately, mechanical gyroscopes exhibit substantial torque drift with time. Since these error torques are often large, some means of separating the drift of the azimuth signals is necessary. Traditionally, this is accomplished by repositioning the gyro's axes in four orientations during each sample sequence by moving the gyro on the two gimbal mounts, as shown in the functional representation in Figure 23-40b. This technique is known as four-position gyrocompassing. The repositioning creates a set of observation equations that, when combined, eliminates the average gyro drift errors. However, this same process introduces significant mechanical and electrical errors due to the act of repositioning the gyro. These errors represent the limit of accuracy for the four-positioning gyrocompassing technique, and until recently were the limiting factor for inertial azimuth measurement.

23.5.3.2 Wheel Speed Modulation. Positioning errors no longer limit inertial azimuth measurement accuracy since recent technology advancement has eliminated the requirement for continuous gyro repositioning. A state-of-the-art mechanical gyroscope has been built with a permanent magnet (PM) wheel which exhibits highly repeatable torque output at any wheel speed. This characteristic makes a technique known as Wheel Speed Modulation a feasible alternative to four-position gyrocompassing [Shearer et al., 1980].

WSM is simply an iterative variation of the gyro's wheel speed, and thus its angular momentum H, between two discrete values. The gyro is maintained in a fixed orientation. For this special case, the gyro torque at each wheel speed is proportional to the respective value of angular momentum, and one can then write two observation equations for the gyro torque in the form

$$M_1 = H_1 \omega_{IA} + \varepsilon$$
$$M_2 = H_2 \omega_{IA} + \varepsilon + \varepsilon_{WSM} .$$

Substracting these two equations yields a new equation which is free of the traditional gyro torque errors ε without repositioning.

$$\Delta M = \Delta H \omega_{IA} + \varepsilon_{WSM} .$$

where ε_{WSM}, the torque change due to the speed change, is quite repeatable for a permanent magnet wheel gyroscope and can be readily calibrated.

Sub-arcsecond azimuth measurement uncertainty has been achieved using WSM of a PM gyro [Shearer, 1982].

23.5.3.3 Ring Laser Gyroscope. Finally, recent technological developments in ring laser gyroscope (RLG) technology show promise for accurate azimuth measurement. For this technique, light from a pulsating laser is sent through a beam splitter. The resulting two identical signals are then sent in opposite directions around a closed path back to a signal detector. If no earth rate or other corrupting rates are sensed by the gyro input axis (IA East or West in the horizontal plane), the two signals will arrive at the signal detector exactly in phase with each other. However, if the gyro is sensing a component of earth rate due to its azimuth misalignment from E-W, the effect will be an apparent change in path length between the two directions, thereby resulting in a phase shift between the two signals at the signal detector. This phase shift is directly proportional to the azimuthal component of the earth rate being sensed by the gyro.

To date, high accuracy has not been achieved with Ring Laser Gyroscopes. Some of the practical problems with Ring Laser Gyroscopes, such as lock-in near zero torque and size versus accuracy tradeoffs, are now being addressed by on-

going research. However, these gyros have several desirable attributes. First, they can be sampled at very fast rates to obtain a rapid azimuth measurement. This estimate can be further enhanced by smoothing, averaging or other signal processing techniques. Second, they do not require long warm up periods, making them ideal for use in dormant systems. Finally, since they have no moving parts, high reliability is likely.

23.5.4 Geophysical and Local Environment Effects on Inertial Instruments

Generally, geophysical phenomena can be classified as global or local [Williams, 1970]. Global phenomena, such as polar motion, changes in the day length (Chapter 24), and plate motions (Section 23.2), although severe at times, generally have less effect in inertial instruments than local effects. Figure 23-41 shows the frequencies of some of the more important local motions. Table 23-7 indicates the parameters most affected by these motions and the expected level of disturbance [McKinley, 1975].

Large, even distant earthquakes (Section 23.3.2) can cause significant perturbations in the output of a gyro since ground accelerations sensed by the gyro generate error torques. Inertially derived azimuth data obtained during the passage of earthquake waves will therefore be degraded by the ground motions. However these degradations are generally short-term phenomena and schemes for extrapolating across earthquake events are commonplace.

Long period crustal motions (Section 23.4) of concern in inertial hardware testing and operation are those caused by atmospheric and hydrostatic loading, fault displacements, and areal strain accumulation. Although these motions can cause sizable rotations and displacements [Gray et al., 1972], they generally have no instantaneous effect on inertial measurement instrumentation, and any cumulative errors can be accurately evaluated by repeated measurements and their effects removed computationally.

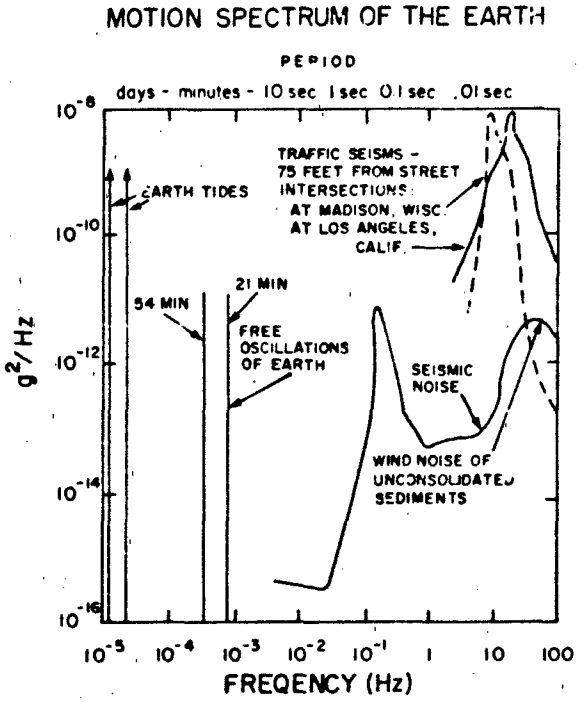


Figure 23-41 Typical natural and cultural ground motion spectrum [Weinstock, 1968]

Despite the obvious contribution from these geophysical phenomena, variations in the local micro-environment are still the major error source for all inertial instruments. Daily temperature cycles, such as those shown in Figures 23-42 and 23-43, and localized thermal gradients can introduce significant time-varying base tilt motions. These tilts and

Table 23-7 Common gyro test parameters affected by motion environment [McKinley, 1975].

Parameter	Earthquake Waves 5 s–1 min to 100 μgal	Tides 12 - 24 hrs 0.25 μgal	Microseisms 1-0.1 s 10 μgal	Pole Motion 12 - 24 months 0.5 s (latitude)	Cultural/Thermal 0.5 μgal to 10 mgal 0.1 to 10 s
Accelerometer					
Scale Factor	X	X	X		X
Bias	X	X	X		X
Non-Linearity	X	X	X		X
Gyro Scale Factor				X	
Bias				X	
G-Sensitive	X	X			X
G^2-Sensitive	X	X			X
Gyrocompass				X	X

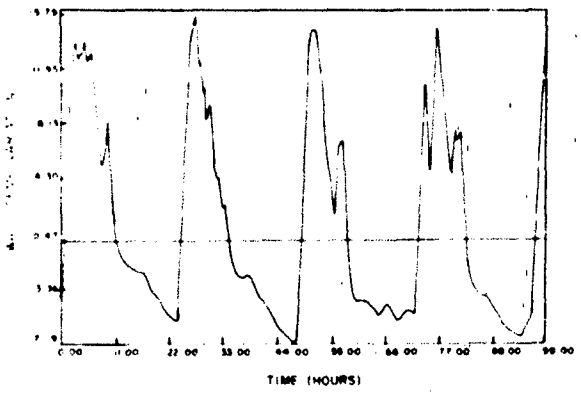


Figure 23-42 Typical temperature variation of test laboratory outside wall surface

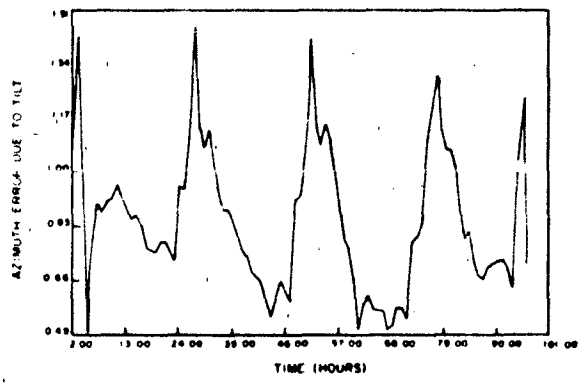


Figure 23-44 Azimuth error due to temperature-induced tilt

the associated tilt rates can cause azimuth measurement errors of up to several arcseconds (Figure 23-44) and must be compensated for. For example, at 45° latitude, 1 arcsecond of tilt of the gyro IA results in a 1 arcsecond azimuth measurement error. A tilt rate about IA of less than 0.2 arcseconds per hour induces a 1 arcsecond azimuth error.

Two fundamental approaches have been developed for minimizing motion-induced errors in the testing and operation of inertial instrumentation. The most commonly used technique has been an attempt to isolate the test instrument or operational system from disturbing motions by either a passive or active isolation system. An alternate approach is to measure and characterize the perturbing noise field, estimate its contribution to component or system behavior, and remove the error term by signal processing. The characteristic of the noise field will determine the degree of sophistication required in estimating and removing the effects of the disturbing motions. A significant amount of research has been conducted to accurately measure and compensate for base tilt rates and additional research is required to successfully compensate inertial azimuth measurements for tilt rates to the present sub-arcsecond accuracy objectives.

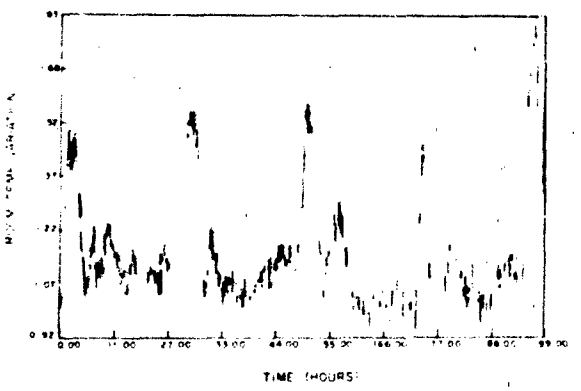


Figure 23-43 Ambient room temperature variation near gyro test station

23.4.5 Azimuth Data Base

A number of agencies maintain a time history of the average motions of azimuth references in the form of azimuth data bases. These azimuth data bases provide important information about the long period motions which affect all inertial test facilities and which can severely degrade test results.

One of the most extensive data bases now being maintained is at the Aerospace Guidance and Metrology Center (AGMC) in Newark, Ohio. Similar data bases described by Roberts [1980] include the Advanced Inertial Test Laboratory at Holloman AFB, the Northrop Test Laboratory Azimuth Facility at Hawthorne, California, and the Charles Stark Draper Laboratory in Cambridge, Mass. However, because of the amounts of data available, the AGMC facility will be described in some detail.

AGMC maintains an azimuth history of a master cube as depicted in Figure 23-45. Three different azimuth references are measured at AGMC. The azimuth of the master cube is determined by periodic observation of Polaris and transfer to the master cube, as previously described. The azimuth of a long line monument is also determined by transferring the azimuth from Polaris to the monument. Finally, azimuth is transferred from the master cube to cubes mounted on long isolation blocks to be used at selected test stations. On the average, ten azimuths are determined per month.

Figure 23-45 shows a plot of repeated azimuth estimates for the master cube azimuth data base for the years 1975–78. Each data point typically represents the average of 16 sightings on Polaris and the master cube. Two trends are immediately apparent, a sinusoidal annual variation and a decreasing steady linear drift. Annual climatic variations and

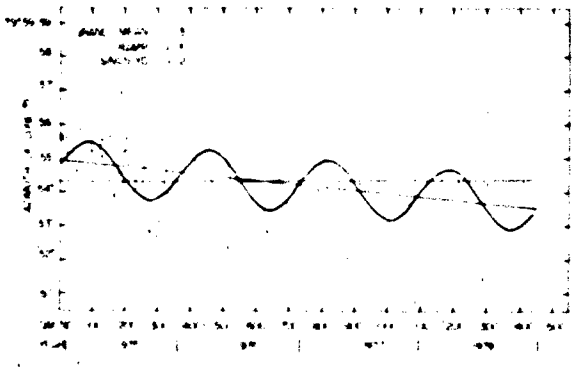


Figure 23-45 Air Force Guidance and Metrology Center master reference

a 1.2 year Chandler period variation caused by the earth's elasticity are possible causes of the sinusoidal pattern. The linear drift is probably the result of actual motion of the cube due to thermally-, geophysically-, and/or structurally-induced stresses.

In addition to the master cube, the azimuths of three other cubes and a long line monument were maintained during the four-year period (1975-78). Their average estimated annual drift rates are shown in Figure 23-46. It is interesting to note the one arcsecond/yr relative drift between cubes 1 and 2. This may be accounted for by differential stress on opposite sides of the isolation mounting block. The block is 24 m long. Thus a change of 0.025 mm per year on one side is enough to produce the resulting apparent azimuth change.

Figure 23-47 shows the annual periodic motions for the AGMC facility as a whole, as well as for the four cubes shown in the previous diagram. Possible causes of the estimated facility motion of 0.7 arc s/yr include annual variations in precipitation and runoff patterns, seasonal variations in solar heating between opposite sides of the building, and freezing and thawing of the ground. Temperature gradients across the width of the support blocks can also account for the periodic motions of the cubes, especially cubes 2 and 3. A temperature change of 0.16 degree is enough to generate these types of motions.

A similar inertial azimuth measuring device was used by Dieselman, et al [1970] to monitor a reference cube at an inertial component test facility similar to AGMC. Tilt and temperature data were collected and are plotted with the azimuth estimates in Figure 23-48. The reference cube shows long period annual variations similar to the ones for AGMC. Note that the long period azimuth variations are very well correlated with the long period tilt and temperature data. This data suggests that the rotation shown are most likely attributable to thermal and insolation effects on the local topography as well as the building (Section 23.4).

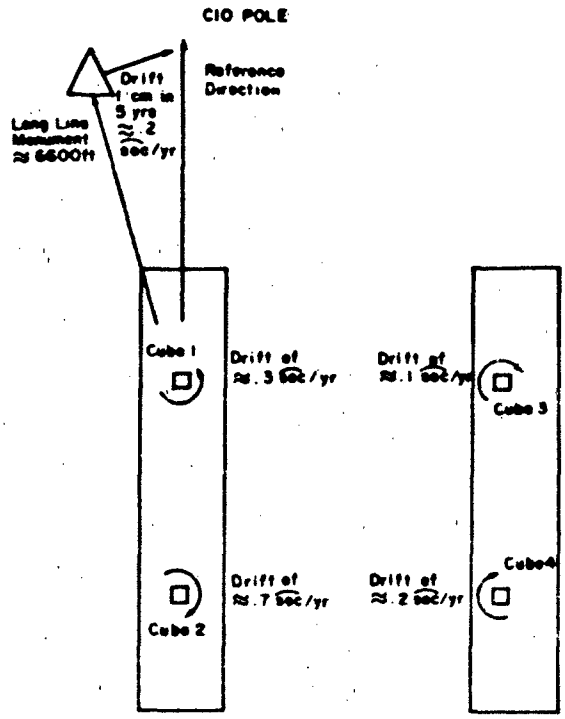


Figure 23-46 Average annual motion of AGMC reference cubes

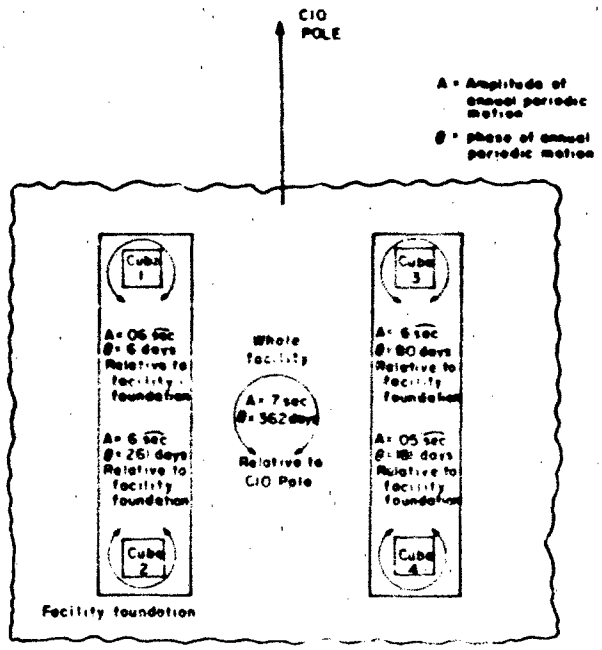


Figure 23-47 Mean annual periodic motions for the years 1975 to 1978, AGMC

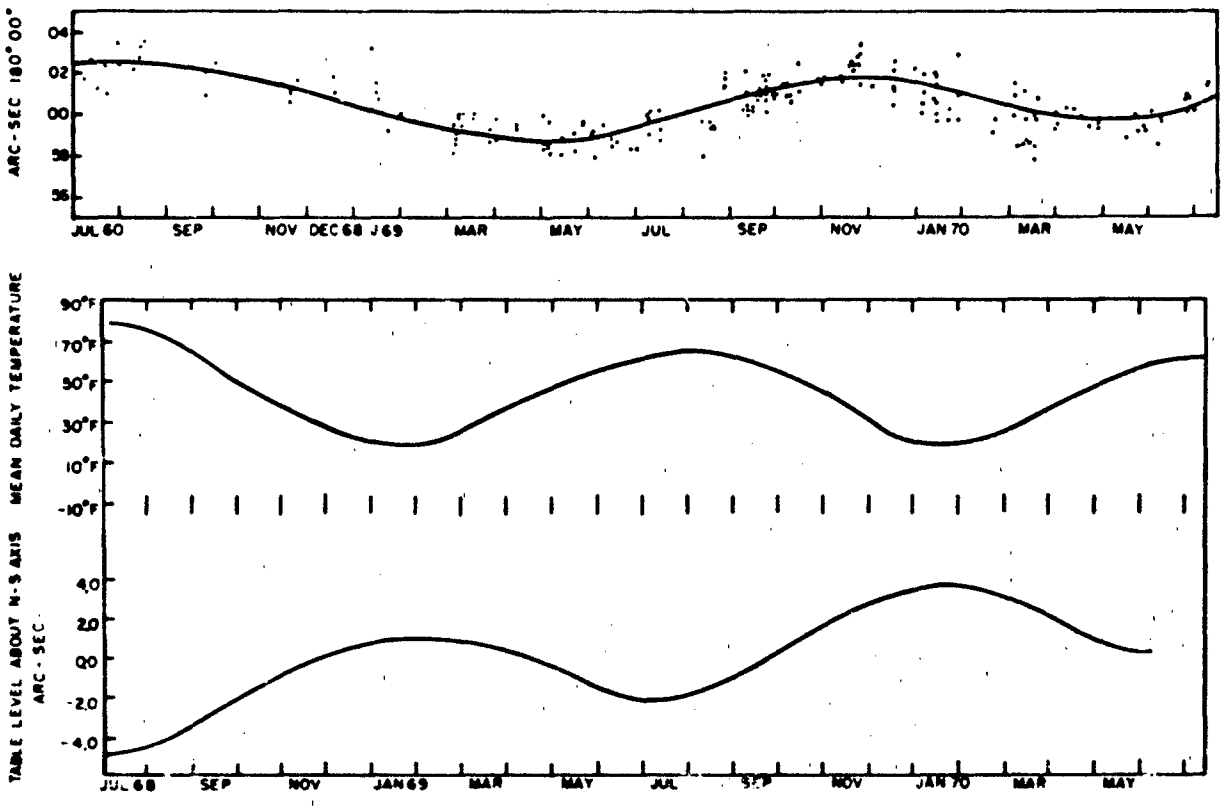


Figure 23-48 Long period variations of master reference cube, AGMC

23.5.6 Conclusions

In summary, as modern systems become increasingly sophisticated, the effects of the ambient motion environment on their performance are also dramatically increasing. The ability to accurately measure and compensate for these environmental influences is complicated by the fact that the very motions being measured tend to degrade the performance of the measurement tool. Thus, what was once thought to be a simple static problem of classical geodesy is now a dynamic concern of modern geophysics.

Looking toward the future, one can only predict that the increasing demand for better measurement accuracy required to support new systems will continue to amplify the impact of the dynamic environment in which we operate. Continued research in rotational geophysics is required to ensure parity of the measurement technology with the operational requirements.

REFERENCES

Agnew, D.C., *Strain Tides at Pinon Flat: Analysis and Interpretation*, Ph.D. Thesis, University of California, San Diego, 1979.

Aki, K. and P.G. Richards, *Quantitative Seismology*, **1**, Freeman, San Francisco, 1980.

Algermissen, S., *Seismic Risk Studies in the United States*, 4th World Conference on Earthquake Engineering, Santiago, Chile, 1969.

Anderson, D.L. and R.S. Hart, "An Earth Model Based on Free Oscillations and Body Waves," *J. Geophys. Res.*, **81:** 1461–1475, 1976.

Anderson, E.G., "Modeling of Physical Influences in Sea Level Records for Vertical Crustal Movement Detection," Proceedings 9th Geodesy Solid Earth and Ocean Physics (GEOP) Conference, *Report 280*, Dept. Geodetic Science, Ohio State University, Columbus, 145–152, 1978.

Baker, T., "What Can Earth Tide Measurements Tell Us about Ocean Tide or Earth Structure?," Proceedings 9th Geodesy Solid Earth and Ocean Physics (GEOP) Conference, *Report 280*, Dept. Geodetic Science, Ohio State University, Columbus, 299–307, 1978.

Bath, M., *Introduction to Seismology*, Wiley, New York, 1973.

Batts, J., "Regional Modification of Acceleration Attenuation Functions," *Bull. Seismol. Soc. Am.*, **71:** 1309–1321, 1981.

Beaumont, C., "Linear and Nonlinear Interactions between the Earth Tide and a Tectonically Stressed Earth," Proceedings 9th Geodesy Solid Earth and Ocean Physics (GEOP) Conference, *Report 280*, Dept. Geodetic Science, Ohio State University, Columbus, 313–318, 1978.

Beaumont, C. and J. Berger, "Earthquake Prediction Modification of the Earth Tide: Tilts and Strains by Dilatancy," *Geophys. J. Royal Astron. Soc.*, **39:** 111–121, 1974.

Beaumont, C. and A. Lambert, "Crustal Structure from Surface Load Tilts Using a Finite Element Model," *Geophys. J. Royal Astron. Soc.*, **29:** 203–226, 1972.

Ben-Menahem, A. and S.J. Singh, *Seismic Waves and Sources*, Springer-Verlag, New York, 1981.

Berger, J. and C. Beaumont, "An Analysis of Tidal Strains from the USA: II. The Inhomogeneous Tide," *Bull. Seismol. Soc. Am.*, **66:** 1821–1846, 1976.

Berger, J. and J. Levine, "The Spectrum of Earth Strain from 10^{-4} to 10^{-1} Hz," *J. Geophys. Res.*, **79:** 1210–1214, 1974.

Berger, J. and R. Lovberg, "Earth Strain Measurements with a Laser Interferometer," *Science*, **170:** 296–303, 1970.

Bird, J.M. and B. Isacks (eds.), *Plate Tectonics*, selected papers from the *J. Geophys. Res.*, 1972.

Brazee, R., *Analysis of Earthquake Intensities With Respect to Attenuation, Magnitude, and Rate of Recurrence*, NOAA, 1976.

Brown, L.D., "Recent Vertical Crustal Movement Along the East Coast of the United States," *Tectonophysics*, **44:** 205–231, 1978.

Brune, J., "Surface Waves and Crustal Structures," *The Earth's Crust and Upper Mantle*, edited by P. Hart, AGU, Monograph 13, Washington, D.C., 1969.

Brune, J. and J. Oliver, "The Seismic Noise of the Earth's Surface," *Bull. Seismol. Soc. Am.*, **49:** 349–353, 1959.

Bullen, K.E., *An Introduction to the Theory of Seismology*, Cambridge University Press, Cambridge, U.K., 1965.

Cabaniss, G.H., "The Measurement of Long Period and Secular Deformation with Deep Borehole Tiltmeters," Proceedings 9th Geodesy Solid Earth and Ocean Physics (GEOP) Conference, *Report 280*, Dept. Geodetic Science, Ohio State University, Columbus, 165–169, 1978.

Cabaniss, G.H. and D.H. Eckhardt, "The AFCRL Earth Tide Program," AFCRL-TR-73-0084, 1973.

Cartwright, D., "The IOS Program of Tidal Definition in the North Atlantic Ocean," *Proceedings 9th Earth Tide Symposium*, Columbia University, New York, 1982.

Cartwright, D. and A. Edden, "Corrected Tables of Tidal Harmonics," *Geophys. J. Royal Astron. Soc.*, **33:** 253–264, 1973.

Castle, R.O., J.P. Church, and M.R. Elliot, "Aseismic Uplift in Southern California," *Science*, **192:** 251–253, 1976.

Chilton, E.G., "Design Requirements for a Testing Laboratory for Inertial Guidance Systems and Components," Air Transport and Space Meeting, New York, 1963.

Chinnery, M.A., "The Frequency of Very Large Earthquakes," *Science*, **190**(4222), 1975.

Clark, S.P. and A.E. Ringwood, "Density Distribution and Constitution of the Mantle," *Rev. Geophys.*, **2:** 35–88, 1964.

Committees on Geodesy and Seismology, National Research Council, "Geodetic Monitoring of Tectonic Deformation — Toward a Strategy," National Academy of Sciences, Washington, D.C., **109**, 1981.

Cox, A. (ed.), *Plate Tectonics and Geomagnetic Reversals*, Freeman, San Francisco, 1973.

Crowford, R., *The Air Force Manual for Design and Analysis of Hardened Structures*, AFWL-TR-74-102, Air Force Weapons Laboratory, Kirtland AFB, N. Mex., 1974.

Crowley, F.A. and H.A. Ossing, "An Analysis of the Vibration Environment with Applications to Single-Degree-of-Freedom Gyroscope Performance Tests," AIAA Paper No. 70-951, 1970.

Currie, D.G. and R. Salvemoser, "The Array-Eyepiece System for the Wild T-4 Theodolite," *Program Review and Field Results*, University of Maryland, 1981.

Department of Defense, *Glossary of Mapping, Charting and Geodetic Terms*, Defense Mapping Agency Topographic Center, Washington, D.C., June 1973.

Dieselman, J., R.L. McLaughlin, K. Mullo, and L.B. Thompson, "Recent Developments in the Testing of High Performance Gyrocompass Devices," *Proceedings 5th Guidance Test Symposium*, **2**, 1970.

Dietz, R.S. and J.C. Holden, "Reconstruction to Pangaea: Breakup and Dispersion of Continents, Permian to Present," *J. Geophys. Res.*, **75:** 4939–4956, 1970.

Donovan, N., B. Bolt, and R. Whitman, *Technology Review*, Cambridge, Mass., May 1977.

Dorman, J., "Seismic Surface Wave Data on the Upper Mantle," *The Earth's Crust and Upper Mantle*, edited by P.J. Hart, AGU, Monograph 13, Washington, D.C., 1969.

Everndon, J.F., "Gravity Profiles in Southern California," *Earthquake Prediction: An International Review*, AGU, Washington, D.C., 485–496, 1981.

Ewing, C.F., and M.M. Mitchell, *Introduction to Geodesy*, Elsevier, New York, 1970.

Farrell, W.E., "Deformation of the Earth by Surface Loads," *Rev. Geophys. Space Phys.*, **10:** 761–797, 1972.

Fix, J., "Ambient Earth Motions in the Period Range from 0.1 to 2560 Sec," *Bull. Seismol. Soc. Am.*, **62:** 1753–1760, 1972.

Garland, G.D., *Introduction to Geophysics*, W.B. Saunders Co., Philadelphia, 1971.

Gass, I.G., "Ophiolites," *Sci. Am.*, August 1982.

Goad, C.C., "Gravimetric Tidal Loading Computed from Integrated Green's Functions," *J. Geophys. Res.*, **85:** 2679–2683, 1980.

Gray, R.A., G.H. Cabaniss, F.A. Crowley, H.A. Ossing, T.S. Rhoades, and L.B. Thompson, "Earth Motions and Their Effects on Inertial Instrument Performance," AFCRL-72-0278, AD748274, 1972.

Greenstelder, R., "Maximum Credible Rock Acceleration from Earthquakes in California," California Division of Mines and Geology, Map Sheet 23, 1974.

Harrison, J.C., "Cavity and Topographic Effects in Tilt and Strain Measurement," *J. Geophys. Res.*, **81:** 319–328, 1976.

Harrison, J.C., "Implications of Cavity, Topographic, and Geologic Influences on Tilt and Strain Measurements," Proceedings 9th Geodesy Solid Earth and Ocean Physics (GEOP) Conference, *Report 280*, Dept. Geodetic Science, Ohio State University, Columbus, 283–287, 1978.

Harrison, J.C., and K. Herbst, "Thermoelastic Strains and Tilt Revisted," *Geophys. Res. Lett.*, **4:** 535–537, 1977.

Herbst, K., "Interpretation of Tilt Measurements in the Period Range above that of the Tides," AFGL-TR-79-0003, ADA074525 (Translation No. 1080), 1979.

Holzer, T.L., S.N. Davis, and B.F. Ludgren, "Faulting Caused by Groundwater Extraction in South Central Arizona," *J. Geophys. Res.*, **84:** 603–612, 1979.

Hugget, G.R., J.F. Slater, and J. Langbein, "Fault Slip Episodes near Hollister, California: Initial Results Using a Multiwavelength Distance Measuring Instrument," *J. Geophys. Res.*, **82:** 3361–3368, 1977.

Isacks, B., J. Oliver, and L. Sykes, "Seismology and the New Global Tectonics," *J. Geophys. Res.*, **73:** 5181, 1968.

Iyer, H., *The History of Science of Microseisms*, VESIAC Report 4410-64-X, University of Michigan, Ann Arbor, 1964.

Jackson, D.D., and W.B. Lee, "The Palmdale Bulge — An Alternate Interpretation," *Eos Trans. AGU*, **60,** 810, 1979.

Jaenke, M.G., "A Proposed Model for the Rotational Motions of a Test Pad," AIAA Paper No. 74-855, 1974.

Johnston, J.C., Z.A. Cybriwsky, and G.K. Leblanc, "The Regulatory Guide 1.60 Its Content and Applicability" abstract, *Earthquake Notes*, **51**(3), 1980.

Kanamori, H., "Velocity and Q of Mantle Waves," *Phys. Earth Planet. Int.*, **2:** 259–275, 1970.

Kanamori, H., and J. Cipar, "Focal Process of Great Chilean Earthquake," May 22, 1960, *Phys. Earth Planet. Int.*, **9:** 128–136, 1974.

Kanamori, H., "The Energy Release in Great Earthquakes," *J. Geophys. Res.*, **82**(20), 1977.

Kayton, M., "Fundamental Limitations on Inertial Measurements," *Guidance and Control*, Academic Press, New York, 1963.

Kuo, J.T., "Static Response of a Multilayered Medium Under Inclined Surface Loads," *J. Geophys. Res.*, **74:** 3195–3207, 1969.

Kuo, J.T., "The Inverse Problem: Ocean Tides Derived from Earth Tide Observations," Proceedings 9th Geodesy Solid Earth and Ocean Physics (GEOP) Conference, *Report 280*, Dept. Geodetic Science, Ohio State University, Columbus, 319–326, 1978.

Lambert, A., and C. Beaumont, "Nano-Variations in Gravity Due to Seasonal Groundwater Movements: Implications for the Gravitational Detection of Tectonic Movements," *J. Geophys. Res.*, **82,** 297–306, 1977.

Lambert, A., and J.O. Liard, "A Search for Long-Term Earthquake Precursors in Gravity Data in the Charlevoix Region, Quebec," *Earthquake Prediction: An International Review*, AGU, Washington, D.C., 473–483, 1981.

Lennon, G.W., and T.F. Baker, "The Earth Tide Signal and Its Coherency," *Quart. J. Roy. Astron. Soc.*, **14:** 161–182, 1973.

Levine, J., "Strain-Tide Spectroscopy," *Geophys. J. Royal Astron. Soc.*, **54:** 27–33, 1978.

Levine, S.A., and A. Gelb, "Geodetic and Geophysical Uncertainties Fundamental Limitations on Terrestrial Inertial Navigation," AIAA Guidance Control and Flight Dynamics Conference, Pasadena, Calif., 1968.

Lewkowicz, J.F., and G.H. Cabaniss, "Near-Field Static Tilts from Surface Loads," AFGL-TR-80-0221, ADA094123, 1980.

Lindh, A.G., and D.M. Boore, "Control of Rupture by Fault Geometry During the 1966 Parkfield Earthquake," *Bull. Seismol. Soc. Am.*, **71,** 95–116, 1981.

Longman, I.M., "A Green's Function for Determining the Deformation of the Earth Under Surface Mass Loads, 1. Theory," *J. Geophys. Res.*, **67:** 845–850, 1962.

McGuire, R., *FORTRAN Computer Program for Seismic Risk Analysis*, U.S. Geological Survey, Open File Report, 76-67, 1976.

McKinley, H.L., "Geokinetic and Geophysical Influence on 1980's ICBM Guidance," AIAA Guidance and Control Conference, Boston, Mass., 1975.

Melchior, P., *The Earth Tides*, Pergamon Press, Oxford, 1966.

Minster, J.B., and T.H. Jordan, "Present-day plate motions," *J. Geophys. Res.*, **83:** 5331–5354, 1978.

Mogi, K., "Earthquake Prediction in Japan," *Earthquake Prediction: An International Review*, AGU, Washington, D.C., 635–666, 1981.

Morgan, W.J., "Rises, Trenches, Great Faults, and Crustal Blocks," *J. Geophys. Res.*, **73**: 1959, 1968.

Morgan, W.J., "Plate Motions and Deep Mantle Convection," *Studies in Earth and Space Sciences*, Geol. Soc. Amer. Mem. 132, 1972.

Mueller, Ivan I., *Spherical and Practical Astronomy as Applied to Geodesy*, Ungar, New York, 1969.

Nuttli, O., "The Mississippi Valley Earthquakes of 1811 and 1812: Intensities, Ground Motion and Magnitudes," *Bull. Seismol. Soc. Am.*, **63**: 227–248, 1973.

Nuttli, O., *The Relation of Sustained Maximum Ground Acceleration and Velocity to Earthquake Intensity and Magnitude*, U.S. Army Engineer Waterways Experimentation Station, Miscellaneous Paper S-73-1, 1979.

Ossing H. and R. Gray, *A Survey of the Ambient Motion Environment in the Southwestern United States*, Interim Scientific Report AFGL-TR-78-0052, ADA056872, 1978.

Parke, M.E., "Open Ocean Tide Modelling," Proceedings 9th Geodesy Solid Earth and Ocean Physics (GEOP) Conference, *Report 280*, Dept. Geodetic Science, Ohio State University, Columbus, 289–297, 1978.

Predicting Earthquakes, Panel on Earthquake Prediction, National Academy of Sciences, Washington, D.C., 1976.

Ranslet, R.B. and J.C. Fatz, "Stability of Stable Areas," AIAA Paper No. 74-859, 1974.

Reilinger, R. and L.D. Brown, "Neotectonic deformation, near surface movements and systematic errors in U.S. releveling measurements: Implications for earthquake prediction," *Earthquake Prediction: An International Review*, AGU, Washington, D.C., 422–440, 1981.

Richter, C., *Elementary Seismology*, Freeman, San Francisco, 1958.

Roberts, G.E., "MX Laboratory azimuth data base study (Status report)," Geodynamics Corporation, Santa Barbara, CDR-0008 (SR380E-1), 1980.

Savage, J.C., W.H. Prescott, M. Lisowski, and N.E. King, "Strain Accumulation in Southern California 1973–1980," *J. Geophys. Res.*, **86**: 6991–7001, 1981.

Schureman, P., "Manual of Harmonic Analysis and Prediction of Tides," *Spec. Pub.* 98, U.S. Coast and Geodetic Survey, 1958.

Schwiderski, E.W., "Ocean tides, Part I: Global Tidal Equations," *Marine Geodesy*, **3**: 161–187, 1979.

Schwiderske, E.W., "On Charting Global Ocean Tides," *Rev. Geophys. Space Phys.*, **18**: 38–61, 1980.

Schwiderski, E.W., "Ocean tides, Part II: A Hydrodynamic Interpolation Model," *Marine Geodesy*, **3**: 219–255, 1980.

Seismic Noise Survey, Geotech Technical Report 63-45, **1**: Garland, Texas, 1963.

Seismic Noise Survey, Geotech Technical Report 63-45, **2**: Garland, Texas, 1963.

Shearer, J.A., "Evaluation of an Advanced Gyro and Wheel Speed Modulation for Highly Accurate Azimuth Determination," AFGL-TR-82-0016, AD114640, 1982.

Shearer, J.A., "Error Analysis and Performance Data from an Automated Azimuth Measuring System," AIAA Paper No. 80-1793, 1980.

Simmons, B.J., R.M. Hanes, and F.S. Heming, "Dynamic Analysis of Test Platform," AIAA Paper No. 81-1816, 1981.

Simpson, D.W. and P.G. Richards (eds.), *Earthquake Prediction: An International Review*, AGU, Washington, D.C., 1981.

Slemmons, D., *Faults and Earthquake Magnitude*, U.S. Army Engineers Waterways Experimentation Station, Miscellaneous Paper S-73-1, 1977.

Smith, D.E., "Spaceborn Ranging System," Proceedings 9th Geodesy Solid Earth and Ocean Physics (GEOP) Conference, *Report 280*, Dept. Geodetic Science, Ohio State University, Columbus, 59–64, 1978.

Smith, D.E., R. Kolenkiewicz, P.J. Dunn, and M.H. Torrence, "The Measurement of Fault Motion by Satellite Laser Ranging," *Tectonophysics*, **52**: 59–67, 1979.

Stein, R.S., "Discrimination of Tectonic Displacement from Slope Dependent Errors in Geodetic Leveling from Southern California, 1953–1979," *Earthquake Prediction: An International Review*, AGU, Washington, D.C., 441–456, 1981.

Strange, W., "The Impact of Refraction Correction on Leveling Interpretations in Southern California," *J. Geophys. Res.*, **86**: 2809–2824, 1981.

Takeuchi, H., "On the Earth Tide," *J. Fac. Sci.* University of Tokyo, Sec II-7: Part II, 1–7, 1951.

Tsutsumi, K., "A Program of Study on Long-Term Earth Tilts," AFCRL-70-0268, AD724868, 1970.

Verhoogen, J., F.J. Turner, L.E. Weiss, C. Wahrhaftig, and W.S. Fyfe, *The Earth: An Introduction to Physical Geology*, Holt, Rinehart & Winston, New York, 1970.

Wahr, J.M., *The Tidal motions of a Rotating, Elliptical, Elastic, and Oceanless Earth*, Ph.D. Thesis, University of Colorado, 1979.

Walcott, R.I., "Late Quaternary Vertical Movements in Eastern North America: Quantitative Evidence of Glacio-Isostatic Rebound," *Rev. Geophys. Space Phys.*, **10**: 849–884, 1972.

Weinstock, H., "Design of Precision Tilt and Rotational Vibration Isolation System for Inertial Sensor Testing," AIAA Paper No. 68-894, 1968.

Whitcomb, J.H., "New Vertical Geodesy," *J. Geophys. Res.*, **81**: 4937–4944, 1976.

Williams, O.W., "A Review of Geophysical Effects and Suggested Testing Conditions Related to Advanced Inertial Components" Paper presented to USAF Scientific Advisory Board Ad Hoc Committee on Geophysical Effects on Inertial Instruments, 1970.

Willmore, P.L. (ed.), *Manual of Seismological Observatory Practice*, World Data Center A for Solid Earth Geophysics, Report SE-20, 1979.

Wilson, J.T., "A New Class of Faults and their Bearing on Continental Drift," *Nature* **207**: 343–347, 1965.

Witty, J.R., "Description of an Inertial Test Facility Incorporating a Passively Isolated and Actively Stabilized Platform," AIAA Paper No. 69-863, 1969.

CHAPTER 23

Wyatt, F., "Displacement of Surface Monuments: Horizontal Motion," *J. Geophys. Res.*, **87:** 979–989, 1982.

Wyatt, F., G.H. Cabaniss, and D.C. Agnew, "A Comparison of Tiltmeters at Tidal Frequencies," *Geophys. Res. Ltrs.*, **9:** 743–746, 1982.

Wyllie, P.J., *The Dynamic Earth*, Wiley, New York, 1971.

Zienkiewicz, O.C., *The Finite Element Method in Structural and Continuum Mechanics*, McGraw-Hill, New York, 1967.

Zschau, J., "Tidal Sea Load Tilt of the Crust, and Its Application to the Study of Crustal and Upper Mantle Structure," *Geophys. J. Royal Astron. Soc.*, **44:** 577–593, 1976.

Chapter 24

GEODESY AND GRAVITY

C. Jekeli

Geodesy is a science as old as man's curiosity about the earth—its shape, position, and relation to the stars and planets; yet it has remained a new and timely science by successfully adapting to rapidly changing requirements and novel techniques to answer these requirements. Simply defined, geodesy is the branch of applied mathematics that determines, by observation and measurement, the size and shape of the earth and the precise positions of points on the earth's surface (geometric geodesy) as well as the direction and magnitude of the earth's external gravity field (physical geodesy). The continuing improvements in observational accuracy lend increasing importance to the dependence of each of these quantities on time, and the definition of geodesy now includes the determination of the temporal variations in positions and the gravity field [Mather, 1973]. More and more, geodesy is called upon to provide ancillary methods and data in the study of global, regional, and local geodynamics, as well as other geophysical phenomena. Geodetic measurements and observations contribute to the understanding of tectonic motions, earthquakes, the dynamic evolution of the ocean floor, and the earth's internal structure, as well as its motion in space. To achieve its goals, geodesy, in turn, relies on the continuing efforts in a wide spectrum of terrestrial physics of which it thus forms an integral part.

The following sections provide the nonspecialist with a cursory survey of modern geodesy, emphasizing definitions, results, achievable accuracies, and existing data, but also including a superficial discussion of methodology, when appropriate. The introduction of artificial satellites, modern computers, and other technological engineering advances have broadened the field of geodetic applications to the extent that it is virtually impossible to mention all aspects in a summary such as this. The geodetically connected fields of (plane) surveying, cartography, and photogrammetry are omitted entirely since they have attained prominence in their own right, particularly with their incorporation of new and unique instrumentation and computer capabilities. The textbooks of Mueller and Ramsayer [1979], the Manual of Photogrammetry [1980], and the annual proceedings of the American Congress on Surveying and Mapping and of the American Society of Photogrammetry are excellent sources of detailed information for the interested reader. A description of the diverse instrumentation, vital to geodesy, must also be relegated to other sources. Brief accounts of principles and accuracy can be found in textbooks such as Bomford [1971], Torge [1980], and Mueller [1969]. Other textbooks expounding the theory of geodesy, in addition to these, are by Moritz [1980], Groten [1979], Heiskanen and Moritz [1967], Kaula [1966], and Vanicek and Krakiwsky [1982].

24.1 GEOMETRIC GEODESY

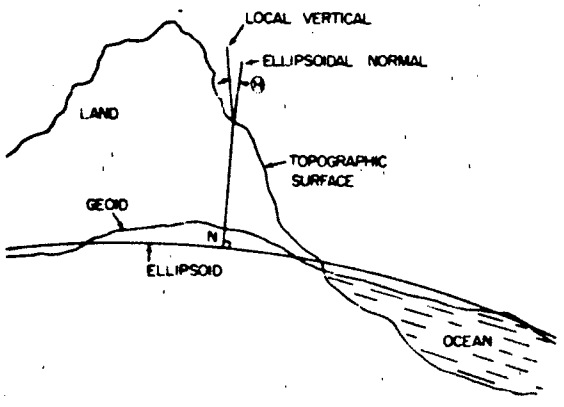
The two reference surfaces in geometric geodesy are the geoid, to which all height observations are referred (vertical control), and the reference ellipsoid, where horizontal geodetic positions are determined (horizontal control). The geoid is the equipotential surface in the earth's gravity field that coincides most closely with the undisturbed mean sea level extended continuously under the continents. The direction of gravity is perpendicular to the geoid at every point. On the earth's surface, this direction is defined [Helmert, 1884] by two angles, the astronomical coordinates: The astronomical latitude is the angle Φ, $-90° \le \Phi \le 90°$, that the gravity vector forms with the equatorial plane; and the astronomical longitude is the angle Λ, $0° \le \Lambda \le 360°$ and positive eastwards, that the plane defined by this vector and the celestial pole forms with the Greenwich meridional plane. Astronomic coordinates are obtained directly from star observations using, for example, a theodolite, its vertical axis being aligned with the local gravity vector.

The reference ellipsoid is a simple mathematical figure that closely approximates the geoid, historically on a regional basis, but on a global basis for modern requirements. It is a surface of revolution formed by rotating an ellipse about its minor (vertical) axis resulting in an oblate (flattened at the poles) ellipsoid. Geodetic ellipsoids are defined by the parameters a = equatorial radius (semimajor axis) and $f = (a - b)/a$ = (polar) flattening, where b is the ellipsoid's semiminor axis. The minor axis is supposed to be parallel to the rotational axis of the earth. Points on the ellipsoid are specified by geodetic coordinates – the geodetic latitude is the angle ϕ, $-90° \le \phi \le 90°$, that the normal to the ellipsoid at the point in question forms with the equator; and the geodetic longitude is the angle λ, $0° \le \lambda \le 360°$ and positive eastwards, that the ellipsoidal meridional plane forms with the plane through the Greenwich meridian. The

IAPTER 24

angles ϕ, λ therefore also define the direction of the psoidal normal.

Because of the irregular distribution of the earth's inal masses, the geoid is an irregular surface which cannot described by a simple geometric figure. It is determined h respect to a particular ellipsoid in terms of the geoid ght, or geoid undulation, which is the separation between two surfaces at any point (Figure 24-1). Another geoid-



are 24-1 The relationship between the geoid and the ellipsoid
N geoid undulation
Θ astrogeodetic deflection of the vertical

ipsoid relationship is the astrogeodetic deflection of the rtical. This is the angle between the local vertical and the rmal to the ellipsoid through the point on the earth's face [Helmert, 1884]. (Figure 24-1). The deflection of vertical is resolved into west-east and south-north comnents, which then provide the functional relationship beeen geodetic and astronomic coordinates.

1.1.1 Horizontal Control

Geodetic surveying generally falls into one of three catories: (1) triangulation networks, whereby points on the rth's surface are connected by a network of triangles whose gles are measured instrumentally and whose sides are mputed with the aid of one or more precisely measured selines, (2) trilateration networks, in which the observations consist primarily of distance measurements of the angle sides, and (3) the traverse, a succession of distance d direction measurements connecting a sequence of points.

The high-precision traverse in the U.S., that is, the anscontinental Traverse (TCT), is accurate to 1 part in and runs in essentially unidirectional east-west and northuth directions (Figure 24-2). It provides the principal mework, scale and orientation, of the National Horizontal ntrol Network. Additional strength in the determination scale and orientation is obtained by incorporating satellite ppler stations (Section 24.3.1) into these traverses. Horntal control within the areas outlined by the TCT is accomplished by densification networks consisting of first-, second-, and third-order triangulations (Table 24-1), as well as lower accuracy local traverses, depending on economic and other needs of the area. Table 24-1 gives a general overview of the accuracies to be achieved and the purposes, uses, and configuration constraints for the several classes of horizontal control.

In most countries the reference ellipsoid, to which the horizontal control refers, is fixed to the earth's surface at an origin point in the network, where the deflection of the vertical and geoid undulation are specified by definition. All measurements to subsequent points in the network are reduced to the same ellipsoid. This entails knowing the height above the geoid (elevation), the geoid undulation, and the deflection of the vertical at each point. The ellipsoid and its relation to the geoid as specified at the origin point define a datum. The orientation of the datum with respect to the earth is determined by three rotation angles—two of these are the geodetic coordinates of the origin point, obtained from the observed astronomic coordinates and the (defined) deflection of the vertical. The remaining degree of freedom, namely a rotation about the ellipsoid normal at the origin point, is fixed by the geodetic azimuth to some other point. It is computed from its relationship to the astronomic azimuth and the deflection of the vertical which incorporates the parallelism of the minor axis of the ellipsoid and the rotational axis of the earth. Because of measuring errors that propagate through the network, additional astronomic azimuths (Laplace azimuths) and longitudes must be determined at selected points to enforce the proper orientation. The position of the center of a geodetic ellipsoid with respect to the earth's center of mass is generally unknown. Table 24-2 lists the most important datums in use throughout the world (Figure 24-3). The relative orientation, position, and scale of several datums can be obtained if the coordinates of a sufficient number of points in each relative system are also known in the other systems, or in an absolute (geocentric) system. Table 24-3 lists the relationships of several regional datums to a global datum. The global datum, in this case, is designated WN14 and is the result of a global solution of geometric observations to satellites, namely ranging and camera observations. Δx, Δy, Δz refer to the displacements of the datum center from the satellite system center with respect to a right-handed Cartesian x, y, z coordinate system (z-axis through the Conventional International Origin pole, x-axis on the equator through the zero meridian, Section 24.4). δ is the relative scale of the two systems, and α, ξ, η are the relative rotations at the origin point of the datum: $-\alpha$ is the rotation in azimuth; ξ, η in latitude and longitude.

The U.S. National Geodetic Survey of the National Ocean Survey/National Oceanic and Atmospheric Administration (NOAA) in concert with corresponding organizations of Canada, Mexico, the republics of Central America, and Denmark are currently completing the new adjustment of the North American geodetic control network. It will en-

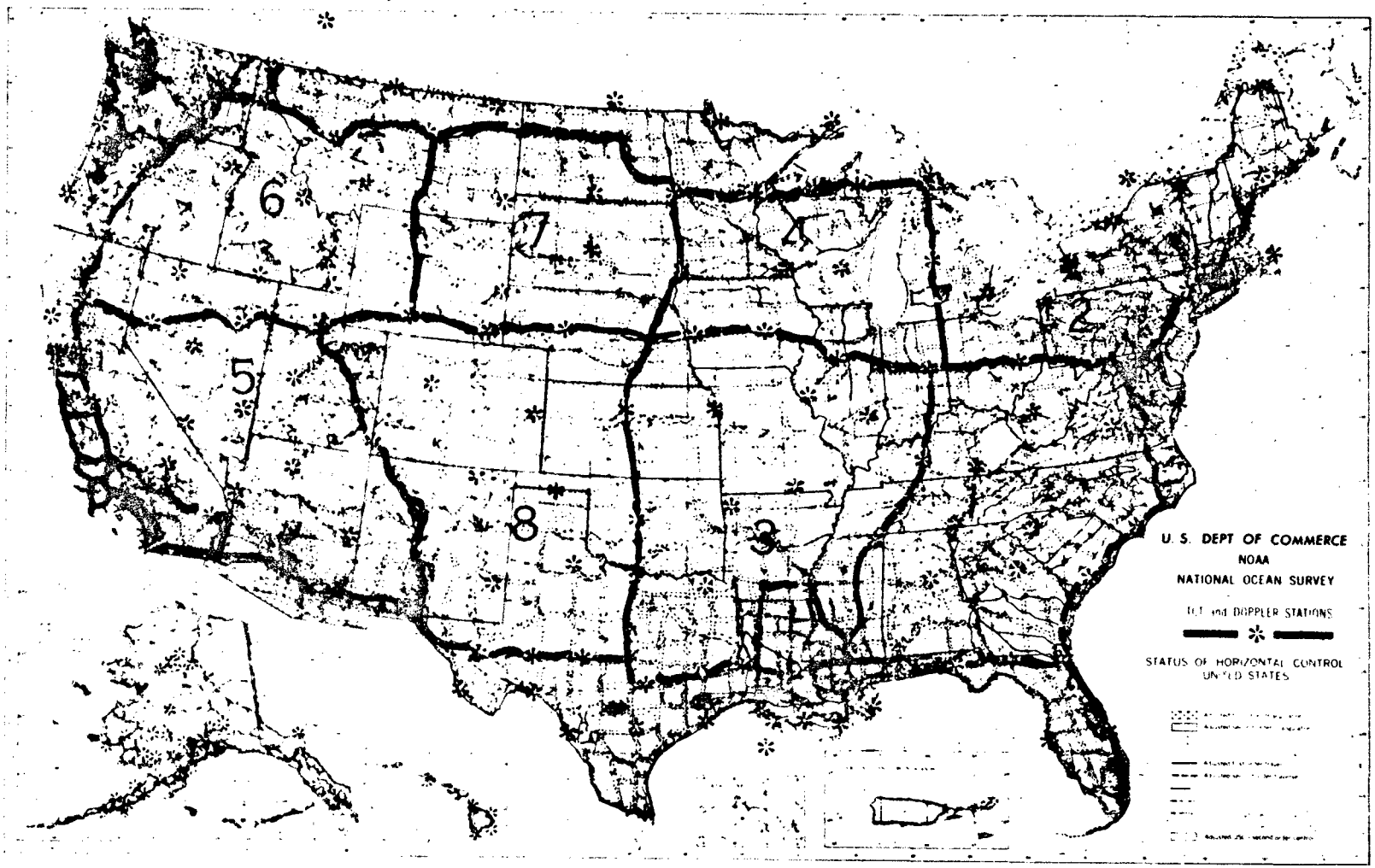


Figure 24-2. Status of horizontal control in the United States [NRC, 1978].

24-1. Synopsis of horizontal control classifications [NOAA, 1980b].

ributes	Order of Surveys and Classes of Accuracy				
	Super First-Order	First-Order	Second Order Class I	Second-Order Class II	Third-Order Class I, II
eral de	Transcontinental control.	Primary horizontal control.	Secondary horizontal control.	Supplemental horizontal control.	Local horizontal control.
pose	Transcontinental traverses. Satellite observations. Lunar ranging. Interferometric surveying.	Primary arcs. Metropolitan area surveys. Engineering projects.	Area control. Detailed surveys in areas of very high land value.	Area control. Detailed surveys in areas of high land value.	Area control. Detailed surveys in areas of moderate and low land value.
work esign	Control develops the national network.		Control strengthens the national network.	Control contributes to the national network.	Control referenced to the national framework.
uracy	1:1 (000 000)	1:100 000	1:50 000	1:20 000	1:10 000 1:5 000
cing	Traverses at 750 km. Spacing-stations at 15 to 30 km or greater. Satellite as required.	Arcs not in excess 100 km. Stations at 15 km. Metropolitan area control 3–8 km.	Stations at 10 km. Metropolitan area control at 1–2 km.	As required.	As required.
mples f use	Positioning and orientation of North American continent. Continental drift and spreading studies.	Surveys required for primary framework. Crustal movement. Primary metropolitan area control	Metropolitan area densification. Land subdivision. Basic framework for densification.	Mapping and charting. Land subdivision. Construction.	Local control. Local improvements and developments.

pass all of North America including Greenland. The m, however, will be a global, or absolute, datum. Its m (geocenter) and shape and scale (flattening and semi-or axis) will be obtained by gravimetric and dynamic date methods, as well as from Doppler observations at 250 stations. The orientation of this datum will be rarily controlled by Very Long Baseline Interferometry BI) observations (Section 24.4.2) which are ultimately to an astronomic coordinate system. The new adjust-t will result, in the U.S., in positions of over 180,000 ons with accuracies of about 1 part in 10^6 [Bossler,]. The deflections of the vertical required for the data ction to the reference ellipsoid were determined at every m the network, computed in the U.S. from about 5000 geodetic deflections in conjunction with gravimetric ections of the vertical (Section 24.2.2). The gravimetric ections were estimated to an accuracy of approximately 1" to 2" from about one million gravity anomalies [C.R. Schwarz, 1978].

In an attempt to construct a unified reference system for military purposes, the World Geodetic System (WGS) has evolved primarily out of improvements, through satellite techniques [NASA, 1977], in the estimation of the size and shape of the earth and its gravity field. In January 1974 the current global datum WGS72 was published by the Department of Defense and represents its most comprehensive study of the earth and its gravity field to date. Station positions can now be determined to an accuracy of 5 to 10 m [Western Space and Missile Center, 1981], thus establishing a truly worldwide geocentric system. One of the features of WGS72 is that it combined vast amounts of data consisting of different data bases (satellite, surface gravity, astrogeodetic, etc.) into a homogeneous unified solution by a large least squares adjustment.

Table 24-2. The most important datums of the world, their associated ellipsoids, and their origin points [extracted from NASA, 1973]

Datum	Reference ellipsoid	Semimajor axis [m]	1/flattening	Station	Origin point coordinates ϕ [°′″]	λ [°′″]
Australian Geodetic	Australian National	6378160	289.25	Johnston Geodetic Station	−25 56 54.55	133 12 30.08
Arc-Cape (South Africa)	Clarke (1880)	6378249.145	293.465	Buffelsfontein	−33 59 32.000	25 30 44.622
European 1950	International (1909)	6378388	297.	Helmertturm	52 22 51.45	13 03 58.24
Indian	Everest (1830)	6378206.4	300.8017	Kalianpur	24 07 11.26	77 39 17.57
North American Datum 1927	Clarke (1866)	6378206.4	294.978698	Meades Ranch	39 13 26.686	261 27 29.494
Pulkova 42	Krassovaki (1940)	6378245	298.3	Pulkovo Observatory	59 46 18.55	30 19 42.09
S. American 69	S. American (1969)	6378160	298.25	Chua	−19 45 41.653	311 53 55.936
Tokyo	Bessel (1841)	6377397.155	299.152813	Tokyo Observatory (old)	35 39 17.51	139 44 40.50

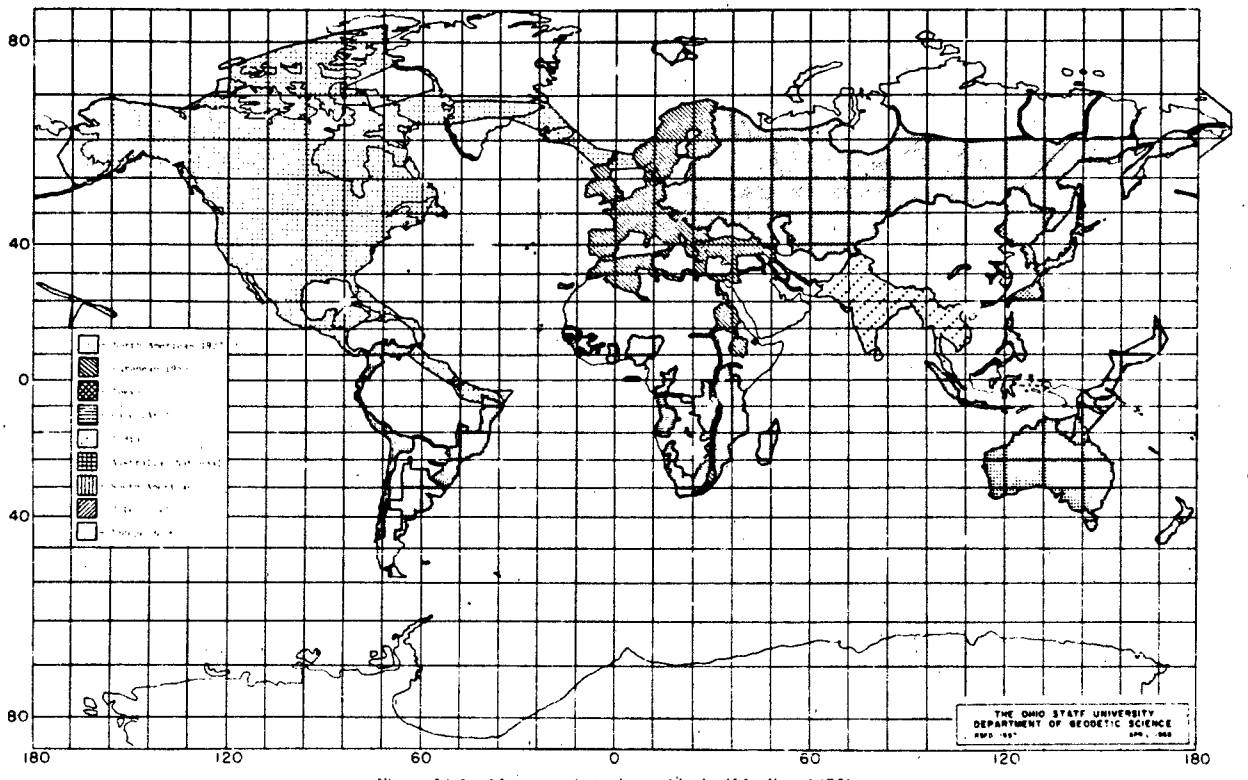


Figure 24-3. Major geodetic datum blocks [Mueller, 1975].

Table 24-3. Values of transformation parameters: Datum to Satellite WN14 System (satellite datum); parenthetical values are standard deviations [extracted from Mueller, 1975].

Transformation parameters \ Datum	Australian National (3 stations)	European 1950 (16 stations)	NAD27 (21 stations)	S. American 69 (10 stations)
ΔX [m]	15/ (3.2)	−99.4 (5.0)	−31.7 (1.7)	−96.6 (4.0)
ΔY [m]	−59.1 (3.2)	−132.0 (5.0)	142.3 (1.6)	−13.7 (4.3)
ΔZ [m]	131.2 (3.6)	−116.0 (4.8)	177.3 (1.5)	−29.4 (4.1)
δ (10^{-6})	1.14 (1.83)	6.06 (2.83)	−0.96 (0.98)	−6.67 (1.41)
ε ["]	−0.35 (0.38)	0.30 (0.65)	−0.33 (0.18)	−0.02 (0.29)
ξ ["]	0.49 (0.64)	−0.13 (0.98)	−0.34 (0.35)	−0.03 (0.35)
η ["]	1.31 (0.52)	0.26 (0.70)	0.84 (0.21)	−0.66 (0.43)

Values used to define the WGS72 ellipsoid and its associated parameters include the semimajor axis (a), the product of the gravitational constant (G) and the earth's mass (M), the earth's angular velocity (ω), and the second degree zonal harmonic coefficient of the geopotential ($C_{2,0}$) (Section 24.2.1).

24.1.2 Vertical Control

Vertical control is generally established by leveling networks that are tied to mean sea level. The networks consist of loops of leveling lines, the misclosure of the loops (the sum of the height differences around the loop should be zero) being indicative of the measurement errors. Since the level surfaces of the earth's gravity field are not parallel, the heights above the geoid depend also on gravity, and for precise work, cannot be deduced solely from leveling data. The height of the topographic surface above the geoid is known as the orthometric height. The orthometric correction, calculable with the aid of gravity measurements and applied to leveled heights, accounts for the nonparallelism of the level surfaces. The leveling networks are classified as first-, second-, and third-order according to the spacing, length, and method of the leveling run. First-order surveys are the most accurate and establish the primary vertical control on a national basis; see Table 24-4 for a brief summary of accuracy standards and recommended uses of the various survey classes.

The leveling networks are tied to tidal stations, these providing the only physical access to the geoid. At these stations, mean sea level is ascertained by averaging sea level recordings over several years to average both short and long period tidal fluctuations. In the U.S. and Canada over 100 000 km of first-order leveling lines, constrained by mean sea level determined at 26 tidal stations along the Pacific and Atlantic coasts and the Gulf of Mexico, were used to establish the National Geodetic Vertical Datum of 1929 [NOAA, 1980]. The readjustment of the geodetic vertical control for North America is currently in progress with an expected completion date of 1987. The data from existing surveys plus the planned new levelings combine to form a total of over 800 000 km of first- and second-order leveling lines in the U.S. alone (Figure 24-4). The gravity support for the vertical control is guided by the requirement that the effect of gravity on the observed heights should be determined to within 1 cm [Whalen, 1980].

A problem of some concern, and which has yet to be satisfactorily resolved, is the establishment of the datum. With modern accuracy standards, the presumption that mean sea level coincides everywhere, at any time, with the geoid must yield to the realization that differences of up to 60 to 70 cm may exist in mean sea level between the Atlantic and Pacific coasts as well as 1.5 mm/yr changes in the geoid due to the worldwide change in mean sea level [Castle and Vanicek, 1980]. With the difficulty of the physical accessibility to the geoid at this level of accuracy (tens of cm), the establishment of a World Vertical Datum awaits improvements in the accuracy and extent of global and local gravity data.

24.1.3 Inertial Positioning

Inertial positioning (or survey) systems evolved in the early to mid 1970s from inertial navigation systems, used primarily for military purposes, when geodetic accuracies were warranted by the use of a stabilized platform to isolate the measuring device from the irregular motions of the vehicle. A system of gyroscopes maintains the orientation of the platform which is suspended by gimbal supports. A set of accelerometers is used to measure the accelerations of the survey vehicle in the three mutually orthogonal directions. One integration of these data yields the velocity and

Table 24-4 Standards for the Classification of Geodetic Vertical Control and principal uses [extracted from NOAA, 1980a]

Classification	Vertical Control				
	First-Order		Second-Order		Third-Order
	Class I	Class II	Class I	Class II	
Relative accuracy between directly connected points or benchmarks (standard error)	0.5 mm $\sqrt{K}$	0.7 mm $\sqrt{K}$	1.0 mm $\sqrt{K}$	1.3 mm $\sqrt{K}$	2.0 mm $\sqrt{K}$
	(K is the distance in kilometers between points)				
Recommended uses	Basic framework of the National Network and metropolitan area control. Regional crustal movement studies. Extensive engineering projects. Support for subsidiary surveys.		Secondary framework of the National Network and metropolitan area control. Local crustal movement studies. Large engineering projects. Tidal projects. Tidal boundary reference. Support for lower order surveys.	Densification within the National Network. Rapid subsidence studies. Local engineering projects. Topographic mapping.	Small-scale topographic mapping. Establishing gradients in mountainous areas. Small engineering projects. May or may not be adjusted to the National Network.

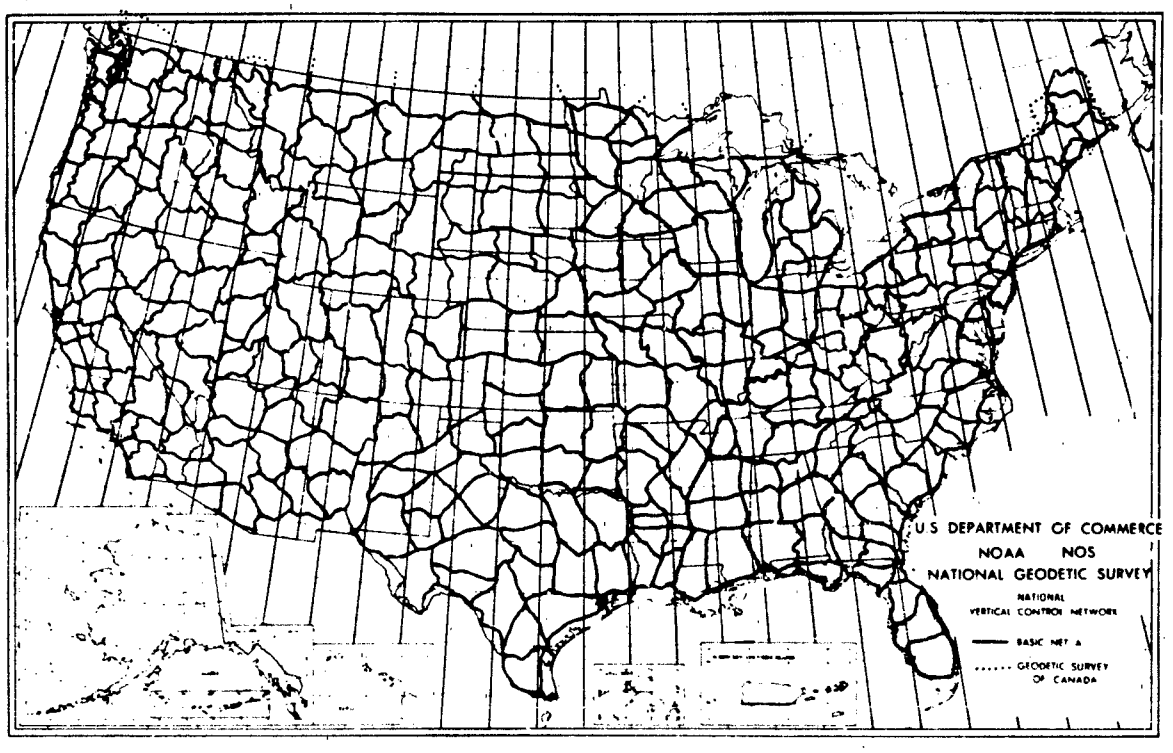


Figure 24-4 National Geodetic Vertical Control Network [NOAA, 1980b].

a second integration yields the distance traveled. Obviously, the acceleration due to gravity must be taken into account. The vehicle (for example, truck or helicopter) carrying the instrument must come to rest every 2, 3, or 4 minutes in order to impose a zero velocity on the velocity output, thus reducing the growth of systematic instrumental errors. For traverse distances of about 75 km, present instrumentation and analyses produce about 50 cm accuracy in horizontal and 40 cm in vertical position. For horizontal control, this is first-order accuracy (Table 24-1). Further improvements in hardware, as well as software, will undoubtedly increase the potential of inertial surveying for geodetic applications [Bose and Huddle, 1981].

24.2 GRAVIMETRIC GEODESY

24.2.1 The Normal Gravity Field

Just as the earth's surface is not a perfect sphere, or even an ellipsoid, the earth's gravity field is not the field generated by a homogeneous ball. Because of the overall ellipticity in the earth's shape, as well as the centrifugal force produced by its rotation, the values of gravity increase as one travels from the equator to either pole. This general trend is conveniently described by a so-called normal gravity field which then serves as the reference against which the actual gravity field is measured. The normal field is defined, in part, by requiring one of its equipotential (or level) surfaces to be an ellipsoid oriented in an earth-fixed coordinate system. Stokes proved that four quantities—the surface potential, the size of the ellipsoid, its shape, and the rotational speed (ω)—suffice to uniquely define the (exterior) normal gravity field. The size and shape of the ellipsoid are specified, for example, by its semimajor axis (a) and its flattening (f). The definition of the field retains its uniqueness if the value of the (constant) surface potential is replaced by the value of the earth's total mass (the combination GM is more accurately determined than the mass M alone, where G is the gravitational constant, $G = 6.67 \times 10^{-8}$ cm^3 g^{-1} s^{-2}), or by the value of equatorial gravity, γ_e. Similarly, f can be replaced by $C_{2,0}$, the second-degree zonal harmonic coefficient in the spherical harmonic expansion of the earth's gravity potential. Once these parameters are specified, the "normal gravity" on the ellipsoid can be calculated as accurately as desired.

The concerted effort to establish a reference ellipsoid and associated normal field that would find international acceptance and application did not materialize until well into this century culminating in the adoption by the International Union of Geodesy and Geophysics [IUGG] in 1930 of the following parameter values:

$$a = 6378388 \text{ m}$$

$$f = 1/297$$

$$\gamma_e = 9.780490 \text{ ms}^{-2}$$

$$\omega = 7.292115 \times 10^{-5} \text{ s}^{-1}$$

The inevitable improvements in observation techniques and instrumentation (particularly with the advent of satellite methods) prompted a redefinition of the reference system in 1967, and again in 1980, when the IUGG agreed on the following [Chovitz, 1981]:

$$a = 6378137 \text{ m} \qquad (\pm 2 \text{ m})$$

$$C_{2,0} = 0.00108263 \qquad (\pm 1 \times 10^{-8})$$

$$GM = 3.986005 \times 10^{14} \text{ m}^3 \text{ s}^{-2} \quad (\pm 4 \times 10^7 \text{ m}^3 \text{ s}^{-2})$$

$$\omega = 7.292115 \times 10^{-5} \text{ s}^{-1} \quad (\pm 2 \times 10^{-12} \text{ s}^{-1})$$

Parenthetical values are standard deviations. This constitutes the Geodetic Reference System 1980. GM includes the mass of the atmosphere and $C_{2,0}$ excludes the permanent tidal deformation. The gravity formula based on these parameters is given by

$$\gamma = 9.780327 (1 + 0.00530224 \sin^2 \phi - 0.000058 \sin^2 2\phi) \text{ms}^{-2}$$

where ϕ is the geodetic latitude; it yields normal gravity accurate to six digits. The method of computing the coefficients of this formula can be found in International Association of Geodesy (IAG), [1971].

24.2.2 Disturbing Gravity Field

The normal gravity, as given by the formula above, approximates the true gravity of the earth to about 5 parts in 10^5. This is hardly adequate for geodetic and modern geophysical and navigational needs, for which the variations in the gravity field at the 10^{-4} to 10^{-6} ms^{-2} (1 part in 10^5 to 10^7) level have a significant influence. In this respect, the direction of the gravity vector is as important as its magnitude.

For calculational convenience, the true gravity field is referenced to the normal field, the difference being the "disturbing" or "anomalous" field. Thus, the disturbing potential is the difference between the actual potential and the normal potential and the gravity disturbance is the difference between magnitudes of the true gravity and the normal gravity. The gravimetric (or absolute) deflection of the vertical is the angle between the true gravity vector and the normal gravity vector; it differs from the astrogeodetic deflection in that the perpendicular to the geodetic ellipsoid generally does not coincide with the direction of normal gravity, especially if the datum to which the astrogeodetic deflection refers is not geocentric. The geoid (the equipotential surface of the true gravity field that closely approximates mean sea level) is separated from the ellipsoid by the geoid undula-

tion, or geoid height. And finally, the gravity anomaly is the difference between the earth's gravity magnitude on the geoid and the normal gravity magnitude on the ellipsoid, the two points being connected by the perpendicular to the ellipsoid. This definition assumes the absence of masses external to the geoid. In practice, the computation of geoidal gravity anomalies requires (through various techniques) a reduction of surface measurements to the geoid, thus accounting for the intervening masses. The gravity anomaly differs from the gravity disturbance in that the normal gravity is computed on the ellipsoid instead of the geoid, the latter computation being impossible if the undulation is not known.

Assuming that the geoid and reference ellipsoid enclose the same total mass and have identical centers of mass and surface potentials, the geoid undulation is directly proportional to the disturbing potential, according to Bruns's formula [Heiskanen and Moritz, 1967]. Its global average is then zero and the global root mean square (rms) value is approximately 30 m. The sources for the geoidal variations are primarily density inhomogeneities below the earth's crust. Table 24-5 summarizes the relative distribution of the sources for the variations in the geoid (also in gravity and its gradient) according to an earth model comprising various density layers at prescribed depths. Therefore, the major features of the geoid have a long-wavelength structure, and its large-scale variations are generally not correlated with the earth's topography (Figure 24-5). On the other hand, the short wavelength constituents of the disturbing potential (hence the geoid) exhibit a definite correlation with topography (Figure 24-6). In addition to its importance in geodetic positioning, the geoid undulation plays an essential role in oceanographic studies where a knowledge of deviations of the actual sea surface from the geoid contributes to the understanding of ocean surface circulation and ocean tides. The so-called sea surface topography has magnitudes of at most 1 to 2 m, and an undulation accuracy of about 10 cm at 100 km (or lower) resolution is required to definitely discern these features [NRC, 1979]. The satellite-borne altimeter is the complementary facet to the problem of determining the sea surface topography. It provides the ocean surface height above the reference ellipsoid (Section 24.3.4). From the satellites GEOS-3 and SEASAT-1 a new wealth of geoidal data was obtained over most of the earth's ocean surface yielding an oceanic geoid at or below the 1 m accuracy level (corrections for the sea surface topography were either modeled or calculated using global tidal models, or they were neglected).

The gravity anomaly, being essentially the derivative of the disturbing potential, is a more local phenomenon than the geoid undulation. Gravity anomalies are generated primarily by shallow density variations, that is, in the earth's crust. Considered over the entire earth, they have an rms value of about 40 to 45 mgal (1 mgal = 10^{-3} gal = 10^{-5} ms^{-2}). Classical formulas exist that relate global sets of gravity anomaly data to both the geoid undulation and the (gravimetric) deflection of the vertical (Stokes and Vening-Meinesz formulas). In the last decade new, alternative estimation techniques (least squares collocation) have been fully developed to provide speed, efficiency, and versatility to the problem of determining undulations and deflections from gravimetry. Aside from purely geodetic interests, gravity anomaly data have found wide application in geological studies ranging from the determination of geologic strata configurations, including isostatic compensation models, to the investigations of tectonic motions and mantle convective processes. The increase in the quality and quantity gravity measurements at sea provides an improved understanding of the origin, properties, and continuing evolution of the oceanic lithosphere. The accuracy currently required for these applications is generally less than 10 mgal with a resolution of about 100 km [NRC, 1979]. A detailed global map of gravity data has been constructed by the Defense Mapping Agency, Aerospace Center [DMAAC, 1973] which has compiled a 1° × 1° mean anomaly field from various sources, comprising a total of 27 441 (as of 1973) values and their estimated accuracies. However, gravity anomalies in unsurveyed areas, particularly in the southern hemi-

Table 24-5. Contributions of each density layer to variances of undulation, anomaly and gradient [Jordan, 1978].

Depth of Density Contrast [km]	Undulation Model		Free-Air Gravity Anomaly Model		Vertical-Vertical Gravity Gradient Model	
	rms Value [m]	Percent of Total Variance	rms Value [mgal]	Percent of Total variance	rms Value [E]	Percent of Total Variance
0.15	0.4	. . .	24.4	33	28.7	99
	3.2	1	28.5	44	3.2	1
	7.2	6	14.7	12	0.1	. . .
350	20.4	44	12.9	9	0.2	. . .
2880	21.5	49	5.4	2	. . .	. . .
	30.7†		42.7†		28.9†	

†Root-sum-square values

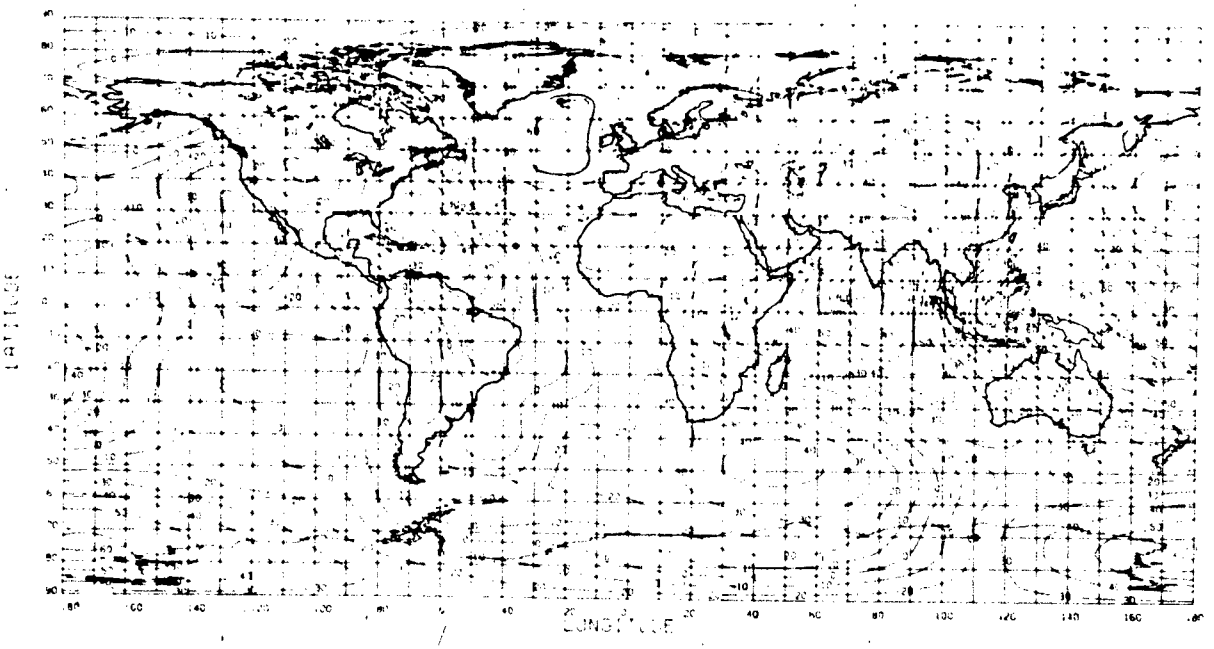


Figure 24-5. Long-wavelength geoid undulation contours (contour interval = 10 m) from (14,14) solution using weighted GEM 10 coefficients and altimetric data; reference flattening = 1/298.25 [Hadgigeorge, 1981].

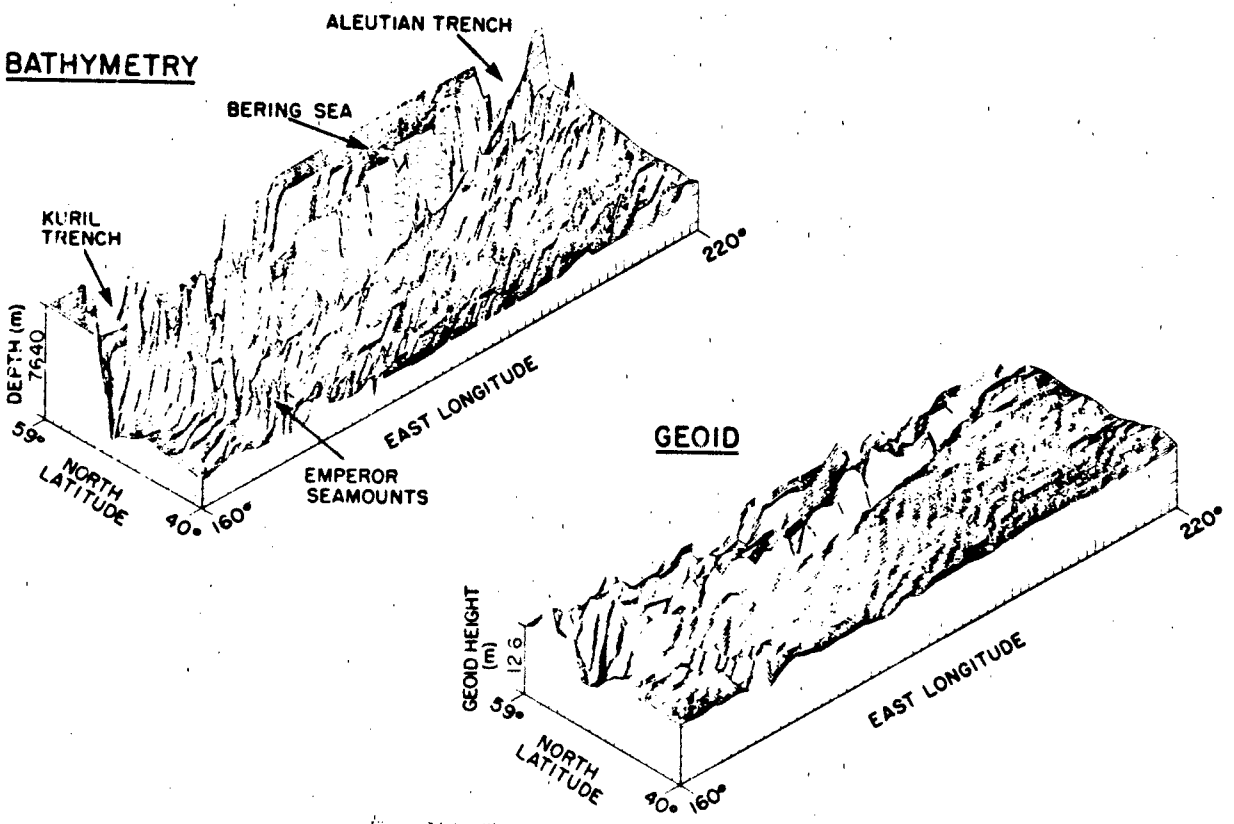


Figure 24-6. Short wavelength geoid versus bathymetry.

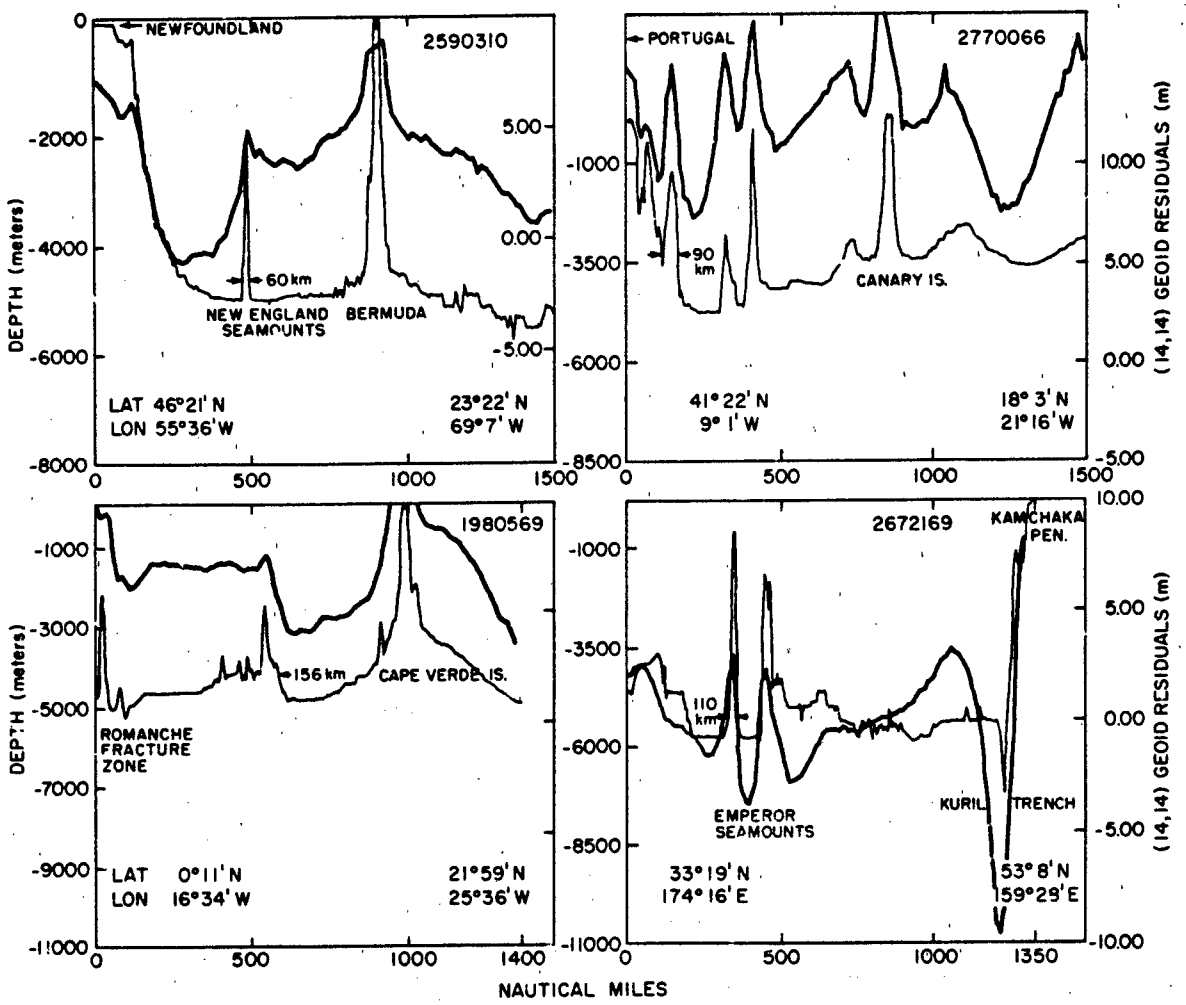


Figure 24-6. (*Continued*)

sphere, in inaccessible areas of the major continents, and in the polar regions, have a relatively large uncertainty since in these areas they are derived through special estimation techniques including the assumed correlation of gravity with geologic formations. The accuracies in the entire set range from a few mgal to tens of mgal.

The principal geodetic application of the gravimetric deflection of the vertical is in the reduction of horizontal control data (specifically directions) to the geocentric reference ellipsoid (absolute datum: Section 24.1.1). The deflection can be computed from gravimetric data to an accuracy of about 1" to 3", or better, depending on the density, accuracy, and extent of available data, the estimation technique, and the terrain [Lachapelle, 1978; Tscherning and Forsberg, 1978]. Since the vertical deflection defines the direction of the true gravity vector with respect to the normal gravity vector, due consideration of the deflection is required in any precise inertial navigation as well as positioning system where the platform is aligned with respect to the local vertical. Conversely, differences in the deflection of the vertical can be determined accurately (anticipated accuracy $< \pm 1$") using inertial positioning data [Schwarz, K.-P., 1978].

For global descriptions of the earth's potential field, it is convenient to use an infinite series of spherical harmonic functions in which the coefficients, regarded as parameters of the field, have been determined by observing the perturbations of satellites in their orbits. For example, one such solution is the Goddard Earth Model 9 (GEM 9), complete to degree and order 20 [Lerch et al., 1979]. Additional terrestrial data were used, including altimetry, to develop GEM 10B, yielding coefficients complete to degree and order 36. With the near global and very dense coverage of altimetry data, solutions to degree and order 180 and higher

have been developed [Rapp, 1979]. Not including inaccuracies in the determined coefficients, a rough estimate of the error in the corresponding geoid undulation models is given by the rule 64/n [meters], where n is the highest degree of the coefficients, and represents the neglect of higher-degree information (omission error). For example, the GEM 9 geoid undulations, apart from a formal standard error of about 1.9 m in the expansion to degree 20, contain an omission error of about 3 m.

24.2.3 Gravimetry

Gravity measurements are characterized as either relative or absolute. Absolute measurements result in the total magnitude of the gravity vector and involve the determination of length and time. For example, the acceleration due to gravity is determined from the length of a pendulum arm and its period of oscillation. With the best pendulum instruments and technique, this method gives a few tenths of a mgal accuracy [Torge, 1980]. In the modern, more sophisticated, and accurate methods, the apparatus consists of a proof mass which moves (falls) in the gravity field. The distance through which it moves is measured by interferometric methods (Michelson interferometer), while the corresponding time is determined electronically with an atomic frequency standard. The accuracy today is about 10 μgal (1 μgal = 10^{-3} mgal = 10^{-8} ms^{-2}) for laboratory instruments, but formal precisions of 1 to 3 μgal have been reported [Sakuma, 1976].

Because of the exacting demands in stability of the measurement environment, absolute determinations are not practical or economical for widespread surveys, although transportable instruments have been developed by various organizations [Hammond and Iliff, 1978]. Instead, with a much simpler apparatus and less time one can measure the relative gravity, that is, the change in gravity between stations. In principle, only one absolute site is necessary to fix the "origin" of an entire network of relative measurements. The measurement principle of the gravimeter can be stated as determining the displacement of a mass attached to a spring while the instrument is moved from one station to the next. The best achievable accuracy in measuring gravity differences with today's instruments is several μgal [Harrison and LaCoste, 1978] and depends primarily on correct calibration models.

In 1971 the International Gravity Standardization Network (IAG 71) was completed and adopted by the IUGG. It comprises gravity values at 1854 stations around the world; at ten stations absolute measurements were made. After a final adjustment of the entire network, the standard errors of the gravity values were estimated to be about 0.1 mgal. Further regional densifications and improvements are in progress using transportable absolute gravimeters.

24.2.4 Gradiometry

Gradiometry involves the measurement of gradients of components of the gravity vector, that is, the change in the gravity components with distance. The principle of measurement is based either on determining the difference in outputs of two accelerometers separated by a known distance, or sensing the ensuing torquing action on specific known proof mass configurations. Precise gradiometers have been in the development stage for many years, but only since 1982 have they begun to prove their viability in routine geodetic applications. The essential technological difficulties include eliminating external vibrations and thermal noise. Also, because the gradiometer senses primarily shortwavelength gravity variations, the determination of the total gravity field requires supplemental long-wavelength information.

The vertical gradient of normal gravity is about 0.31 mgal/m; however, of greater interest is the gradient of the gravity anomaly. This is an extremely local phenomenon, and it is unavailing to attempt to quantify it using a global rms value. Any component can vary from a few Eötvös (1 Eötvös = 1 E = 0.1 mgal/km = 10^{-9} s^{-2}) to several hundred or even a thousand Eötvös, depending on the terrain and the underlying mass anomalies.

Some development and analysis has been oriented toward a satellite-borne gradiometer to be used in a global gravity mapping mission; however, the realization of this concept lies some years in the future. A renewed effort is underway to develop the airborne gradiometer for the purpose of providing, in an efficient manner, a detailed survey of the gravity field in local areas. This device should be operational, attaining an expected precision of a few E, within the next years. In order to account for the varying effects of gravity, the gradiometer has also been proposed as a means of improving the accuracy of inertial navigation systems [Doukakis, 1979; Wells, 1981].

24.3 SATELLITE GEODESY

The two basic methods of tracking artificial earth satellites from earth stations are the dynamic method and the geometric method. The purpose of the dynamic method is to establish an ephemeris for the satellites; that is, to enable the prediction of the satellite's position in its orbit as a function of time. This entails solving (from worldwide tracking data, such as Doppler Ranging) for the parameters of the perturbing forces, such as the gravitational force, atmospheric drag, lunar-solar attraction, and solar radiation pressure of which the orbital elements are a function.

The perturbing forces cause secular and long- and short-periodic variations in the normal (elliptical) Keplerian orbit that would be described by the satellite if the earth were a point mass in an otherwise empty universe. Table 24-6 lists

Table 24-6. Disturbing forces on earth-orbiting satellite (Blitzer, 1970).

Source	Disturbing force per unit mass [m/s²] Altitude 150 km	Altitude 750 km	Altitude 1500 km	Synchronous Altitude
Central Gravity	9.35	7.85	6.42	0.22
Earth Oblateness (J_2)	30×10^{-3}	20×10^{-3}	14×10^{-3}	160×10^{-7}
(J_3)	0.09×10^{-3}	0.06×10^{-3}	0.04×10^{-3}	0.08×10^{-7}
(J_4)	0.07×10^{-3}	0.04×10^{-3}	0.02×10^{-3}	0.01×10^{-7}
Principal Tesseral Harmonic (J_{22})	0.09×10^{-3}	0.07×10^{-3}	0.04×10^{-3}	0.5×10^{-7}
Atmospheric Drag Vanguard I*	3×10^{-3}	10^{-7}	—	—
Echo I**	1300×10^{-3}	500×10^{-7}	—	—
Lunar-Solar Attraction ***	10×10^{-7}	10×10^{-7}	10×10^{-7}	70×10^{-7}
Solar Radiation Pressure Vanguard I*	10^{-7}	10^{-7}	10^{-7}	10^{-7}
Echo I**	500×10^{-7}	500×10^{-7}	500×10^{-7}	500×10^{-7}

*Area/Mass = 2.12×10^{-2} m²/kg
**Area/Mass = 10.2 m²/kg
***This is not the direct attraction but the effective disturbing force.

a coarse, comparative summary of typical values of the various perturbing effects (central gravity is the force per unit mass due to the earth considered as a point mass—it is not a perturbing element). J_2, J_3, J_4, and $J_{2,2}$ are coefficients of the expansion into spherical harmonic functions of the earth's gravitational potential. Particularly noteworthy is the importance of the area-to-mass ratio of a satellite with respect to atmospheric drag and solar radiation pressure, the latter predominating with increasing altitude. The overall design of a satellite orbit is therefore contingent on the purpose of the satellite mission. High orbits of small, but massive, satellites (such as LAGEOS) are designed for station positioning or for the determination of polar motion and variation of earth rotation, while low orbits are required to detect variations in the earth's gravitational field or to study other specific perturbing effects such as atmospheric drag.

With the geometric method, the tracking data are obtained simultaneously from several stations and used to determine relative station positions (satellite triangulation); accurate knowledge of the satellite's orbit is not required. Various intermediate methods exist: for example, the short arc method, where observations are generally limited in duration (approximately 1/4 orbit), and either a fixed (nonperturbed) or computed orbit is used. For this method, the relative station positions become a part of the set of unknowns to be determined. In point positioning, observations from a single station determine its position in a coordinate system defined by the satellite ephemeris.

24.3.1 Doppler Positioning

Of the several new techniques of observing satellites, the Doppler method has been the most successful in achieving worldwide and routine application. By measuring the change in frequency (corresponding to the Doppler shift) of a signal transmitted from a satellite with respect to a reference frequency on the ground, one can deduce the range and range rate to the satellite. Usually two signals at different frequencies are transmitted simultaneously to essentially eliminate the frequency dependent effect of the ionosphere on the propagation of the signal. Most satellites now transmit in the hundreds of MHz, but to further reduce the ionospheric effect this is increased (for example, for NAVSTAR) into the GHz range.

A system of six satellites, the Navy Navigation Satellite System (NNSS) in operation since 1971, has provided a global positioning capability within two hours of observation. These satellites transmit a "broadcast ephemeris," which provides the observer with the satellite's position as a function of time. This information is updated and loaded into the satellite's memory once or twice a day and is used for

real-time positioning. A "precise ephemeris" is available for more accurate (but delayed in time) position fixes. The three-dimensional spatial position and the time synchronization offset are obtained by the simultaneous observation of four satellites with a corresponding positional accuracy of about 25 m, sufficient for some, but certainly not all, navigational requirements. With longer periods of observation (six to eight days), combining data from many passes, and using the precise ephemeris, accuracies below the 1 m level can be obtained [Brown, 1979]. Even greater accuracy comes from the use of the short arc method to deduce relative positions (baseline accuracy of 20 to 30 cm). The NNSS is to be superseded by the NAVSTAR/GPS (Global Positioning System) in the late 1980s. This system will consist of 18 satellites, 6 each in 3 orbital planes, providing a four-satellite configuration observable anywhere on the earth at almost all times. The accuracy for point positioning using one hour's observations and the broadcast ephemeris is an estimated 10 m [Fell, 1980], while observations spanning five days encompassing many passes in conjunction with the precise ephemeris yield geodetically useful accuracies of about 50 cm. The main source of error is the uncertainty in the satellite's position owing, in part, to an inadequate knowledge of the earth's external gravitational field. Baseline determinations of 300 km or less in length, for which inaccuracies in the determination of the satellite's orbit are less detrimental, are anticipated to be accurate to less than 10 to 20 cm, approaching the 1 cm level with improved tropospheric error modeling [Anderle, 1980].

The Doppler principle also has applications in space, providing the means for one satellite to track another either in a high-low configuration (the high [geosynchronous] satellite tracks the low satellite) or a low-low configuration where one satellite follows the other in the same orbit. The analysis of satellite-to-satellite tracking has thus far been oriented toward determinations of the gravity field. The basic measurement is the relative velocity between the two satellites which is indicative of the difference in gravitational potentials [Rapp and Hajela, 1979].

24.3.2 Satellite Interferometry

Radio interferometry using satellite signals is one of the newest techniques to make geodetic linear measurements. The length of baseline vectors between pairs of survey marks ranges from 0.1 to 200 km (short-baseline interferometry) and is determined by radio-interferometric observations of the signals broadcast by satellites of the GPS. The signals are much stronger than the quasar sources used in VLBI (Section 24.4.2). Errors can be reduced to the 1 cm level, or less, by properly accounting for the tropospheric effects [Rogers et al., 1978; Counselman and Shapiro, 1979; MacDoran, 1979; Anderle, 1981].

24.3.3 Laser Ranging

Just emerging from the exploratory and experimental stage, satellite laser ranging promises to become a valuable asset for geodetic applications. Basically, the ranges are deduced from the measured travel time of laser pulses. The overriding advantage of the laser is its tremendous accuracy, achieved by utilizing shortpulsed ($\sim$ 1 ns) lasers to enhance the definability of the returning pulse. Also, the satellite need only be equipped with retroreflectors, while for Doppler ranging, an on-board power supply is required. In contrast to Doppler ranging, however, laser ranging is weather-dependent. Ranging data to the satellite LAGEOS have been used to determine baseline lengths; for example, Kolenkiewicz and Ryan [1981] (using observations from a three month period) report a 6 cm agreement with the VLBI determined length of a 3930 km baseline connecting Westford, Mass. and Owens Valley, Calif. The analysis of LAGEOS data has also yielded determinations of polar motion and length-of-day variations with respective formal accuracies of about 0".01 and 0.3 ms [Tapley, 1982]. As with all transmissions through the atmosphere, the random effects of the troposphere and ionosphere limit the achievable accuracy. Improvements in laser ranging will result from better atmospheric density models and by upgrading the basic measurement (travel-time of pulse) accuracy.

24.3.4 Satellite Altimetry

The satellite borne radar altimeter has had an exceptional impact on the continuing efforts to improve our knowledge of the earth's gravity field. The distance of the satellite above the ocean surface is obtained simply by measuring the travel time of a reflected radar pulse. This distance and the given position of the satellite (its orbit must be either known or solved for) in a geocentric coordinate system yield directly the distance of the ocean surface from the geocenter, or equivalently, from the reference ellipsoid (sea surface height). That is, one obtains an estimate for the geoid undulation if the sea surface topography (at most 1 or 2 cm) is neglected. The altimeter measuring accuracy was better than 1 m for the GEOS-3 satellite and its orbital position was obtained to within a few meters. Imposing the constraint of equal altimeter heights at track crossings (track crossing adjustments) and accounting for atmospheric and sea surface conditions, the GEOS-3 geoid could be determined to an optimal accuracy of 60 cm to 1 m.

The subsequent (short-lived) satellite SEASAT-1 carried an altimeter emitting a shorter pulse at a more rapid rate with a consequent 5 to 10 cm sensitivity. These data are currently being processed with expected accuracies of about 20 cm or better on a regional basis [Marsh et al., 1981; Anderle, 1980]. Figure 24-7 shows the oceanic geoid recovered from 1° × 1° mean SEASAT geoid heights. See

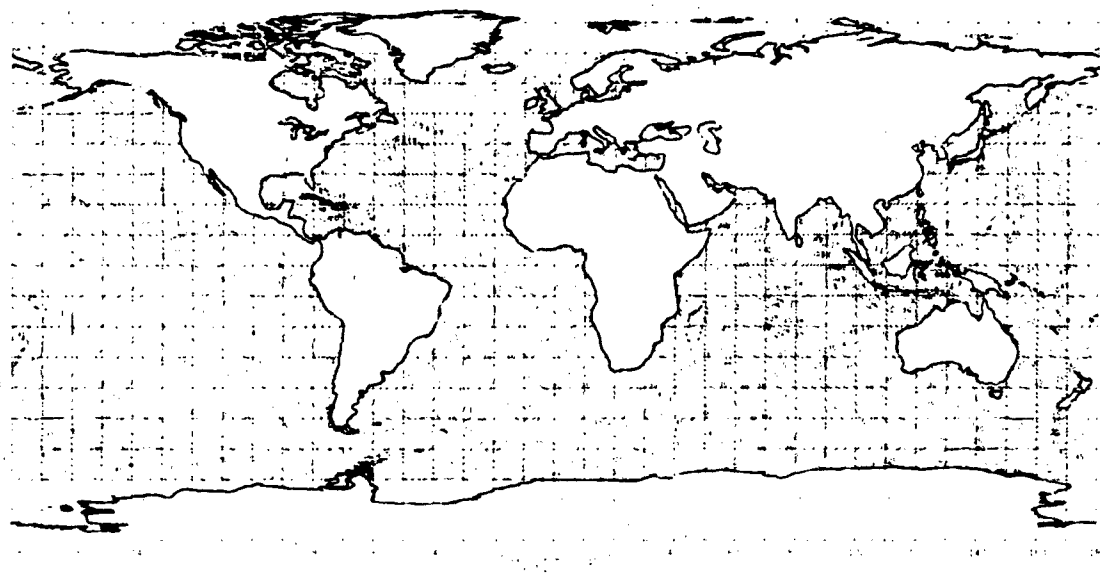


Figure 24-7. Global Seasat 1° × 1° observed geoid (contour interval = 5m); reference flattening = 1/298.25.

also Rooney et al. [1980]. The satellite GEOSAT carrying a similar altimeter is to be launched by DoD in 1985.

One of the ultimate objectives of satellite altimetry is the capability to determine and monitor ocean circulation. In this case, the sea surface topography is the object of measurement. This requires prior knowledge of the geoid, unless only temporal changes in sea surface topography are to be ascertained. The proposed five-year NASA satellite mission for this project is TOPEX; its anticipated error budget is given in Table 24-7.

Table 24-7. TOPEX Error Budget [extracted from NASA, 1981].

Source of Error	Cause of Error	Uncorrected Error [cm]	Corrected Error [cm]	Wavelength [km]
Altimeter	Altimeter noise		1.5	20
Altimeter	Ocean waves	4	1.0	1000
Troposphere	Mass of air	240	0.7	1000
Troposphere	Water vapor	20	1.0	50–500
Troposphere	Rain	—	—	
Ionosphere	Free electrons	10	0.5	50–10 000
Orbital error	Gravity	5 (km)	0.7	10 000
Orbital error	Drag	5	3.6	10 000
Orbital error	Solar radiation	30	7.0	10 000
Orbital error	Earth radiation	3	1.0	10 000
Orbital error	Station location	10	3.0	10 000
Orbital error	Timing		0.2	10 000
Sea level	Weight of air	50	3.0	200–1000
Sea level	Geoid	100 (m)	1.5	200

The table summarizes the errors in satellite measurements of ocean surface topography. This brief summary assumes (a) a dense spherical satellite orbiting at a height of 1300 km tracked by a realistic network of laser stations, (b) accurate measurements of earth's gravity from the GRAVSAT mission, (c) additional data from a dual-frequency radiometer and dual-frequency altimeter on the spacecraft, (d) surface pressure with an accuracy of 3 millibars from global weather charts, (e) a spacecraft clock having an accuracy of 100 μs, (f) an average wave height on the sea surface of two meters and a wave skewness of less than 0.1, and (g) no data collected in heavy rain. The corrected error is the error in measurements made along a single satellite pass crossing the ocean basin. We expect that long-term averages of many passes will substantially reduce the random error.

24.4 GEODETIC ASTRONOMY

The position of the earth's rotational axis varies in time with respect to inertial space, as well as with respect to the earth's crust. The variations in inertial space are usually termed nutations, specifically, either forced or force-free nutations, depending on whether or not they are generated by the gravitational attraction of an extra-terrestrial disturbing body such as the sun or moon. Similarly, the variation of the spin axis with respect to the earth's crust, called polar motion, is either forced of force-free motion. The mechanisms for polar motion are generally of a geophysical nature, such as the elasticity of the mantle and the liquidity of the outer core. The force-free polar motion of the instantaneous rotation axis is also known as wobble; the predominant variation is called the Chandler wobble and has a period of about 435 days. If the rotational axis serves as the third axis in both an earth-fixed coordinate system, as well as in the celestial (inertial) system, then nutations change the declinations of stars and polar motion affects the latitudes of terrestrial stations. This permits the observational separability of the two types of variations in the spin axis. Tables 24-8 and 24-9 list the amplitudes, periods, and causes of the various motions. With the increase in accuracy of astronomical observations, the earth cannot be considered a rigid body. One must account for the effects on the motions of the rotational axis that are attributable to the earth's elasticity, liquid core, redistributions of internal mass, and plate tectonics.

The variations of latitude (polar motion) are monitored by the International Polar Motion Service (IPMS) (formerly the International Latitude Service (ILS)) and the Bureau

Table 24-8 Spectrum of changes in earth's rotation [Rochester, 1973].

A. Inertial Orientation of Spin Axis	B. Terrestrial Orientation of Spin Axis (Polar Motion)	C. Instantaneous Spin Rate ω about Axis
1. Steady precession: amplitude 23°5; period = 25 700 years.	1. Secular motion of pole: irregular = 0".2 in 70 years.	1. Secular acceleration: $\dot{\omega}/\omega \simeq -5 \times 10^{-10}$/yr.
2. Principal nutation: amplitude 9".20 (obliquity); period 18.6 years.	2. 'Markowitz' wobble: amplitude = 0".02(?); period 24–40 years(?).	2. Irregular changes: (a) over centuries $\dot{\omega}/\omega \leq \pm 5 \times 10^{-10}$/yr; (b) over 1–10 years $\dot{\omega}/\omega \leq \pm 80 \times 10^{-10}$/yr; (c) over a few weeks or months ('abrupt'), $\dot{\omega}/\omega \leq \pm 500 \times 10^{-10}$/yr.
3. Other periodic contributions to nutation in obliquity and longitude, amplitude < 1"; periods 9.3 years, annual, semiannual, and fortnightly.	3. Chandler wobble: amplitude (variable) = 0".15; period 425–440 days; damping time 10–70 years(?).	
4. Discrepancy in secular decrease in obliquity: 0"./century(?).	4. Seasonal wobbles: annual, amplitude = 0".09; semiannual, amplitude = 0".01.	3. Short-period variations: (a) biennial, amplitude = 9 ms; (b) annual, amplitude = 20–2 ms; (c) semiannual, amplitude = 9 ms; (d) monthly and fortnightly, amplitudes = 1 ms.
	5. Monthly and fortnightly wobbles: (theoretical) amplitudes = 0".001.	
	6. Nearly diurnal free wobble: amplitude ≤ 0".02(?); period(s) within a few minutes of a sidereal day.	
	7. Oppolzer terms: amplitudes = 0".02; periods as for nutations.	

Table 24-9. Mechanisms with effects now distinguishable on the earth's rotation [after Rochester, 1973].

Mechanism	Effect*
Sun	A,B7,C1,C3c
Gravitational torque	C2c(?)
Solar wind torque	
Moon	A,B7,C1,C3c-d
Gravitational torque	
Mantle	B1,B3-4,C1-2a,C3c-d
Elasticity	B1,B3
Earthquakes	B3(?),C1
Solid friction	C2a
Viscosity	
Liquid core	A2-3,B2,B6
Inertial coupling	C2b-c(?)
Topographic coupling	A4(?),B3,C2
Electromagnetic coupling	
Solid inner core	B2(?)
Inertial coupling	
Oceans	B1,B3,B5,C2a
Loading and inertia	B3(?),C1
Friction	
Groundwater	B4
Loading and inertia	
Atmosphere	B4
Loading and inertia	C2c,C3a-c
Wind stress	C1
Atmospheric tide	

*Numbers refer to Table 24-8.

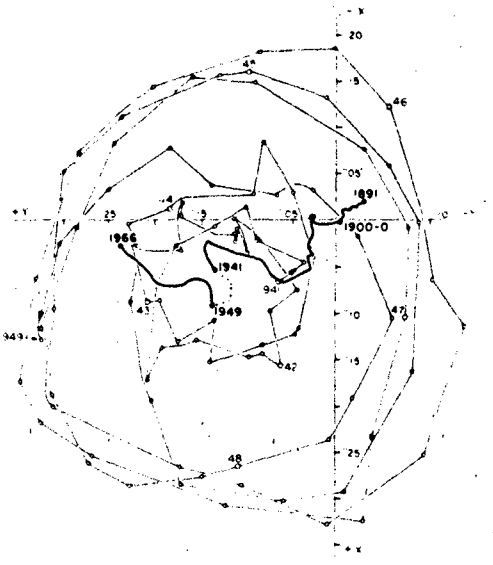


Figure 24-8. Polar wander: 1891–1966
Polar wobble: 1941–1949

International de l'Heure (BIH). These organizations coordinate the observations obtained throughout the world using classical methods (for example, zenith telescopes), as well as implementing the satellite Doppler, VLBI, LLR, etc., techniques. The results are published annually by both organizations.

Figure 24-8 shows the motion of the pole of rotation over a period of 8 years, as well as the motion (polar wander) of the average pole over a period of 75 years. The coordinate system with respect to which polar motion is measured is centered at the Conventional International Origin (CIO). This is the mean position of the pole of 1903 as defined by the latitudes of the five observatories of the ILS. Figure 24-9 shows the most recent motions (1978–1980) of the pole as determined by the BIH from worldwide observations. The origin in this case is defined by the BIH on the basis of the latitudes of all contributing observatories.

Time determination, coordination and dissemination are, on an international scale, three of the principal functions of the BIH. Because of the excellent long-term stability of atomic frequency standards (1 part in 10^{13}), atomic time (AT) has virtually replaced ephemeris time (ET) as the basic time scale. ET refers to the time variable in the theory of motion of the sun, moon, and planets and is less accurate and more difficult to determine through observations [Winkler and VanFlandern, 1977]. Several observatories around the world maintain their own atomic time; in the U.S. it is maintained by the U.S. Naval Observatory (USNO) and the National Bureau of Standards (NBS). These are pooled and intercompared by the BIH resulting in International Atomic

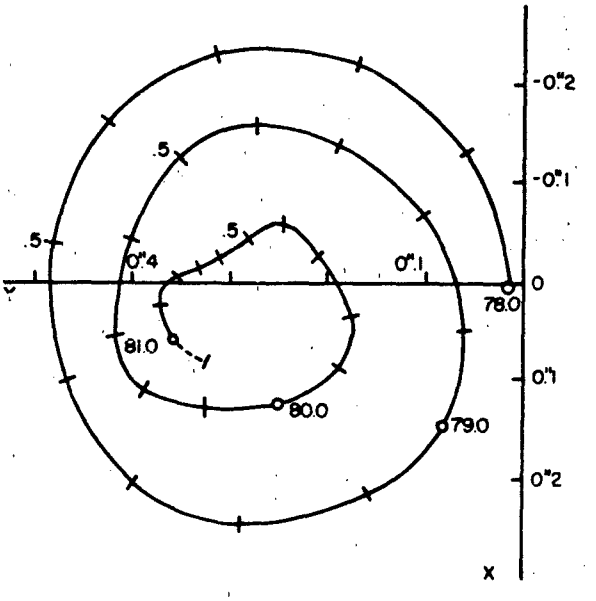


Figure 24-9. Path of the pole from 1978.0 to 1981.0 [BIH, 1981, p. B-50].

Table 24-10. Difference between the earth rotation parameters by various services [extracted from Guinot, 1981].

Services	Interval	A_x	B_x	A_y	B_y
ILS-BIH	62–78	36 ± 4	−0.2 ± 0.9	−2 ± 3	1.0 ± 0.5
	67–78	39 ± 4	2.3 ± 1.3	−4 ± 3	−0.2 ± 0.8
IPMS-BIH	62–78	29 ± 2	0.3 ± 0.3	2 ± 2	−1.1 ± 0.5
	67–78	29 ± 1	0.0 ± 0.4	1 ± 3	−0.4 ± 0.7
DMA-BIH	72–79	−15 ± 2	−1.6 ± 0.6	−2 ± 3	1.6 ± 1.1

Service—BIH = A + B (T − 1975.0)
Units are: 0".0001 for x and y.
The variations B are yearly rates.

Time (TAI). Coordinated Universal Time (UTC) is based on the same atomic time scale as AT, but is offset occasionally, by so-called leap seconds, to maintain a close relation to the actual spin rate of the earth, called UT1. UT1 is Universal Time referring to the CIO, while UTO refers to the rotational rate of the instantaneous rotation axis. Several observatories throughout the world determine UTO through astronomical observations. The longitudes of these stations are used to define the Mean Greenwich Meridian which together with the CIO provides the orientation of the average terrestrial coordinate system. Table 24-10 compares the BIH determinations of polar motion (x and y components) with the determinations of several other organizations, IPMS, ILS, and DMA which used satellite Doppler observations.

24.4.1 Lunar Laser Ranging

Retroreflectors were placed on the moon by the Apollo 11, 14, and 15 and the Soviet Luna 17 and 21 missions. Ranging to all of these reflectors began shortly after each mission, and is still continuing, primarily and regularly at the McDonald Observatory, Fort Davis, Texas. The ranging is done with high energy pulsed lasers fired through large aperture telescopes which also serve as the receiving optics for the reflected return images. Photo-electric devices record the photon receptions against a precise time record. The resulting accurate linear measurements (± 15 cm) between the observing site and the reflectors on the lunar surface are well suited to studies of lunar rotation as well as irregularities in the earth's rate of rotation (determination of UTO) and polar motion (however, the determination of the polar motion component perpendicular to the meridian requires more than one observing station). Accuracies in the time determination of about 0.5 to 0.7 ms have been stated [McCarthy, 1979], but compared to other determinations, such as by the BIH, the root mean square difference can be 1 to 2 ms [Langley et al., 1981].

24.4.2 Very Long Baseline Interferometry

The primary observable in VLBI for geodetic and astrometric studies is the measured time interval between the arrival of a radio signal, from some extragalactic source such as a quasar, at one end of the interferometer and its arrival at the other end. This interval is called the delay and its derivative is the delay rate. From a sufficient set of these data as functions of time, the geometry of the interferometer baseline and the position of the observed radio sources can be determined. Figure 24-10 shows, in simplified form, a typical conventional interferometer with two antenna-receiver systems. At each antenna, the radio frequency signal from the observed source is converted to a lower "intermediate" frequency by mixing with a local oscillator (LO) signal. The LO signals are supplied to the mixers at both antennas independently, thus allowing large separations between the receiving stations. The resulting signals at each

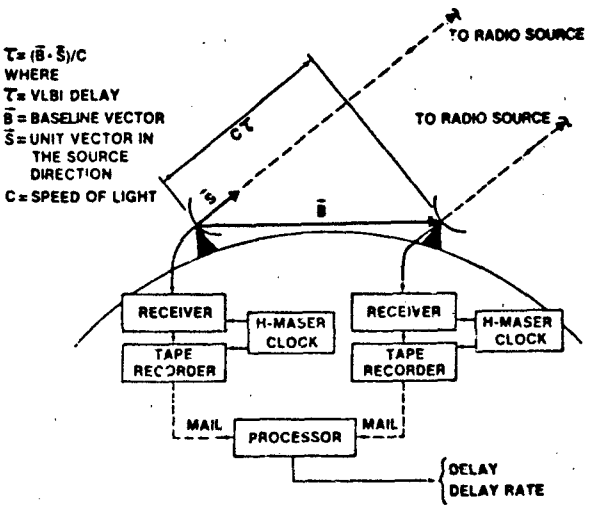


Figure 24-10. Very Long Baseline Interferometry (VLBI) observation geometry [Ryan et al., 1978].

station are recorded and time-tagged on magnetic tape. The two tapes are later cross-correlated to obtain the phase delay, which translates into time delay apart from the indeterminable multiple of $2\pi\lambda$, where λ is the wavelength of the signal. This ambiguity is resolved by combining data at several frequencies if the source has a polychromatic spectrum (bandwidth synthesis technique) [Rogers, 1970].

The state-of-the-art VLBI system incorporates the Mark III recorder, developed by the Northeast Radio Observatory Corporation (Haystack Observatory), which has an increased frequency bandwidth recording capability compared to earlier devices, and thus decreases the effect of random errors. Another parameter affecting system noise is the efficiency of collecting the signal; this depends on the size and accuracy of the antenna surface, as well as the signal strength. Transportable systems require smaller antennas and the resulting degradation in accuracy can be compensated for by increasing the recording rates. Other errors are caused by clock instabilities; inaccurate definition, location, and lack of strength of the extragalactic signal; random atmospheric disturbances; and imperfect (or lack of) modeling of geophysical effects, such as solid earth tides, crustal motion, and earth rotation. Precisely because these latter contribute to the error budget of VLBI (that is, instrumental and other environmental factors do not predominate), the continual monitoring of these geodynamic phenomena with VLBI will improve the associated models. For example, a formal precision of ″.002 in polar motion components and 0.1 ms in UT1 was reported by Ma [1981] using 14 days of observations. The published proceedings of a VLBI conference held at the Massachusetts Institute of Technology, Cambridge, Mass., in 1979 [NASA, 1980] provide a comprehensive documentation of the technology, geodetic applications, and future capabilities of VLBI.

CHAPTER 24

REFERENCES

Anderle, R.J., "Advanced Satellite Systems and Their Applications," *Manuscripta Geodaetica*, **5**:257–281, 1980.

Anderle, R.J., "Applications of NAVSTAR GPS Geodetic Receiver to Geodesy and Geophysics," *Am. Geophys.*, **37**:37–46, 1981.

Blitzer, L., *Handbook of Orbital Perturbations*, TRW Systems Group, Redono Beach, Calif, 1970.

Bose, S.C. and J.R. Huddle, "Regional Adjustment of Inertial Gravity Disturbance Vector Measurements by Optimal Two-Dimensional Smoothing," *Bull. Geodesique*, **55**:340–353, 1981.

Bomford, G., *Geodesy*, Clarendon Press, Oxford, 1971.

Bossler, J.D., "Status of the New Adjustment in the United States," in *Proceedings Second International Symposium on Problems Related to the Redefinition of North American Geodetic Networks*, Arlington, Va, April 24–28, 1978. (avail. from NTIS as PB 299048/4)

Brown, D.C., "Positioning by Satellites," *Rev. Geophys.* **17**:1424–1430, 1979.

Bureau Internationale de l'Heure, *Annual Report for 1980*, Paris, 1981.

Castle, R.O. and P. Vanicek, "Interdisciplinary Considerations in the Foundation of the New North American Vertical Datum," in *Proceedings Second International Symposium on Problems Related to the Redefinition of North American Vertical Geodetic Networks*, Ottawa, Canada, May 26–30, 1980, Canadian Institute of Surveying, Ottawa, 1980.

Chovitz, B.H., "Modern Geodetic Earth Reference Models," *Eos*, Trans. AGU, **62** (7), Feb. 17, 1981.

Counselman, C.C. and I.I. Shapiro, "Miniature Interferometer Terminals for Earth Surveying," *Bull. Geodesique*, **53**:139–163, 1979.

DMAAC, "1° × 1° Mean Free-Air Gravity Anomalies," Reference Publication 73-0002, DMAAC, St. Louis, Mo, Dec. 1973.

Doukakis, E., "Evaluation of the Performance of Gradiometer-Assisted Inertial Navigation Systems," Technical Translation DMAAC-TC-2892K, DMAAC, St. Louis, Mo, 1982.

Fell, P.J., "Geodetic Positioning Using a Global Positioning System of Satellites," Report 299, Department of Geodetic Science, Ohio State University, 1980.

Groten, E., *Geodesy and the Earth's Gravity Field*, F. Dummler, Bonn, 1979.

Guinot, B., "Comments on the Terrestrial Pole of Reference, the Origin of Longitudes, and on the Definition of UT1," in *Reference Coordinate Systems for Earth Dynamics*, edited by E.M. Gaposchkin and B. Kolaczek, D. Reidel, Dordrecht, Holland, 1981.

Hadgigeorge, G., "Seasat Altimeter Reductions for Detailed Determinations of the Oceanic Geoid," *Ann. Geophys.*, **37**:123–132, 1981.

Hammond, J.A. and R.L. Iliff, "The AFGL Absolute Gravity Program," in *Proceedings Ninth Geodesy/ Solid Earth and Ocean Physics (GEOP) Research Conference*, Columbus, Ohio, Oct. 2–5, 1978, Report 280, Department of Geodetic Science and Surveying, Ohio State University, 1978.

Harrison, J.C. and L.J.B. LaCoste, "The Measurement of Surface Gravity," in *Proceedings Ninth Geodesy/Solid Earth and Ocean Physics (GEOP) Research Conference*, Columbus, Ohio, Oct. 2–5, 1978, Report 280 Department of Geodetic Science, Ohio State University, 1978.

Heiskanen, W.A. and H. Moritz, *Physical Geodesy*, Freeman, San Francisco, 1967.

Helmert, R.F., *Die Mathematischen und Physikalischen Theorien der Hoheren Geodesie, Vols. 1 and 2*, Teubner, Leipzig, Germany, 1884.

I.A.G., "Geodetic Reference System 1967," Special Publication No. 3, International Association of Geodesy, Paris, 1971.

I.A.G., "The International Gravity Standardization Net 1971," Special Publication No. 4, Bureau Central de l'Association Internationale de Geodesie, Paris, 1971.

Jordan, S.K., "Statistical Model for Gravity, Topography, and Density Contrasts in the Earth," *J. Geophys Res*, **83**:1816–1824, April 10, 1978.

Kaula, W.M., *Theory of Satellite Geodesy*, Blaisdell, London, 1966.

Kolenkiewicz, R. and J.W. Ryan, "Ten Lageos Laser Determined Continental United States Baselines with some VLBI Comparisons," paper presented at the Annual Spring Meeting of the American Geophysical Union, Baltimore, May 25–29, 1981.

Lachapelle, G., "Estimation of Geoid Deflection Components in Canada," in *Proceedings Second Symposium on Problems Related to the Redefinition of North American Geodetic Networks*, Arlington, Va, April 24–28, 1978. (avail. from NTIS as PB 2990484/4)

Langley, R.B., R.W. King, and I.I. Shapiro, "Earth Rotation from Lunar Laser Ranging," *J. Geophys. Res.*, **86**:11913–11918, 1981.

Leick, A., "The Observability of the Celestial Pole and Its Nutation," Report 262, Department of Geodetic Science, Ohio State University, 1978.

Lerch, F.J., S.M. Klosko, R.E. Laubscher, and C.A. Wagner, "Gravity Model Improvement Using GEOS-3 (GEM 9 and 10)," *J. Geophys. Res.*, **84**:3897–3916, 1979.

Ma, C., "Geodesy by Radio Interferometry: Polar Motion and UT1 from MERIT," paper presented at the Annual Spring Meeting of the American Geophysical Union, Baltimore, May 25–29, 1981.

MacDoran, P.F., "Satellite Emission Radio Interferometric Earth Surveying, Series — GPS Geodetic System," *Bull. Geodesique*, **53**: 117–138, 1979.

Manual of Photogrammetry, 4th Edition, edited by C.L Slama, American Society of Photogrammetry, Falls Church, Va., 1980.

Marsh, J.G., R.E. Cheney, V. Grano, and T.V. Martin, "Mean Sea Surface Computations Based Upon Seasat Altimeter Data," paper presented at the Annual Spring Meeting of the American Geophysical Union, Baltimore, May 25–29, 1981.

Mather, R.S., "Four Dimensional Studies in Earth Space," *Bull. Geodesique*, **108**:187–209, 1973.

McCarthy, D.D., "Polar Motion and Earth Rotation," *Rev. Geophys. Space Phys.* **17**:1397–1403, 1979.

Moritz, H., *Advanced Physical Geodesy*, Abacus Press, Kent, UK; Heyden, Philadelphia, 1980.

Mueller, I.I., *Spherical and Practical Astronomy*, Ungar, New York, 1969.

Mueller, I.I., "Tracking Station Positioning from Artificial Satellite Observations," *Geophys. Surveys*, **2**:243–276, 1975.

Mueller, I.I. and K.H. Ramsayer, *Introduction to Surveying*, Ungar, New York, 1979.

NASA, *NASA Directory of Observation Station Locations*, vol.1, 3rd edition, Goddard Space Flight Center, Greenbelt, Md, Nov. 1973.

NASA, "National Geodetic Satellite Program," Special Publication 365, parts I and II, Washington, D.C., 1977.

NASA, "Radio Interferometry Techniques for Geodesy," Proceedings of a conference held at Massachusetts Institute of Technology, Mass., and at the Northeast Radio Observatory, Westford, Mass., June 19–21, 1979, NASA Conference Publication 2115, 1980.

NASA, "Satellite Altimetric Measurements of the Ocean," Report of the TOPEX Science Working Group, Jet Propulsion Laboratory, Pasadena, 1981.

NOAA, "Classifications, Standards of Accuracy, and General Specifications of Geodetic Control Surveys," Federal Geodetic Control Committee, NOS/NOAA, U.S. Department of Commerce, Rockville, Md, 1980a.

NOAA, "Specifications to Support Classifications, Standards of Accuracy, and General Specifications of Geodetic Control Surveys," Federal Geodetic Control Committee, NOS/NOAA, U.S. Department of Commerce, Rockville, Md, 1980b.

National Research Council, *Geodesy: Trends and Prospects*, National Academy of Sciences, Washington, D.C., 1978.

National Research Council, *Applications of a Dedicated Gravitational Satellite Mission*, National Academy of Sciences, Washington, D.C., 1979.

Rapp, R.J., "Potential Coefficient and Anomaly Degree Variance Modelling Revisited," Report 293, Department of Geodetic Science, Ohio State University, 1979.

Rapp, R.H. and D.P. Hajela, "Accuracy Estimates of 1° × 1° Mean Anomaly Determinations from a High-Low SST Mission," Report 295, Department of Geodetic Science, Ohio State University, 1979.

Rochester, M.G., "The Earth's Rotation," *Eos*, Trans. AGU, **54**:763–780, Aug. 1973.

Rogers, A.E.E., "Very-Long-Baseline-Interferometry with Large Effective Bandwidth for Phase Delay Measurements," *Radio Sci.* **5**:1239–1243, 1970.

Rogers, A.E.E., C.A. Knight, H.F. Hinteregger, A.R. Whitney, C.C. Counselman III, I.I. Shapiro, S.A. Guorevitch, and T.A. Clark, "Geodesy by Radio Interferometry: Determination of a 1.24-km Base Line Vector With ~5-mm Repeatability," *J. Geophys. Res.*, **83**:325–334, 1978.

Rooney, T., G. Hadgigeorge, and G. Blaha, "Analysis of Residuals from Short-Arc Adjustment of Seasat Altimetry," paper presented at the Fall Meeting of the American Geophysical Union, San Francisco, Dec. 8–12, 1980.

Ryan, W.J., T.A. Clark, R. Coates, B.E. Corey, W.D. Cotton, C.C. Counselman, H.F. Hinteregger, C.A. Knight, C. Ma, D.S. Robertson, A.E.E. Rogers, I.I. Shapiro, A.R. Whitney, and J.J. Wittels, "Precision Surveying Using Radio Interferometry," *J. Surveying and Mapping Division* (Proc. Am. Soc. Civil Engineers), **104**, (SU1): 25–34, Nov. 1978.

Sakuma, A., "Absolute Measurements of Gravity," Report of Special Study Group 3.18, *Travaux de l'Association Internationale de Geodesie*, Paris, **25**:141–144, 1976.

Schwarz, C.R. "Deflection Computations for Network Adjustment in the United States," in *Proceedings Second International Symposium on Problems Related to the Redefinition of North American Geodetic Networks*, Arlington, Va, April 24–28, 1978.

Schwarz, K.-P., "Accuracy of Vertical Deflection Determination by Present-Day Inertial Instrumentation," in *Proceedings Ninth Geodesy/Solid Earth and Ocean Physics (GEOP 9) Research Conference*, Columbus, Ohio, Oct. 2–5, 1978, Report 280, Department of Geodetic Science, Ohio State University, 1978. (avail. from NTIS as PB 29904/4)

Tapley, B.D., "Polar Motion and Earth Rotation from Lageos Laser Ranging," in *High-Precision Earth Rotation and Earth-Moon Dynamics*, edited by O. Calame, D. Reidel, Dordrecht, Holland, 1982.

Torge, W. *Geodesy*, Walter de Gruyter, Berlin, 1980.

Tscherning, C.C. and R. Forsberg, "Prediction of Deflections of the Vertical," in *Proceedings Second International Symposium on Problems Related to the Redefinition of North American Geodetic Networks*, Arlington, Va, April 24–28, 1978. (avail. from NTIS as PB 29904/4)

Vanicek, P. and E.J. Krakiwsky, *Geodesy: The Concepts*, Elsevier, Amsterdam, 1982.

Wells, E.M., "A Priori and Real Time Use of a Gravity Gradiometer to Improve Inertial Navigation System Accuracy," Report 530, Department of Aeronautics and Astronautics, Stanford University, Calif., 1981.

Western Space and Missile Center, "Western Test Range Geodetic Coordinates Manual," Vandenberg Air Force Base, Calif., Feb. 1981.

Whalen, C.T., "Status of the National Geodetic Vertical Control Network in the United States," in *Proceedings Second International Symposium on Problems Related to the Redefinition of North American Vertical Geodetic Networks*, Ottawa, Canada, May 26–30, 1980, Canadian Institute of Surveying, Ottawa, 1980.

Winkler, G.M.R. and T.C. VanFlandern, "Ephemeris Time, Relativity, and the Problem of Uniform Time in Astronomy," *Astron. J.*, **82**:84–92, Jan. 1977.

Chapter 25

INFRARED ASTRONOMY

S. D. Price and T. L. Murdock

A description of the infrared celestial background is presented in this chapter with special emphasis on the 2 to 30 μm spectral region. This description is based upon comprehensive surveys of the apparently brightest sources of discrete and diffuse emission in the sky plus selective measurements on fainter objects.

25.1 OBSERVATIONAL PARAMETERS

Most astronomical observations are of the celestial position and brightness of a source at a given resolution and spectral region. These quantities may be measured as a function of time to detect variability and/or motion. Information on the spatial extent of the source depends upon the resolution of the measurement, while knowledge of the spectral energy distribution is a function of spectral resolution, and the range of wavelengths covered. Position and brightness can be determined much more accurately than distance for most objects.

25.1.1 Coordinate Systems

The position of a celestial object is measured in terms of the direction to the source projected onto a sphere of arbitrary radius—the celestial sphere. Angular coordinates are used, measured either along a great circle formed by the intersection of the sphere and a plane that includes the center, or a small circle that is formed by a plane which does not include the center.

A projection of the celestial sphere is drawn in Figure 25-1 for an observer at geographic latitude ϕ. The local vertical, defined with a plumb line, intersects the celestial sphere at the zenith z and the nadir. The horizon is the great circle perpendicular to the vertical. The celestial north and south poles are projections of the respective poles of the earth. The meridian is the visible, or upper, half of the great circle which goes through the zenith and poles and intersects the horizon at the north and south points. The celestial equator is a great circle which crosses the horizon at the east and west points.

During the course of a year, the sun appears to move along a great circle in the sky. This circle, called the ecliptic, is a reflex of the earth's orbit projected onto the celestial sphere as shown in Figure 25-2. The ecliptic and celestial equator intersect at the equinoxes. The vernal equinox, in Figure 25-2, is the point where the sun crosses the celestial equator going south to north. The pole of the earth's rotation precesses about the ecliptic pole due to tidal forces of the sun and moon with a 26 000 yr period. This results in a westward progression of ♈ of about 50 arc sec yr along the ecliptic.

Horizon (Alt-Azimuth) System: This system is unique to the observer. The azimuth, A in Figure 25-1, is the angular distance from the north point, eastward along the horizon to the vertical circle that passes through the point being considered. The altitude h is the angle from the horizon along the vertical circle. The complement of the altitude is called the zenith angle.

Celestial Coordinate System: The celestial coordinates of right ascension (α) and declination (δ) are used for stellar positions and catalogs. Since precession slowly changes the zero point of this coordinate system, the positions must be referred to a specific epoch.

The right ascension is measured eastward from the vernal equinox to the hour circle passing through the point in question. An hour circle is a great circle through the celestial poles. The declination is the angle from the celestial equator along an hour circle. Right ascension, customarily, is expressed in units of time where $1^h = 15°$ or $4^m = 1°$. Declinations are positive north of the celestial equator and negative south.

Ecliptic Coordinate System: The ecliptic plane is the plane of symmetry for objects in the solar system. The ecliptic longitude λ is measured from the vernal equinox eastward along the ecliptic. The ecliptic latitude β is the angular distance from the ecliptic along a vertical great circle. Elongation angle ϵ is the angle along a great circle between the sun and the point in question.

Galactic Coordinate System: The galactic plane is the geometric reference for studies of galactic structure. The galactic longitude l^{II} is the angle along the galactic plane from the galactic center. The galactic latitude b^{II} is measured along a great circle perpendicular to the galactic plane. By agreement of the International Astronomical Union, the celestial coordinates of the north galactic pole are $\alpha(1950) = 12^h49^m0$ and $\delta(1950) = +27°24'0$, the ga-

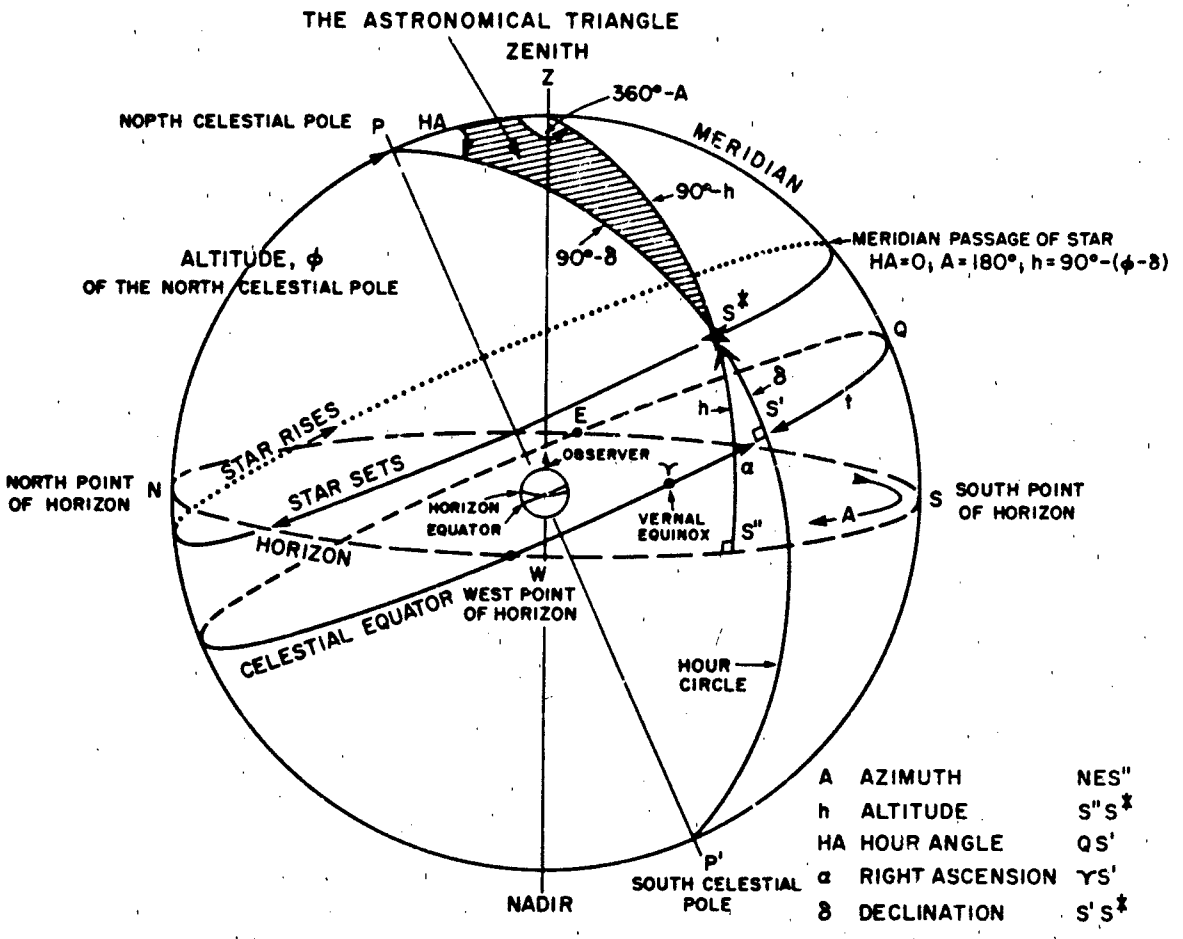


Figure 25-1. Horizon and Equatorial Coordinates. The observer is in the northern hemisphere, which places the north celestial pole at an altitude ϕ = the latitude of the observer. The meridian is the arc NPZS. The arc ZS' defines the vertical circle through star S' intersecting the horizon at S". The azimuth is measured along the horizon eastward from N to S". The altitude is the arc S'S" along the vertical circle through S'. The hour circle PS'P' intersects the celestial equator at S'. The hour angle is measured from the intersection of the meridian and celestial equator, Q, westward along the equator to S'. Right ascension is measured eastward along the celestial equator from ϒ to S'. Declination is measured from the equator along the hour circle. The shaded triangle PZS' has the basic spherical trigonomic relations used to convert from one coordinate system to another.

lactic center at $\alpha(1950) = 17^h42^m.4$ and $\delta(1950) = -28°.55$ with the galactic plane inclined 62.6 to the celestial equator.

25.1.2 Time

The time scales currently in use differ in the definition of the fundamental unit and how passage of time is measured. Ephemeris Time (ET) is a uniform measure used in gravitational theories of the solar system to calculate precise, fundamental ephemerides of the sun, moon, and planets. The fundamental unit is the tropical year defined by the mean motion of the sun in ecliptic longitude at January 0^d12^h ET 1900. Universal Time (UT) is defined by the successive transits of celestial objects and, therefore, is related to the rotation of the earth. The fundamental unit is the mean solar day, the interval between two successive transits of the mean sun. The mean sun is defined to move at a uniform rate along the celestial equator. Sidereal time is defined by the interval between two successive transits of the mean vernal equinox. The apparent motion of the sun with respect to the stars results in the mean sidereal day being $3^m55^s.9095$ seconds shorter than the mean solar day. The International Atomic Time is the most precisely determined time scale in use in astronomy. It is a laboratory reference defined by the transition between two hyperfine levels in the ground state of cesium 133 at a frequency of 9 192 631 770 Hz.

The differences between Ephemeris and Universal Time are tabulated in the "Astronomical Almanac" (U.S. Gov-

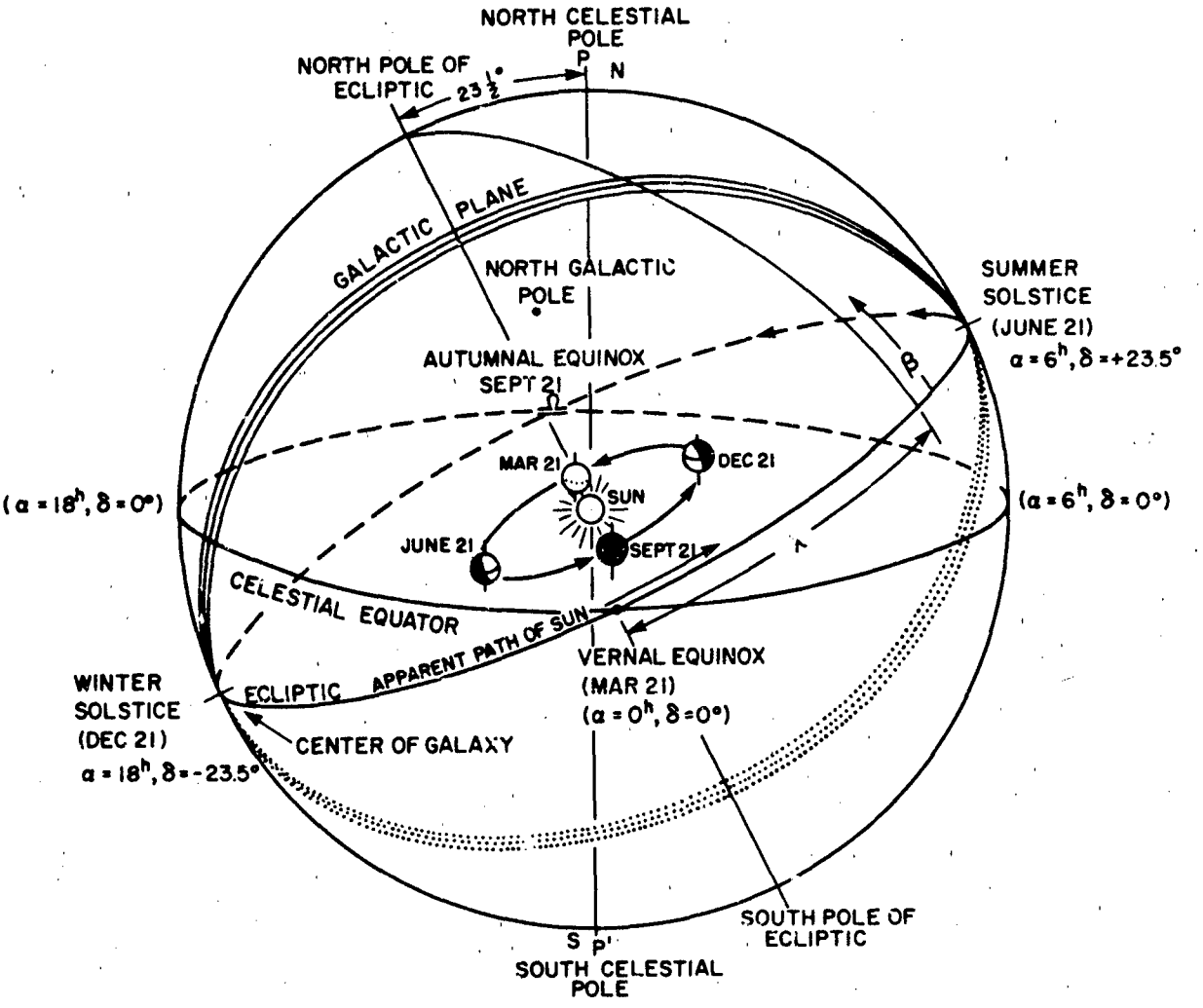


Figure 25-2. Galactic and Ecliptic Coordinate. The orbit of the earth projected onto the celestial sphere defines the ecliptic plane. The ecliptic pole is inclined 23°44 from the celestial pole. By definition, the intersection of the ecliptic and equatorial planes is the zero point for right ascension and ecliptic longitude, λ. The north galactic pole is some 63° from the celestial pole.

ernment Printing Office) as is the difference between International Atomic Time and Coordinated Universal Time (UTC). UTC is the scale available from broadcast signals.

25.1.3 Intensities—Magnitude

Absolute intensity measurements by direct comparison with laboratory standards exist for only a few astronomical objects. Instead, intensity measurements are referred to one or more secondary standard stars. The secondary standard stars are a network of non-variable stars roughly uniform in distribution across the sky which have accurate photometry with respect to each other. These secondary standards are referenced to a small number of primary standards, stars which have absolute flux determinations from laboratory calibrated sources.

Johnson [1965] used the sun as a standard with the obvious problem of transferring the very bright extended source to fainter intensity. Vega, α Lyr, is currently the most common star chosen for a primary standard. Absolute calibration of α Lyr through direct comparison with laboratory standard sources covers the wavelength from 0.35 to 1.08 μm to an accuracy of better than 3% [Hayes and Latham, 1975; Tug, White, and Lockwood, 1977]. Low [1973] has obtained the absolute 11 μm flux on 3 stars to 10% accuracy by comparison with laboratory standards.

Stellar intensities are often expressed in terms of mag-

nitude, a logarithmic scale referenced to an adopted standard. The apparent brightness ratio, I_2/I_1, of two stars of apparent magnitudes m_2 and m_1 is

$$I_2/I_1 = 10^{-0.4(m_2 - m_1)}$$

and

$$m_2 - m_1 = -2.5 \log (I_2/I_1). \qquad (25.1)$$

The absolute magnitude M is defined as the apparent magnitude a star would have at a distance of 10 parsecs (3.085678×10^{19} cm). The absolute and apparent magnitudes are related by

$$M = m + 5 - 5 \log r - A \qquad (25.2)$$

with r = the stellar distance in parsecs and A the interstellar extinction, in magnitudes, to the star.

The measured magnitude also depends on the spectral energy distribution of the star through the spectral sensitivity of the detector. The spectral selectivity of the measurement is usually expressed by the system unique parameters of effective wavelength λ_e and effective bandwidths $\Delta\lambda_e$ defined by

$$\lambda_e = \frac{\int_0^\infty \lambda S_\lambda \, d\lambda}{\int_0^\infty S_\lambda \, d\lambda},$$

and

$$\Delta\lambda_e = \int_0^\infty S_\lambda \, d\lambda \qquad (25.3)$$

with S_λ the system spectral response.

The effective wavelengths, effective bandwidths, and flux for a zero magnitude star within the respective band passes are given in Table 25-1 for a widely used system of photoelectric photometry. The photometric bands in the infrared coincide with atmospheric transmission windows.

25.1.4 Distance

Direct radar returns and radio transmission from interplanetary space probes lead to very accurate distance determination within the solar system. Distances to nearby stars can be calculated from trigonometric parallax, the stellar displacement as viewed from different locations of the earth's orbit around the sun. More distant stars require indirect methods for calculating distance, which depend on the statistical properties of a group of stars and the apparent brightnesses of "standard candles".

Table 25-1. Definition of zero magnitude flux as function of wavelength. [Johnson, 1965; Low, 1973]

Band	λ_e (μm)	$\Delta\lambda_e$ (μm)	Flux ($m(\lambda_e) = 0$) W cm^{-2}μm^{-1}
U	0.365	0.068	4.27×10^{-12}
B	0.44	0.098	6.61×10^{-12}
V	0.55	0.089	3.72×10^{-12}
R	0.70	0.72	1.74×10^{-12}
I	0.88	0.24	8.32×10^{-13}
J	1.25	0.38	3.31×10^{-13}
H	1.65	0.30	1.28×10^{-13}
K	2.22	0.48	4.14×10^{-14}
L	3.6	0.70	6.38×10^{-15}
M	5.0	1.73	1.82×10^{-15}
N	10.6	4.33	8.7×10^{-17}
Q	21	5.8	6.5×10^{-18}

25.1.4.1 Parallax. The trigonometric parallax is defined as the angle, p or π, subtended by the mean radius of the orbit of the earth at the distance of the star. As of 1980 some 11 600 stars have parallaxes measured with varying degrees of accuracy; only 10 stars have parallaxes determined to better than 2% [Upgren, 1980].

A parsec pc is the distance at which the parallax is one arc second. It is numerically the inverse of the parallax and, as such, is an independent measure which scales by the value of an astronomical unit (equals mean distance between the earth and sun). By definition 1 pc = 206264.802 AU From Allen [1973] 1 AU = 1.495979×10^{13} cm; thus 1 pc = 3.085679×10^{18} cm.

Beyond 50 to 100 pc where trigonometric parallaxes are unreliable, distances can be inferred from the observed motion of stars. The measured quantities are the proper motion, the observed angular motion pependicular to the line of sight, and the radial velocity, which is determined from the Doppler shift of the spectral lines. The tangential velocity can be determined from the proper motion and radial velocity if the direction of the total velocity vector is known. Various methods based on statistical arguments for an aggregate of stars are used to estimate the direction of the total velocity vector or the value of v_t directly. These procedures are valid out to 100 to 200 pcs but may be extended to a kiloparsec using a reasonable theory of galactic rotation.

25.1.4.2 Standard Candles. The distance modulus (M-m) can be derived from the observed apparent magnitude and knowledge of the absolute magnitude or intrinsic brightness of a source. An estimate of the interstellar extinction will then produce a distance through Equation (25.2).

Spectroscopic Parallax: The intensity ratios of certain luminosity sensitive spectral lines correlate well with the absolute magnitude of those stars which have accurately determined parallax. For those stars bright enough for a

high dispersion spectrum to be obtained, the spectroscopic parallax is derived from the spectroscopic estimate of the absolute magnitude and the observed apparent magnitude.

Variable Stars: Stars that vary in brightness in a distinctive manner can be seen to large distances and are useful distance indicators when their absolute magnitude can be related to the variability.

The absolute magnitudes of novae at maxima are correlated with the time rates of decay in the light curves. Furthermore, the absolute brightness of the light curves are calibrated from the novae close enough to have expansion parallaxes, that is, the angular expansion rates and radial velocities of the ejected material are measured and a distance derived assuming spherical symmetry.

As a class, the cepheid variables exhibit an empirical relationship between absolute magnitude and period of pulsation. The various cepheid subclasses—classical and dwarf cepheids, RR Lyrae, δ Sct, W Vir, β CMa and β Cep stars—differ in the shape of the light curve, amplitude, and temperature variation during pulsation. The absolute magnitudes for these stars are derived from statistical parallaxes and stellar evolution models applied to the clusters which contain them.

25.1.4.3 Dynamic Distances. A theory of galactic rotation may be used with radial velocity measurements to obtain distances greater than a kiloparsec. This technique has been extensively exploited in radio astronomy particularly for HII regions and molecular clouds.

In the simplest case, galactic constituents are assumed to move in circular orbits with a decreasing angular velocity with increasing distance from the center. At a given longitude in the galactic plane, the largest observed radial velocity is at the point where the line of sight is closest to the galactic center. A rotation curve, in terms of the galactocentric distance of the sun, $\bar{\omega}_0$ is constructed from the maximum observed radial velocity as a function of longitude. This curve can be used to derive distances relative to $\bar{\omega}_0$ from the observed radial velocities. If the measured radial velocity is smaller than the maximum, the positional ambiguity is resolved by assuming that the larger, brighter objects are closer.

25.1.4.4 Extragalactic Distances. Primary indicators of extragalactic distances such as novae and cepheid variables are calibrated by fundamental means within our own galaxy. These so called primary standards are good out to about a Megaparsec (Mpc). Overlapping, and calibrated by the primary indicators, are less reliable indicators such as the brightest stars in a galaxy and the size and brightness of HII regions. These are observable out to about 10 Mpc. At distances greater than 10 Mpc isophotal galactic diameters and galactic luminosity class are used for distance determination.

At very large distances the radial velocity is, on the average, directly proportional to the distance. The constant of proportionality, the Hubble constant H reflect the expansion rate of the universe. If the Hubble constant is known then a measure of the radial velocity of a galaxy results in a distance. Various recent determinations of the Hubble constant range from 50 to 100 km s^{-1} Mpc^{-1}.

25.2 THE SOLAR SYSTEM

Our knowledge and understanding of objects in the solar system have advanced spectacularly over the past decade with the fruition of the space exploration programs and the development of new measurement techniques. This section presents a description of the solar system as seen at wavelengths between visual and 30 μm from a point near the earth but above the atmosphere.

Objects in the solar system scatter and absorb sunlight. In the visible, this scattered sunlight has a spectral distribution similar to that of the sun, a 6000 K blackbody with absorption and emission lines, but modified by the spectral reflectivity of the body's surface. The object is warmed by absorbed sunlight and consequently emits thermal radiation. This thermal re-emission is characterized by the Planck equation appropriate to the temperature and modified by the spectral emissivity of its surface layers. Thermal emission arises from the topmost layers, which are opaque or semi-opaque at the wavelength in question. In the simple case of a body with no atmosphere, the thermal emission spectrum is determined by the temperature, the physical structure, and the chemical composition of the emitting layer. The asteroids, most planetary satellites, Mercury and the moon fall into this category. If the atmosphere of the body is optically thin, the infrared thermal emission is not appreciably altered and measurements of the spectral intensity continues to provide information about the surface layers. At the other extreme, if the atmosphere is completely opaque, infrared emission reveals the character and state of the atmospheric layers which are optically thick at the measured wavelength. Venus, Jupiter, Saturn, Uranus and Neptune are clear examples of objects in this category. The atmospheres of objects intermediate to these extremes are optically thick only at some wavelengths, and the spectral emission is a superposition of reflected sunlight and thermal emission from the surface and atmospheric emission and absorption. Mars and comets are examples of this class. Finally, an optically thin dust cloud is present in the solar system. The zodiacal light is scattered sunlight and thermal emission from the dust grains in this cloud.

25.2.1 Objects with Negligible Atmospheres

To first approximation, a typical object in the solar system does not rotate and is located at a constant distance, R, from the sun. Assume that there are no internal sources

of heat, that the surface cannot conduct heat away, and that the surface is in equilibrium with sunlight. At the subsolar point of such an object, the equilibrium surface temperature is given by the heat balance between absorbed sunlight and radiated thermal emission

$$T_0 = \frac{(1 - A)}{\varepsilon_{IR}\sigma}\left(\frac{E}{R^2}\right)^{1/4} \qquad (25.4)$$

where

A = Bond Albedo (ratio of total incident to reflected light for a spherical body)

E = Solar constant at 1.0 AU = .1368 W/m^2

ε_{IR} = Infrared average emissivitiy

σ = Stefan-Boltzmann Constant = 5.66956 $\times$ 10^{-12} Wcm^{-2}deg^{-4}.

The temperature distribution on a non-rotating spherical surface will be isothermal concentric bands around the subsolar point. If the subsolar point is defined to be the thermal pole, the temperature at thermal colatitude ϕ is

$$T(\phi) = T_0 \cos^{1/4}\phi. \qquad (25.5)$$

For an object with a rotation period P an isotropic surface material with uniform density ρ, specific heat c and thermal conductivity K, the temperature distribution inside the object is determined by the heat conduction equation

$$\rho c \frac{\partial T}{\partial t} = K \nabla^2 T. \qquad (25.6)$$

The boundary condition on the surface in sunlight is

$$\varepsilon_{IR}\sigma T_s^4 - \frac{(1 - A)}{R^2} E \cos\phi = K \nabla T_s, \qquad (25.7)$$

while on the dark side

$$\varepsilon_{IR}\sigma T_s^4 = K \nabla T_s. \qquad (25.8)$$

If the rotational axis lies in a plane perpendicular to the incident sunlight, the thermal pole traverses the rotational equator, and thermal emission from the dark side arises from thin isothermal longitudinal stripes whose temperature falls to a minimum just before local dawn. The minimum temperature is determined by the rotational period and by the thermal properties of the uppermost several meters of the surface material. Typical minimum temperatures for the moon and Mercury are about 100 K, indicating a loosely packed, dusty, dielectric surface material.

For the more realistic case of an atmosphereless object with a solid bedrock overlain by varying depths of dust and broken rock, K, ρ, and c are not constant. The conduction equation must also be modified to include radiation between the dust grains, which is a major heat transport mechanism in this case. Furthermore, rocks which protrude through the dusty surface have been shown to introduce significant discrepancies between the predictions based on the simplified models discussed above and infrared measurements of the lunar terrain during eclipse [Allen and Ney, 1969] and the planet Mercury near inferior conjunction [Murdock, 1974].

If the field of view of a measurement is larger than the apparent solid angle subtended by the body, the resulting spectrum is the superposition of scattered sunlight and thermal emission from the visible surface, each region weighted by the respective apparent projected area. The phase angle i is the angle between the sun and the observer as seen from the center of the object. The fraction of the illuminated side seen by the observer is

$$k = 0.5\,(1 + \cos i).$$

Objects with orbital radii less than that of the observer will go through all phases $0 \leq k \leq 1$. Objects with orbital radii larger than the observer will go through partial phases $k_{min} \leq k \leq 1$. The phase angle may be calculated from

$$\cos i = \frac{R_{eo}^2 + R^2 - R_e^2}{2\,R_{eo}\,R}, \qquad (25.9)$$

where

R_{eo} = true geocentric distance in AU

R_e = true Earth-Sun distance in AU

R = heliocentric distance of the object in AU.

Values of phase angle and the orbital distances for various solar system objects are tabulated in the "Astronomical Almanac."

Color temperature T_c is defined as the temperature of the Planck function $B(\lambda,T)$ that best fits the measured spectral shape. Brightness temperature T_b at any wavelength λ is defined as that temperature which satisfies

$$I(\lambda) = B(\lambda, T_b)\,\Omega_e,$$

where $I(\lambda)$ is the measured spectral intensity and Ω_e is the emitting solid angle. For simple objects in solar orbit, the emitting solid angle required to bring the color and brightness temperatures in the midinfrared into agreement is, to good approximation

$$\Omega_e = k\Omega,$$

where Ω is the total solid angle subtended by the body and k is the fraction of the sunlit surface visible from earth.

Thus, the sunlit portion of the disk dominates the total thermal emission at these wavelengths.

The effective brightness temperature at wavelength λ_e is found from

$$I(\lambda_e) = B(\lambda_e, T_e)\,\Omega_e.$$

The effective brightness temperature can be normalized to the mean sun-object distance R_s, by

$$T_s(\lambda) = T_e(\lambda)\left(\frac{R}{R_s}\right)^{1/2}.$$

The apparent brightness of the object varies with distance from the earth through the relation

$$I(\lambda) = I(R_{eo},\lambda)\left(\frac{R_{eo}}{d}\right)^2,$$

where R_{eo} = true geocentric distance at the time of measurement and d = desired geocentric distance.

25.2.1.1 Mercury. Mercury's orbit is highly eccentric and its infrared spectrum varies not only with phase angle but also with solar distance. The subsolar point temperature variation from aphelion to perihelion is from 575 to 710 K. At phase angles less than 160° the effective emitting area required to bring the brightness temperature into agreement with the color temperature is the solid angle of the visible sunlit portion of the apparent disk. At phase angles greater than 160°, the contributions from hot rocks in the partially lit crescent and the dark side become significant. The warm color temperature of the crescent is due to crater walls and raised rocks on the surface, which remain in sunlight longer than predicted by the $\cos^{1/4}\theta$ variation for a smooth spherical surface. These objects cool at a slower rate than the rest of the surface material. Approximately 3% of the crescent solid angle is required to make the crescent color and brightness temperatures equal. The minimum effective temperature of 106 K observed for the dark side [Murdock and Ney, 1970] is indicative of a dusty, loosely packed surface. Phase curves of the visual, λ = 3.6, 10.8, and 20 μm emission for Mercury at mean solar distance (R = 0.387AU) corrected to a constant geocentric distance of 1.0 AU [Murdock, 1974] is given in Table 25-2.

Table 25-2. Mercury magnitudes at 0.387 AU from the sun, 1.0 AU from the earth.

θ	0.54μm	3.6μm	10.8μm	20μm
0	−2.5	−9.0	−13.0	−12.9
20	−1.8	−8.8	−12.7	−12.7
40	−1.2	−8.3	−12.3	−12.5
60	−0.7	−8.0	−11.9	−12.3
80	−0.1	−7.4	−11.2	−12.1
100	+0.6	−6.2	−10.6	−11.9
120	+1.6	−4.9	− 9.3	−11.4
140	+3.0	−3.1	− 8.0	−11.0
160	+4.9	+0.3	− 5.8	−10.4

25.2.1.2 The Moon. The lunar surface is primarily composed of a loosely packed dielectric dust quite similar in thermal and visual properties to the surface of Mercury. Pictures of the moon and Mercury are remarkably similar as are their visual and infrared phase curves [Ryadov, Furashov, and Sharonov, 1964; Pikkarainen 1969; Murdock 1974]. The lunar albedo is 0.07 [Allen, 1973] and the subsolar point temperature varies between 397 K and 391 K from perihelion to aphelion. Allen [1970] found that approximately 3% of the average lunar surface and as much as 10% of crater areas must be exposed rock in order to account for the lunar thermal anomolies found by Shorthill and Saari [1965]. Thus the infrared appearance of the Moon would be nearly identical to that of the planet Mercury at the same solar distance and the shapes of the phase curves for the two objects can be assumed to be similar. The lunar full disk λ_e = 11.4 μm brightness at i = 0° is M = −22.5.

25.2.1.3 Asteroids and Planetary Moons. As of January 1984, 2992 asteroids are listed with orbital elements in the "Minor Planets Circulars" published on behalf of Commission 20 of the International Union. In general, the asteroids have albedos between 0.065 and 0.23 [Gradie and Tedesco, 1982]. The infrared spectral energy distributions between 5 and 30 μm for asteroids are quite similar to those of gray bodies [Gillett and Merrill, 1975; Hansen, 1977]. A decrease in emissivity from 20 to 85 μm has been observed for some asteroids [LeVan and Price, 1984]. The asteroid distribution is concentrated to the ecliptic plane with an exponential folding height of about 7°.

The planetary moons are an inhomogeneous set of objects. The Mariner, Pioneer, Voyager, and Viking series of exploration experiments have shown that each object has a unique spectral signature. Typical Bond albedos range between 0.2 and 0.7 for ice covered or dust covered satellite surfaces [Hansen, 1977, Smith et al., 1981].

25.2.2 Objects with Optically Thick Atmospheres

Emission from objects with optically thick atmospheres comes from those depths which are opaque at the wavelength in question. The visual and infrared spectra of these objects are complex superpositions of scattered sunlight and molecular vibration and rotation band emission over an underlying continuum which can be approximated by a gray body spectrum. In general, thermal radiation at different infrared wavelengths comes from different atmospheric depths and results in different effective brightness temperatures.

Table 25-3. Planets with optically thick atmospheres, compiled from data given in Stier et al. [1977], Wright [1976], Reed et al. [1978], Ward et al. [1977], Erickson et al. [1978], Tokunga et al. [1977], Gillett and Forest [1974], Hanel et al. [1981], Allen and Murdock [1970] and Courtin et al. [1978].

	A	T_{eq}	T_{eff}	T_{10}	T_{20}	T_{30-250}
Venus	0.77 ± 0.07	228 ± 5	240 ± 5		200–250	245
Jupiter	0.35 ± 0.07	111	125 ± 3	124	125	130
Saturn	0.50 ± 0.03	76 ± 1	94.4 ± 3	110	92	85.6 ± 3
Titan	0.21	86	110	126	110	
Uranus	0.31 ± 0.35	58.5	58.5	133 ± 4 at 5μm	54.7 ± 1.8	58.5
Neptune	0.33	46	59.7	115 at 7.5μm	60	59.7 ± 4

A = Bond Albedo
T_{eq} = Equilibrium Temperature
T_{eff} = Effective Temperature
T_{10} = 10 μm Brightness Temperature
T_{20} = 20 μm Brightness Temperature
T_{30-250} = Brightness Temperature in the 30- 250 μm Region

The visual appearance of these objects varies from the smooth featureless cloud tops of Venus to the brightly banded turbulent clouds of Jupiter. Table 25-3 summarized the visual and infrared properties averaged over the disk for this class of object.

25.2.3 Objects with Semi-Transparent Atmospheres

The third class of objects in the solar system consists of those bodies with semi-transparent atmospheres. The surface emission is modified or blocked at some wavelengths, and the observed spectral intensity originates in both surface layers and the atmosphere.

25.2.3.1 Mars. The Martian atmosphere is optically thin. At wavelengths between 30 and 300 μm the mean transmission is between 85% and 95%. The atmospheric effects are also small at the shorter infrared wavelengths. The predominant result is that the product $(K \rho c)^{1/2}$ of the Martian surface is three times larger than that for Mercury and the Moon. The Martian albedo averaged over the disk is 0.10 and the λ_c = 10 μm effective brightness temperature varies between 230 K and 270 K with phase angle and solar distance variations [Wright, 1976].

25.2.3.2 Comets. Comets appear to originate in deep space beyond the orbit of Pluto. Chance gravitational encounters with nearby stars or each other propel several of these objects per year into the outer solar system. Gravitational interactions with the giant planets send the comets into the inner solar system on parabolic or hyperbolic orbits. Encounters with the planets can also modify these open orbits into closed long period ellipses. Since the comets originally condensed from the solar nebulae at such large distances

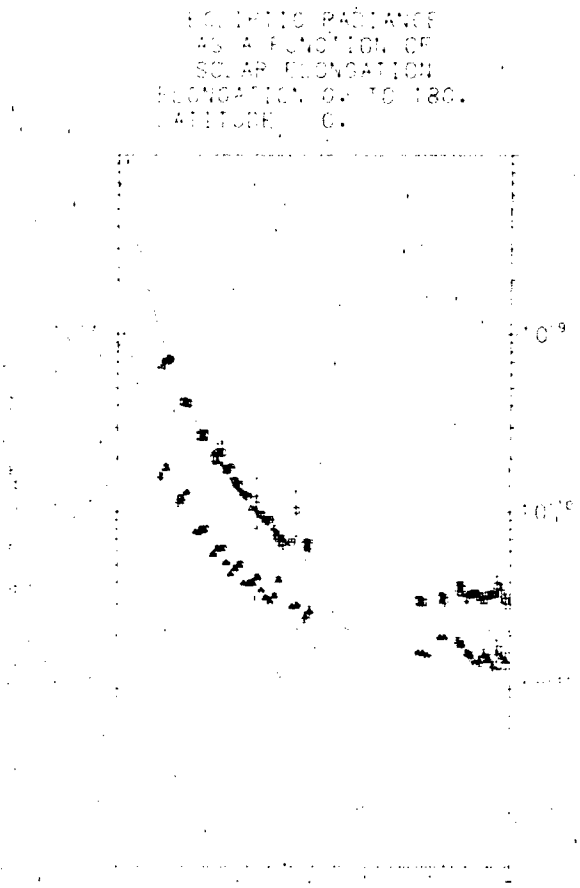


Figure 25-3. Variation of the zodiacal emission at 10.9 μm (squares) and 20.9 μm (triangles) along the ecliptic plane. Solid and dashed lines are the model predictions at the respective wavelengths from Equation (25.10).

from the sun, they appear to have lost very few of their volatile compounds and seem to be light, porous clumps of frozen gases and dust grains. New comets which have never approached the sun should have retained most of their original accretion structure. As the comet moves closer to the sun it begins to warm, sheds its outer layers and probably changes its internal structure by melting. The dust and gas that has been removed from the outer layers escapes from the core and follows its own solar orbit trajectory as the comet tail and anti-tail. Typical radii of the comet nucleus range between 5 and 30 km with albedos about 0.18 [Ney, 1974].

25.2.4 The Zodiacal Light

A low density cloud of dust surrounds the sun. This cloud is concentrated toward the eliptic plane of the solar system and extends inward from the orbit of Jupiter to a minimum solar distance determined by the sublimation temperature of the dust grains. Sunlight scattered from these grains appears as a diffuse background glow in the night sky. This glow is brightest toward the sun and is centered near the ecliptic plane. The zodiacal light has been observed visually since ancient times and systematically studied since the late 1600s. Satellite observations and ground based studies of the visual zodiacal light have been used in conjunction with data from the Pioneer and Helios space probes to define the extent and apparent angular distribution of the dust cloud [Leinert, 1975]. Recently, infrared measurements taken during rocket flights and satellite experiments have returned the first definitive infrared observations of the clouds thermal emission spectrum [Soifer et al., 1971; Price et al., 1980; Murdock and Price, 1984; Hauser et al., 1984].

The observed variation in the 10.9 and 20.9 μm radiance along the ecliptic plane is shown in Figure 25-3. The plane to pole changes in these radiances are shown in Figure 25-4 for a sun centered longitude of 45°. Infrared spectrophotometry of the zodiacal emission at a sun centered longitude of 60° and at the north ecliptic pole in the summer is shown in Figure 25-5.

The dynamics of small particles in the solar system are [Leinhert, 1975] as follows. The solar gravitational acceleration is opposed by the centripetal acceleration from the orbital velocity and outward radiation pressure of incident sunlight. The radiation pressure exceeds the gravitational

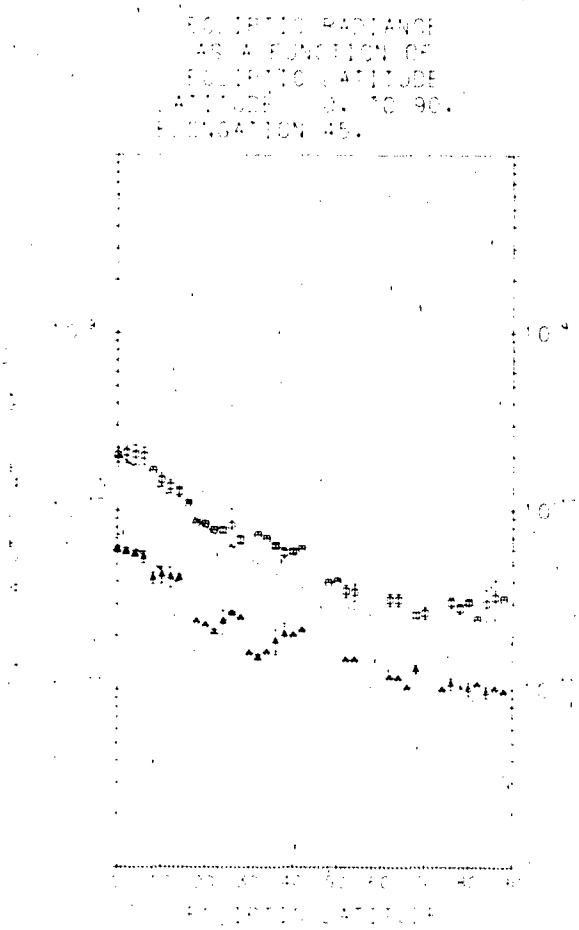


Figure 25-4. Plane to pole variation in the 10.9 and 20.9 zodiacal radiance at an elongation of 45°. Symbols are the same as in Figure 25-3.

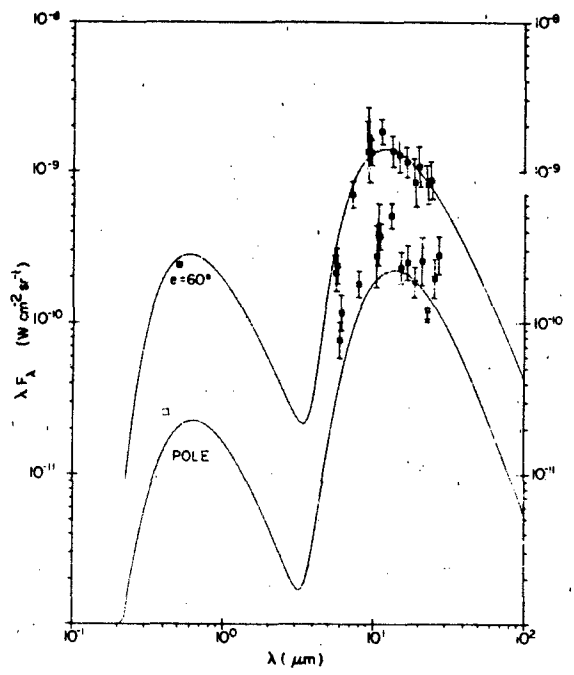


Figure 25-5. Spectrophotometry of the zodiacal light at the ecliptic pole and in the plane at 60° elongation. Visual values are from Allen [1973].

force for particles smaller than 0.2 μm in diameter and these particles are blown out of the solar system. A braking force (Poynting-Robertson effect) due to the small aberration of sunlight decelerates a small particle in the solar system causing it to slowly spiral into the sun. As the dust particle nears the sun it warms up, subliming until it is small enough to be driven out of the solar system by radiation pressure. A typical dust grain has a radius of 1.0 μm and takes about 10^4 years to be ejected from the solar system after injection.

The spectral energy distribution of the zodiacal light is a superposition of scattered sunlight and thermal emission from all the particles along the line of sight. The scattered sunlight at a solar elongation of 90° in the ecliptic plane can be represented by the solar energy distribution diluted by 9×10^{-14} [Allen, 1973]. The infrared energy, $j(\lambda)$, in a solid angle, $d\omega$, from the dust in the solar system is the sum of the thermal emission and scattered sunlight and is given by

$$j(\lambda) = \int_\omega d\omega \int_0^\infty ds\, N(D,Z,s)\, \pi s^2 \left[Q_{em}(\lambda,s) B(\lambda,T_p) + Q_{sc}(\Theta,\lambda,s) F(\lambda) \left(\frac{R_\odot}{D}\right)^2 \right] d\omega \qquad (25.10)$$

dl = increment along the line of sight

Q_{em} = emission efficiency = $\varepsilon(\lambda)$ emissivity

s = particle radius

T_p = particle temperature

N = number density at an in plane distance of D from the sun and z above or below the plane

$F(\lambda)$ = infrared solar energy distribution

D = projection of solar distance, R, of particle into the ecliptic plane

Z = height of particle above the plane

$R^2 = D^2 + Z^2$

$Q(\lambda,s,\Theta)$ = scattering efficiency = $\frac{a(\lambda)}{4\pi} P(\Theta)$ with

$a(\lambda)$ = albedo of the particle.

The integral in Equation (25.10) is heavily weighted for near earth conditions and a mean value of emissivity ($\varepsilon = 0.7$), albedo ($a = 0.3$) and particle area times density ($s^2 N_o = 6.44 \times 10^{-22} cm^{-1}$) reasonably represent the observed data. Separating the spatial density variation into an exponential vertical distribution and a power law in plane distribution Equation (25.10) fits the observed data if

$$j(\lambda) = 1.4 \times 10^{-21} cm^{-1} \int_\omega dl \left[\frac{D}{R^2} \exp\left\{ -2.6 \frac{|Z|}{R} \right\} B(\lambda,T_p) + 0.137 P(\Theta) F(\lambda) \left(\frac{R_\odot}{R}\right)^2 \right] d\omega \qquad (12.11)$$

where $P(\Theta)$ is the scattering coefficient in terms of Legendre polynomials $P(\Theta_o) = P_0(\cos\Theta) + 1.3054\ P_1(\cos\Theta) + 1.3014\ P_2(\cos\Theta)$.

This expression was used to calculate the model fits in Figures 25-3 through 25-5. Table 25-4 lists the observed 10.9 and 20.9 radiances as a function of ecliptic elongation and latitude in 5° bins in each coordinate. Satellite measurements [Hauser et al., 1984] are 40% higher at the ecliptic pole and a factor of two higher in the plane.

Satellite observations [Low et al. 1984] indicate the presence of three roughly parallel bands of excess zodiacal emission located at −10°6, −2°8 and 9° latitude. The 160–200 K color temperature of this excess emission is characteristic of the main asteroid belt. This argues for an asteroidal origin for the zodiacal dust. The infrared spectra of the zodiacal emission is consistent with dust found in comet tails. The origin of the interplanetary dust is still not resolved.

25.3 THE GALAXY

The galaxy is composed of stars and clouds of gas and dust in roughly equal proportion by mass. It is lenticular in shape with an ellipsoidal bulge and a surrounding disk with an almost exponential decrease in brightness and mass with distance from the center. The gas has an exponential scale height of about 4 kpc, the stars 2-3 kpc. There is evidence for a spiral pattern in the disk. A sparse scattering of single stars and clusters of stars (globular clusters) are spherically distributed about the galactic center in the galactic halo. The sun is about 8.5 kpc from the galactic center and travels at a speed of about 225 km/s.

25.3.1 Stars

Stars may be grouped into four general categories according to size, luminosity, and density. Most stars are dwarf or main sequence stars. Within 20 parsecs of the sun 90% of all stars, including the sun, are dwarfs. White dwarfs are stars of low luminosity, high temperature, and very high density. They constitute about 8% of the nearby stars. Giant stars are large objects of low density and high intrinsic luminosity. Although they account for only 0.7% of the stars in the solar neighborhood, they dominate the bright end of the source counts because of their high luminosity. Supergiant stars are extremely bright stars of very low density, up to 100 times more luminous and 10 times larger

Table 25-4a. Zodiacal Radiance at 10.9 μm. Units are 10^{-12} Wcm^{-2}μm^{-1}sr^{-1}.

	Ecliptic Elongation																				
	0°					25°						50°				75°					
	90																				
L	85		35	32	33	34					25	28	26		25			25			
	80	27	29			37	37		27	26								28	28	27	28
A	75	32	33	30	29			34	28						30	29	29	28	28	28	29
	70	29	30	31		31	32	32	43			31	31	31	29	28	28	27	31	35	30
T	65	38	36	33	31	30	32	30	45		32	31	31		33	30	30	32		30	29
	60			36	34	31	38	34	48	50			33	31	32	34	30	31	30	32	33
I	55		38		37	42	34	42	38	52	35		34	31	33	34	32	36	35	35	33
	50			39	34	40	36	50	40	55	36		37	35	35		35	34	34	38	36
T	45			42		43	46	52	49	54	49	48	43	36	37	37	35	46	42	43	40
	40				60	47	50	51	67	50	60	46	41	41	39	49	43	46	43	40	40
U	35				55				65	57	65	50	47	41	53	50	46	49	43	40	39
	30				72	75	66	65	79	65	72	60	49	58	66	56	50	47	43	45	40
D	25					90	69	81	91	79	77	83	65	57	84	72	49	51	48	43	60
	20				150	115	140	106	95	101	87	92	71	86	74	63	56	54	51	54	42
E	15				189		167	144	146	132	114	107	93	88	73	74	64	57	61	44	46
	10					339	244	226	184	160	158	126	122	92	82	73	69	63	49	57	59
	5					507	428	367	250	220	194	165	132	104	89	74	73	68	70	62	65
	0						700	415	260	269	216	190	143	117	106	82	72	68	68	64	

Table 25-4a (continued).

	Ecliptic Elongation																				
	100°					125°						150°									
	90																				
L	85				27				26		22		26		25	25	24				
	80	28	27	28	36	27	28	27	27		29		28	29	29			24	23		
A	75	29	28	29		30	27	27	28	28	27		26	27		29	28	31	24		
	70	30	28	28	29	28	28	28	27	26				28	27	29	30	29	28		
T	65	30	29	31	29	34	29				27					27	29	31			
	60	34	33	30	29	29	31	31	31	30						30	29		34		
I	55	32	33	31	33	33	34	35	32	29	29				32						
	50	40	36	36	34	31	30	30	30	29	28	30	31	35	30						
T	45	38	37	37	37	35	35	31	30	28	32	27	30	28	29	31					
	40	35	36	35	34	31	31	37	34	33	29	29	29	27	29	29	35				
U	35	34	35	33	32	32	32	31	31	33	34	30	29	27	29	29	30	28			
	30	40	47	38	37	36	35		32	32	31	31	35	27	29	29	29	28			
D	25	52	38	36	35	38	38	36	36	36	33	28	36	31	30	29	28	30			
	20	45	42	42	37	34	34	34	35	35	36	30	33	31	32	29	31	33			
E	15	40	34	33	33	34		34	33		35	31	31	34	35	25	32	33			
	10	59					34	32		33	33	35	31	33	34	35	32	32			
	5							33	33	32	31	31	33	33	34	34	36	35			
	0								34	31	36	32	35	34	33	34	38				

Table 25-4b. Zodiacal Radiance at 20.9μm. Units are 10^{-12}Wcm^{-2}μm^{-1}sr^{-1}.

		Ecliptic Elongation																			
		0°				25°				50°				75°							
	90																				
L	85	12	10	11					11	10	10	10	11			11					
	80	11	10			10	9	10	10	10							10	11	11	10	11
A	75	10	10	10	9		10	10					12	11	11	10	9	9	10	9	12
	70	10	10	10	7	10		10	10			11	10		10	9	12	10	11	17	14
T	65	12	12	11	10		10	11	10		11	12	9	10	12	12	9		14	14	14
	60		12	12	12	12	11	11	11			10	9	11		9	14	15	15	17	13
I	55		14		12	12	11	12		11	15	10	15	10	11	18	16		11	11	14
	50		14	14	12	13	12	13	12	14	14	11	16	13	10		15	15	13	15	14
T	45			15		13	14	13	13	14	13	18	15	11	14	16	14	18	16	14	14
	40				25	15	15	16	15	15	17	16	14	16	12	17	14	15	15	13	18
U	35				26		17	17	16	18	17	18	17	13	18	17	17	15	18	18	15
	30				27	25	19	18	20	20	19	21	19	20	20	17	18	17	16	17	16
D	25					28	20	22	25	20	23	20	23	19	23	19	17	20	18	15	17
	20				43	34		30	27	28	23	28	26	25	20	22	21	20	22	21	23
E	15				51	37	47	40	48	34	34	33	31	26	25	26	23	23	20	20	22
	10					84	65		52	44	44	40	34	33	28	27	27	24	23	23	25
	5					113	110	102	69	67	54	49	42	38	34	26	32	30	28	25	24
	0						198	116	76	78	61	54	48	38	43	32	39	29	26	25	

Table 25-4b (*Continued*)

		Ecliptic Elongation																			
		100°					125°					150°									
	90																				
L	85				10						9		12	10	4	9	10	10			
	80	10	10	11	10	10	7	10	10	15	9	11		10				8			
A	75	7	10	10		14	9	7		8	10	10	11	9	10	13	10				
	70	13	13	13	13	12	13	13	12	12			11	10	10	10					
T	65	15	14	11	12	13	10			13	13				11	10	11				
	60	12	13	16	16	16	15	14	14	13						12					
I	55	14	14	13	14	13	13	14	14	14	15	15		16							
	50	14	12	13	12	12	12	12	13	12	12	13	15	15	17						
T	45	13	14	13	12	13	13	11	11	12	13	12	12	14	16	13					
	40	13	14	14	15	12	10	12	12	12	11	13	12	13	13	13	14				
U	35	13	13	12	11	13	13	13	14	11	13	13	12	14	12	14	13				
	30	16	16	17	16	16	16	15	14	14	12	14	13	14	13	13	15				
D	25	19	21	21	18	14	19	15	15	13	14	14	13	15	12	15	12				
	20	18	16	15	13	19	20	20	15	15	14	14	14	16	14	14	13				
E	15	19	14	15	17	16	16	20	20	19	19	16	16	14	15	15	13				
	10	24					16	16		20	17	17	17	15	14	14	13				
	5							17	16	15	17	19	18	18	18	14	14				
	0									16		19	17	17	14	16	15				

than giants. Supergiants are rare but can be seen over great distances.

25.3.1.1 Spectral Classification. The features in a normal stellar spectrum permit the classification of stars in a sequence of decreasing temperature. The main spectral characteristics of the classification scheme are listed in Table 25-5. Each spectral class is divided into 10 or more subintervals.

Table 25-5. Stellar Classification.

Class	Class Characteristics
Oxygen Stars	
O	Hot stars with He II absorption.
B	He I in absorption, H absorption stronger at later sub-classes.
A	H absorption very strong, decreasing at later sub-classes. Ca II increases with sub-class.
F	Ca II increases with sub-class, H weaker than class A, lines of metals developing.
G	Ca II strong, Fe and other metals strong, H weaker, molecular lines appear.
K	Strong metallic lines, molecular bands more pronounced.
M	Cool stars with lines of neutral metals very strong. TiO bands present and increasing with sub-class.
Carbon Stars	
R or C N	Cool stars (≈ M) Strong CN and C_2 bands, no TiO, metallic line spectra similar to K & M stars.
Heavy Metal Stars	
S	Cool stars (≈ M) bands of ZrO, YO and LaO, Zr and Te lines.

Ionization is denoted by Roman numerals. A neutral atom is I, singly ionized II.

A plot of absolute visual magnitude against spectral class (or temperature) for stars results in a well ordered display called the Hertzsprung-Russell diagram. Stars of different luminosity classes occupy different regions of the diagram. Thus, a two dimensional classification scheme is required, one which specifies the temperature, alternatively spectral class, the other the luminosity, or intrinsic brightness. Table 25-6 lists the absolute visual magnitude as a function of spectral and luminosity class for normal stars.

25.3.1.2 Stellar Luminosities, Colors and Temperatures. The total power over all wavelengths emitted by a star is the luminosity, L. The effective temperature T_e of a star of radius R is related to the luminosity by

$$L = 4\pi R^2 \sigma T_e^4$$

$\sigma \equiv$ Stefan-Boltzmann constant $= 5.66956 \times 10^{-12}$ $Wcm^{-2}deg^{-4}$.

The apparent bolometric magnitude, m_{bol}, corresponds with the apparent luminosity of the star at the earth. The bolometric correction B.C. is the value that is added to the visual magnitude to obtain m_{bol}.

The difference in magnitudes for observations of a star in different spectral bands is called the color. The colors for dwarf and giant stars, based on the photometric system defined in Table 25-1, are listed in Tables 25-7 and 25-8, respectively, as a function of spectral type. Also included are the bolometric correction and effective temperature for each class. The absolute magnitude in any of the specified spectral bands can be calculated from the data in Tables 25-6, 25-7 and 25-8.

25.3.1.3 Infrared Stars. The dominant stellar component of the apparent background shifts to the later, and cooler, spectral types as the wavelength of observation is increased. 70% of the stars in the Two Micron Sky Survey (TMSS) of Neugebauer and Leighton [1969] are M, S, or C giants; less than 3% of the stars in the visual spectral region are this cool [Nort, 1950]. At wavelengths longward of 6 μm several new classes of objects dominate over the "normal" stellar component. These sources have low infrared color temperatures due to either an intrinsically low temperature or extinction. The discrete, isolated, starlike sources among these new objects appear to be undergoing rapid mass loss. This mass loss produces extensive dusty envelopes that contain grains of silicate, graphite, SiC, Fe or Fe_3C. The emission from the embedded star is absorbed, in part, and re-emitted in the infrared.

25.3.1.4 M Stars. M stars are recognized by the optical features outlined in Table 25-5 or by low resolution IR spectroscopy for optically faint objects. Main spectral features are the CO fundamental and first overtone at 4.7 and 2.3 μm, H_2O at 2.7 μm, and circumstellar silicates at 9.7 μm and in some cases, 18.6 μm. The silicate feature is generally in weak emission in early M stars. It becomes progressively stronger with advancing spectral subclass, and in extreme cases it is optically thick and appears as a strong absorption. Examples are given in Figure 25-6. Many of these stars exhibit OH maser emission.

Table 25-6. Absolute visual magnitude as a function of spectral and luminosity class. Compiled from data given by Keenan [1963], Walborn [1973], Conti and Alschuler [1981], Humphreys [1978], Jung [1970, 1971], Mikami [1978a,b], Joy and Apt [1974] with spectral type corrections of Wing and Yorka [1979], Keenan [1978], and Egret, Keenan and Heck [1982]. Colon denotes uncertain values.

	Main Sequence	Subgiants	Giants		Supergiants		
Type	V	IV	III	II	Ib	Iab	Ia
O3	−5.5						−7.0:
O4	−5.5		−6.4:			−6.5	−7.0:
O5	−5.5		−6.4		−6.3:	−6.4	−7.0:
O6	−5.5		−5.6		−6.3:	−6.3	−7.0:
O7	−4.8		−5.5	−5.9	−6.2	−6.3	−7.0:
O8	−4.4		−5.5	−5.9	−6.2	−6.5	−7.0:
O9	−4.3	−5.2	−5.5	−5.9	−6.2	−6.7	−7.0:
						I	
B0	−4.2	−4.6	−5.1:	−5.6:		−6.0:	
B1	−3.5	−3.9	−3.3	−5.1:		−6.0:	
B2	−2.2	−3.0	−2.2	−4.4		−6.0:	
B3	−1.3	−2.3	−1.9	−3.9		−6.0:	
B5	−1.6	−1.8	−1.6	−3.7		−6.0:	
B7	−1.5	−1.0	−1.3	−3.6		−6.0:	
B9	0.0	−0.3	−0.7	−3.1		−6.0:	
A0	0.2	0.1	−0.6	−2.8		−6.0:	
A1	0.5	0.3	−0.4	−2.6		−6.0:	
A2	0.6	0.6	−0.2	−2.4		−6.0:	
A3	0.8	0.9	0.0	−2.3		−6.0:	
A5	1.5	1.2	0.2	−2.1		−6.0:	
A7	1.6	1.5	0.6	−2.0		−6.0:	
F0	2.3	2.0	0.6	−2.0		−6.0:	
F2	2.5	2.2	1.4	−2.0		−6.0:	
F5	3.0	2.3	0.2	−2.0		−6.0:	
F6	3.2	2.3	0.0	−2.0		−6.0:	
F8	3.5	2.4	0.6			−6.0:	
G0	4.1	2.9	1.0	−2.0		−6.0:	
G2	4.3	3.0	0.6	−2.0		−6.0:	
G5	4.9	3.1	0.3	−2.0		−6.0:	
G8	4.5	3.2	0.3	−2.1		−6.0:	
K0	5.8	3.2	0.2	−2.1:		−6.0:	
K1	6.0		0.2	−2.1		−6.0:	
K2	6.2		0.1	−2.2:		−6.0:	
K3	6.5		−0.2	−2.3:		−6.0:	
K4	6.9		−0.2	−2.3:		−6.0:	
K5	7.3		−0.5	−2.3:		−6.0:	
K7	8.0		−0.6	−2.3:		−6.0:	
M0	8.3		−0.7	−2.4:		−6.0:	
M1	9.2		−0.9	−2.4:		−6.0:	
M2	10.0		−1.1	−2.4:		−6.0:	
M3	11.0		−1.1	−2.4:			
M4	12.4		−1.0	−2.4:			
M5	15.0		−0.9				
M6	16.7		−0.9:				
M7	18.0		−0.9:				
M8			−0.5:				

Absolute visual magnitude as a function of spectral and luminosity class

Table 25-7. The Main Sequence colors compiled from data given by Johnson [1966], Aaronson, Frogel and Persson [1978], Veeder [1974], and Hayes [1978].

Spt	U-V	B-V	V-R	V-I	V-J	V-H	V-K	V-L	V-M	V-N	BC	T_{eff}
05-7	−1.46	−0.32	−0.15	−0.47	−0.73		−0.94	−1.01			− 3.9	42000
08-9	−1.44	−0.31	−0.15	−0.47	−0.73		−0.94	−1.01			− 3.5	34000
09.5	−1.40	−0.30	−0.14	−0.46	−0.73		−0.94	−1.00			− 3.1	31900
B0	−1.38	−0.30	−0.13	−0.42	−0.70		−0.93	−0.94			−2.96	30300
B0.5	−1.29	−0.28	−0.12	−0.39	−0.66		−0.88	−0.93			−2.83	28600
B1	−1.19	−0.26	−0.11	−0.36	−0.61		−0.81	−0.86			−2.59	25700
B2	−1.10	−0.24	−0.10	−0.32	−0.55		−0.74	−0.77			−2.36	23100
B3	−0.91	−0.20	−0.08	−0.27	−0.45		−0.61	−0.63			−1.94	18900
B5	−0.72	−0.16	−0.06	−0.22	−0.35		−0.47	−0.48			1.44	15300
B6	−0.63	−0.14	−0.06	−0.79	−0.36		−0.41	−0.41			−1.17	14000
B7	−0.54	−0.12	−0.04	−0.17	−0.25		−0.35	−0.34			−0.94	13000
B8	−0.39	−0.09	−0.02	−0.12	−0.17		−0.24	−0.22			−0.61	11500
B9	−0.25	0.06	0.00	−0.06	−0.09		−0.14	−0.11			−0.31	10106
A0	0.00	0.00	0.02	−0.00	−0.01		−0.03	0.00	−0.03	−0.03	−0.15	9410
A2	0.12	0.06	0.08	0.09	0.11		0.13	0.16	0.13	0.13	−0.08	8900
A5	0.25	0.14	0.16	0.22	0.27		0.36	0.40	0.36	0.36	−0.02	8210
A7	0.30	0.19	0.19	0.28	0.35		0.46	0.52	0.46	0.46	−0.01	7920
F0	0.34	0.27	0.30	0.47	0.58		0.79	0.86	0.79	0.79	−0.01	7160
F2	0.39	0.36	0.35	0.55	0.68	0.95	0.93	1.07	0.93	0.93	−0.02	6880
F5	0.45	0.42	0.40	0.64	0.79	1.08	1.07	1.25	1.07	1.07	−0.03	6560
F8	0.55	0.50	0.47	0.76	0.96	1.29	1.25	1.45	1.27	1.27	−0.08	6190
G0	0.63	0.59	0.50	0.81	1.03	1.38	1.35	1.53	1.35	1.44	−0.10	6010
G2	0.79	0.63	0.53	0.86	1.10	1.47	1.44	1.61	1.54	1.69	−0.13	5860
G5	0.89	0.70	0.54	0.89	1.14	1.52	1.49	1.67			−0.14	5780
G8	1.06	0.74	0.58	0.96	1.24	1.60	1.63	1.85			−0.18	5580
K0	1.36	0.89	0.64	1.06	1.38	1.80	1.83	2.00			−0.24	5260
K2	1.60	0.95	0.74	1.22	1.57	2.09	2.15	2.25			−0.35	4850
K5	2.28	1.18	0.99	1.62	2.04	2.71	2.82	2.84			−0.66	4270
K7	2.52	1.37	1.16	2.00	2.43	3.13	3.25	3.40			−0.93	4030
M0	2.67	1.43	1.28	2.26	2.85	3.46	3.60	3.78			−1.21	3880
M1	2.70	1.49	1.42	2.56	3.21	3.83	4.00	4.15			−1.49	3720
M2	2.69	1.53	1.51	2.76	3.45	4.08	4.27	4.47			−1.75	3600
M3	2.70	1.58	1.61	2.83	3.72	4.36	4.57	4.85			−1.90	3480
M4	2.70	1.62	1.71	3.30	3.99	4.64	4.87	5.40			−2.28	3370
M5	2.80	1.68	1.84	3.49	4.34	5.00	5.25	5.80			−2.59	(3260)
M6	2.99	1.76	2.01	3.87	4.69	5.46	5.75	7.8			−2.93	(3140)
M7	3.24	1.87	2.26	4.42	5.34	6.16	6.50				−3.46	(2880)
M8	(3.50)	1.99	2.52	4.90	6.25	6.85	7.25	9.6			−4.00	(2620)

Most of the late M type stars exhibit nearly periodic variations of one or more magnitudes in the infrared. The period of variability correlates with the strength of the silicate emission feature [DeGiora-Eastwood et al., 1981] and with the amplitude of variability. An extreme example is AFGL 2205 with a 4 magnitude amplitude of variability at 2.2 and 3.4 μm and a 4 year period.

25.3.1.5 Carbon Stars. In addition to the features listed in Table 25-5, carbon stars are distinguished in the infrared by an absorption at 3.1 μm due to various hydrocarbons and a broad SiC dust feature near 11 μm. Figure 25-7 shows these features for progressively cooler carbon stars. The combination strong molecular blanketing in the visual for the cooler carbon stars and circumstellar absorption make these stars optically faint. Consequently, the relative proportion of carbon stars as compared to late M stars at a given magnitude increases with increasing wavelength.

25.3.1.6 Planetary Nebulae and Related Objects. Infrared spectra of planetary nebulae show, in general, emission bands superimposed on featureless continua characteristic of heated dust. The emission bands, shown in Figure 25-8 for the planetary NGC 7027, are probably small hy-

Table 25-8. Colors for the Giant Stars, compiled from data given by Johnson [1966], Aaron, Frogel and Persson [1978], Lee [1971], Mendoza and Johnson [1965], Hayes [1978], Ridgway, Joyce, White and Wing [1980].

Spt	U-V	B-V	V-R	V-I	V-J	V-H	V-K	V-L	V-M	V-N	BC	T_{eff}
G5	1.55	0.92	0.69	1.17	1.52		2.08	2.18	2.02	2.05		
G8	1.64	0.94	0.70	1.18	1.56		2.18	2.27	2.09	2.12		4930
K0	1.92	1.04	0.77	1.30	1.71		2.33	2.47	2.23	2.28	−0.42	4790
K1	2.13	1.10	0.81	1.37	1.80		2.48	2.61	2.36	2.39	−0.48	4610
K2	2.32	1.16	0.84	1.42	1.87		2.64	2.73	2.45	2.48	−0.53	4450
K3	2.74	1.30	0.96	1.61	2.12		2.94	3.07	2.75	2.80	−0.60	4270
K4	3.07	1.41	1.06	1.81	2.36		3.25	3.39	3.05	3.11	−0.90	4095
K5	3.34	1.54	1.20	2.10	2.71		3.67	3.83	3.47	3.54	−1.19	3980
M0	3.42	1.56	1.24	2.14	2.77		3.74	3.89	3.59	3.65	−1.28	3895
M1	3.48	1.58	1.29	2.24	2.89	3.57	3.90	4.06	3.72	3.78	−1.36	3810
M2	3.52	1.61	1.37	2.45	3.12	3.72	4.16	4.33	3.91	3.97	−1.52	3730
M3	3.46	1.62	1.52	2.82	3.53	3.96	4.63	4.81	4.39	4.45	−1.91	3640
M4	3.29	1.62	1.79	3.38	4.19	4.39	5.34	5.54	5.10	5.14	−2.55	3560
M5	3.04	1.85	2.13	4.06	4.96	5.07	6.20	6.44	6.00	6.00	−3.34	3420
M6	4.15	1.70	2.70	4.90	5.90	5.90	7.20	7.50			−4.26	3250
M7		1.80	3.95	7.00	8.30	6.87	9.85	10.38			−7.7	
M8		1.93	4.95	8.51	10.25		11.92	12.60			−8.6	
M9		2.53	5.95	10.02	12.20		14.05	14.88			−10.8	

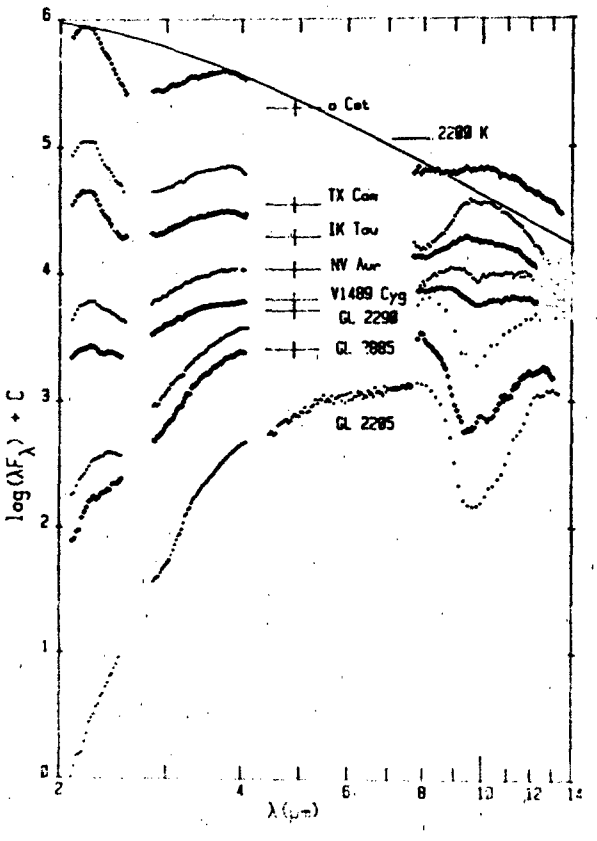


Figure 25-6. Infrared spectrophotometry of M stars illustrating the range of optical depth of the silicate feature in circumstellar envelopes [Ney and Merrill, 1980].

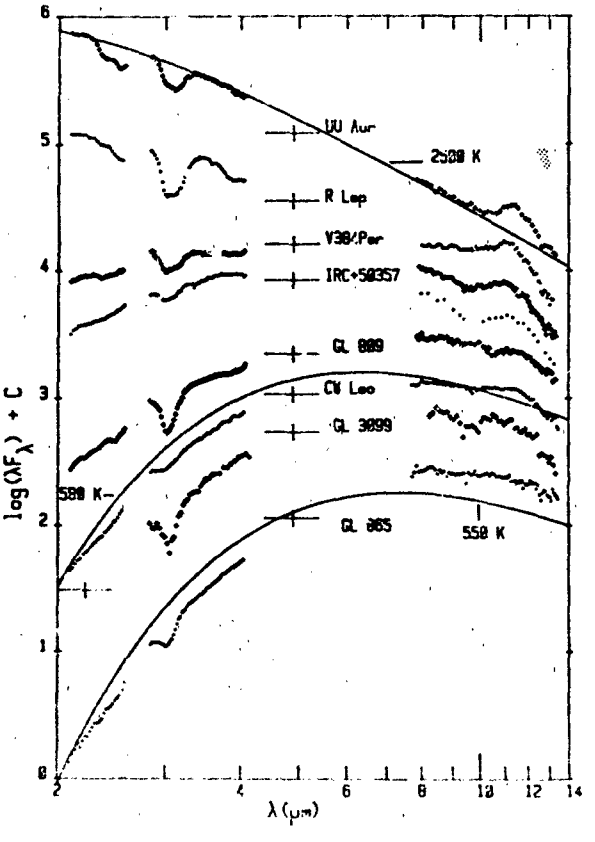


Figure 25-7. Infrared spectrophotometry of carbon stars illustrating the range of optical depths in the circumstellar envelopes associated with these objects [Ney and Merrill, 1980].

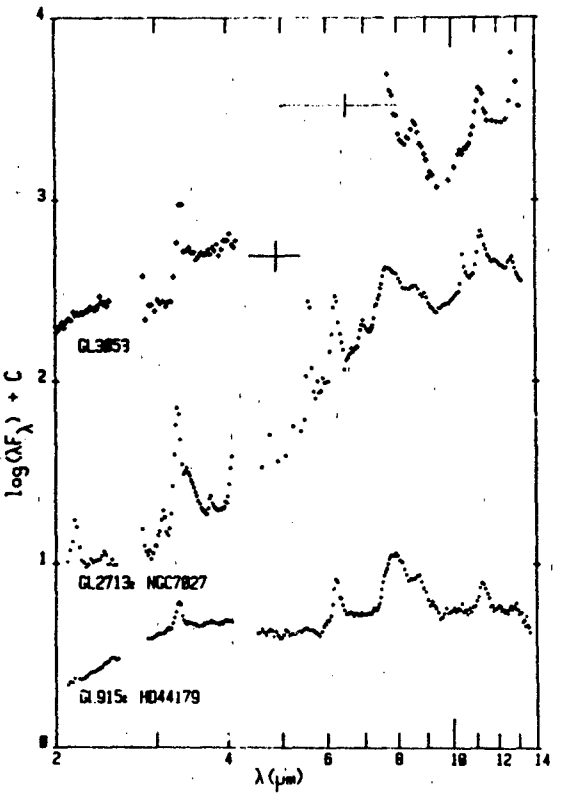


Figure 25-8. Spectrophotometry of objects with band emission from dust. NGC 7027 is identified as a planetary nebula; AFGL 915 is a disk object and AFGL 3053 is a source inside the HII region Sharpless 159 [Ney and Merrill, 1980].

drogenated carbon grains heated by UV photons [Leger and Puget, 1984]. These bands are present in the spectra of a variety of objects including compact and diffuse HII regions and the nuclei of active galaxies.

The 9.7 μm silicate band is not present in the spectra of planetary nebulae. This is consistent with the observed C/O abundance ratio in planetaries being much greater than unity. The condensates forming the dust are expected to be SiC and graphite, or possibly amorphous carbon.

Wolf-Rayet stars in the carbon sequence typically show broad continua due to heat, dust, and free-free emission of thermal electrons in their circumstellar envelopes.

Objects such as GL 915 (see Figure 25-8) exhibit disc shaped geometry, that is the circumstellar dust is confined to a plane. They are bright infrared sources but the visual properties of the exciting star depend on the geometry of the observation. GL 618 and 2688 are infrared sources with the central star obscured by the disk of dust. Optical observation of GL 2688 nebulosity above and below the dust lane show a highly reddened spectrum of an F5-Ia star. GL 915 is seen more nearly pole on and the spectra of the B9II-III star can be measured directly.

25.3.1.7 Luminous Stars with Infrared Excesses. Low excitation emission-line objects, hot supergiants and other types of stars earlier than M are included in this category. Be stars with $T \geq 13\,000$ K often show an excess of infrared due to free-free emission in a hot ($T \geq 10\,000$ K) circumstellar plasma shell [Gehrz et al., 1974]. The infrared excess, in general, decreases at the later spectral types. An optically thin dust shell enhances the infrared emission around some hot supergiants. The shell condenses from the mass lost by the star.

25.3.2 Gas and Dust

The galactic gas and dust is confined to the galactic plane with a thickness of about 200 parsecs. The distribution is patchy, with concentrations of matter in interstellar and molecular clouds. The interstellar gas is composed primarily of hydrogen, which is detected by the 21 cm emission in the hyperfine structure of neutral atomic hydrogen. The gas clouds tend to be along the spiral arms of the galaxy and mapping at 21 cm is used to define this structure. A number of other elements in the interstellar gas are observed by their atomic lines in the visual and ultraviolet and molecular lines in the radio region.

Dust is mixed with the interstellar gas. The dust may be composed of bare graphite grains plus grains of silicates and/or iron with water ice mantles. In some of the denser clouds, water ice absorption at 3.07 μm and the silicate feature at 9.7 μm has been observed. The interstellar absorption of starlight is almost entirely due to dust.

25.3.2.1 Interstellar Extinction. In Section 25.1.3 it was seen that the interstellar absorption, in magnitudes, was required to relate the observed apparent magnitude of a star to its absolute magnitude. From Table 25-9, the amount of absorption depends on the wavelength, in general decreasing with wavelength. This wavelength dependence is called in-

Table 25-9. Average interstellar absorption, in magnitudes, normalized to the visual with $A_v = 1.0$. Data from Johnson [1968] and Becklin et al. [1978].

λ(μm)	A_λ/A_v	λ(μm)	A_λ/A_v
0.365 (U)	1.58	1.65 (H)	0.16
0.4	1.45	2.22 (K)	0.09
0.44 (B)	1.32	3.5 (L)	0.05
0.5	1.13	4.8 (M)	0.03
0.553 (V)	1.00	8.7	0.06
0.67	0.67	9.5	0.13
0.9 (I)	0.46	11.2	0.09
1.0	0.38	12.5	0.04
1.25 (J)	0.26	20.0 (Q)	0.03

Colors from Table 25-1 that correspond to the listed wavelengths are indicated in parentheses.

terstellar reddening; a star will appear redder than it is due to the interstellar absorption.

Near the sun the average out of the plane extinction is [Milne and Aller, 1980]

$$C = k_v w(1 - e^{-r\sin|b|/w}) \operatorname{cosec} |b| \qquad (25.12)$$

b = galactic latitude,

k_v = 1.75 visual magnitudes per kiloparsec,

r = radial distance from the sun, and

w = scale height of absorbing material

= 0.119 kpc $b > 0°$

0.086 kpc $b < 0°$.

Large scale extinction in the galaxy can be estimated from the observed gas distribution and the fact that interstellar reddening is correlated with the total column density of hydrogen [Savage and Mathis, 1979].

25.3.2.2 HII Regions. An HII region is an optically bright gaseous nebula excited by an O or early B type star or a cluster of such stars ($T_e > 15\,000$ K). The stellar ultraviolet radiation photoionizes the surrounding gas; hydrogen is almost completely ionized, helium predominately singly ionized, and other elements singly or doubly ionized. The ionized electrons are collisionally thermalized to a kinetic temperature of about 10 000 K. Recombination of the thermal electrons and ions produces emission line spectra including forbidden lines.

Compact HII regions are the smallest gaseous nebulae and are considered to be young objects. Diffuse HII regions are an order to magnitude larger in size than the compact HII regions and considerably more rarified. Extended HII regions are very large and have low density; they overlap to a considerable degree forming a hot, tenuous plasma in the galactic plane. All of these objects emit copious infrared radiation. A major fraction of the ultraviolet luminosity of these sources is absorbed by the dust in and around the HII region and then is converted into the infrared. Generally, the infrared intensity peaks at about 70 μm but spectral distribution is broader than dust emitting at a single temperature. The dust and gas are well mixed as the radio flux from free-free emission from the ionized HII region correlates with the reradiated 11 μm flux. HII regions are among the strongest infrared sources in the galaxy at wavelengths longer than 10 μm.

25.3.3.3 Molecular Clouds. Molecular clouds are dense, cool clouds of interstellar gas and dust in which a number of molecular lines have been detected. The clouds may be divided into those which have a peak gas temperature measured by CO emission of $T_k < 20$ K throughout the region and those which are warmer. The second group ($T_k > 20$ K) require additional heat sources to explain the observed temperature, the first does not. The second group also has evidence of recent star formation such as compact HII regions and infrared sources.

Typical spectra for infrared sources in molecular clouds show red continua with strong ice and silicate absorption as seen in Figure 25-9. The emission of the cloud as a whole is characteristic of a much cooler temperature, peaking near 100 μm.

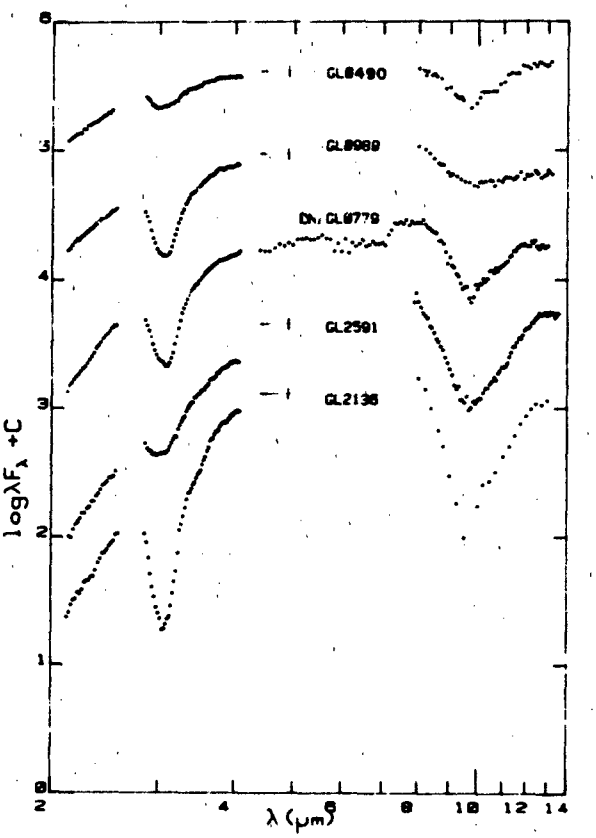


Figure 25-9. Spectrophotometry of source in molecular clouds. The prominent features are "ice" at 3.15 μm and the 10 μm silicate absorption [Ney and Merrill, 1980].

25.4 EXTRAGALACTIC OBJECTS

Galaxies are usually classified according to morphology. Ellipticals are elliptical in shape and arranged in a sequence of apparent ellipticity, from circular to most elongated. Spiral galaxies are in a sequence according to the relative sizes of the nuclear bulge and disk and details of the spiral structure in the disk. Early spiral galaxies have a large nuclear bulge and a small disk with tightly wound spiral arms. As

the classification scheme advances, the bulge decreases, the disk increases, and the arms open. At the end of this sequence are the irregular galaxies with little or no central condensation and little or no order to the disk.

Apart from this standard classification there are galaxies that are more condensed than normal systems. These objects can be classified along a continuum as to degree of compactness, the open Seyfert galaxies at one end and quasi-stellar objects (QSOs) at the other. The Seyferts are almost normal galaxies with an intensely bright and compact core. By definition QSO appear stellar on photographs on a plate scale used for surveys.

Infrared emission is a general feature in the nuclei of Seyfert galaxies. The infrared flux accounts for most of the total luminosity for many of these galaxies, and a significant portion of the QSO and BL Lac objects are intrinsically bright in the infrared. The emission from these objects can in general be represented by a power law through the visual and well into the infrared. This indicates a non-thermal cause, possibly noncoherent synchrotron emission, for these sources.

An unbiased infrared survey [de Jong et al., 1984] found that most of the infrared emission from galaxies is longward of 50 μm. The infrared emission is strongly correlated with the type of galaxy. Elliptical galaxies and lenticulars (SO) have little or no infrared emission in excess of the normal stellar component. The infrared excess increases with advancing galactic type, with late type spirals emitting as much as five times more infrared than visual energy. An infrared to visual energy ratio as large as 50 has been found for some optically faint unidentified galaxies [Soifer et al., 1984]. The more infrared luminous galaxies tend to be warmer ($T_{IR} \sim 50$ K) than weak infrared galaxies ($T_{IR} \sim 25$ K).

25.5 THE INFRARED CELESTIAL BACKGROUND

The infrared celestial background can be divided into three physical structures—the diffuse zodical emission, the sources in our galaxy, and external galaxies. The relative importance of these constituents depends on the wavelength. Currently, the observational data base is reasonably complete to a magnitude $M(\lambda) \leq 4$–4.5 over the spectral region from 2–30 μm.

25.5.1 2–8 μm

Late type giant stars dominate the background at these wavelengths. The zodiacal background is at a minimum being in the spectral crossover region between reflected sunlight and thermal emission. There are about 7300 stars brighter than m (2.2 μm) = 3 in the sky and roughly 26 000 objects brighter than m (2.2 μm) ≤ 4 [Hughes, 1970]. At m (2.2 μm) = 4 a typical M giant can be seen to a distance of about 1 kpc and a late K giant star out to 0.4 kpc. These distances exceed the 0.3 kpc scale height of the galactic disk. Thus, the concentration of stars to the galactic plane increases as the limiting magnitude is increased as most of the brighter giants have been detected toward the galactic poles by m (2.2 μm) ≤ 4. The contrast in areal densities in the galactic plane within 30° of the center and the pole is about 6 for m (2.2 μm) < 3 and 100 for m (2.2 μm) < 6.8.

For a field of view larger than 1 μsr the large stellar density near the galactic plane produces a diffuse background due to the aggregate of all the stars along the line of sight. This background is almost constant out to 30° from the center at 6×10^{-10} Wcm^{-2}μm^{-1}sr^{-1} at 2.4 μm and 1.5×10^{-10} Wcm^{-2} μm^{-1}sr^{-1} at 4.2 μm. The brightness falls off almost exponentially at greater latitudes. The latitudinal cross section of this emission is almost gaussian with a full width at half maximum of 3°5 at 2.4 μm and 2°5 at 4.2 μm.

The stellar source counts to reasonably faint levels (M ≤ 9) in the 2–5 μm spectral region can be modeled by a disk distribution of stars [Elias, 1978; Jones et al., 1981] with components divided by spectral and luminousity classes. This model has the density of the i^{th} component at R kpc from the sun in the (l, b) direction varying exponentially with an in plane distance, x, from the galactic center with a scale height of 2.5 kpc and perpendicular to the plane, |z|, with a scale height Z_i. Normalized to the solar neighborhood, the density is

$$\rho_i = \exp\left[(8.5 - x)/2.5 - |z|/Z_i\right]$$

$$r = (x^2 + z^2)^{1/2}$$

$$= \text{the galacto-centric distance of a point } (R, l, b) \text{ from the sun}$$

$$= R \sin b$$

$$x^2 = 8.5^2 + R^2 \cos^2 b - 17\, R \cos b \cos l$$

the sun is 8.5 kpc from the galactic center.

The luminosity function (No. of stars kpc^{-3}mag^{-1}) of the giant and supergiant stars is assumed to be gaussian

$$\phi_i(M) = \frac{N_i}{(2\pi)^{1/2}\sigma} \exp\left[-(M - \overline{M}_i)/2\sigma_i^2\right]$$

N_i = source density of the i^{th} component in the solar neighborhood

$\overline{M}_i$ = mean absolute magnitude component

σ_i = dispersion in absolute magnitude component.

Values for these parameters [Elias, 1978; Jones et al., 1981] are

Spectral Class	$\overline{M_i}$ (mag)	σ_i (mag)	log (N_i) (No. stars kpc^{-3})	Z_i (kpc)
F8-G2 III	0.45	1.0	4.65	0.50
G5 III	−0.08	1.0	4.65	0.50
G8 III	−0.56	0.8	5.20	0.25
K0,1 III	−0.80	0.7	5.23	0.30
K2,3 III	−1.66	0.7	5.23	0.30
K4,5 III	−3.36	0.6	4.28	0.30
M0 III	−4.16	0.6	3.48	0.30
M1 III	−4.40	0.6	3.13	0.30
M2 III	−4.76	0.6	3.13	0.30
M3 III	−5.23	0.6	3.13	0.30
M4 III	−6.04	0.6	3.00	0.30
M5 III	−6.90	0.5	3.00	0.30
M6 III	−7.90	0.5	2.45	0.30
M7 III	−8.90	0.5	2.09	0.30
M8+III	−9.90	0.5	1.65	0.20
A-G I-II	−7.00	2.0	1.53	0.5
K-M2 I-II	−9.50	1.0	1.49	0.5
M3-M4 I-II	−11.00	1.0	1.10	0.5

Dwarf stars [after Mamon and Soneira, 1982]

$$\phi_d \begin{cases} = 0 & M < -10 \\ = 10^7 * 10^{0.01M}\,[1 + 10^{-0.19(M-2)}]^4 & -10 \leq M \leq 9 \\ = 3 \times 10^{-7} \exp[-(M/9)^2] & M \geq 9 \end{cases}$$

with $Z_i = 0.09$ kpc $M \leq 1.5$

0.325 kpc $M \geq 3.5$

and linear interpolation between.

The number of stars brighter than apparent magnitude m in solid angle $d\omega$ in the (l,b) direction is

$$N(m,l,b)d\omega = d\omega \int_{-\infty}^{m} dm' \int_{o}^{\infty} R^2[\Sigma\, \rho_i(r,M)\phi_i(M)]dR$$

$$M = m - 5\log_{10}(R) - A(R) - 10. \qquad (25.13)$$

The interstellar extinction A(R) is given by (see also Equation (25.11))

$$\delta A(R)/\delta R = 0.07\,(2.2/\lambda)^2 \exp[(8.5 - x)/3.7 - |z|/0.1].$$

The in plane scale height of the dust is 3.7 kpc and the exponential thickness is 0.2 kpc.

Table 25-10. Number of sources per square degree brighter than a given magnitude (from the model in Section 25.5).

Mag/(l,b)	(0,0)	(20,0)	(60,0)	(180,00)	(30,0)	pole
			2.2 μm			
1	0.2	0.2	0.2	0.1	0.01	0.008
2	1.2	1.1	0.7	0.2	0.1	0.03
3	4.4	4.4	2.3	0.6	0.2	0.07
4	14.2	14.5	6.7	1.6	0.5	0.2
5	42.4	42.7	17.9	4.0	1.2	0.5
6	122.7	116.3	45.9	9.9	2.9	1.1
7	359.3	297.7	113.4	24.6	7.1	2.4
8	975.1	730.0	274.7	61.8	16.0	4.8
9	2489.5	1750.1	670.0	156.9	33.3	9.6
10	6339.0	4238.4	1688.0	396.5	65.7	20.7
			11 μm			
1	33.8	29.1	4.6	0.4	0.02	0.009
2	120.5	101.9	13.4	0.9	0.08	0.03
3	286.0	242.0	27.7	1.7	0.2	0.07
4	496.0	420.9	45.0	3.1	0.5	0.2
5	768.0	625.2	68.6	5.9	1.2	0.5
6	1278.7	967.9	112.7	12.4	3.1	1.2
7	2445.4	1723.4	212.6	28.7	7.6	2.8
8	5499.9	3563.1	447.4	70.4	17.7	6.4
9	11423.7	7408.2	1006.2	179.8	40.0	15.9
10	24420.6	15996.2	2477.3	465.6	91.1	45.8

The emission observed at 2.4 μm from the stellar component for the nuclear bulge has been modeled [Martiara et al., 1978] as

$$I(r) = 6.2 \times 10^{34} \text{ w } \mu\text{m}^{-1} \text{ kpc} \frac{\exp(-r/2.5)}{(r^2 + 0.14^2)}$$

with

$$r = x^2 + (z/0.52)^{2\ 1/2}. \qquad (25.14)$$

Source counts predicted by this model are given in Table 25-10.

25.5.2 8–30 μm

The thermal emission from the zodiacal dust, the asteroids, and most planets peak in this spectral region. The asteroids and planets without atmospheres have grey body spectral energy distributions with equilibrium temperatures characterized by their distance from the sun (Section 25.2). The zodiacal emission (Section 25.2.4) is the brightest component of the diffuse background.

The brightest classes of galactic objects at these wavelengths are stars embedded in circumstellar dust shells and compact HII regions. These objects absorb the visual and ultraviolet emission of the embedded stars and reradiated the energy in the infrared. These objects are intrinsically very luminous and are detectable over large distances. As a class the circumstellar dust shell source has a mean absolute 11 μm magnitude of −12, a one magnitude dispersion about the mean, a density of 90 sources per kpc^3 in the polar neighborhood and a perpendicular scale height of 300 pc [Grasdalen et al., 1983]. A 400 K color temperate for the emission provide reasonable extrapolation to other wavelengths.

The HII regions can be characterized by a mean absolute 11 μm magnitude of −18, a 2.0 magnitude dispersion, a density of 10 sources per kpc^3 in the solar neighborhood.

There are about 350 000 stars brighter than 4.5 magnitudes in the 8–30 μm region. The source density at this magnitude limit out of the galactic plane is about 1 star per square degree. About half of these objects are identified with stars in various catalogs most of which have K and M spectral classifications [Rowan-Robinson et al., 1984].

The predicted source counts based on the model in the previous section and including HII regions and star with circumstellar dust shells are given in Table 25-10.

CHAPTER 25

REFERENCES

Aaronson, M., J.A. Frogel, and S.E. Persson, Photometric Studies of Composite Stellar Systems. II. Observations of H_2O Absorption and the Coolest Stellar Component of E and SO Galaxies," *Astrophys. J.*, **220:** 442, 1978.

Allen, C.W., *Astrophysical Quantities*, Athlone Press, 1973.

Allen, D.A., "Infrared Studies of the Lunar Terrain," PhD Thesis, University of Cambridge, 1970.

Allen, D.A. and T.L. Murdock, "Infrared Photometry of Saturn, Tatan and the Rings," *Icarus*, **14:** 1–2, 1971.

Allen, D.A. and E.P. Ney, "Lunar Thermal Anomalies: Infrared Observations," *Science*, **164:** 419, 1969.

Becklin, E.E., K. Mathews, G. Neugebauer, and S.P. Willner, "Infrared Observations of the Galactic Center. I. Nature of the Compact Sources," *Astrophys. J.*, **219:** 121, 1978.

Conti, P.S. and W.R. Alschuler, "Spectroscopic Studies of O-Type Stars. I. Classification and Absolute Magnitudes," *Astrophys. J.*, **170:** 325, 1971.

Courtin, R., D. Gautier, and A. Lacombe, "On the Thermal Structure of Uranus from Infrared Measurements," *Astron. Astrophys.* **63:** 97–101, 1978.

DeGiora-Eastwood, K., J.A. Hackwell, G.L. Grasdalen, R.D. Gehrz, "A Correlation Between Infrared Excess and Period for Mira Variables," *Astrophys. J. (Ltrs)*, **245:** L75, 1981.

Egret, D., P.C. Keenan, and A. Heck, "The Absolute Magnitudes of G5-M3 Stars near the Giant Branch," *Astron. Astrophys.*, **106:** 115, 1982.

Elias, J. H., "A Study of the IC 5146 Dark Cloud Complex," *Astrophys. J.*, **223:** 859, 1978.

Erickson, E. F., D. Goorvitch, J. P. Simpson, and D. W. Strecker, "Far Infrared Spectrophotometry of Jupiter and Saturn," *Icarus*, **35:** 61–73, 1978.

Gehrz, R.D., J.A. Hackwell, and G.L. Grasdalen, "Infrared Studies of AFGL Sources," AFGL TR-79-0274, ADA 084713, 1979.

Gehrz, R.D., J.A. Hackwell, and G.L. Grasdalen, "Infrared Observations of Be Stars from 2.3 to 19.5 Microns," *Astrophys. J.*, **191:** 675, 1974.

Gillett, F.C. and W.J. Forrest, "The 7.5–13.5 μm Spectrum of Saturn," *Astrophys. J.*, **187:** L37–L38, 1974.

Gillett, F.C. and K.M. Merrill, "The 7.5–13.5 μm Spectra of Ceres and Vesta," *Icarus*, **26:** 358–360, 1975.

Gillett, F.C. and G.H. Rieke, "5-20 Micron Observations of Uranus and Neptune," *Astrophys. J.*, **218:** L141–144, 1977.

Gradie, J. and E. Tedesco, "Compositional Structure of the Asteroid Belt," *Science*, **216:** 1405–1407, 1982.

Grasdalen, G.L., R.D. Gehrz, J.A. Hackwell, M. Castelaz and C. Gullinson, "The Stellar Component of the Galaxy as seen by the AFGL Infrared Sky Survey," *Astrophys. J. Suppl.*, **53:** 413, 1983.

Hanel, R., B. Conrath, F.M. Flaser, V. Kunde, W. Maguire, J. Pearl, J. Pirraglia, R. Samuelson, L. Herath, M. Allison, D. Cruikshank, D. Gautier, P. Gierasch, L. Horn, R. Koppany, and C. Ponnamperuma, "Infrared Observations of the Saturnian System from Voyager 1," *Science*, **212:** 192–200, 1982.

Hansen, O., "An Explication of the Radiometric Method for Size and Albedo Determination," *Icarus*, **31:** 456–482, 1977.

Hauser, M.G., F.C. Gillett, F.J. Low, T.N. Gautier, C.A. Beichmann, G. Neugebauer, H.A. Aumann, B. Baud, N. Boggess, J.P. Emerson, J.R. Houck, B.T. Soifer, and R.G. Walker, "IRAS Observations of the Diffuse Infrared Background," *Astrophys. J. (Ltrs)*, **279:** L15, 1984.

Hayes, D.S., "The Absolute Calibration of the HR Diagram: Fundamental Effective Temperatures and Bolometric Corrections," IAU Symposium 80 *The HR Diagram*, edited by A.G. Davis Philip and D.S. Hayes, 65, 1968.

Hayes, D.S. and D.W. Latham, "A Rediscussion of the Atmospheric Extinction and the Absolute Spectral Energy Distribution of Vega," *Astrophys. J.*, **197:** 593, 1975.

Humphreys, R.M., "Studies of Luminous Stars in Nearby Galaxies. I. Super-giants and O Stars in the Milky Way," *Astrophys. J. Suppl*, **38:** 309, 1978.

Ingersoll A.P., G.S. Orton, G. Münch, G. Neugebauer, and S.C. Chase "Pioneer II Infrared Radiometer Experiment: The Global Heat Balance of Jupiter," *Science*, **188.** 472, 1975.

Ingersoll et al., "Pioneer Saturn Infrared Radiometer: Preliminary Results," *Science*, **207:** 439, 1980.

Johnson, H.L., "Interstellar Extinction in the Galaxy," *Astrophys. J.*, **141:** 923, 1965.

Johnson, H.L., "Astronomical Measurements in the Infrared," *Ann. Rev. Astron. Astrophys.*, **4:** 193, 1966.

Johnson, H.L., "Nebulae and Interstellar Matter: Interstellar Extinction," edited by B.M. Middlehurst and L.H. Allen, Ch. 5, 167, 1978.

Jones, T.J., M. Ashley, A.R. Hyland and A. Ruelas-Mayorga, "A Search for the Infrared Counterpart of Type II OH Masers-I. A Model for the IR Background Source Confusion," *Mon. Not. R. Astr. Soc.*, **197:** 413, 1981.

Joy, A.H. and H.A. Abt, "Spectral Types of M Dwarf Stars," *Astrophys. J., Suppl.*, **28:** 1, 1974.

Jung, J., "The Derivation of Absolute Magnitudes from Proper Motions and Radial Velocities and the Calibration of the HR Diagram II," *Astron. Astrophys.*, **4:** 53, 1970,

Keenan, P.C., "Classification of Stellar Spectra," *Basic Astronomical Data*, Ch. 8, edited by K. A. Strand, University of Chicago Press, 78, 1963.

Keenan, P.C., "A More Accurate HR Diagram for the Cooler Stars," IAU Symposium 80, "The HR Diagram," edited by A.G. Davis Philip and D.S. Hayes, 13, 1978.

Lee, T.A., "Photometry of High-Luminosity M-Type Stars," *Astrophys. J.*, **162:** 217, 1970.

Leger, A. and J.L. Puget, "Identification of the Unidentified IR Emission Features of Interstellar Dust?" Astron. Astrophys., **137:** L5, 1984.

Leinert, C., "Zodiacal Light—A Measurement of the Interplanetary Medium," *Space Sci. Rev.*, **18:** 281–339, 1975.

Levan, P.D. and S.D. Price, "85 μm Fluxes from Asteroids: 2 Palles, 7 Iris, 15 Eunamia, and 45 Eugenia," *Icarus*, **57:** 35, 1984.

Low, F.J., "Ground Based Infrared Measurements," AFCRLTR-73-0371, AD768620 1973.

Low, F.J., D.A. Beintoma, T.N. Gautier, F.C. Gillett, C.A. Beckman, G. Neugebauer, E. Young, H.H. Aumann, N. Boggess, J.D. Emerson, H.J. Habing, M.G. Hauser, J.R. Houck, M. Rowan-Robinson, B.T. Soifer, R.G. Walker and Wesselius, "Infrared Cirrus: New Components of the Extended Infrared Emission," *Astrophy. J. (Ltrs).* **278:** L19, 1984.

Maihara, T., N. Oda, T. Sugujama, and H. Okuda, "2.4 — Micron Observation of the Galaxy and Galactic Structure," *Publ. Astron. Soc. Japan*, **30:** 1, 1978.

Mamon, G.A. and R.M. Soneira, "Stellar Luminosity Functions in the R, I, J and K Bands Obtained by Transformation from the Visual Band," *Astrophys. J.*, **225:** 181, 1982.

Mendoza V., E.E. and H.L. Johnson, "Multicolor Photometry of Carbon Stars," *Astrophys. J.*, **141:** 161, 1965.

Mikami, T., "Absolute Magnitudes of M-Type Stars in the Solar Neighborhood," *Publ. Astron. Soc. Japan*, **30:** 191, 1978a.

Mikami, T., "Absolute Magnitudes of F-, G- and K-Type Stars in the Solar Neighborhood," *Publ. Astron. Soc. Japan*, **30:** 207, 1978b.

Milne, D.H. and L.H. Aller, "An Average Model for Galactic Absorption," *Astron. J.*, **85:** 17, 1980.

Murdock, T.L., "Mercury's Infrared Phase Effect," *Astron. J.*, **79:** 1457–1464, 1964.

Murdock, T.L. and E.P. Ney, "Mercury: The Dark Side Temperature," *Science*, **170:** 535–537, 1970.

Murdock, T.L. and S.D. Price, "Measurements of Infrared Zodiacal Light: *Astron. J.*, in press, 1985.

Neugebauer, G. and R.B. Leighton, "Two Micron Sky Survey: A Preliminary Catalog," NASA SP-3047, 1969.

Ney, E.P., "Multiband Photometry of Comets Kahoutek, Bennett, Bradfield and Encke," *Icarus* **23:** 551–560, 1974.

Ney, E.P. and K.M. Merrill, "Study of Sources in AFGL Rocket Infrared Study," AFGL TR-80-0050, ADA 084098, 1980.

Nolt, I.G., A. Tokunaga, F.C. Gillett, and J. Caldwell, "The 22.7 Micron Brightness of Saturns Rings Versus Declination of the Sun," *Ap. J.*, **219:** L63, 1978.

Nort, H., "The Stars of Six Spectral Classes and the Milky Way," *Bull. Astron. Neth.*, **11:** 181, 1950.

Pikkarainen, T., "The Surface Structures of the Moon and Mercury Derived from Integrated Photometry," *Ann. Acad. Sci. Fenn.*, **A6:** 316, 1969.

Price, S.D., L.P. Marcotte, and T.L. Murdock, "Infrared Observation of the Zodiacal Dust Cloud," *Astron. J.*, **85:** 765–771, 1980.

Reed, R.A., W.J. Forrest, J.R. Houck, and J.B. Pollack, "Venus: The 17-38 Micron Spectrum," *Icarus*, **33:** 554–557, 1978.

Ridgeway, S.T., R.R. Joyce, N.M. White, and R.F. Wing, "Effective Temperatures of Lefe-Type Stars: The Field Giants from K0 to M6," *Astrophys. J.*, **235:** 126, 1980.

Rowan-Robinson, M., P. Clegg, C. Beichmann, G. Neugebauer, B.T. Soifer, H.H. Aumann, D. Beintema, N. Boggess, J. Emerson, N. Gautier, F. Gillett, M. Hauser, F. Low, and R. Walker, "The IRAS Minisurvey" *Astrophys. J. (Ltrs)*, **278:** L7, 1984.

Ryadov, U.Y., N.I. Furashov, and G.A. Sharonov, "Measurements of the Moon's Natural Infrared Thermal Emission," *Astron. Zh.*, **41:** 112, 1964.

Savage, B.D. and J.S. Mathis, "Observed Properties of Interstellar Dust," *Ann Rev. Astron. Astrophys.*, **73:** 1979.

Shorthill, R.W. and J.M. Saari, "Non-uniform Cooling of the Eclipsed Moon," *Science*, **150:** 210–212, 1965.

Smith, B.A., L. Soderblom, R. Beebe, J. Boyce, G. Briggs, A. Bunker, S.A. Collins, C.J. Hansen, T.V. Johnson, J.L. Mitchell, R.J. Terrile, M. Carr, A.F. Cook II, J. Cuzzi, J.B. Pollack, G.E. Danielson, A. Ingersoll, M.E. Davies, G.E. Hunt, H. Masursky, E. Shoemaker, D. Morrison, T. Owen, C. Sagan, J. Veverka, R. Stone, V.E. Suomi, "Encounter with Saturn: Voyager 1 Imaging Science Results," *Science*, **212:** 163–191, 1981.

Soifer, B.T., J.R. Houck and M. Harwit, "Rocket Infrared Observations of the Interplanetary Medium," *Ap. J. Ltrs.*, **168:** L73, 1971.

Stier, M.T., W.A. Traub, G.G. Fazio, E.L. Wright, and F.J. Low, "Far Infrared Observations of Uranus, Neptune and Cere," *Astrophys. J.*, **226:** 347–349, 1978.

Tokunaga, A., Knacke, R.F., and Owen, T., "17-25 μm Spectra of Jupiter and Saturn," *Astrophys. J.*, **213:** 569–574, 1977.

Tug, H., N.M. White, and G.W. Lockwood, "Absolute Energy Distributions of α Lyrae and 109 Virginis from 3295 Å to 9040 Å," *Astron. Astrophys*, **61:** 679, 1977.

Upgren, A.P., "New Parallaxes for Old: A Coming Improvement in the Distance Scale of the Universe," *Pub. Astron. Soc. Pac.*, **9:** 143, 1980.

Veeder, G.J., "Luminosities and Temperatures of M Dwarf Stars from Infrared Photometry," *Astron. J.*, **79:** 1056, 1974.

Walborn, N.R., "Space Distribution of the O Stars in the Solar Neighborhood," *Astron. J.*, **78:** 1067, 1973.

Ward, D.B., G.E. Gull, and M. Harwit, "Far Infrared Spectral Observations of Venus, Mars and Jupiter," *Icarus*, **30:** 295–300 1977.

Wing, R.F. and S.B. Yorka, "Comparisons of Temperature Classes Obtained by Spectroscopic and Photometric Technique: M Dwarfs, Giants and Supergiants," IAU Colloquium 47, *Spectral Classification of the Future*, edited by McCarthy, Philip, Coyne, Vatican Obs., 1979.

Wright, E.N., "Recalibration of the Far Infrared Brightness Temperatures of the Planets," *Astrophys. J.*, **210:** 250–262, 1976.

APPENDIX

UNITS, CONSTANTS, AND CONVERSION FACTORS

THE INTERNATIONAL SYSTEM OF UNITS (SI)[1]

Table A–1. SI base units.

Name	Symbol	Definition
meter	m	"The meter is the length equal to 1 650 763.73 wavelengths in vacuum of the radiation corresponding to the transition between the levels $2p_{10}$ and $5d_5$ of the krypton-86 atom."
kilogram	kg	"The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram [a cylinder of platinum-iridium]."
second	s	"The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom."
ampere	A	"The ampere is the constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross section, and placed 1 meter apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per meter of length."
kelvin	K	"The kelvin is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water." The "degree Celsius" is defined by the equation $t = T - T_0$, where T is the thermodynamic temperature in kelvins and $T_0 = 273.15$ K.
mole	mol	"The mole is the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kilogram of carbon 12."
candela	cd	"The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of 1/683 watt per steradian."

Table A–2. SI derived units.

Name	Symbol (dimensions)	Definition	Value in cgs units[a]
Absorbed dose	Gy ($m \cdot s^{-2}$)	The *gray* is the absorbed dose when the energy per unit mass imparted to matter by ionizing radiation is one joule per kilogram. (The gray is also used for the ionizing radiation quantities: specific energy imparted, kerma, and absorbed dose index, which have the SI unit joule per kilogram.) 1 rad = 10^{-2} Gy.	10^4 ergs/gm
Activity	Bq (s^{-1})	The *becquerel* is the activity of a radionuclide decaying at the rate of one spontaneous nuclear transition per second. 1 Ci (curie) = 3.7×10^{10} Bq.	1 sec^{-1}
Dose equivalent	Sv ($m^2 \cdot s^{-2}$)	The *sievert* is the dose equivalent when the absorbed dose of ionizing radiation multiplied by the dimensionless factors Q (quality factor) and N (product of any other multiplying factors) stipulated by the International Commission on Radiological Protection is one joule per kilogram.	10^4 ergs/gm
Electric capacitance	F ($m^{-2} \cdot kg^{-1} \cdot s^4 \cdot A^2$)	The *farad* is the capacitance of a capacitor between the plates of which there appears a difference of potential of one volt when it is charged by a quantity of electricity equal to one coulomb.	8.988×10^{11} esu
Electric conductance	S ($m^{-2} \cdot kg^{-1} \cdot s^3 \cdot A^2$)	The *siemens* is the electric conductance of a conductor in which a current of one ampere is produced by an electric potential difference of one volt.	8.988×10^{11} esu (cm/sec)

1. From *Physics Vade Mecum*, edited by H.L. Anderson, American Institute of Physics, New York, 1981.

Table A–2 (*continued*)

Name	Symbol (dimensions)	Definition	Value in cgs units[a]
Electric inductance	H ($m^2 \cdot kg \cdot s^{-2} \cdot A^{-2}$)	The *henry* is the inductance of a closed circuit in which an electromotive force of one volt is produced when the electric current in the circuit varies uniformly at a rate of one ampere per second.	10^9 emu (cm)
Electric potential difference, electromotive	V ($m^2 \cdot kg \cdot s^{-2} \cdot A^{-1}$)	The *volt* (unit of electric potential difference and electromotive force) is the difference of electric potential between two points of a conductor carrying a constant current of one ampere, when the power dissipated between these points is equal to one watt.	$(1/2.998) \times 10^{-2}$ esu ($cm^{1/2} \cdot gm^{1/2} \cdot sec^{-1}$)
Electric resistance	Ω ($m^2 \cdot kg \cdot s^{-3} \cdot A^{-2}$)	The *ohm* is the electric resistance between two points of a conductor when a constant difference of potential of one volt, applied between these two points, produces in this conductor a current of one ampere, this conductor not being the source of any electromotive force.	$(1/8.988) \times 10^{-11}$ esu ($cm^{-1} \cdot sec$)
Energy	J ($m^2 \cdot kg \cdot s^{-2}$)	The *joule* is the work done when the point of application of a force of one newton is displaced a distance of one meter in the direction of the force.	10^7 ergs ($cm^2 \cdot gm \cdot sec^{-2}$)
Force	N ($m \cdot kg \cdot s^{-2}$)	The *newton* is that force which, when applied to a body having a mass of one kilogram, gives it an acceleration of one meter per second squared.	10^5 dyn ($cm \cdot gm \cdot sec^{-2}$)
Frequency	Hz (s^{-1})	The *hertz* is the frequency of a periodic phenomenon of which the period is one second.	cycles/sec (sec^{-1})
Illuminance	lx ($cd \cdot sr \cdot m^{-2}$)	The *lux* is the illuminance produced by a luminous flux of one lumen uniformly distributed over a surface of one square meter.	
Luminous flux	lm ($cd \cdot sr$)	The *lumen* is the luminous flux emitted in a solid angle of one steradian by a point source having a uniform intensity of one candela.	
Magnetic flux	Wb ($m^2 \cdot kg \cdot s^{-2} \cdot A^{-1}$)	The *weber* is the magnetic flux which, linking a circuit of one turn, produces in it an electromotive force of one volt as it is reduced to zero at a uniform rate in one second.	10^8 Mx ($cm^{3/2} \cdot gm^{1/2} \cdot sec^{-1}$)
Magnetic flux density	T ($kg \cdot s^{-2} \cdot A^{-1}$)	The *tesla* is the magnetic flux density given by a magnetic flux of one weber per square meter.	10^4 Gs ($cm^{-1/2} \cdot gm^{1/2} \cdot sec^{-1}$)
Power	W ($m^2 \cdot kg \cdot s^{-3}$)	The *watt* is the power which gives rise to the production of energy at the rate of one joule per second.	10^7 ergs/sec ($cm^2 \cdot gm \cdot sec^{-2}$)
Pressure or stress	Pa ($m^{-1} \cdot kg \cdot s^{-2}$)	The *pascal* is the pressure or stress of one newton per square meter.	10 dyn/cm^2 ($cm^{-1} \cdot gm \cdot sec^{-2}$)
Quantity of electricity	C	The *coulomb* is the quantity of electricity transported in one second by a current of one ampere.	2.998×10^9 esu ($cm^{3/2} \cdot gm^{1/2} \cdot sec^{-1}$)

[a]For more precise work use 2.997 924 58 for 2.998 and 8.987 551 79 for 8.988.

Table A–3. SI prefixes.

Factor	Prefix	Symbol	Factor	Prefix	Symbol
10^{18}	exa	E	10^{-1}	deci	d
10^{15}	peta	P	10^{-2}	centi	c
10^{12}	tera	T	10^{-3}	milli	m
10^9	giga	G	10^{-6}	micro	μ
10^6	mega	M	10^{-9}	nano	n
10^3	kilo	k	10^{-12}	pico	p
10^2	hecto	h	10^{-15}	femto	f
10^1	deka	da	10^{-18}	atto	a

Table A–4. Conversion to SI units.

1. Acceleration

The gal is a special unit employed in geodesy and geophysics to express the acceleration due to gravity.

1 ft/s^2	= 0.30480* m/s^2	1 gal	= 0.01000* m/s^2
Standard gravity (g)	= 9.8067 m/s^2	Sun's surface	= 274.0 m/s^2
		1 AU from Sun	= 0.005931 m/s^2

2. Angle

1 degree (°)	= 1.7453 × 10^{-2} rad	1 second (")	= 4.8481 × 10^{-6} rad
1 minute (')	= 2.9089 × 10^{-4} rad		

3. Area

The darcy is a unit for measuring the permeability of porous solids.

1 acre	= 4046.9 m^2	1 in.^{2}	= 6.4516* × 10^{-4} m^2
1 are	= 100.00* m^2	1 square mile (international)	= 2.5900 × 10^6 m^2
1 barn (b)	= 1.0000* × 10^{-28} m^2	1 square mile (statute)'	= 2.5900 × 10^6 m^2
1 circular mil	= 5.0671 × 10^{-10} m^2	1 square (building)	= 9.2903 m^2
1 darcy	= 9.8692 × 10^{-13} m^2	1 square rod (rd^2), square pole, or square perch	= 25.293 m^2
1 ft^2	= 0.092 903 m^2	1 square yard (yd^2)	= 0.83613 m^2
1 hectare	= 10 000* m^2		

4. Density

1 grain/gal (U.S. liquid)	= 0.017118 kg/m^3	1 ton (short)/yd^3	= 1186.6 kg/m^3
1 oz (avoirdupois)/in.^{3}	= 1730.0 kg/m^3	Density of water (4°C)	= 999.97 kg/m^3
1 lb/ft^3	= 16.018 kg/m^3	Density of mercury (0°C)	= 13595 kg/m^3
1 lb/in.^{3}	= 27680 kg/m^3	Solar mass/cubic parsec	= 6.770 × 10^{-20} kg/m^3
1 lb/gal (U.S. liquid)	= 119.83 kg/m^3	STP gas density for molecular weight M_0	= 0.044615M_0 kg/m^3
1 ton (long)/yd^3	= 1328.9 kg/m^3		

5. Electricity and magnetism

A = ampere, C = coulomb, F = farad, H = henry, Ω = ohm, S = siemens, V = volt, T = tesla, Wb = weber, * = exact value.

1 abampere	= 10.000* A	1 ohm centimeter	= 1.0000* × 10^{-2} Ω·m
1 abcoulomb	= 10.000* C	1 ohm circular-mil per foot	= 1.6624 × 10^{-9} Ω·m
1 abfarad	= 1.0000* × 10^9 F	1 statampere	= 3.3356 × 10^{-10} A
1 abhenry	= 1.0000* × 10^{-9} H	1 statcoulomb	= 3.3356 × 10^{-10} C
1 abmho	= 1.0000* × 10^9 S	1 statfarad	= 1.1127 × 10^{-12} F
1 abohm	= 1.0000* × 10^{-9} Ω	1 stathenry	= 8.9876 × 10^{11} H
1 abvolt	= 1.0000* × 10^{-8} V	1 statmho	= 1.1127 × 10^{-12} S
1 ampere hour	= 3600* C	1 statohm	= 8.9876 × 10^{11} Ω
1 emu of capacitance	= 1.0000* × 10^9 F	1 statvolt	= 299.79 V
1 emu of current	= 10.000* A	1 unit pole	= 1.2566 × 10^{-7} Wb
1 emu of electric potential	= 1.0000* × 10^{-8} V	Potential of electron at 1st Bohr orbit	= 27.212 V
1 emu of inductance	= 1.0000* × 10^{-9} H	Ionization potential from 1st Bohr orbit	= 13.606 V
1 emu of resistance	= 1.0000* × 10^{-9} Ω	Nuclear electric field at 1st Bohr orbit	= 5.140 × 10^{11} V/m
1 esu of capacitance	= 1.1127 × 10^{-12} F	Current in 1st Bohr orbit	= 1.054 × 10^{-3} A
1 esu of current	= 3.3356 × 10^{-10} A	Dipole moment of nucleus and electron in 1st Bohr orbit	= 0.8478 × 10^{-29} C·m
1 esu of electric potential	= 2.9979 × 10^2 V	Magnetic field, atomic unit	= 1715 T
1 esu of inductance	= 8.9876 × 10^{11} H	Field at nucleus due to electron in 1st Bohr orbit	= 9.9551 × 10^6 A/m
1 esu of resistance	= 8.9876 × 10^{11} Ω	Magnetic moment, atomic unit	= 2.542 × 10^{-21} J/T
1 faraday (based on ^{12}C)	= 9.6487 × 10^4 C	Earth magnetic moment	= 7.98 × 10^{22} J/T
1 faraday (chemical)	= 9.6496 × 10^4 C		
1 faraday (physical)	= 9.6522 × 10^4 C		
1 gamma	= 1.0000* × 10^{-9} T		
1 gauss	= 1.0000* × 10^{-4} T		
1 gilbert	= 7.9577 × 10^{-1} A (amp. turns)		
1 maxwell	= 1.0000* × 10^{-8} Wb		
1 mho	= 1.0000* S		
1 oersted	= 79.577 A/m		

6. Energy

Btu = British thermal unit (thermochemical), 1 Btu (International Table) = 1.000 67 Btu (thermochemical); cal = calorie (thermochemical), 1 cal (International Table) = 1.000 67 cal (thermochemical); J = joule; W = watt.

1 Btu	= 1054.4 J	1 foot-poundal	= 0.042140 J
1 Btu (mean)	= 1055.9 J	1 kilowatt hour (kW·h)	= 3.6000* × 10^6 J
1 Btu (39°F)	= 1059.7 J	1 therm	= 1.0551 × 10^8 J

Table A-4 (continued)

1 Btu (60°F)	= 1054.7 J	1 ton (nuclear equivalent of TNT)	= 4.184×10^9 J
1 calorie	= 4.1840* J	1 watt hour (W·h)	= 3600* J
1 calorie (mean)	= 4.1900 J	1 watt second (W·s)	= 1.0000* J
1 calorie (15°C)	= 4.1858 J	Energy of unit wave number (hc)	= 1.9865×10^{-25} J
1 calorie (20°C)	= 4.1819 J	Mass energy of unit atomic weight	= 1.4924×10^{-10} J
1 kilocalorie	= 4184.0* J		
1 electron volt (eV)	= 1.6022×10^{-19} J		
1 erg	= $1.0000^* \times 10^{-7}$ J		
1 foot-pound (ft·lbf)	= 1.3558 J		

Note: 1 quad = 10^{15} Btu

1 quad per year = 0.472 million barrels of oil per day (1 barrel = 42 gallons)
= 1 trillion cubic feet of gas per year
= 44.4 million tons of coal per year (for medium heating value coal at 22.5 Btu/ton)
= 33.4 million kilowatts of electricity
= 293 billion kilowatt-hours of electricity per year at 100% efficiency
= 95.2 billion kilowatt-hours of electricity per year at 32.5% efficiency

See Chaps. 10.00 (Energy Demand) and 11.00 (Energy Supply).

7 Force

1 dyne	= $1.0000^* \times 10^{-5}$ N	1 lbf/lb (thrust/weight [mass] ratio)	= 9.8067 N
1 kilogram-force	= 9.8067 N	1 poundal	= 0.13825 N
1 kip (1000 lbf)	= 4448.2 N	1 ton-force (2000 lbf)	= 8896.4 N
1 ounce-force	= 0.27801 N	Proton-electron attraction at distance a_0	= 8.238×10^{-8} N
1 pound-force (lbf)	= 4.4482 N		

8 Frequency

1 hertz (Hz)	= 1 cycle/s	Frequency of free electron in magnetic field H	= $2.7993 \times 10^{10} H$ Hz·T^{-1}
1 kayser	= 3×10^{10} Hz		
Rydberg frequency (cR_∞)	= 3.2898×10^{15} Hz	Plasma frequency associated with electron density N_e	= $8.979 N_e^{1/2}$ Hz (N_e in m^{-3})
Frequency of 1st Bohr orbit ($2cR_\infty$)	= 6.5797×10^{15} Hz		

9 Heat

Btu = British thermal unit (thermochemical). 1 Btu (International Table) = 1.000 67 Btu (thermochemical). cal = calorie (thermochemical). 1 cal (International Table) = 1.000 67 cal (thermochemical). J = joule. K = kelvin. W = watt; h = hour.

Thermal conductivity k

1 Btu·ft/h·ft^2·°F	= 1.7296 W/m·K	1 Btu·in./s·ft^2·°F	= 518.87 W/m·K

Thermal conductance C

1 Btu/h·ft^2·°F	= 5.6745 W/m^2·K

Heat capacity

1 Btu/lb·°F	= 4184.0* J/kg·K	1 cal/s	= 4.1840* W
1 cal/g·°C	= 4184.0* J/kg·K		

Thermal resistance R

1 °F·h·ft^2/Btu	= 0.176 23 K·m^2/W	Thermal resistivity	
1 clo	= 0.20037 K·m^2/W	1 °F·h·ft^2/Btu·in.	= 6.9381 K·m/W

Thermal diffusivity

1 ft^2/h	= 2.5806×10^{-5} m^2/s

10 Length

1 angstrom (Å)	= $1.0000^* \times 10^{-10}$ m	1 microinch	= $2.5400^* \times 10^{-8}$ m
1 atomic unit (a_0)	= $0.529\,18 \times 10^{-10}$ m	1 micron (μm)	= $1.0000^* \times 10^{-6}$ m
1 astronomical unit (AU)	= 1.4960×10^{11} m	1 mil	= $2.5400^* \times 10^{-5}$ m
1 cable's length	= 219 m	1 mile (int. nautical)	= 1852.0* m
1 chain	= 20.117 m	1 mile (U.S. nautical)	= 1852.0* m
1 electron radius (r_e)	= 2.818×10^{-15} m	1 mile (international)	= 1609.3 m
1 fathom	= 1.8288 m	1 mile (U.S. statute)	= 1609.3 m
1 fermi (femtometer) (fm)	= $1.0000^* \times 10^{-15}$ m	1 parsec (pc)	= 3.0857×10^{16} m
1 foot (ft)	= 0.30480* m	1 pica (printer's)	= 4.2175×10^{-3} m
1 foot (U.S. survey)	= 0.30480 m	1 point (printer's)	= 3.5146×10^{-4} m
1 furlong	= 201.17 m	1 rod	= 5.0292 m
1 hand	= 0.10160 m	1 solar radius ($R_\odot$)	= 6.960×10^8 m
1 inch (in.)	= 0.02540* m	Wavelength of 1-eV photon (hc eV)	= 1.2399×10^{-6} m
1 league (land)	= 4828 m		

Table A-4 (continued)

1 light year (ly)	$= 9.4606 \times 10^{15}$ m	1 x unit	$= 1.002 \times 10^{-13}$ m
1 link (Gunther's or surveyors')	= 0.20117 m	1 yard	= 0.91440* m

11 Light

cd = candela, lm = lumen, [1 lumen = flux from (1/60π) cm² of blackbody at 2044K], lx = lux.

1 apostilb	= 1 lm/m² for perfectly diffusing surface	1 lambert	= 3183.1 cd/m²
		1 lumen of maximum-visibility radiation (5550 Å)	$= 1.470 \times 10^{-3}$ W
1 cd/in.²	= 1550.0 cd/m²	1 stilb (sb)	= 10 000* cd/m²
1 foot-candle	= 10.764 lx	1 phot	= 10 000* lx
1 foot-lambert	= 3.4263 cd/m²		

12 Mass

1 atomic unit (electron) (m_e)	$= 9.1095 \times 10^{-34}$ kg	1 ounce (avoirdupois)	$= 2.8350 \times 10^{-2}$ kg
1 atomic mass unit (^{12}C scale) (amu)	$= 1.660\,57 \times 10^{-27}$ kg	1 ounce (troy or apothecary)	$= 3.1103 \times 10^{-2}$ kg
1 carat (metric)	$= 2.0000* \times 10^{-4}$ kg	1 pennyweight (troy)	$= 1.5552 \times 10^{-3}$ kg
1 dram, apothecary	$= 3.8879 \times 10^{-3}$ kg	1 pound (lb avoirdupois)	= 0.45359 kg
1 dram, avoirdupois	$= 1.7718 \times 10^{-3}$ kg	1 pound (troy or apothecary)	= 0.37324 kg
1 gamma	$= 1.0000* \times 10^{-9}$ kg	1 quintal (q)	= 100.00 kg
1 grain	$= 6.4799 \times 10^{-5}$ kg	1 scruple	$= 1.2960 \times 10^{-3}$ kg
1 hundredweight (gross or long)	= 50.802 kg	1 slug	= 14.594 kg
		1 ton (assay)	= 0.029167 kg
1 hundred weight (net or short)	= 45.359 kg	1 ton (long, 2240 lb)	= 1016.0 kg
1 kgf·s²/m	= 9.8067 kg	1 ton (short, 2000 lb)	= 907.18 kg
		1 ton (metric ton)	= 1.0000* kg

13 Mass per unit length

1 denier	$= 1.1111 \times 10^{-7}$ kg/m	1 tex	$= 1.0000* \times 10^{-6}$ kg/m

14 Mass per unit time

1 perm (0°C)	$= 5.7214 \times 10^{-11}$ kg/Pa·s·m²	1 lb/hp·h	$= 1.6897 \times 10^{-7}$ kg/J
1 perm-in (0°C)	$= 1.4532 \times 10^{-12}$ kg/Pa·s·m	1 ton (short)/h	= 0.25200 kg/s
1 lb/h	$= 1.2600 \times 10^{-4}$ kg/s		

15 Power

1 Btu (int.)/h	= 0.29307 W	1 horsepower (boiler)	= 9809.5 W
1 Btu (int.)/s	= 1055.1 W	1 horsepower (electric)	= 746.00* W
1 Btu (thermochem.)/h	= 0.292 88 W	1 horsepower (metric)	= 735.50 W
1 cal (thermochem.)/s	= 4.1840* W	1 horsepower (water)	= 746.04 W
1 force de cheval	= 735.5 W	1 horsepower (U.K.)	= 745.70 W
1 erg/s	$= 1.0000* \times 10^{-7}$ W	1 ton (refrigeration)	= 3516.8 W
1 ft·lbf/h	$= 3.7662 \times 10^{-4}$ W	Star, $M_{bol} = 0$ radiation	$= 2.97 \times 10^{26}$ W
1 horsepower (550 ft·lbf/s)	= 745.70 W	Solar luminosity	$= 3.826 \times 10^{26}$ W

16 Pressure or stress (force per unit area)

1 atmosphere (standard)	= 101 325* Pa	1 inch of water (39.2°F)	= 249.08 Pa
1 atmosphere (technical = 1 kgf/cm²)	= 98 066.5* Pa	1 kgf/cm²	= 98 066.5* Pa
		1 kip/in.² (ksi)	$= 6.8948 \times 10^{6}$ Pa
1 bar	= 100 000* Pa	1 millibar	= 100.00* Pa
1 cm Hg(0°C)	= 1333 2 Pa	1 newton/cm²	= 10000* Pa
1 centimeter of water (4°C)	= 98.064 Pa	1 poundal/ft²	= 1.4882 Pa
1 dyne/cm²	= 0.100 00* Pa	1 lbf/ft²	= 47.880 Pa
1 foot of water (39.2°F)	= 2989 0 Pa	1 lbf/in.² (psi)	= 6894.8 Pa
1 gf/cm²	= 98 0665* Pa	1 torr (mm Hg, 0°C)	= 133.32 Pa
1 inch of mercury (32°F)	= 3386.4 Pa		

17 Temperature

Degree Celsius	$T_K = t_C + 273.15$	Triple point of natural water	= 273.16 K
Degree Fahrenheit	$t_C = (t_F - 32)/1.8$	Elementary temperature ($c\hbar/r_e k$)	$= 8.1262 \times 10^{11}$ K
Degree Fahrenheit	$T_K = (t_F + 459.67)/1.8$		
Degree Rankine	$T_K = T_R/1.8$	Temperature of 1 eV	= 11 605 K
Kelvin	$t_C = T_K - 273.15$		

18 Time

1 day	= 86400* s	1 year (sidereal)	$= 3.1558 \times 10^{7}$ s
1 day (sidereal)	= 86164 s	1 year (tropical)	$= 3.1557 \times 10^{7}$ s
1 hour	= 3600.0* s	1 atomic second (s_A)	= 9192631770 ^{133}Cs cycles
1 hour (sidereal)	= 3590.2 s		

Table A-4 (*continued*)

1 minute	= 60.000* s	1 atomic unit (1st Bohr orbit/2π) (τ_0)	= 2.4189 × 10^{-17} s
1 second (sidereal)	= 0.99727 s	Jordan's elementary time (r_e/c)	= 9.3996 × 10^{-24} s
1 year (365 days)	= 3.1536* × 10^7 s		

19. Torque

1 dyne·cm	= 1.0000* × 10^{-7} N·m	1 lbf·in	= 0.11298 N·m
1 kgf·m	= 9.8067 N·m	1 lbf·ft	= 1.3558 N·m
1 ozf·in	= 0.0070616 N·m		

20. Velocity

1 ft/s	= 0.30480* m/s	1 AU per year	= 4.7406 km/s
1 in/s	= 0.02540* m/s	1 parsec per year	= 9.7781 × 10^6 m/s
1 km/h	= 0.27778 m/s	Electron in Bohr orbit	= 2.1877 × 10^6 m/s
1 knot (international)	= 0.51444 m/s	1-eV electron	= 5.9309 × 10^5 m/s
1 mi/h (international)	= 0.44704* m/s	Angular velocity of Earth on its axis	= 7.2921 × 10^{-5} rad/s
1 mi/s (international)	= 1609.3 m/s	Mean angular velocity of Earth in its orbit	= 1.9910 × 10^{-7} rad/s
1 mi/h (international)	= 1.6093 km/h		
Velocity of light (c)	= 2.9979 × 10^8 m/s		

21. Viscosity

1 centipoise	= 1.0000* × 10^{-3} Pa·s	1 lb/ft·s	= 1.4882 Pa·s
1 centistokes	= 1.0000* × 10^{-6} m^2/s	1 lbf·s/ft^2	= 47.880 Pa·s
1 ft^2/s	= 0.092030 m^2/s	1 lbf·s/in^2	= 6894.8 Pa·s
1 poise	= 0.10000* Pa·s	1 rhe	= 10.000* (Pa·s)$^{-1}$
1 poundal·s/ft^2	= 1.4882 Pa·s	1 slug/ft·s	= 47.880 Pa·s
1 lb/ft·h	= 4.1338 × 10^{-4} Pa·s	1 stokes	= 1.0000* × 10^{-4} m^2/s

22. Volume

1 acre-foot	= 1233.5 m^3	1 ft^3	= 0.028 317 m^3
1 barrel (oil, 42 gal)	= 0.15899 m^3	1 gallon (Canadian liquid)	= 4.5461 × 10^{-3} m^3
1 barrel (bbl), liquid	= 31 to 42 gallons	1 gallon (U.K. liquid)	= 4.5461 × 10^{-3} m^3
1 barrel (bbl), standard for fruits, vegetables and other dry commodities except cranberries	= 7056 cubic inches = 105 dry quarts = 3.281 bushels, struck measure	1 gallon (U.S. dry) 1 gallon (U.S. liquid) 1 gill (U.K.) 1 gill (U.S.)	= 4.4049 × 10^{-3} m^3 = 3.7854 × 10^{-3} m^3 = 1.4207 × 10^{-4} m^3 = 1.1829 × 10^{-4} m^3
1 barrel (bbl), standard, cranberry	= 5826 cubic inches = $86\frac{45}{64}$ dry quarts = 2.709 bushels, struck measure	1 in^3 1 liter 1 ounce (U.K. fluid) 1 ounce (U.S. fluid)	= 1.6387 × 10^{-5} m^3 = 1.0000* × 10^{-3} m^3 = 2.8413 × 10^{-5} m^3 = 2.9574 × 10^{-5} m^3
1 board foot	= 2.3597 × 10^{-3} m^3	1 cubic parsec	= 2.9380 × 10^{+99} m^3
1 bushel (U.S.)	= 3.5239 × 10^{-2} m^3	1 peck (U.S.)	= 8.8098 × 10^{-3} m^3
1 bushel (bu), struck measure (U.S.)	= 2150.42* in^3	1 pint (U.S. dry) 1 pint (U.S. liquid)	= 5.5061 × 10^{-4} m^3 = 4.7318 × 10^{-4} m^3
1 bushel, heaped (U.S.)	= 1.278 bushels, struck measure	1 quart (U.S. dry) 1 quart (U.S. liquid)	= 1.1012 × 10^{-3} m^3 = 9.4635 × 10^{-4} m^3
1 bushel (bu), struck measure (British Imperial)	= 1.032 U.S. bushels, struck measure	1 stere Solar volume ($4\pi R^3/3$)	= 1.0000* m^3 = 1.4122 × 10^{27} m^3
1 cord (cd) (firewood)	= 128* ft^3	1 tablespoon	= 1.4787 × 10^{-5} m^3
1 cup (measuring)	= 2.3659 × 10^{-4} m^3	1 teaspoon	= 4.9289 × 10^{-6} m^3
1 dram (U.S. fluid)	= 3.6967 × 10^{-6} m^3	1 ton (register)	= 2.8317 m^3
1 drachm (U.K. fluid)	= 3.5516 × 10^{-6} m^3	1 yd^3	= 0.764 55 m^2
1 fluid ounce (U.S.)	= 2.9574 × 10^{-5} m^3		

Table A-5. Precise physical constants [1]

Quantity	Symbol (expression)	Value in SI (cgs) units	Error (ppm)
1 Speed of light in vacuum	c	2 997 924 58 × 10^{8} m·s^{-1} (10^{10} cm·sec^{-1})	0.004
2 Elementary charge	e	1 602 189 2 × 10^{-19} C (10^{-20} emu)	2.9
		(4 803 242 × 10^{-10} esu)	2.9
3 Planck's constant	h	6 626 176 × 10^{-34} J·s (10^{-27} erg·sec)	5.4
	$\hbar = h/2\pi$	1 054 588 7 × 10^{-34} J·s (10^{-27} erg·sec)	5.4
4 Electron rest mass	m_e	0 910 953 4 × 10^{-30} kg (10^{-27} gm)	5.1
5 Avogadro constant	N_A	6 022045 × 10^{23} mol^{-1} (10^{23} mol^{-1})	5.1
recent value		6 022 097 8 × 10^{23} mol^{-1} (10^{23} mol^{-1})	1.0
6 Molar gas constant	R	8 314 41 × 10^{0} J·mol^{-1}·K^{-1} (10^{7} erg·mol^{-1}·K^{-1})	31
7 Boltzmann constant	$k = R/N_A$	1 380 662 × 10^{-23} J·K^{-1} (10^{-16} erg·K^{-1})	32
8 Gravitational constant	G	6 672 0 × 10^{-11} N·m^2·kg^{-2} (10^{-8} dyn·cm^2·gm^{-2})	615
9 Molar volume (273.15 °K, p = 1 atm)	$V_m = RT_0/p_0$	22.413 83 × 10^{-3} m^3·mol^{-1} (10^{3} cm^3·mol^{-1})	31
10 Faraday constant	$F = N_A e$	9 648 456 × 10^{4} C·mol^{-1} (10^{3} emu·mol^{-1})	2.8
11 Rydberg constant	$R_\infty = [4\pi\varepsilon_0]^{-2}(m_e e^4/4\pi\hbar^3 c)$	1 097 373 177 × 10^{7} m^{-1} (10^{5} cm^{-1})	0.07
recent value		1 097 373 147 6 × 10^{7} m^{-1} (10^{5} cm^{-1})	0.0003
12 Fine structure constant	$\alpha^{-1} = [4\pi\varepsilon_0](\hbar c/e^2)$	137 036 04	0.11
recent value		137 035 963	
13 Classical electron radius	$r_e = [4\pi\varepsilon_0]^{-1}(e^2/m_e c^2)$	2 817 938 0 × 10^{-15} m (10^{-13} cm)	2.5
14 Specific electron charge	e/m_e	1 758 804 7 × 10^{11} C·kg^{-1} (10^{7} emu·gm^{-1})	2.8
15 Electron Compton wavelength	$\lambda_c = \hbar/m_e c = \alpha^{-1} r_e$	3 861 590 5 × 10^{-13} m (10^{-11} cm)	1.6
16 Bohr radius	$a_0 = \alpha^{-2} r_e$	0 529 177 06 × 10^{-10} m (10^{-8} cm)	0.82
17 Magnetic flux quantum	$\phi_0 = [c]^{-1}(hc/2e)$	2 067 850 6 × 10^{-15} T·m^2 (10^{-7} Gs·cm^2)	2.6
	h/e	4 135 701 × 10^{-15} J·s·C^{-1} (10^{-7} erg·sec·emu^{-1})	2.6
18 Quantum of circulation	$h/2m_e$	3 636 945 5 × 10^{-4} J·Hz^{-1}·kg^{-1} (10^{0} erg·sec·gm^{-1})	1.0
19 Atomic mass unit	1 u = gm·mol^{-1}/N_A	1 660 565 5 × 10^{-27} kg (10^{-24} gm)	5.1
20 Proton rest mass	m_p	1 672 648 5 × 10^{-27} kg (10^{-24} gm)	5.1
		1 007 276 470 u (amu)	0.011
	m_p/m_e	1836 151 52	0.38
21 Neutron rest mass	m_n	1 674 954 3 × 10^{-27} kg (10^{-24} gm)	5.1
		1 008 665 012 u (amu)	0.027
22 Electron g factor	$\frac{1}{2}g_e = \mu_e/\mu_B$	1 001 159 656 7	0.0035
recent value		1 001 159 652 200	0.0004
23 Bohr magneton	$\mu_B = [c](e\hbar/2m_e c)$	9 274 078 × 10^{-24} J·T^{-1} (10^{-21} erg·Gs^{-1})	3.9
24 Nuclear magneton	$\mu_N = [c](e\hbar/2m_p c)$	5 050 824 × 10^{-27} J·T^{-1} (10^{-24} erg·Gs^{-1})	3.9
25 Electron magnetic moment	μ_e	9 284 832 × 10^{-24} J·T^{-1} (10^{-21} erg·Gs^{-1})	3.9
26 Proton magnetic moment	μ_p	1 410 617 1 × 10^{-26} J·T^{-1} (10^{-23} erg·Gs^{-1})	3.9
	μ_e/μ_p	658 210 688 0	0.010
27 Proton gyromagnetic ratio	γ_p	2 675 198 7 × 10^{8} rad·s^{-1}·T^{-1} (10^{4} rad·sec^{-1}·Gs^{-1})	2.8
28 Stefan-Boltzmann constant	$\sigma = (\pi^2/60)k^4\hbar^3c^2$	5 670 32 × 10^{-8} W·m^{-2}·K^{-4} (10^{-5} erg·sec^{-1}·cm^{-2}·K^{-4})	125
29 First radiation constant	$c_1 = 2\pi hc^2$	3 741 832 × 10^{-16} W·m^2 (10^{-5} erg·sec^{-1}·cm^2)	5.4
30 Second radiation constant	$c_2 = hc/k$	1 438 786 × 10^{-2} m·K (10^{0} cm·°K)	31

Energy equivalents

Quantity	Symbol (expression)	Value	Error (ppm)
Atomic mass unit	u	931 501 6 MeV	2.8
Proton mass	m_p	938 279 6 MeV	2.8
Neutron mass	m_n	939 573 1 MeV	2.8
Electron mass	m_e	0 511 003 4 MeV	2.8
Electron volt	1 eV/k	11 604 50 K	31
	1 eV/hc	8 065 479 cm^{-1}	2.6
	1 eV/h	2 417 969 6 × 10^{14} Hz	2.6
	1 eV	1 602 189 2 × 10^{-12} ergs	2.9
Planck's constant	$\hbar$	6 582 173 × 10^{-22} MeV sec	2.6
	$\hbar c$	1 973 285 8 × 10^{-11} MeV cm	2.6
	$(\hbar c)^2$	0 389 385 7 GeV^2 mb	5.2
Rydberg constant	$R_\infty hc$	13 605 804 eV	2.6
Voltage-wavelength product	$V\lambda$	12 398 520 eV Å	2.6
Gas constant	R	1 987 19 cal·mol^{-1}·K^{-1}	31

Table A-6. Mathematical constants.[1]

Constant	Number	Log
π	3 141 592 653 6	0 497 149 872 7
2π	6 283 185 307 2	0 798 179 868 4
4π	12 566 370 614 4	1 099 209 864 0
π^2	9 869 604 401 1	0 994 299 745 4
$\sqrt{\pi}$	1 772 453 850 9	0 248 574 936 3
e	2 718 281 828 5	0 434 294 481 9
mod $M = \log e$	0 434 294 481 9	1 637 784 311 3
$1/M = \ln 10$	2 302 585 093 0	0 362 215 688 7
2	2 000 000 000 0	0 301 029 995 7
$\sqrt{2}$	1 414 213 562 4	0 150 514 997 8
$\sqrt{3}$	1 732 050 807 6	0 238 560 627 4
$\sqrt{10}$	3 162 277 660 2	0 500 000 000 0
$\ln\pi$	1 144 729 885 8	0 058 703 021 2
e^{π}	23 140 692 632 8	1 364 376 353 8
Euler constant γ	0 577 215 664 9	1 761 338 108 8
1 radian	57 295 779 513 1	1 758 122 632 4
	3437 746 770 78	3 536 273 882 8
	206 264 806 25	5 314 425 133 2
1′	0 017 453 292 5	2 241 877 367 6
1″	0 000 290 888 2	4 463 726 117 2
1‴	0 000 004 848 1	6 685 574 866 8

Square degrees on a sphere $129\,600/\pi = 41\,252\,961\,24$
Square degrees in a steradian $32\,400/\pi^2 = 3282\,806\,35$
For Gaussian distribution $(1/\sigma\sqrt{2\pi})\exp(-x^2/2\sigma^2)$
Probable error/standard error $r/\sigma = 0\,674\,489\,750\,2$
Probable error/average error $r/\eta = 0\,845\,347\,539\,4$
$\sigma/\eta = 1\,253\,314\,137$ $\mu = (\pi/\sigma)\sqrt{2} = 0\,476\,936\,276\,2$

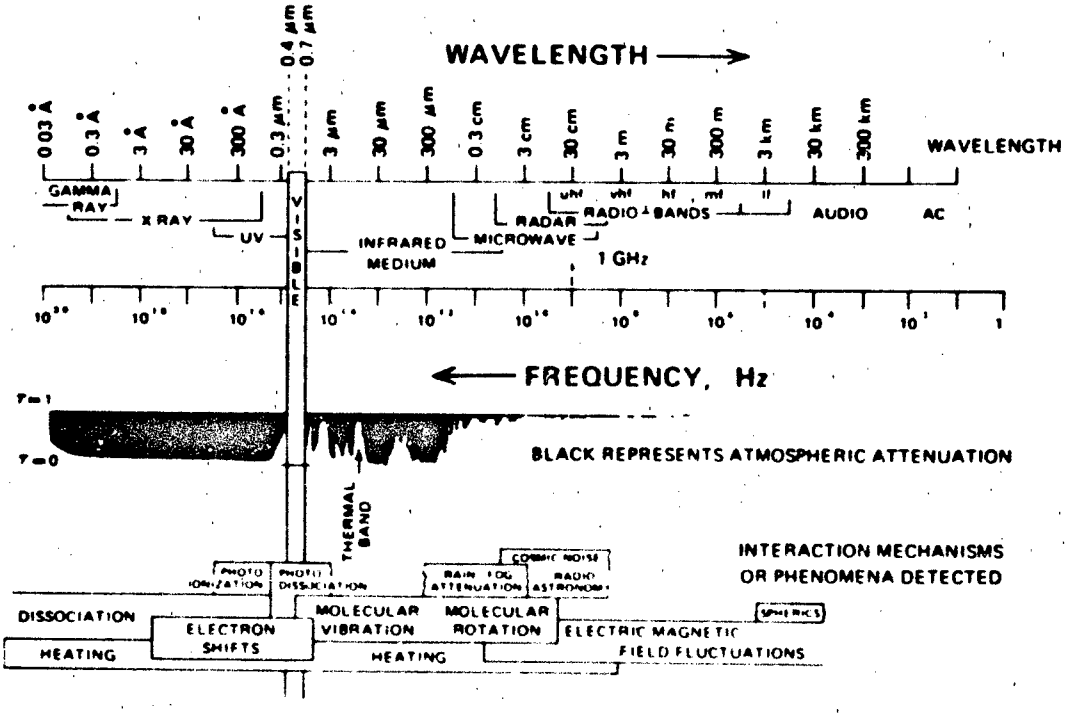


Figure A-1. Electromagnetic spectrum

INDEX

INDEX

INDEX

INDEX

INDEX

INDEX