

AD-A142 441

OBJECTIVE ANALYSIS AND PREDICTION TECHNIQUES - 1983(U)

1/3

SYSTEMS AND APPLIED SCIENCES CORP VIENNA VA

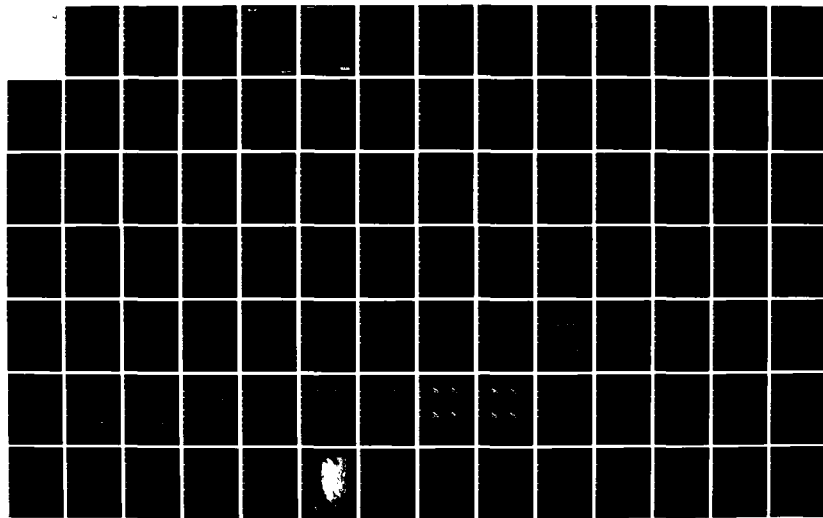
A M GERLACH 30 NOV 83 SCIENTIFIC-3 AFGL-TR-83-0333

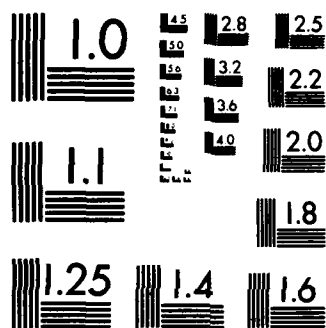
UNCLASSIFIED

F19628-82-C-0023

F/G 4/2

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

AFGL-TR-83-0333

OBJECTIVE ANALYSIS AND PREDICTION TECHNIQUES - 1983

AD-A142 441

Alan M. Gerlach (ed.)

Systems and Applied Sciences Corporation (SASC)  
1577 Springhill Road  
Vienna, VA 22180

REF. FILE COPY

November 30, 1983

Scientific Report No. 3

Approved for public release, distribution unlimited

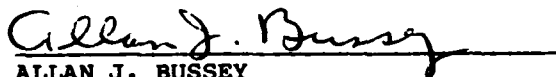
AIR FORCE GEOPHYSICS LABORATORY  
AIR FORCE SYSTEMS COMMAND  
UNITED STATES AIR FORCE  
HANSCOM AFB, MASSACHUSETTS 01731

JUN 22 1984

84 06 22 009

This report has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

"This technical report has been reviewed and is approved for publication"

  
ALLAN J. BUSSEY  
Contract Manager

FOR THE COMMANDER

  
ROBERT A. McCLATCHEY  
Director, Atmospheric Sciences Division

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify AFGL/DAA/LYC Hanscom AFB, MA 01731. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

# UNCLASSIFIED

# AD-A142441

SECURITY CLASSIFICATION OF THIS PAGE

## 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)		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFGL-TR-83-0333										
6a. NAME OF PERFORMING ORGANIZATION Systems and Applied Sciences Corporation (SASC)	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Geophysics Laboratory										
6c. ADDRESS (City, State and ZIP Code) 1577 Springhill Road Vienna, VA 22180		7b. ADDRESS (City, State and ZIP Code) Hanscom AFB, MA 01731 Monitor/LYT/Allan J. Bussey										
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Geophysics Laboratory	8b. OFFICE SYMBOL (If applicable) LY	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER Contract No. F19628-82-C-0023										
8c. ADDRESS (City, State and ZIP Code) Hanscom AFB, MA 01731		10. SOURCE OF FUNDING NOS. <table border="1"><thead><tr><th>PROGRAM ELEMENT NO</th><th>PROJECT NO</th><th>TASK NO</th><th>WORK UNIT NO</th></tr></thead><tbody><tr><td>62101F</td><td>6670</td><td>00</td><td>AC</td></tr></tbody></table>		PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT NO	62101F	6670	00	AC	
PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT NO									
62101F	6670	00	AC									
11. TITLE (Include Security Classification) Objective Analysis and Prediction Techniques - 1983												
12. PERSONAL AUTHOR(S) Gerlach, Alan M. (ed.)												
13a. TYPE OF REPORT Scientific No. 3	13b. TIME COVERED FROM 82/12/1 to 83/11/30	14. DATE OF REPORT (Yr., Mo., Day) 83/11/30	15. PAGE COUNT 219									
16. SUPPLEMENTARY NOTATION												
17. COSATI CODES <table border="1"><thead><tr><th>FIELD</th><th>GROUP</th><th>SUB GR.</th></tr></thead><tbody><tr><td>0401</td><td></td><td></td></tr><tr><td>0402</td><td></td><td></td></tr></tbody></table>		FIELD	GROUP	SUB GR.	0401			0402			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Objective analysis      Interactive forecasting (McIDAS) Initialization      Mesocyclone detection/classification VAS data processing      Satellite cloud analysis (nephanalysis)	
FIELD	GROUP	SUB GR.										
0401												
0402												
19. ABSTRACT (Continue on reverse if necessary and identify by block number) → This report summarizes weather research conducted in several technical areas: in numerical weather prediction, development of a multivariate, global statistical analysis system, generalization of the Machenbauer initialization method, objective analysis with Hough vectors; in mesoscale forecasting, error analysis of LFM-II forecasts, development of forecast guidance and enhanced manually digitized radar (MDR) displays; in boundary layer meteorology, specialized computer programs for studies of wind shear, atmospheric attenuation effects on remote sensing, atmospheric chemical dispersion modeling, fog modeling, microwave refractive structure determination; in radar meteorology, study of single-Doppler detection of synoptic-scale wind, relative storm motions, detection of gust fronts, sub-storm scale motions, development of mesocyclone detection and classification algorithm, construction of Co-Doppler lidar; in satellite meteorology, development of VAS data processing software, implementation on McIDAS of Air Force operational nephanalysis, study correlating MDR data												
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED										
22a. NAME OF RESPONSIBLE INDIVIDUAL Allan J. Bussey		22b. TELEPHONE NUMBER (Include Area Code) (617) 861-2977	22c. OFFICE SYMBOL LY									

# UNCLASSIFIED

**UNCLASSIFIED**

SECURITY CLASSIFICATION OF THIS PAGE

Box 18

Doppler radar wind detection  
lidar

Box 19

with satellite imagery, development of microprocessor-based satellite data ingest system, maintenance and operation of AFGL McIDAS.



A-1

**UNCLASSIFIED**

SECURITY CLASSIFICATION OF THIS PAGE

## FOREWORD

This interim scientific report details the results and status of research performed by Systems and Applied Sciences Corporation (SASC) under Contract F19628-82-C-0023 with the Atmospheric Sciences Division, Air Force Geophysics Laboratory, for the period December 1, 1982 to November 30, 1983.

Presentations vary in technical detail according to their subject matter, ranging from narrative descriptions of work accomplished to technical papers.

The contract calls for effort in five discrete areas, reported upon in the monthly R&D Status Reports under these headings:

- Numerical Weather Prediction
- Mesoscale Forecasting
- Boundary Layer Meteorology
- Radar Meteorology
- Satellite Meteorology.

This report is organized similarly.

Personnel associated with these areas at any time during the reporting period were:

Numerical Weather Prediction	- Donald C. Norquist, M.S.
	Isidore Halberstam, Ph.D.
	Shu-Lin Tung, M.S.
	Chris Johnson, B.S.
	Joan-Marie Freni, B.A.
Mesoscale Forecasting	- Randy Schechter, M.S.
	Charles F. Ivaldi, Jr., B.S.
	Michael E. Niedzielski, B.S.
	James H. Willand
Boundary Layer Meteorology	- Joan M. Ward, B.A.
	Joan-Marie Freni, B.A.
	James H. Willand
	Clyde L. Lawrance
Radar Meteorology	- F. Ian Harris, Ph.D.
	Glenn R. Smythe, M.S.
	James G. Wieler, M.S.
	Ralph J. Donaldson, Jr., S.M.
	Alison L. Godfrey, B.S.

Satellite Meteorology

- Michael E. Niedzielski, B.S.  
Gary B. Gustafson, B.S.  
Randy Schechter, M.S.  
Charles F. Ivaldi, Jr., B.S.  
Barry A. Mareiro

Principal Investigator was Alan M. Gerlach, Ph. D.

Subcontracts with Harris Corporation Computer Systems Division and Perkin-Elmer Customer Service Division provided on-site maintenance for the McIDAS and Weather Radar Data Processing System, respectively.

Reports were prepared by the scientists, engineers, and mathematicians identified in the Table of Contents.

## TABLE OF CONTENTS

REPORT DOCUMENTATION PAGE	1
FOREWORD	3
TECHNICAL REPORTS	9
I. NUMERICAL WEATHER PREDICTION	9
A. Development and Testing of a Multivariate, Global Statistical Analysis System - D. C. Norquist	10
1. Introduction	10
2. Surface Pressure Analysis	10
3. Upper Air Analysis	20
a. Type 1 - Conventional Upper Air Data	22
b. Type 2 - Aircraft Data	23
c. Type 4 - Satellite Sounding Data	24
d. Type 6a - Satellite Temperature and Wind Data	28
4. Experimental Method	38
5. Improvements Proposed for the ASAP Codes	43
Appendix I-A. NMC Subroutine PTOSIG	44
Appendix I-B. NMC Subroutine GETPS	47
B. Generalization and Applications of Machenhauer Initialization Scheme - S.-L. Tung	49
C. Objective Analysis with Hough Vectors - I. Halberstam, S.-L. Tung	51
II. MESOSCALE FORECASTING	55
A. An Error Analysis of LFM-II Forecasts During the 1982-83 Winter Season - R. Schechter	56
1. Introduction	56
2. Data Collection Procedures	57
3. Results	61
4. Discussion	71
5. Conclusions	80
B. Forecast Guidance Displays - R. Schechter	82
C. Mesoscale Forecasting Experiment IV - R. Schechter	87
D. Enhanced Manually Digitized Radar (MDR) Displays - C. F. Ivaldi, Jr.	89
Appendix II-A. Documentation of Command Option in GRDMDR	94

III. BOUNDARY LAYER METEOROLOGY	96
A. Wind Shear Research - J. H. Willand	97
1. Shear Specification Study	97
2. Acoustic Doppler Sounder Data Base	103
B. Atmospheric Attenuation Effects on Remote Sensing - J. H. Willand	105
C. Atmospheric Chemical Dispersion Modeling - J. M. Ward	110
D. Fog Prediction Modeling - J. M. Ward	112
E. Weather Instrumentation Evaluation - J. M. Ward	115
F. WFT Maintenance and Operation - C. L. Lawrance	116
G. Microwave Refractive Structure of the Lower Atmosphere - J.-M. Freni	117
IV. RADAR METEOROLOGY	121
A. Single-Doppler Radar Detection of Synoptic-Scale Wind Anomalies - R. J. Donaldson, Jr.	122
B. Study of Relative Storm Motions - R. J. Donaldson, Jr., F. I. Harris	127
C. Radar Detection of Gust Fronts - G. R. Smythe, F. I. Harris	130
D. Mesocyclone Detection and Classification Algorithm - J. G. Wieler	132
1. Introduction	132
2. The Detection Algorithm	132
3. Preliminary Results	135
4. Conclusions	135
E. Sub-Storm Scale Motions from Single Radar - G. R. Smythe, F. I. Harris	138
F. Systems Support - A. L. Godfrey, F. I. Harris	139
1. Chromatics-Related Software Development	139
2. Radar Evaluation Software	140
3. Computer Operations	143
G. CO <sub>2</sub> Doppler Lidar - S. B. Alejandro, F. I. Harris	144
1. Introduction	144
2. Trailer Preparation	144
3. CO <sub>2</sub> Laser System	144
4. CO <sub>2</sub> Lidar Milestones	145
5. Lidar Beam Modeling	145

a.	Introduction	145
b.	Theory	146
c.	Computational Method	147
d.	Computational Sensitivity	147
1)	Resolution in $r$ for Integration	148
2)	Resolution in $\sigma$	148
e.	Results	148
f.	Circular Versus Annular	149
g.	Truncation Effects	149
h.	Summary	158
V.	SATELLITE METEOROLOGY	165
A.	VAS Data Processing Software - M. E. Niedzielski	166
B.	Satellite-Manually Digitized Radar Feasibility Study - C. F. Ivaldi, Jr.	172
	Appendix V-A. Command Documentation for Program FRQDIS	184
	Appendix V-B. Command Documentation for Program FRQHIS	187
	Appendix V-C. Command Documentation for Program PRBDIS	188
C.	Experimental Cloud Analysis Program - G. B. Gustafson	189
1.	Introduction	189
2.	RDNEPH	189
3.	Algorithm Evaluation	192
4.	Sharpness	197
5.	Accuracy	199
D.	Microprocessor-Based Satellite Data Ingest System - M. E. Niedzielski	201
1.	Introduction	201
2.	Software System Specifications	201
3.	Data Base Management	202
4.	Scheduler Routine	203
5.	Terminal Communications	204
6.	Summary	205
	Appendix V-D. Data File Structures Description	206
E.	McIDAS Software Enhancement - M. E. Niedzielski, C. F. Ivaldi, Jr., R. Schechter	210
1.	Three-Dimensional Graphics	210
2.	Image Frame Identification	210
3.	Graphics Drawing	210

4. Software Documentation	211
5. Map Specification Compatibility in Nowcasting	211
6. Data Plotter Enhancement	212
7. Input Parameter Processing	212
Appendix V-E. Documentation of Three-Dimensional Drawing for Program ANNOT	214
Appendix V-F. Command Documentation for Program LISENT	215
F. McIDAS Operations and Enhancement - B. A. Mareiro	216
1. Northeast Area Remote Sensing System (NEARSS) Development	216
2. McIDAS Enhancement	216
3. McIDAS Operation and Maintenance	217

# I. NUMERICAL WEATHER PREDICTION

## A. Development and Testing of a Multivariate, Global Statistical Analysis System

### 1. Introduction

The AFGL global modeling group recently conducted successful testing of a variable resolution global spectral numerical weather prediction model. In conjunction with the development of the AFGL global spectral model (GSM) (Brenner et al., 1982<sup>1</sup>), two objective analysis methods were developed to provide initial fields of the model's prognostic variables. In addition to a method (described in Section B, below) that attempts to fit data to basis functions, a statistical scheme based on the method of optimum interpolation (OI) was developed by SASC. The model and a normal mode initialization (NMI) procedure will be tested in data assimilation cycles using both analysis procedures in order to examine the relative advantages and shortcomings of each. This section details the formulation of the statistical surface pressure and upper air analyses, which are designated as the AFGL Statistical Analysis Programs (ASAP). Methods used to prepare various types of observed data for input to the ASAP codes are discussed, and results of using the procedures to update a forecast field from the GSM are discussed.

Multivariate optimum interpolation procedures formed the basis for both the surface pressure analysis and the upper air analysis. The surface pressure analysis scheme is based on the method described by McPherson et al. (1979),<sup>2</sup> while the procedure used in formulating the upper air analysis follows the method outlined by Bergman (1979).<sup>3</sup> The following will outline the respective analysis codes designed in this project.

### 2. Surface Pressure Analysis

The surface pressure analysis scheme is multivariate in surface pressure and winds. Because of greater frictional effects over land, the scheme uses no land-based wind observations. Also, because of the geostrophic nature

1. Brenner, S., C.H. Yang, and S.Y.K. Yee, 1982: The AFGL Spectral Model of the Moist Global Atmosphere: Documentation of the Baseline Version. AFGL-TR-82-0393, Air Force Geophysics Laboratory, Hanscom AFB, MA.

2. McPherson, R.D., K.H. Bergman, R.E. Kistler, G.E. Rasch, and D.S. Gordon, 1979: The NMC operational global data assimilation system. Mon. Wea. Rev., 107, 1445-1461.

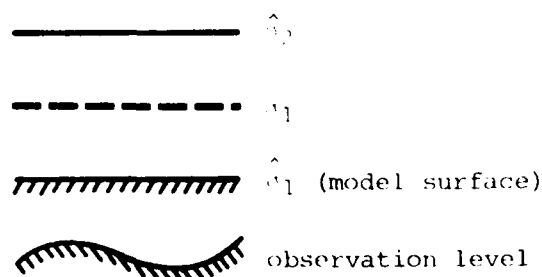
3. Bergman, K.H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. Mon. Wea. Rev., 107, 1423-1444.

of the method's structure functions, it does not use any wind observations that lie within  $20^{\circ}$  latitude of the equator. The geostrophic relationship relates the pressure and wind fields, and the statistical structure functions used in the optimum interpolation procedure are identical with the structure functions used for height and winds in Bergman (1979).<sup>3</sup> The only exception is the use of a height-difference vertical structure function in place of the pressure-difference function used in the temperature and wind analysis. The surface pressure forecast fields from the GSM in spectral coefficients (using a rhomboidal truncation J) provide the first guess, and various types of surface pressure and wind observations are used in the optimum interpolation to correct the first guess surface pressure values at regularly spaced grid intervals (obtained by evaluating surface pressure from predicted spectral coefficients at the grid point sites).

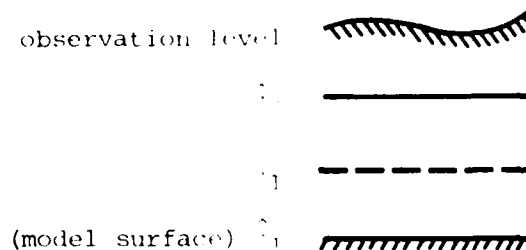
The first step in the surface pressure analysis is to obtain the observation-minus-first guess "residuals" at the location of each of the observations in the data set. This requires obtaining representations of the observations on the model terrain surface (the surface on which the first guess pressures lie) at their latitude, longitude location, as well as representation of the first guess pressure and winds at the observation sites. The latter values are obtained by evaluating surface pressure, wind divergence, and vorticity in the lowest model layer of observation sites using the predicted spectral coefficients valid at the analysis time. No attempt is made to vertically interpolate the first guess wind components (u,v) to the model terrain surface - they are corrected for friction only using the method described by Bergman (1979).<sup>3</sup> The following procedure used to reduce observed surface pressure to the model surface approximately preserves the horizontal pressure gradient; thus, the observed surface winds are changed only to correct for frictional effects.

The procedure for obtaining an estimate of observed surface pressure on the model surface is the same as the procedure used in the National Meteorological Center's (NMC) surface pressure analysis. The procedure requires a knowledge of the pressure and temperature at a known height at or near the surface. This observation level may lie above or below the level of the model surface. The two cases are illustrated as follows:

## Case I



## Case II



In both cases the hydrostatic equation in the form

$$p(\text{m.s.}) = p(\text{o.l.}) \exp \left\{ -g[z(\text{m.s.}) - z(\text{o.l.})] / RT \right\} \quad (1)$$

is used to obtain a representation of observed surface pressure at the model surface, where  $g$  is the acceleration due to gravity and  $R$  is the gas constant for dry air. Given values of pressure  $p(\text{o.l.})$  and temperature  $T(\text{o.l.})$  at a known elevation  $z(\text{o.l.})$ , and using the known terrain height  $z(\text{m.s.})$  on the model surface (obtained from evaluating spectral coefficients of model terrain as used in the forecast model at the observation site), it remains to obtain a value for the mean temperature between the model surface and the observation level. This is done as follows:

Case I: A linear interpolation in  $z$  is performed on temperature between the observation level and the lowest model sigma layer ( $\sigma_1$ ) to obtain a temperature differential between the observation level and the model surface:

$$\Delta T_{\text{o.l.} - \text{m.s.}} = \left[ T(\sigma_1) - T(\text{o.l.}) \right] \frac{z(\text{m.s.}) - z(\text{o.l.})}{z(\sigma_1) - z(\text{o.l.})}$$

where for this calculation  $z(\sigma_1) = \left[ \Delta z(\sigma_1)/2 \right] + z(\text{m.s.})$  where

$$\Delta z(\sigma_1) = T(\sigma_1) \cdot (R/g) \ln \left[ \sigma_2^{-1} \right]$$

since  $p_{f.g.}(\text{m.s.}) / p_{f.g.}(\sigma_2) = \sigma_2^{-1}$ , using the definition of  $\sigma(p) = p/p_{\text{sfc}}$ .

Half of this temperature differential is added to the observation level temperature to get an estimate for  $\bar{T}$ .

Case II: A value for  $T(m.s.)$  is estimated by assuming that a standard atmosphere lapse rate ( $.0065^{\circ}\text{K/m}$ ) exists between the first sigma layer and the model surface:

$$T(m.s.) = T(\sigma_1) + .0065 [z(\sigma_1) - z(m.s.)]$$

where  $z(\sigma_1)$  is obtained in the same way as in Case I. Then  $\bar{T}$  is half of the sum of  $T(o.l.)$  and  $T(m.s.)$ .

A possible drawback of using this method for obtaining  $p(m.s.)$  is that both observed and first guess values are used together to obtain a representation of observed pressure at the model terrain level. By using first guess values in this calculation, the observed pressure value that is to be used to correct the first guess surface pressure has been biased toward that value already. This could mean that the full impact of the correction from the observations would not be felt on the model first guess values. This is an area where further study is called for to come up with a more suitable alternative for obtaining  $p(m.s.)$ .

In many types of surface observations the pressure  $p(o.l.)$  may be reported at a level  $z(o.l.)$  different from the station level where the temperature  $T_{sta}$  was measured. For example, the First Global Atmospheric Research Program (GARP) Global Experiment (FGGE) observation sets<sup>4</sup> used in this study contain a parameter that might be referred to as PZ, along with a code KPCI, to indicate the observed pressure/height, as in Table 1.

TABLE 1. SURFACE PRESSURE-HEIGHT OBSERVATION PARAMETER

KPCI	Description (Contents of PZ)
0	sea level pressure
1	station pressure
2	pressure at 1000 m
3	pressure at 2000 m
4	pressure at 3000 m
5	pressure at 4000 m
6	geopotential height at 850 mb
7	geopotential height at 700 mb
8	geopotential height at 500 mb
9	unknown

4. Obtained from Department of the Air Force, OL-A, USAF Environmental Technical Applications Center (MAC), Federal Building, Asheville, NC 28801.

In order to obtain a pressure at the height of the model terrain height  $z$ (m.s.) as described above at a particular observation site, values of observed (or representations of observed)  $p$ ,  $z$ , and  $T$  must be available all at the same level. These values will be designated by the subscript "o.l." For all of the values of KPCI shown above (except KPCI = 9, in which case the observation is rejected), the values of  $p_{o.l.}$  and  $z_{o.l.}$  are known immediately. For KPCI = 1, the observation level is just the actual elevation of the station above sea level, so  $T_{o.l.}$  is the reported temperature. In all the other cases, the reported observation level represented by PZ is potentially at a different altitude from the actual station elevation (EL) where observed temperature was measured. For these cases, a representation for  $T_{o.l.}$  must be derived.

Case I.  $z_{o.l.} < EL$  (PZ observation level is below the ground)

In this case, the NMC "Shuell" method is used to obtain the temperature value that is used to reduce station pressure at elevation EL to a pressure at a desired level  $z_{o.l.}$  below the ground surface. The most common example of this is the case of pressure reduced to mean sea level ( $z_{o.l.} = 0$ ). Setting  $p_{o.l.}$  and  $z_{o.l.}$  from the given information, and representing the given temperature and station elevation (EL) by  $T_{sta}$  and  $z_{sta}$ , the first step is to extrapolate temperature downward using the standard atmosphere lapse rate from  $z_{sta}$  to  $z_{o.l.}$ :

$$T'_{o.l.} = T_{sta} + (-.0065) (z_{o.l.} - z_{sta}).$$

The  $T'_{o.l.}$  is modified as follows: if  $T_{sta} > 290.66^{\circ}K$ ,  $T'_{o.l.} = 290.66 + .005 \times (T_{sta} - 290.66)^2$ ; if  $T_{sta} \leq 290.66$ ,  $T'_{o.l.} = 290.66$ . Next, a lapse rate  $\gamma$  is calculated as follows:

$$\gamma = \frac{T_{sta} - T'_{o.l.}}{z_{sta} - z_{o.l.}} \text{ when } z_{sta} - z_{o.l.} > 75 \text{ m,}$$

$$\gamma = 0 \text{ when } z_{sta} - z_{o.l.} \leq 75 \text{ m.}$$

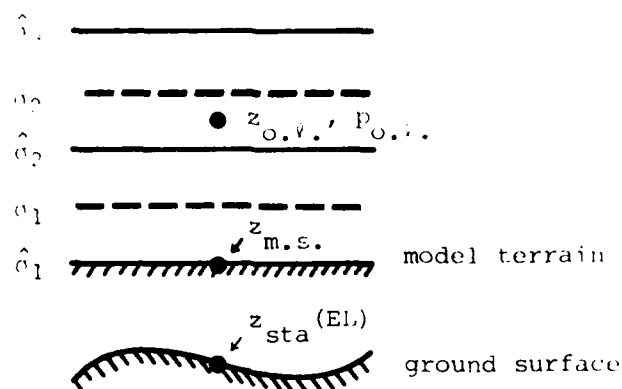
Next, we use  $\gamma$  to obtain a value for  $T_{o.l.}$ :

$$T_{o.l.} = T_{sta} + \gamma(z_{o.l.} - z_{sta}).$$

Notice that if  $z_{sta} - z_{o.l.} > 75$  m,  $T_{o.l.} = T'_{o.l.}$ , and if  $z_{sta} - z_{o.l.} \leq 75$  m,  $T_{o.l.} = T_{sta}$ .

Case II.  $z_{o.l.} \geq EL$  (PZ observation level is at or above the ground)

Consider the following diagram:



In this case, the values  $z_{o.l.}$ ,  $p_{o.l.}$ ,  $T_{sta}$ , and  $z_{sta}$  (EL) are given, and we wish to find a value  $T_{o.l.}$  corresponding to  $z_{o.l.}$ ,  $p_{o.l.}$ . The height  $z_{\sigma_k}$  of the first model layer that is greater than or equal to  $z_{o.l.}$  is estimated as follows:

For  $k = 1$ :  $z_{\sigma_1} = 0.5(\Delta z_{\sigma_1}^{\wedge} + z_{m.s.})$  where  $\Delta z_{\sigma_1}^{\wedge} = T_{\sigma_1} (R/g) \ln (\sigma_2)^{-1}$

For  $k > 1$ :  $z_{\sigma_k} = z_{\sigma_{k-1}} + \Delta z_{\sigma_k}^{\wedge}$  where  $\Delta z_{\sigma_k}^{\wedge} = 0.5(T_{\sigma_k} + T_{\sigma_{k-1}}) (R/g) \ln (\sigma_{k-1}/\sigma_k)$ .

When  $z_{\sigma_k} \geq z_{o.l.}$  is found, a check is made to determine the altitude difference between  $z_{\sigma_k}$  and  $z_{sta}$ . If  $z_{\sigma_k} - z_{sta} \leq 1000$  m, a linear interpolation in height is performed between  $z_{\sigma_k}$  and  $z_{sta}$  to obtain  $T_{o.l.}$  from  $T_{sta}$ :

$$T_{o.l.} = T_{sta} + \frac{T_{\sigma_k} - T_{sta}}{z_{\sigma_k} - z_{sta}} (z_{o.l.} - z_{sta}).$$

If  $z_{\sigma_k} - z_{sta} > 1000$  m, the linear interpolation is carried out between  $z_{\sigma_k}$  and  $z_{\sigma_{k-1}}$ :

$$T_{O,\ell} = T_{O,k-1} + \frac{T_{O,k} - T_{O,k-1}}{z_{O,k} - z_{O,k-1}} (z_{O,\ell} - z_{O,k-1}).$$

In both cases detailed above, the  $T_{O,\ell}$  value obtained is used to solve Eq. (1) for  $p(m.s.)$ . Here again, a mixture of observation and first guess values is used to obtain an observation level temperature that is supposed to be representative of the observed temperature only. This may further contribute to the bias of the observed value of pressure at the model terrain surface toward the first guess surface pressure value. A better method of obtaining  $T_{O,\ell}$  should be included in the study of a more suitable approach for determining a value for  $p(m.s.)$ .

Residuals of surface pressure and zonal and meridional wind components on the model terrain surface at all observation locations are calculated, and then submitted to a check against gross error limits. Currently, the limits are set at 50 mb in pressure and  $20 \text{ ms}^{-1}$  in the wind components. Observations in which either PZ or T is missing are discarded, as are observations that are outside the  $\pm 3$  hour window about the analysis time. In cases of observations at different times at a given location, only the observation taken nearest in time to the time of the analysis is accepted. After all residuals are calculated, the observation sites where the residuals are located are grouped around each grid point of the analysis grid. The analysis grid for both the surface pressure analysis and the upper air analysis is set by choosing the smallest even integer number of Gaussian latitudes that is greater than  $(4J + 1)/2$ , where  $J$  is the rhomboidal truncation (highest wave number) of the forecast model for which the analysis is being prepared. At each of the Gaussian latitudes,  $2(J + 1)$  is chosen as the number of evenly spaced longitude points. In the present study,  $J = 30$  so that a  $62 \times 62$  analysis grid was used.

The grouping around each grid point begins by finding all observation sites lying within a distance equivalent to  $2.5^\circ$  of latitude of the grid point. If less than ten observations are found in this first scan, the scan radius is increased in  $2.5^\circ$  latitude-distance equivalent increments until at least ten observations are found. If after  $15^\circ$  scan ten observations are still not found, a check for at least three observations is conducted. If less than three observations are found, the grid point will not receive a correction - it will retain its first guess value. As many as 100

observations are allowed to correct the first guess value, but there are rarely that many observations within  $2.5^{\circ}$  latitude distance of any grid point, so in practice this rarely occurs.

The error flagging procedure for p,u,v is carried out exactly as is discussed in Bergman (1979),<sup>3</sup> using his Eq. (7.1). However, in this case  $a = 4$ ,  $b = 2$  for p and  $a = 8$ ,  $b = 4$  for u,v, and for all three  $\sigma$  is set to a constant value of 1.25. Quality control indicators (QCI) from 0-9 are assigned to p and u,v, using values from the FGGE II-B horizontal quality control indicators and are set as stated in Table 2. The higher the QCI value, the more likely the observation is to receive an error flag.

TABLE 2. HORIZONTAL QUALITY CONTROL INDICATORS

<u>QCI</u>	<u>Description from FGGE II-B Data Set</u>
0	observed value found correct during quality control check
1	quality control check has not been made
2	observed value found erroneous - reconstituted value inserted
3	original value missing - reconstituted value inserted
4	value consistent with present and past observations (ships only)
5	quality control check made - action unrecorded, but likely value entered
6	observed value found suspect during quality control check
7	value inconsistent with present and past observations (ships only)
8	observed value found erroneous during quality control check
9	missing

These quality control indicators are used only to determine which observation gets a flag when the difference in residuals forces the inequality in Bergman's Eq. (7.1) to be violated. Land based observations are assigned quality control values as listed above, and when the inequality is not satisfied the ocean based observations are assigned a quality control ten units higher (lower in quality) than indicated above just for the purpose of determining which observation gets the error flag. In this way, a ship observation can never impose a flag on a land observation. Also, no attempt is made in the current analysis to assign flags indicating that the observation should be kept.

Because of the large number of observations that may be clustered around a particular grid point, the iterative rejection cycle that eliminates observations with the most flags removes not only the observation that has the most flags, but also any other observations that have that same number when their value is divided by ten and the decimal value is dropped. By removing only the observation with the largest number of flags each time, the code was requiring too much time in the error flagging procedure. The observation rejection scheme is carried out separately for pressure and winds, and as in Bergman (1979)<sup>3</sup> all observations flagged by the rejected observation will have that flag removed before the next rejection check begins. When the highest number of flags for any observation is less than ten, the rejection procedure begins removing only the observation with the greatest number of error flags. The iterations continue until all remaining observations have at most one flag, or until ten iterations have been completed. In the latter case, any observations with more than one flag are still considered for potential use in the correction of the grid point value of  $p$ .

The next step is to calculate the  $p - p$ ,  $u - p$ , and  $v - p$  correlations of each remaining observation with the grid point. These correlations  $\rho_{ig}^{pp}$ ,  $\rho_{ig}^{up}$ ,  $\rho_{ig}^{vp}$  are the right-hand-side (RHS) values of Bergman's Eq. (2.12):

$$\sum_{\ell=1}^m \sum_{j=1}^{n_{\ell}} \left( \rho_{ij}^{k\ell} + \eta_{ij}^{k\ell} \epsilon_{ik} \epsilon_{j\ell} \right) a_{j\ell}^{\ell} = \rho_{ig}^{kr}$$

where  $k = p, u, v$  and  $r = p$  only in the surface pressure analysis. The  $\rho$  correlations are a product of a horizontal correlation based on distance and a vertical correlation based on height difference. In the selection process to follow which finds the ten largest correlations, the model terrain height at the grid point and the station elevation at the observation site are used in the vertical correlation:

$$v_{ig}^{uu} = \left\{ 1 + CP_{ss} \cdot CP_w [z_{sta}(i) - z_{m.s.}(q)]^2 \right\}^{-1} / (1 + \epsilon_i^2)$$

where  $v$  is the vertical correlation used in place of the pressure separation form of the vertical correlation,  $CP_{ss} = 1.68 \times 10^{-8} m^{-2}$ ,  $CP_w = 25$  as in the NMC surface pressure analysis, and  $i$  represents the index of the observation that might potentially be used to correct the value of  $p$  at the grid point  $q$ .

The normalized observation errors  $\varepsilon_i$  are assigned using Bergman's Table 2, so they vary with variable and observation types but not in space or time. On the basis of the values computed for the  $\rho$  correlations using this vertical correlation, the ten single observations (considering each of  $p$ ,  $u$ , and  $v$  as separate observations) with the largest values of  $\rho$  are those that are used in the correlation of  $p$  at  $g$ . All of the chosen observations must have correlations greater than 0.1 - if fewer than ten meet this criterion, only those that do are used in the correction.

When the ten or fewer observations with the largest  $\rho$  values are identified, their value of  $\rho$  is recalculated using the vertical correlation value that has the model terrain elevation at both the observation site and the grid point:

$$v_{ig}^{uu} = \left\{ 1 + CP_{ss} \cdot CP_w [z_{m.s.}(i) - z_{m.s.}(g)]^2 \right\}^{-1}.$$

Also, in this recalculation the normalized observation errors are ignored. Then all of the cross-correlations making up the coefficients of  $a'_{jz}$  in Bergman's Eq. (2.12) for  $r = p$  using these selected observations are calculated. The observation error correlations  $\eta$  are not used, because in the current version of the analysis all horizontal and vertical observation error correlations are assumed to be zero. The values of the elements of the (at most) 10 X 10 matrix of coefficients in Bergman's Eq. (2.12) are simply the correlations between the selected observations. This system of equations is solved for the non-dimensional weights  $a'_{jl}$ , which are converted to the coefficients  $a_{ik}$  of Bergman's Eq. (2.5) as follows:

$$a_{ju} = \left( \frac{\overline{p_g^2}}{\overline{u_j^2}} \right)^{1/2} a'_{ju} = \frac{10^{-2} \rho_{g j} (k_p/k_h)^{-1/2}}{G_j} a'_{ju}$$

$$a_{jv} = \left( \frac{\overline{p_g^2}}{\overline{v_j^2}} \right)^{1/2} a'_{jv} = \frac{10^{-2} \rho_{g j} (k_p/k_h)^{-1/2}}{G_j} a'_{jv}$$

$$a_{jp} = \left( \frac{\overline{p_g^2}}{\overline{p_j^2}} \right)^{1/2} a'_{jp} = \frac{\rho_{g j}}{p_j} a'_{jp}$$

where  $\rho_q$  here represents density (calculated using first guess temperature from the lowest two model layers extrapolated to the model surface, along with the uncorrected first guess surface pressure) at the grid point; and  $f_j = 2\omega \sin \phi_j$  is the coriolis factor in which  $\omega$  is the earth's rotation rate and  $\phi_j$  is the latitude of the observation site  $j$ . Values used for  $k_p$  and  $k_h$  in this analysis [see Bergman's Eqs. (A22, A23)] are  $k_p = 5$  (dimensionless) and  $k_h = 1.96 \times 10^{-12} \text{ m}^{-2}$ , and  $G_j$  is the geostrophic decoupling factor:  $G_j = 1 - \exp(-0.05|\phi_j|)$  where  $\phi_j$  is in degrees. In the case of pressure observations correcting grid point pressure, the density is assumed to be essentially spatially invariant, so  $\rho_q = \rho_j$  and no correction is imposed to convert  $a'_{jp}$  to  $a_{jp}$ . The products of the observation-minus-first guess residuals and the  $a_{jp}$  for each of the observations used in the correction are summed to form the correction applied to the first guess surface pressure at  $q$ .

When the correction has occurred at all grid points where a sufficient number of nearby observations were found, the grid point values of updated surface pressure are used to define the new sigma structure for the first guess field. Thus, grid-point first guess values of temperature, winds, and specific humidity at all model layers as defined by the original first guess surface pressure are interpolated linearly in the natural logarithm of pressure to the new model layers as defined by the corrected surface pressure. Actually, this vertical interpolation is not necessary if the first guess values are thought of as being associated with a sigma value rather than a pressure value - they remain at the same sigma layers regardless of the change in surface pressure. This viewpoint will be adopted in future work. After this vertical interpolation is performed, all grid point fields ( $T, u, v, q, p_{sfc}$ , where  $q$  represents specific humidity) are converted to spectral coefficients of temperature, vorticity, divergence, and specific humidity to serve as first guess fields in the upper air analysis.

### 3. Upper Air Analysis

The upper air analysis scheme is multivariate in geopotential height ( $z$ ) and winds. As in the case of the surface pressure analysis, the geostrophic relationships of Bergman relating  $z$ ,  $u$ , and  $v$  are used to construct the statistical structure functions. The update for  $q$  is univariate using

Bergman's temperature structure functions. Various types of upper air observations, such as rawinsondes, pilot balloons, aircraft observations, satellite soundings, and satellite single-level temperatures and winds are used to correct the grid point values of  $z$ ,  $u$ ,  $v$ , and  $q$  obtained from the spectral coefficients. The first guess height values are calculated at layer interfaces (levels) from the first guess temperature both at the grid points and the observation sites to form residuals using:

$$z_{\sigma_{k+1}}^{\wedge} = z_{\sigma_k}^{\wedge} + (R/g) T_{\sigma_k} \ln(\hat{\sigma}_k / \hat{\sigma}_{k+1}); k = 1, K$$

where  $K$  is the number of layers in the model atmosphere. In all cases where this procedure is used,  $\hat{\sigma}_{K+1} = 1.0/p_{sfc}$ , which means that the pressure at the top of the model atmosphere is set to 1.0 mb arbitrarily for this calculation. In the upper air analysis scheme, height values are corrected at sigma levels, whereas wind and specific humidity grid values are corrected in the sigma layer positions.

The first guess representations of  $z$ ,  $u$ ,  $v$ ,  $q$  at the observation sites are obtained through evaluation of the spectral coefficient functions of  $T$ , relative vorticity ( $\xi$ ), divergence ( $d$ ), and  $q$  at all layers at the observation sites, then converting the physical space values of  $T$  to level  $z$  using the above relationship and  $z_{\sigma_1}^{\wedge} = z_{m.s.}$ . Then to calculate the residuals at the sigma layers (and for  $z$ , levels), the observations must be vertically interpolated to the same vertical positions in pressure. The methods used to perform these interpolations vary with observation type. The following paragraphs present a description of the methods used to obtain representations of the observations at model sigma layers and levels. In all cases, the pressure values at the sigma layers and levels result from the chosen sigma structure of the atmosphere and the corrected surface pressure (in spectral coefficients in the form of a physical space value) at the observation location. The basic scheme used in the vertical interpolations for all data types is: (1) obtain representations of layer mean temperatures from the observations, and where possible, values of wind and humidity interpolated to the model layer pressures, and (2) use the local hydrostatic equation (above) to convert layer mean temperatures to level heights. The FGGE II-B data tapes used in this study as a source of global observations

contain many different types of data, but four general types were selected for use in the upper air analysis. Because of the diverse nature of the observational data, the procedure is carried out differently for each data type. The following paragraphs describe the procedures used to secure representations of observational heights at model levels, and observational temperature, winds, and specific humidity for model layers for each data type.

a. Type 1 - Conventional Upper Air Data

The heart of the vertical interpolation process for the upper-air conventional observations is contained in subroutine PTUSIG, which is a routine used in the preprocessing code of the NMC global spectral model. This procedure is detailed in Appendix I-A. Model layer mean temperatures are obtained from sounding values of  $p$ ,  $T$ , and  $z$  by first identifying the two observation levels (having all three parameters  $p$ ,  $T$ , and  $z$  present) lying immediately above and below each of the model levels at the top and bottom of the layer. Assuming that geopotential height varies quadratically in  $\ln p$ , a value for the height is calculated from the surrounding observation levels for both model levels that border the layer. Then using the hydrostatic relation  $\bar{T} = g(\Delta z) / [R \ln (p_l/p_u)]$  where  $p_l$ ,  $p_u$  refer to the level pressures at the bottom and top of the layer respectively, the layer mean temperature  $\bar{T}$  is found.

The two observation levels used to obtain model level height must surround the model level, but must not be separated by more than 300 mb in pressure. This is done for all model levels, including the bottom of the lowest layer. If the pressure of this level (i.e., the updated first guess surface pressure evaluated at the observation site) is greater than the pressure of the lowest observation level containing both  $z$  and  $T$ , no level height is calculated and therefore no layer temperature is calculated for that layer. If the highest complete observation level with values for both  $z$  and  $T$  lies below a particular model level, the observation temperature is used as the layer mean temperature for the layer topped by that level provided that (1) the observation level has a pressure less than or equal to 70 mb, and (2) the pressure of the observation level is no more than ten mb greater than the model layer pressure. The layer mean temperatures derived in this way are used as  $T_{\sigma_k}$  in the local hydrostatic equation where now

$z_{01}^{\wedge} = z_{m.s.}$ . In layers where no observation layer mean temperature is obtainable, the first guess temperature is used in this calculation to preserve the first guess thickness of that layer. Thus the  $z$  residual at the top of that layer would reflect only the differences between observation thickness and first guess thickness of the layers below that level, and would not be due to the thickness of that layer.

The wind components and the specific humidity values are interpolated using a linear  $\ln p$  interpolation between two surrounding observation levels separated by not more than 300 mb. For the sigma layer  $k$ , the interpolation has the form:

$$u_{o_k} = \frac{1}{2} (u_l + u_u) + \frac{u_l - u_u}{\ln p_l - \ln p_u} [\ln p_{o_k} - \frac{1}{2} (\ln p_l + \ln p_u)]$$

where the  $l$  and  $u$  subscripts indicate the observation values immediately below and above the sigma layer pressure  $p_{o_k}$ . For all layers, the two observations used in the interpolation must surround the sigma layer pressure - no extrapolation is allowed. If the sigma layer pressure is less than the pressure level of the highest observed wind, the observation representation for that sigma layer is assigned as missing. This procedure is also followed for specific humidity.

#### b. Type 2 - Aircraft Data

Aircraft observations consist of single point observations of temperature and wind at a known flight altitude. In the FGGE II-B data set, several wind observations may be included in the report, each with its own latitude and longitude location. It was assumed that the altitude applies to all wind observations in the report.

The aircraft reports, although allowing space for it in the record, do not contain a pressure value at flight altitude. Because pressure is the vertical coordinate in the system, this makes it difficult to locate where the observation should be placed with respect to the model atmosphere. NMC has provided formulas which they use to assign a pressure to an aircraft flight altitude, as follows:

$$p(z) = 226.3 \exp [1.576106 \times 10^{-4} (11000 - z)]; \quad z > 11 \text{ km}$$

$$p(z) = 1013.5 [(288 - 0.0065 z)/288]^{5.256}; \quad z \leq 11 \text{ km.}$$

These formulas are apparently based on standard atmosphere formulations. A pressure is assigned to the flight altitude  $z$  in the following way. First, the heights of the model levels surrounding the observation  $z$  are obtained, and the one to which  $z$  is closest is identified (call this  $z_{0k}^*$ ). Then pressures are obtained for both  $z$  and  $z_{0k}^*$  using the same formula above (the larger of  $z$  and  $z_{0k}^*$  is compared with 11,000 m and the appropriate equation is used for both). Then  $p(z_{0k}^*)$  is compared with the model level pressure  $p_{0k}^*$ , and the difference is applied as a correction to  $p(z)$ .

At this point, the model layer pressures surrounding  $p(z)$  are identified, and the first guess temperatures and winds at these layers are linearly interpolated in  $\ln p$  to  $p(z)$ . The observation-minus-first guess residuals individually for temperature and wind components are calculated at this level, and these residuals are added to the model layer first guess values whose pressure is closest to  $p(z)$ . The respective sums constitute an estimate of the observation value in that model layer for temperature and winds. A value for height at the top of that layer is calculated from first guess height at the bottom of that layer and this temperature using the local hydrostatic equation. This same procedure is also carried on for any extra wind observations using the same flight altitude.

#### c. Type 4 - Satellite Sounding Data

Satellite sounding data come in four different forms:

ITYP = 1: layer thickness between reference pressure level and standard pressure level.

ITYP = 2: layer precipitable water between a reference pressure level and a standard pressure level.

ITYP = 3: layer mean temperature between two non-standard pressure surfaces.

ITYP = 4: layer precipitable water between two non-standard pressure surfaces.

For ITYP = 1 and ITYP = 3, the following scheme is used to obtain representations of observed geopotential height on sigma level surfaces.

First, determine the "observation layer" mean temperature. For ITYP = 3, this is given directly. For ITYP = 1, the observation layer mean temperature is given by

$$\bar{T}_O = \frac{g(\Delta z)}{R \ln \left( \frac{p_\ell}{p_u} \right)}$$

where  $\Delta z$  is the (positive) thickness in geopotential meters between the lower (in altitude) pressure surface  $p_\ell$  and upper pressure surface  $p_u$ . Next, once the mean temperature is known for each observation layer, the pressure for the observation layer may be estimated as follows. Using the notation to indicate observation levels, obtain the  $\hat{\sigma}$  value for each observation level from the observation level pressures  $\hat{p}_o$  and the corrected first guess pressure:

$$\hat{\sigma}_k = \frac{\hat{p}_o(k)}{p_{sfc}(f.g.)}$$

which gives the relative  $\sigma$  location of each observation level in the model atmosphere. Then calculate the observation layer  $\sigma$  from

$$\sigma_o(k) = \left[ \frac{\hat{\sigma}_k^{1+\kappa} - \hat{\sigma}_{k+1}^{1+\kappa}}{(1+\kappa)(\hat{\sigma}_k - \hat{\sigma}_{k+1})} \right]^{1/\kappa}$$

where  $\hat{\sigma}_1 = 1.0$ ,  $\hat{\sigma}_N = 0.0$ ,  $\kappa = R/c_p$ , and  $N$  = number of observation layers. From these  $\sigma_o(k)$  we can obtain the observation layer pressures

$$p_o(k) = \sigma_o(k) \cdot p_{sfc}(f.g.)$$

which gives the relative  $\sigma$  location of each observation layer in the model atmosphere. Then use the  $\bar{T}_o$  and  $p_o$  for each observation layer along with the  $p(\sigma_\ell)$ ,  $\ell = 1, ML$ , which are the model layer pressures, to obtain the model layer mean temperatures. This is done by linearly interpolating (in  $\ln p$ ) the observation layer mean temperatures to model layer pressures:

$$\bar{T}(\sigma_\ell) = \frac{\bar{T}_{o_\ell} + \bar{T}_{o_u}}{2} + \frac{\bar{T}_{o_\ell} - \bar{T}_{o_u}}{\ln p_{o_\ell} - \ln p_{o_u}} \left[ \ln p(\sigma_\ell) - \frac{\ln p_{o_\ell} + \ln p_{o_u}}{2} \right]$$

No extrapolation is allowed. Therefore, if  $p(\sigma_\ell)$  for the lowest  $\sigma$  layer has a greater value (lower in altitude) than  $p_o(k=1)$ , the lowest observation layer pressure, then  $\bar{T}(\sigma_\ell)$  is set equal to  $\bar{T}_{o_\ell}$ , the lowest observation layer

mean temperature, and only when  $p(\sigma_\ell)$  is within 50 mb of  $p_0$  ( $k = 1$ ). The same is true at the top of the atmosphere where  $p(\sigma_\ell)$  for the highest model  $\sigma$  layer has a lesser value (higher in altitude) than  $p_0$  ( $k = N$ ), the highest observation layer pressure. A 20 mb interval provision holds at the top.

This scheme yields estimates of model  $\sigma$  layer mean temperatures,  $\bar{T}(\sigma_\ell)$ . These are then used along with the model surface terrain height to obtain estimates of the observational geopotential height at each model  $\sigma$  level using the local hydrostatic equation where  $z_{\sigma_1}^\wedge = z_{m.s.}$ .

If there are no values of  $\bar{T}(\sigma_k)$  for the lower model layers, then  $z_{\sigma_k}^\wedge$  is calculated for the first model layer for which an estimate of  $\bar{T}(\sigma_k)$  is available using first guess values of  $z$  for levels below this layer. Note that since no observational estimates are available for these lower levels, the  $z$  values for these levels will be assigned as missing. For levels above the topmost calculated  $z_{\sigma_k}^\wedge$  level, the existing first guess layer mean temperatures for those layers will be used along with the  $z_{\sigma_k}^\wedge$  just below to obtain the  $z_{\sigma_k}^\wedge$  values above. In this way, the first guess thicknesses are preserved in these layers.

For ITYP = 2 and ITYP = 4, the following scheme is used to obtain representations of observed specific humidity on sigma layers.

First, determine the layer precipitable water for each observation layer. For ITYP = 4, this is given directly. For ITYP = 2, we use the integral nature of the definition of layer precipitable water given by

$$W = \frac{1}{g\rho_0} \int_{p_u}^{p_\ell} q \, dp$$

and note that the integral over the entire series of layers is equal to the sum of the individual layer precipitable water values. Thus, for example,

$$W_{50}^{1000} = W_{850}^{1000} + W_{750}^{850} + W_{500}^{700} + \dots + W_{70}^{100} + W_{50}^{70}.$$

Suppose  $W_{850}^{1000}$  is known, then if  $W_{700}^{1000}$ ,  $W_{500}^{1000}$ , . . . ,  $W_{50}^{1000}$  are given,

$$W_{700}^{850} = W_{700}^{1000} - W_{850}^{1000}$$

$$W_{500}^{700} = W_{500}^{1000} - W_{700}^{1000}$$

$$W_{400}^{500} = W_{400}^{1000} - W_{500}^{1000}$$

⋮

$$W_{50}^{70} = W_{50}^{1000} - W_{70}^{1000}$$

In this way, the layer precipitable water value for each observation layer is known. A value for  $q(p_u)$  at the top of the highest layer is obtained by a linear interpolation in  $\ln p$  on the first guess  $q$  values to  $p_u$ . Then the  $q(p_\ell)$  values at the bottom of this layer and then each successive lower layer are estimated as follows. Assuming that specific humidity varies linearly with  $\ln p$  between two pressure surfaces  $p_\ell$  and  $p_u$  according to

$$q(p) = \frac{q(p_\ell) + q(p_u)}{2} + \frac{q(p_\ell) - q(p_u)}{\ln p_\ell - \ln p_u} \left[ \ln p - \frac{\ln p_\ell + \ln p_u}{2} \right],$$

then if we substitute this expression into the integral definition of  $W$  above, we obtain the following expression for  $q(p_\ell)$ , the specific humidity at the bottom of the layer:

$$q(p_\ell) = \frac{Wg\rho_0 - q(p_u) [\frac{1}{2}(p_\ell - p_u) + A]}{\frac{1}{2}(p_\ell - p_u) - A}$$

where

$$A(p_\ell, p_u) = \frac{p_\ell - p_u}{\ln p_\ell - \ln p_u} - \frac{1}{2} (p_\ell + p_u).$$

Thus, knowing the layer precipitable water (in  $10^{-1}$  m) and the top and bottom pressures of the layer (in mb), and using  $g = 9.8 \text{ ms}^{-2}$  and the density of water  $\rho_0 = 1 \text{ g cm}^{-3}$ , then  $q$  will have units of  $\text{g g}^{-1}$ . If in any lower layer the observed  $W$  value is not available, the  $q_{ob}$  value for the bottom of that layer is then non-available. A first guess  $q$  value is found for that pressure level using surrounding  $q_{f.g.}$  values and a linear interpolation in  $\ln p$ . Then

if a W value is available for the layer below that, this  $q_{f.q.}$  is used in the calculation of a  $q_{ob}$  for the bottom of that lower layer.

When representations of  $q_{ob}$  are found at all possible observation interfaces (i.e., at the bottom of all layers in which an observation value of W is available), these are then linearly interpolated using the formula above for  $q(p)$  to each sigma layer pressure.

#### d. Type 6a - Satellite Temperature and Wind Data

Satellite temperature and cloud motion wind data are single level observations like the aircraft data. In this case, the pressure of the observation  $p(z)$  is reported instead of the altitude. Therefore, the surrounding sigma layer pressures can be found immediately, and as in the case of the aircraft data, first guess temperature and wind components are linearly interpolated in  $\ln p$  to  $p(z)$ . The observation-minus-first guess residuals are calculated at this level, and these residuals are added to the first guess values whose pressure is closest to  $p(z)$ . Unlike the aircraft observations, at most one wind observation is included in the satellite report.

At this point in the processing, sets of observation-minus first guess residuals exist at all observation locations for each of K sigma levels for z and at K sigma layers for u, v, and q. Each residual is submitted to a gross error check according to the limits: z, 400 m; u, v,  $20 \text{ ms}^{-1}$ , q,  $20 \text{ g kg}^{-1}$ . Any missing values at any of the layers/levels are filled in by place-holding numbers which are used to identify the value as missing. The number of actual values present for a given observation site can vary from K residuals for z, u, v, and somewhat fewer values of q for rawinsondes, to one level residual of z and the layer's residuals of u, and v for single level satellite observations and aircraft observations. According to the number of residuals present for a particular observation, a number of "points" is assigned to the observation taken as a whole. Assignment of points for each layer and its level immediately above it is as follows: 1 point if the z level residual and u, v, q layer residuals are present; 0.75 points for z, (u, v) residuals present; 0.50 points for z or (u, v) and q residuals present; and 0.25 points if z or (u, v) residuals are present. Thus, a rawinsonde can garner as many as K points if humidity data were

available at all layers (very rare), while an aircraft or single level satellite can be assigned at most 0.75 points for its  $z$ ,  $(u, v)$  residuals at one level/layer.

The search for observations located around each analysis grid point is then conducted, using the same analysis grid as in the surface pressure analysis. A predetermined maximum number of observations designated by NS is chosen for each grid point. At present, the analysis selects the sets of residuals (all level/layer residuals in an observation are taken together as one "observation" in this selection process) used to correct the grid values at all levels/layers in the following manner: The Type 1 observations are searched until the first NS closest observations (using horizontal distance  $D$  in  $\mu = \exp(-k_h D^2)$ ) to find observations with largest values of  $\mu \geq 0.1$  are found. Then the search continues through the rest of the Type 1 observations. However, as soon as the NS observations with  $\mu \geq 0.1$  are found, the inverse distance factor is multiplied by the number of "points" assigned to the observation, and this value for each remaining Type 1 observation is compared to the like value of the NS values already accepted. If it is greater than any of them, it takes the place of the observation with the  $\mu$  times points value just less than it. This observation and the others previously accepted below it are "bumped" down one place in priority, and the NS observation is excluded from consideration. This procedure continues until all Type 1 observations have been considered. If less than NS observations have been found from the complete set of Type 1 observations, the entire set of Type 4 observations is compared against the accepted Type 1's in the same manner. If NS surrounding observation sets still have not been found, the Type 2 and Type 6a sets are considered in order. Thus the search continues through the four data types in this order, discontinuing after considering the first data type set in which the collection of NS observations has been completed. The entire set of NS observations is a candidate for participation in the procedure that corrects that grid point's first guess values at all model layers/levels. In the present version of the analysis,  $NS = 8$ , but a correction can take place even if only one observation has a  $\mu \geq 0.1$ . If no observations satisfy the  $\mu \geq 0.1$  condition, no correction is calculated.

The purpose for assigning points to the observation sets (where "set" indicates an entire observation at all levels and layers) is to

indicate how much that observation set could contribute to correct the grid point values at all layers/levels. Based on the current point system used, a complete rawinsonde observation for a  $K = 12$  model structure would have 12 points, while an aircraft or single level satellite would have only 0.75 points. This means that a complete rawinsonde observation located 16 times farther from the grid point than a complete aircraft or single level satellite observation would have equal likelihood of being chosen for that grid point. This is because it can contribute 48 residuals to the correction of the entire stack of grid point first guess values, while the single level observation could only contribute three. Types of data are processed in the indicated order because of the perceived accuracy of the observations and the number of residuals generally present in each type. Conventional upper air soundings are seen as the most accurate and as having the most residuals. Satellite soundings are taken next because of the complete set of height/temperature residuals they provide, even though they may be somewhat less accurate than the single level aircraft observations. Satellite winds are thought to be the least accurate of all; thus they are considered last in the selection process. Observation error values upon which this processing order is based are taken from Table 2.1.1 from Gustafsson (1981).<sup>5</sup>

The sets of residuals selected for potentially correcting each grid point's first guess values are passed on to the code that performs the optimum interpolation update. Because the specific humidity is corrected univariately, the correcting procedure for  $q$  at all layers for each grid point is done after the correcting is carried out for  $z$ ,  $u$ , and  $v$  at all levels/layers. Except for the fact that the procedure uses only  $q$  residuals to correct first guess  $q$ , all other steps carried out in the  $q$  correction follow in the same order as for the  $z$ ,  $u$ , and  $v$  correction. The first step is the error flagging procedure. As in Bergman's inequality (7.1), all residuals of one variable are compared against all other residuals of that same variable at all other levels/layers in all other observations, with flags imposed when the inequality (7.1) is violated. The value of forecast error standard deviations [ $\sigma$  in inequality (7.1)] are the same as those

5. Gustafsson, N., 1981: A review of methods for objective analysis. Dynamic Meteorology: Data Assimilation Methods, Vol. 36, Applied Mathematical Sciences Series, L. Bengtsson, M. Ghil, and E. Källén, eds., Springer-Verlag, 330pp.

used in the NMC optimum interpolation analysis (Dey, 1983<sup>6</sup>) and are reprinted here in Table 3. Each residual takes with it its own quality control value, taken from the QCI values given in the FGGE II-B data sets. These QCI values are based on vertical consistency checks done on data as part of the procedure of archiving the data in FGGE. The QCI values listed in Table 4 are the ones assigned in the code.

TABLE 4. VERTICAL QUALITY CONTROL INDICATORS

<u>QCI</u>	<u>Description from FGGE II-B Data Set</u>
0	value found correct during vertical control check
1	vertical control check has not been made
2	observed value found erroneous during vertical control check; calculated value inserted
3	original value missing - calculated value inserted
4	original value missing - value assigned
5	vertical control check made - action unrecorded, most likely value entered
6	value found suspect during vertical control check
7	value found erroneous during check against certain limits
8	value found erroneous during vertical control check
9	original value missing - no value entered

If the inequality is violated for a particular pair of residual values, one or both must be flagged. The one flagged is the residual whose modified observation quality control indicator has the larger value (indicating the lower quality). The modification of the assigned QCI value is done only for the purpose of comparing the overall quality of one residual against the other to decide which one gets the flag. The QCI values 0-8 for Type 1 data are increased by one (1-9), while the modified QCI (MQCI) values for Types 2, 4, and 6a are given by  $MQCI = 3 (QCI + 3) + ITBI$ , where  $ITBI = 1$  for Type 2, 2 for Type 4, and 3 for Type 6a. For example, a QCI value of 0 will correspond to  $MQCI = 1$  for Type 1, 10 for Type 2, 11 for Type 4, and 12 for Type

6. Dey, C.H., 1983: The NMC optimum interpolation procedure. Presented at NMC workshop on vector processing and the statistical analysis of meteorological data, Camp Springs, MD.

TABLE 3. NMC FORECAST ERROR STANDARD DEVIATIONS  $\sigma$  VALUES (from Dey, 1983<sup>6</sup>)

Latitude	90S-10S			10S-10N			10N-30N			30N-50N			50N-90N			90S-90N		
	Pressure (mb)	z (m)	u (m sec <sup>-1</sup> )	v (m sec <sup>-1</sup> )	T (°K)	z	u	v	z	u	v	z	u	v	z	u	v	T (°K)
1000		25.4	4.2	4.1	14.7	3.5	3.3	3.3	19.4	4.1	3.9	19.7	4.6	4.9	21.6	4.4	4.3	4.0
850		25.4	3.9	3.7	15.5	3.2	2.8	2.8	16.7	3.7	3.4	19.6	4.1	4.3	19.4	3.6	3.6	4.0
700		26.8	4.2	4.1	19.3	3.6	3.3	3.3	18.7	4.1	3.6	19.7	4.2	3.8	20.8	3.6	3.6	4.0
500		33.5	4.8	4.6	25.4	4.3	3.5	3.5	25.6	4.8	4.2	23.2	4.8	4.6	25.8	4.4	4.4	4.0
400		39.2	5.5	5.7	29.6	4.1	4.0	4.0	32.4	5.3	4.9	27.2	5.6	5.3	29.9	5.3	5.0	4.0
300		47.0	6.6	6.8	40.1	5.8	5.3	5.3	42.3	6.5	6.1	34.1	7.1	6.4	36.5	5.7	5.5	4.0
250		50.6	9.5	7.3	49.4	5.8	5.2	5.2	47.3	7.3	6.9	37.4	6.6	6.6	39.2	5.4	5.3	4.0
200		53.5	7.3	7.0	55.3	7.8	5.8	5.8	56.3	7.6	7.6	41.6	6.7	6.2	42.0	4.4	4.3	4.0
150		57.3	6.5	6.7	61.4	8.2	5.7	5.7	67.4	7.3	7.1	46.7	5.9	4.9	48.3	3.7	3.7	4.0
100		69.8	7.0	6.4	78.0	8.8	7.1	7.1	80.1	7.7	6.2	55.2	5.1	4.1	59.4	3.7	3.6	4.0
70		77.7	6.6	5.5	93.2	7.7	4.8	4.8	100.0	5.7	5.3	64.2	7.1	4.9	71.8	5.5	5.1	4.0
50		90.1	8.3	8.0	108.0	9.9	8.3	8.3	101.0	8.7	8.6	78.3	8.2	6.7	88.1	7.6	7.4	4.0

6a. In this way, the conventional upper air data will never be flagged by observation types.

The observation rejection iterative procedure works very much like that for pressure and winds in the surface pressure analysis. The number of flags assigned to each residual is summed, and separately for heights, winds, and specific humidity the residual with the largest number of flags is identified. This residual, along with all others that have within four of that maximum number, is eliminated, and the flags that they imposed on the remaining residuals are removed. The remaining residual of each variable with the largest number of flags is identified; it and others within four flags of it are eliminated, the imposed flags on other residuals are removed, and so on. The four flag check is removed once the highest number of flags falls below four. This continues until all residuals of each of height, winds, and (in its own part of the code) specific humidity have at most one flag remaining, or ten iterations have been carried out, whichever comes first. The remaining residuals are candidates for correction of the grid point first guess values of  $z$ ,  $u$ ,  $v$ , and (separately)  $q$  at all levels/layers for that particular grid point.

In the multivariate  $z$ ,  $u$ ,  $v$  correction, the variables are corrected one at a time, one level at a time, beginning with  $z$ . The auto- and cross-correlations  $\rho_{ig}^{zz}$ ,  $\rho_{ig}^{uz}$ , and  $\rho_{ig}^{vz}$  for the grid point  $g$  and each candidate residual are calculated, and the resulting values are divided by  $1 + \epsilon_i^2$ , where  $\epsilon_i$  is the normalized observation error for observation  $i$ . Here and in the construction of the matrix of observational correlations, the values of the normalized observation errors are held fixed in space and time, taken from Bergman's Table 2. When a suitable method of prediction-error cycling is developed, this will be implemented to obtain  $\epsilon$  by means of Bergman's Eq. (2.10d). The grid point-observation correlations are sorted by size, and the largest (up to a maximum number designated by MS) are identified. Only correlations greater than 0.1 are accepted in this set - if fewer than MS correlations have magnitudes greater than 0.1, only those that do are accepted for performing the correction. Currently MS = 10 for all four variables. Only for the process of selecting the MS correlations used in the correction, cross-correlations are multiplied by 1.75 in order to insure that some observations of the variables of the type not being corrected

have a better chance for inclusion in the correction. This enhances the multivariate nature of the  $z, u, v$  correction at the risk of increasing the analysis error by choosing less correlated observations for the correction.

When the MS residuals to be used in the correction are identified, the MS X MS correlation matrix is formulated by calculating the correlation of each residual with all other residuals. The resulting symmetric matrix and the right-hand-side correlation  $\rho_{ig}^{kr}$  of each of the MS observations with the grid point variable being corrected define the system of equations given by Bergman's Eq. (2.12). This system is solved using the method of conjugate gradients (Beckman, 1960)<sup>7</sup> to yield the non-dimensional weights  $a'_{ik}$ . These are used to calculate a non-dimensional analysis error (Bergman's Eq. (2.13)). The weights are dimensionalized to form the weights  $a_{ik}$  using the ratios of  $\overline{z_i^2} = R^2 C / g^2$ ,  $\overline{u_i^2} = \overline{v_i^2} = 2 C k_h R^2 G_i^2 / f_i^2$  ( $C$  is a constant with dimensions  $\text{deg}^2$  that cancels out in each ratio) applied in the form

$$a_{ju} = \left( \frac{\overline{z_g^2}}{\overline{u_j^2}} \right)^{1/2} a'_{ju} = \left( \frac{\overline{v_g^2}}{\overline{u_j^2}} \right)^{1/2} a'_{ju} = \left( \frac{\overline{u_g^2}}{\overline{u_j^2}} \right)^{1/2} a'_{ju}$$

$$a_{jv} = \left( \frac{\overline{z_g^2}}{\overline{v_j^2}} \right)^{1/2} a'_{jv} = \left( \frac{\overline{v_g^2}}{\overline{v_j^2}} \right)^{1/2} a'_{jv} = \left( \frac{\overline{u_g^2}}{\overline{v_j^2}} \right)^{1/2} a'_{jv}$$

$$a_{jz} = \left( \frac{\overline{z_g^2}}{\overline{z_j^2}} \right)^{1/2} a'_{jz} = \left( \frac{\overline{v_g^2}}{\overline{z_j^2}} \right)^{1/2} a'_{jz} = \left( \frac{\overline{u_g^2}}{\overline{z_j^2}} \right)^{1/2} a'_{jz}.$$

The product of the dimensional weights  $a_{ik}$  and the corresponding residuals are summed to form the correction, which is added to the first guess value to form the corrected value. The multivariate procedure is repeated for each

7. Beckman, F.S., 1960: The solution of linear equations by the conjugate gradient method. Mathematical Methods for Digital Computers, Vol. 1, A. Ralston and H.S. Wilf, eds., Wiley, 62-72.

remaining level of first guess values of  $z$ , then in turn for all layer values of  $u$ , then all layer values of  $v$ . This means that potentially a different set of observations is selected to correct each of the 3K first guess values in the multivariate correction. Consideration is being given to selecting the same MS residuals for correcting  $z$ ,  $u$ , and  $v$  at each level as is the practice at NMC (Morone and Dey, 1982<sup>8</sup>). This has the benefit of not only reducing computing time, since the same covariance matrix would be used at each level to correct all three variables, but also further enhancing the geostrophic balance between the corrected mass and wind values by insuring that the same residuals are used to correct all three variables.

The formulations used to calculate the observation error covariances [ $\eta$  in Bergman's Eq. (2.12)] are the same as are used for calculating the forecast error covariances  $\rho$ , except that different values of  $k_h$  and  $k_p$  are used in the horizontal and vertical observation error correlations. The vertical error correlation functions for rawinsonde temperatures (heights) and winds shown in Bergman's Fig. 4 were analyzed to yield approximate values for  $k_p$  of 8.3 and 5.2 for temperatures (heights) and winds respectively, and vertical satellite temperature error correlation functions values in the form of a table (Table 5) were obtained from NMC (Morone, 1982, personal communication). In the case of this table of values for satellite temperatures, a linear  $\ln p$  interpolation was used to obtain the value for a residual at a particular pressure level. In this code, satellite wind observation errors are assumed to be uncorrelated in the horizontal and vertical, and for conventional soundings any two observations of the same variable in the same sounding are taken to be a correlation of unity. The value used for  $k_h$  for the horizontal satellite temperature observation error correlation function is  $1.13 \times 10^{-11} \text{ m}^2$ , following Bergman.

As soon as the  $z$ ,  $u$ ,  $v$  updates have been completed at all levels/layers for the grid point being considered, corrected level heights are converted to layer temperatures using the inverse form of the local hydrostatic equation with  $z_{\sigma_1}^{\wedge} = z_{\text{m.s.}}$ . At this point, the specific humidity correction

8. Morone, L.L., and C.H. Dey, 1983: The NMC global data assimilation system. Proceedings, Sixth Conference on Numerical Weather Prediction, AMS, Boston, 115-119.

TABLE 5. VERTICAL ERROR CORRELATION FUNCTION VALUES FOR SATELLITE T, z  
(from Morone, 1982, personal communication)

X	$\frac{tt}{\eta_{ij}}$
1	.99
2	.88
3	.61
4	.36
5	.18
6	.05
7	-.05
8	-.12
9	-.15
10	-.14
11	-.11
12	-.08
13	-.05
14	-.03
15	-.02
16	-.02
17	-.01
18	-.01
19	0.0
20	0.0

where  $X = 20 [\ln (p_i/p_j)]^2 + 1$

$p_i$  = pressure level of location i

$p_j$  = pressure level of location j

is carried out for all layers at the same grid point. After the corrected value for each layer is computed, it is compared against the saturation specific humidity calculated using the corrected temperature value at the same layer. If the corrected  $q$  value exceeds  $q_s$ , it is set equal to  $q_s$ . If it is less than ten percent of  $q_s$ , it is set equal to  $0.1(q_s)$ . In this way, the corrected humidity field is not allowed to exceed 100 percent nor fall below ten percent relative humidity.

When corrected grid point values of  $T$ ,  $u$ ,  $v$ ,  $q$  are obtained over the entire grid lattice, the values of each variable at each sigma layer are expanded in spectral coefficients (using the same rhomboidal truncation as the forecast model) of  $T$ ,  $\xi$  (converted to absolute vorticity),  $d$ , and  $q$ . These are written on a file along with the earlier obtained spectral coefficients of the corrected surface pressure field, and are thus ready for processing by the non-linear normal mode initialization (NMI) scheme (Gerlach, 1982<sup>9</sup>) which is the scheme developed by Ballish (1980)<sup>10</sup> for NMC. The subsequent initialized fields would then be input to the GSM to conduct forecasts. At the present time, no data assimilation cycle tests over extended periods of time have been conducted. Plans for such tests are being formulated and will be conducted subject to revision of the ASAP codes, especially to improve program efficiency and to allow for prediction error cycling.

Before leaving the discussion of the upper air analysis, a note of explanation concerning the rationale of carrying heights on the sigma levels is called for. Once the decision was made to use  $z$  rather than  $T$  as the mass variable that is corrected in the optimum interpolation procedure, it became necessary to convert forecast values of  $T$  at sigma layers to values of  $z$  at corresponding vertical locations to act as the first guess height values. Originally, sigma layer  $T$  values were converted to sigma layer  $z$  values using the same form of the finite difference hydrostatic equation used to perform this same task in the GSM (see Eqs. (5-8) in Sela, 1980<sup>11</sup>) set up in the form

9. Gerlach, A.M., ed., 1982: Objective Analysis and Prediction Techniques. AFGL-TR-82-0394, Contract F19628-82-C-0023, Systems and Applied Sciences Corporation.

10. Ballish, B.A., 1980: Initialization, Theory, and Application to the Spectral Model. Ph. D. thesis, University of Maryland, 151 pp.

11. Sela, J.G., 1980: Spectral modeling at the National Meteorological Center. Mon. Wea. Rev., 108, 1279-1292.

of Brenner et al. [1982; see their Eqs. (81), (82) ]. However, after the values of  $z$  were corrected, it was necessary to convert these values back to the equivalent  $T$  values, since  $T$  is the prognostic variable in the forecast model. When this form of the finite difference hydrostatic equation was inverted and corrected  $z$  values at sigma layers were used to calculate corresponding  $T$  values, highly erratic temperature values resulted. This was found to be due to the highly unstable nature of the matrix  $A$  in Sela's Eq. (7), which when inverted to solve for  $T$  from  $z$  proved to be very sensitive to very small changes in  $z$  (i.e., the corrections applied to  $z$  in the OI correction). An alternative formulation of the finite difference hydrostatic equation proposed by Yang (1982)<sup>12</sup> was tested under the same conditions. This method proved to have the opposite problem: conversion of sigma layer  $z$  to sigma layer  $T$  yielded reasonable results, but the inverse procedure gave wildly varying answers. Since nothing in the optimum interpolation theory or in the geostrophic theory used to specify the relationship between the mass and motion structure functions required that both mass and wind variables be carried at the same location in the vertical (as long as the isobaric gradient of height is not changed), it was decided to correct  $z$  at layer interfaces (levels) to avoid this problem. Thus the simple form of the local hydrostatic equation (that is, the form requiring only the height at the level below and the layer temperature) rather than the Sela or Yang formulation (which require a knowledge of all layer temperatures or heights) is used to obtain level heights from layer temperatures. Its inverse form is used to convert corrected heights at sigma levels to corresponding sigma layer temperatures.

#### 4. Experimental Method

An experimental test of the ASAP procedure was conducted to examine the quality of the analysis it generates. Because this analysis scheme was developed for the purpose of providing initial conditions for the GSM, it was decided that the best test of its quality is not in the analysis itself, but in the quality of the forecast generated from the analysis. Therefore, it is less important that the scheme provide an analysis that fits the

12. Yang, C.-H., 1982: On the solution of the hydrostatic relation in the spectral model of the National Meteorological Center. Mon. Wea. Rev., 110, 1100-1102.

observations used in the analysis than it is that the ensuing forecast verify with observations at the forecast time. This is the philosophy that was used in the design of the experimental method. Another important consideration in the initial experimental testing of the OI analysis methods developed for this project is the fact that these are the first tests of newly developed codes. The purpose at the outset should be to see if the analysis scheme gives reasonable forecasts when compared with the forecast results using other proven analysis methods. Any obvious logical or theoretical error which may have crept into the coding (even though the codes were all thoroughly checked by doing independent calculations to verify their correctness) may show up quite vividly in the results of such a comparison of forecasts.

With this philosophy in mind, a FGGE III-A global analysis for 1/7/79 at 00Z was selected as a starting point for a test of the ASAP scheme in conjunction with the NMI and GSM. The reason FGGE III-A analyses were used instead of FGGE III-B analyses was that the grid point topography values used in the III-B analyses were not available at the time of the experiment. A less important reason for this choice was that the III-A analyses for this date and the following dates were prepared at NMC using a version of the optimum interpolation on which the ASAP codes were based, whereas the III-B analyses were formulated using an optimum interpolation procedure developed at the European Centre for Medium Range Weather Forecasts (Bengtsson et al., 1982<sup>13</sup>). Values of  $z$ ,  $T$ ,  $u$ ,  $v$  at 12 mandatory pressure levels (1000 mb - 50 mb) and relative humidity ( $r$ ) at the lower six levels (1000 mb - 300 mb) from the III-A analyses available on a  $2.5^\circ$  latitude-longitude grid were linearly interpolated to 62 Gaussian latitudes. Then every other longitude point around each latitude circle was selected to form 62 X 72 latitude-longitude grids of the values at each level. Using the PTOSIG procedure detailed in Appendix I-A, 12 sigma layer values of  $T$ ,  $u$ ,  $v$  values are obtained from the mandatory level values of  $z$ ,  $T$ ,  $u$ ,  $v$ . Specific humidity  $q$  is formed from  $T$  and  $r$  at levels 1000 mb - 300 mb, and these values are extrapolated up to 50 mb assuming  $q$  decreases linearly in  $\ln p$  between 300 mb

13. Bengtsson, L., M. Kanamitsu, P. Källberg, and S. Uppala, 1982: FGGE 4-dimensional data assimilation at ECMWF. Bull. Am. Meteorol. Soc., 63, 29-43.

and 50 mb. These  $q$  values are then interpolated to sigma layers in the same way the wind components were.

Values of first guess surface pressure that were used to determine the sigma structure for 62 X 72 sigma layer  $T$ ,  $u$ ,  $v$ ,  $q$  values were obtained using the procedure GETPS. Like PTOSIG, this is an NMC pre-processing procedure that assumes that geopotential varies quadratically in  $\ln p$ . The terrain height ( $z_{m.s.}$ ) values on the  $2.5^\circ$  grid were interpolated as above to the 62 X 72 grid; then for each grid point the mandatory level  $z$  values immediately above and below the  $z_{m.s.}$  grid value were identified. Then a value for  $p_{m.s.}$  was obtained using the GETPS procedure, described in Appendix I-B. After the 62 X 72 grid values of  $T$ ,  $u$ ,  $v$ ,  $q$  at 12 model layers [ $\sigma_k = 1.0, .925, .8, .65, .5, .375, .3, .25, .2, .15, .1, .05, .0$ , with  $\sigma_k$ , the layer sigma value, defined in terms of  $\hat{\sigma}_k$  following Sela (1980), Eq. (10)] are obtained, the fields along with  $p_{m.s.}$  are expanded in spectral coefficients of  $T$ ,  $\xi$  (converted to absolute vorticity),  $d$ ,  $q$ , and  $\ln p_{m.s.}$  with a rhomboidal truncation of  $J = 30$ . The spectral coefficients were initialized using the NMI in two iterations using four vertical modes (the external mode and the first three internal modes), then submitted to the GSM for the generation of a 24 h forecast valid 1/8/79, 00Z, and a 48 h forecast valid 1/9/79, 00Z (designated 48F). The spectral coefficients representing the 24 h forecast were used as the first guess in a complete ASAP correction procedure (surface pressure analysis, then upper air analysis) using FGGE II-B observations. The results of this analysis were then initialized using the same NMI procedure as stated above, then submitted to the GSM to produce a 24 h forecast valid 1/9/79, 00Z (designated 24A). Finally, a set of spectral coefficients was generated from the FGGE III-A analysis for 1/8/79, 00Z using the same procedure used to obtain these for the 1/7/79, 00Z starting fields. These were initialized following the same NMI procedure, and a 24 h forecast valid at 1/9/79, 00Z (designated 24F) was generated.

In light of recent studies conducted on the interaction between analysis and initialization in the global data assimilation procedure (see for example Williamson et al., 1981,<sup>14</sup> and Williamson and Daley, 1983<sup>15</sup>), the

14. Williamson, D.L., R. Daley, and T.W. Schlatter, 1981: The balance between mass and wind fields resulting from multivariate optimal interpolation. Mon. Wea. Rev., 109, 2357-2376.

15. \_\_\_\_\_, \_\_\_\_\_, 1983: A unified analysis-initialization technique. Mon. Wea. Rev., 111, 1517-1536.

effect of the initialization on the analyses used in 24A and 24F was examined. The mass-motion imbalances introduced in the optimum interpolation analysis through the use of geostrophic relationships in defining the height-wind covariances from the height-height covariances (e.g., modeling straight line flow in areas of curvature) are removed in the initialization, but this removal usually results in large changes in the initial conditions themselves. That portion of the analysis fields removed by the initialization represents the imbalance in the analysis fields, and is called the fast component of the analysis because it is responsible for the high frequency modes in the initial conditions that can cause high-frequency oscillations in the forecast. The remaining part of the analysis (that left by the initialization) is called the slow component. Since each analysis is initialized before it is introduced as initial conditions for the forecast, the slow component should carry the major portion of the change imposed by the analysis; that is, the analysis must not adjust the first guess closer to the observations locally at the expense of the required atmospheric mass-motion balance. The quality of the analysis used in preparing the initial conditions is not measured by how closely it fits the observations it uses, but rather how close the ensuing initialized fields are to those observations. If the fields actually submitted as the initial state for the forecast are not representative of the observed initial state, the resulting forecast cannot be expected to verify with observations at the forecast time. An analysis that locally fits the observations but on a larger scale violates the required mass-wind balances will consist primarily of the fast component, and will be altered significantly by the initialization in order to reinstate those necessary balances. Therefore it is of interest to examine how large a fraction of the analysis is relegated to the fast component. If too much of the information contributed by the analysis to the first guess is removed in the initialization, it may indicate a serious defect in the analysis procedure that is creating undue imbalances.

To study the analysis-initialization interaction in the case of the ASAP codes, the 24 hour forecast valid 1/8/79, 00Z was taken as the first guess, and its ASAP analyzed counterpart was taken as the analysis. The change imposed on the first guess by the analysis was calculated by computing the global rms difference between the grid point values of height

and wind (computing rms vector differences) at the respective levels/layers and for surface pressure for the two fields. After initialization of the analysis, the rms difference between the analysis and the initialized fields was obtained. If the first guess field had no fast component, then this latter rms difference in its entirety would represent the fast component of the analysis. In fact, the first guess field does contain a fast component due to imbalances introduced in the forecast in spite of the normal mode initialization. This was shown by calculating the rms differences between an initialization of the forecast and the forecast fields themselves. Differences between the two were small but appreciable. However, the fast component introduced during the forecast should be small compared to the fast component introduced during the analysis. RMS values for the total change (forecast vs. analysis) and the fast component (analysis vs. initialized analysis) were not yet available at the time of this report. For comparison purposes only, the same calculations are being carried out when the FGGE III-A analysis for 1/8/79, 00Z is used as the analysis and the same forecast field is used as the first guess (thus this analysis is thought of as an analysis of this first guess). Then the initialization of this analysis is carried out to isolate the fast component of this analysis. Obviously, the FGGE III-A analysis was constructed using a different forecast model and a different optimum interpolation procedure. The purpose of this comparison is simply to compare the rms values from the ASAP procedure with those obtained from an established OI scheme to determine what order of magnitude rms differences to expect.

Tables showing results of these calculations are not as yet available. The computations were proceeding at the time of this writing, but are expected soon. Of perhaps even greater interest are the rms differences between the respective forecasts 48F, 24A, and 24F and the observations (FGGE II-B observations for 1/9/79, 00Z) that correspond to the forecast time for all three forecasts. These are anticipated soon, as are the rms differences between these three forecasts and the FGGE III-A analysis for this same date and time. The first set of results should show the relative ability of the ASAP OI procedures in producing quality forecasts as they are verified against observations, while the second set should show how the ASAP OI forecast differed from the FGGE III-A forecast when compared to a common grid point field representative of conditions at the forecast time.

## 5. Improvements Proposed for the ASAP Codes

In their present configuration, the codes run too slowly. An effort has already begun to make the codes more efficient with as few compromises on quality as possible. For example, it is hoped that the spectral interpolation of first guess values to observation sites can be retained even though it is a fairly time consuming procedure (as compared to a less accurate but faster interpolation method). To do this, the codes are being scrutinized for places where inefficient coding or input-output procedures are being used.

Several improvements will be attempted in an effort to obtain analyses which when initialized more closely match the corresponding observations (that is, have a smaller slow component error). First, an appropriate method of calculating normalized observation errors will be developed and employed. Also, to save running time and to enhance the multivariate nature of the analysis, the same correlation matrix will be used for correcting  $z$ ,  $u$ , and  $v$  in a particular layer. This will necessitate converting the code back to correcting  $z$  on layers, which in turn will require identifying a workable layer  $T$  - layer  $z$  and layer  $z$  - layer  $T$  conversion. Finally, a flagging procedure done for the entire globe will precede the data selection section of the codes in order to improve the chance of individual observations being used at continuous grid points to provide better continuity in the correction fields.

# Appendix I-A. NMC Subroutine PTOSIG

The subroutine generates estimates of values of dynamic variables such as temperature  $T$ , wind speeds  $u$ ,  $v$ , and specific humidity  $q$ , all of which are defined as layer variables. The underlying assumptions are:

- (1) The geopotential of a pressure surface is a quadratic function of  $\ln p$ .
- (2) The winds and specific humidity are linear functions of  $\ln p$ .
- (3) The temperature  $T$  in the layer is estimated from  $\phi$  at the surrounding model levels using the hydrostatic equation

$$\partial \phi / \partial \ln p = -RT.$$

## A. Calculation of Layer Temperature

The first step is to calculate the geopotential at the model layer interfaces (levels) from geopotential  $\phi_1, \phi_2, \dots, \phi_L$  and temperature  $T_1, T_2, \dots, T_L$  known at the pressure levels  $p_1, p_2, \dots, p_L$ . Assuming  $\phi$  is quadratic in  $\ln p$ , we have

$$\phi(\ln p) = \phi(\overline{\ln p}) + A(\ln p - \overline{\ln p}) + \frac{B}{2}(\ln p - \overline{\ln p})^2 \quad (A1)$$

for  $p_2 \leq p \leq p_1$ . Here  $A$  and  $B$  are constants and

$$\overline{\ln p} = \frac{1}{2}(\ln p_1 + \ln p_2). \quad (A2)$$

By the hydrostatic assumption, we have

$$\partial \phi / \partial \ln p = -RT = A + B(\ln p - \overline{\ln p})$$

$$\partial^2 \phi / \partial (\ln p)^2 = -R(\partial T / \partial \ln p) = B.$$

The last relation in the second equation states that  $T$  must be a linear function in  $\ln p$ ; hence

$$\partial T / \partial \ln p = (T_1 - T_2) / (\ln p_1 - \ln p_2) \quad \text{in } p_2 \leq p \leq p_1$$

and

$$B = -R(T_1 - T_2) / (\ln p_1 - \ln p_2). \quad (A3)$$

Next, by substituting  $\phi_1, \phi_2$  in Eq. (A1) we obtain

$$\phi_1 = \phi(\overline{\ln p}) + A(\ln p_1 - \overline{\ln p}) + \frac{B}{2}(\ln p_1 - \overline{\ln p})^2 \quad (A4)$$

$$\phi_2 = \phi(\overline{\ln p}) + A(\ln p_2 - \overline{\ln p}) + \frac{B}{2}(\ln p_2 - \overline{\ln p})^2. \quad (A5)$$

Adding these two and noting the definition of  $\overline{\ln p}$  in Eq. (A2), we find

$$\begin{aligned} \frac{1}{2}(\phi_1 + \phi_2) &= \phi(\overline{\ln p}) + \frac{B}{2} \left\{ [\ln p_1 - \frac{1}{2}(\ln p_1 + \ln p_2)]^2 + \right. \\ &\quad \left. [\ln p_2 - \frac{1}{2}(\ln p_1 + \ln p_2)]^2 \right\} = \phi(\overline{\ln p}) + \frac{B}{8}(\ln p_1 - \ln p_2)^2; \end{aligned}$$

that is,

$$\phi(\overline{\ln p}) = \frac{1}{2}(\phi_1 + \phi_2) - \frac{B}{8}(\ln p_1 - \ln p_2)^2. \quad (A6)$$

Subtracting Eq. (A5) from Eq. (A4), on the other hand, we find

$$A = (\phi_1 - \phi_2)/(\ln p_1 - \ln p_2). \quad (A7)$$

Eqs. (A3), (A6), and (A7) define all the parameters for evaluation of  $\phi(\ln p)$  in Eq. (A1), where  $p_2 \leq p \leq p_1$ . Once  $\phi(\ln p)$  is found for levels bordering (above and below) the layer for which  $T$  is desired, the local hydrostatic form

$$T = (\phi_2 - \phi_1)/[R \ln (p_1/p_2)]$$

is used to calculate layer temperature.

#### B. Calculation of Wind Components and Specific Humidity

A linear in  $\ln p$  assumption is used to obtain a value  $\chi(\ln p)$  at a desired pressure  $p$  from values of  $\chi_1, \chi_2, \dots, \chi_L$  at pressure levels  $p_1, p_2, \dots, p_L$ , where  $\chi = u, v, q$ . Assume then that

$$\chi(\ln p) = \chi(\overline{\ln p}) + A(\ln p - \overline{\ln p})$$

where  $A$  is a constant in  $p_2 \leq p \leq p_1$  and  $\overline{\ln p} = \frac{1}{2}(\ln p_1 + \ln p_2)$ . Since  $\chi$  is linear in  $\ln p$ ,

$$\partial \chi / \partial \ln p = A = \text{constant} = (\chi_1 - \chi_2)/(\ln p_1 - \ln p_2).$$

Then

$$\chi_1 = \chi(\overline{\ln p}) + A[\ln p_1 - \frac{1}{2}(\ln p_1 + \ln p_2)]$$

and

$$\chi_2 = \chi(\overline{\ln p}) + A[\ln p_2 - \frac{1}{2}(\ln p_1 + \ln p_2)].$$

Hence

$$\chi(\overline{\ln p}) = \frac{1}{2}(\chi_1 + \chi_2)$$

and

$$\chi(\ln p) = \frac{1}{2}(\chi_1 + \chi_2) + \frac{\chi_1 - \chi_2}{\ln p_1 - \ln p_2} [\ln p - \frac{1}{2}(\ln p_1 + \ln p_2)].$$

This last equation is the form used to interpolate between known values of  $\chi_1, \chi_2$  for  $p_2 \leq p \leq p_1$  to obtain  $\chi$  at  $p$ , for  $\chi = u, v, q$ .

# Appendix I-B. NMC Subroutine GETPS

This subroutine estimates the value of  $\ln p_*$ , where  $p_*$  is the surface pressure, from data on known pressure levels. The geopotential of the level where  $p_*$  is desired, the surface topography, is given and denoted by  $\phi_*$ .

Assume we are given  $\phi_1, \phi_2, \dots, \phi_L, T_1, T_2, \dots, T_L$  at the known pressure levels  $p_1, p_2, \dots, p_L$ . Also, assume that  $\phi$  is a quadratic function of  $\ln p$  in an individual layer  $p_k \leq p \leq p_{k-1}$  where here the indices increase upward in the atmosphere. Thus, we assume the form

$$\phi(\ln p) = \phi(\overline{\ln p}) + A(\ln p - \overline{\ln p}) + \frac{B}{2}(\ln p - \overline{\ln p})^2 \quad (B1)$$

that satisfies the hydrostatic equation

$$\partial \phi / \partial \ln p = -RT = A + B(\ln p - \overline{\ln p})$$

so that

$$\partial^2 \phi / \partial (\ln p)^2 = -R(\partial T / \partial \ln p) = B. \quad (B2)$$

The first step is to find the levels  $\phi_k, \phi_{k-1}$  that surround the surface topography geopotential  $\phi_*$ . Then letting

$$DH = \ln p_{k-1} - \ln p_k > 0$$

$$A = (\phi_{k-1} - \phi_k) / DH < 0$$

$$B = -R(T_{k-1} - T_k) / DH \begin{cases} < 0 & \text{lapse} \\ = 0 & \text{isothermal} \\ > 0 & \text{inversion} \end{cases}$$

$$\phi_o \equiv \phi(\overline{\ln p}) = \frac{1}{2}(\phi_{k-1} + \phi_k) - \frac{B}{8}(DH)^2.$$

If we let  $\chi_* = \ln p_*$ ,  $\chi_o = \overline{\ln p}$ , then rewriting Eq. (B1) gives

$$\frac{B}{2}(\chi_* - \chi_o)^2 + A(\chi_* - \chi_o) + \phi_o - \phi_* = 0. \quad (B3)$$

Using the quadratic equation to solve for  $\chi_* - \chi_o$  yields

$$\chi_{*} - \chi_0 = - \frac{A \pm \sqrt{D}}{B}$$

where  $D \equiv A^2 - 2B(\phi_0 - \phi_*) \geq 0$ . Choosing the + sign gives

$$\chi_{*} - \chi_0 = - \frac{A + \sqrt{D}}{B}.$$

Rationalizing the numerator, using the definition of  $D$ , and noting that  $-A = |A|$  since  $A < 0$  in all cases, yields

$$\chi_{*} - \chi_0 = 2(\phi_0 - \phi_*) / (|A| + \sqrt{D}).$$

Solving for  $\chi_{*}$  and using  $p_{*} = \exp(\chi_{*})$  gives the surface pressure.

## B. Generalization and Applications of Machenhauer Initialization Scheme

At the beginning of the year, the NMC initialization scheme was implemented on the AFGL CDC 6600 system. Details of the scheme are outlined in last year's report (Gerlach, 1982).<sup>9</sup> The scheme has been adapted and tested for variable resolutions. It can now accept any vertical and horizontal resolution and solve for the corresponding eigenvalues and eigenvectors, subject only to computer space limitations. The high resolution initialization, similar to the high resolution AFGL model, can only be realistically run on the Air Force Weapons Laboratory (AFWL) Cray-1 system.

The scheme linearizes the model equations for the dependent spectral coefficients  $\hat{\zeta}_n^s$  (vorticity),  $\hat{D}_n^s$  (divergence), and  $\hat{\Gamma}_n^s$  (composite vertical variable including surface pressure and temperatures). These are expanded in both vertical and horizontal modes. The vertical modes  $\hat{Z}_I(\sigma_k)$   $k = 1, 2, \dots, K$  are obtained by finding the eigenvectors of the eigenvalue problem  $G\hat{Z}_I = \gamma_I \hat{Z}_I$ ,  $I = 1, K$ , where  $G$  is a matrix that depends on the structure of the model and is computed on the basis of the input information on the model resolution. For each vertical mode  $I$  and zonal wavenumber  $s$ , a separate eigenvalue equation exists that can be solved for the meridional modes. The resulting tendency equation for each  $I$  and each  $s$  then has the form

$$\frac{\partial}{\partial t} \hat{X}_\ell^s = iL_S \hat{X}_\ell^s$$

$$\frac{\partial}{\partial t} \hat{Y}_\ell^s = iL_A \hat{Y}_\ell^s$$

where  $\ell$  is the eigenvalue or frequency index,  $L_S$  and  $L_A$  are the meridional mode matrices representing the symmetric and antisymmetric components, respectively, and

$$\hat{X}_\ell^s = \text{Col} (\zeta_{s+1}^s, \zeta_{s+3}^s, \dots, D_s^s, D_{s+2}^s, \dots, \Gamma_s^s, \Gamma_{s+2}^s, \dots)$$

$$\hat{Y}_\ell^s = \text{Col} (\zeta_s^s, \zeta_{s+2}^s, \dots, D_{s+1}^s, D_{s+3}^s, \dots, \Gamma_{s+1}^s, \Gamma_{s+3}^s, \dots).$$

If the resolution of the model is truncated at even wavenumber  $J$ , then the square matrix  $L_S$  has dimension  $(3J/2 + 2)$  and yields  $(J + 2)$  gravity

modes.  $L_A$ , also a square matrix, has dimension  $(3J/2 + 1)$  and yields  $J$  gravity modes. If  $J$  is odd, however, then both have dimensions  $3(J+1)/2$  and yield  $(J + 1)$  gravity modes. The expansion coefficients are all pre-stored on file for use in the iterative scheme.

The Machenhauer scheme is then implemented by expanding the field and the tendencies in terms of the normal modes. Each of all gravity modes with period less than two days is adjusted to make its tendency vanish. The field is then reconstructed and the tendencies recomputed and reexpanded through an iteration procedure. The number of iterations, as well as the number of vertical modes selected, depends on the resolution. For the low resolution (16 X 16) with six layers, only two vertical modes are used. For the higher resolutions (e.g., 24- and 30- rhomboidal) and with 12 layers, four modes are applied.

Two sets of FGGE III-A data were initialized with the adapted code for the higher resolution of 24 X 24 and 12 vertical layers. Both sets then served as initial conditions and forecasts were produced from them as well as from the uninitialized analysis. Results indicate that initialization improved the analysis for both cases. This despite the fact that four vertical modes resulted in a divergent solution after two iterations. When two vertical modes were used, the solutions converged, but the forecasts based on those initializations were still inferior to the four-mode initialization. Recently a set of optimally interpolated (OI) analyses was made available with a resolution of 30-rhomboidal horizontal waves and 12 vertical layers. The data were initialized using four vertical modes and the solution converged after two iterations. The changes in the velocity divergence due to the initialization were found to be greater than those found when FGGE III-A data that had been based on the Flattery analysis (Cooley, 1974)<sup>16</sup> were used.

16. Cooley, D.S., 1974: A Description of the Flattery Global Analysis Method - No. 1. NOAA NWS Tech. Proc. Bull. No. 105.

## 4. Objective Analysis with Hough Vectors

Most of the past year's accomplishments can be found in the paper "Objective analysis using Hough vectors evaluated at irregularly spaced locations" by Isidore Halberstam and Shu-Lin Tung, which will appear in Monthly Weather Review. Preprints are available from the authors or from SASC Lexington Operations. This summary emphasizes the additional work completed subsequent to preparation of the paper.

To prove the value of the objective analysis scheme described in the above-mentioned paper, a full forecast and analysis cycle is being attempted. The analyses make use of the residuals, i.e., the differences between the forecasts and the observations. They are fitted to Hough functions through the finality principle described in the paper, and the analyzed residuals subtracted from the forecast values at the grid points to produce the final analysis. The forecast in question here is a 12 h forecast of the AFGL spectral, global model beginning with the FGGE III-A analysis at 1/15/78, 00Z and valid at 12Z for the same date. The analysis is produced at 12Z and another 12 h forecast generated for 00Z on 1/16/78. This forecast is used to arrive at an analysis for 1/16/78, 00Z from which another 12 h forecast is generated verifying at 12Z on 1/16/78. Observations at 12Z will be combined with forecast values to produce a final analysis at 12Z on 1/16/78. This analysis will be compared with the 12 h forecast produced for that time. It will also be used as verification for a 24 h forecast starting with the analysis at 12Z on 1/15/78. Paralleling this analysis-forecast cycle are forecasts made from FGGE III-A data. A 12 h forecast beginning with the FGGE III-A analysis will be produced for both 00Z and 12Z on 1/16/78. Also a 24 h forecast will be generated for 12Z on 1/16/78. These forecasts will be verified against FGGE III-A data valid for the respective verification times. An intercomparison of results will indicate whether the Hough analysis was viable, at least as compared with FGGE III-A analyses. Comparisons will also be made against observations. Because of the lack of objective quality control on the current procedure, the analyses will have some built-in error sources that FGGE III-A did not have and are expected to fare worse than FGGE III-A.

The problem of vertical interpolation was also encountered in this effort. The AFGL model produces a forecast for wind divergence, vorticity,

temperatures, and specific humidity. These can be converted to geopotential heights, wind velocity components, and relative humidity. The forecast variables are produced normally on  $\sigma$ -layers and are meant to represent the layer mean values. Geopotential heights are ordinarily predicted at  $\sigma$ -levels, but Sela's (1980)<sup>11</sup> form of the hydrostatic equation has no difficulty in producing layer heights from temperatures. Observations, however, appear at mandatory levels in the forms of wind velocities, geopotential heights, and relative humidity. To perform an analysis on a  $\sigma$ -layer requires interpolation from the mandatory levels to the  $\sigma$ -layers. If one assumes that temperatures, velocities, and humidities are linearly dependent on  $\ln p$ , where  $p$  is pressure, such an interpolation presents little difficulty. But that implies that heights are quadratic in  $\ln p$ , and, although interpolation of heights to  $\sigma$ -layers can be done easily, the derivation of temperatures from heights after the analysis is no longer a trivial matter. Sela's formula does not function well when asked to produce temperatures from heights at  $\sigma$ -layers. As pointed out by Yang (1982),<sup>12</sup> an oscillation occurs in the error which accumulates in the vertical computation of temperature. Yang proposes an alternate form of the hydrostatic equation, but that produces oscillations in the opposite direction; that is, in the computation of heights from temperatures. In order to avoid these problems, it was decided to perform the analyses on mandatory levels and to derive winds, temperatures, and humidity at the  $\sigma$ -layers from them. The forecast values can be interpolated to the mandatory levels with the model's post-processor. Its pre-processor can then interpolate winds and humidity from the mandatory levels to the  $\sigma$ -layers after completion of the analysis. But in order to interpolate heights to the  $\sigma$ -layers, analyzed temperatures are required at the mandatory levels. BASU's current Hough objective analysis scheme, however, does not provide an analysis of temperature. It only provides heights and wind velocity multivariately based on the Laplace tidal equations. It was therefore decided to derive temperatures from heights differently from the pre-processor.

The current post-processor supplies information on the mandatory levels. The analysis is then performed on the mandatory levels, yielding wind velocities and heights. The analyzed heights furnish mean temperatures of the layers between mandatory levels by a local application of the hydrostatic equation. The temperatures are then linearly interpolated to

the  $\sigma$ -layer as in the pre-processor. Velocities are also interpolated linearly to the  $\sigma$ -layers by the pre-processor. There is a slight problem in applying the temperature interpolation to the lowest layer, where a mean temperature is usually required between the lowest mandatory level (1000 mb) and the surface. Once the 1000 mb level has been analyzed, the surface terrain heights may be used in conjunction with the analyzed heights to obtain the lowest mean temperatures. Invoking the hydrostatic relationship would yield those temperatures, but that would require a corrected surface pressure first. The pre-processor in fact derives a surface pressure from the analyzed heights on the mandatory levels. But here temperatures are required a priori. This chicken and egg problem was solved by deriving the surface pressure through a quadratic fit of the three lowest mandatory levels. Results of the fitting procedure are both realistic and accurate at least for the test atmospheres on which the procedure was tried. With the temperatures and winds interpolated to the  $\sigma$ -layers, the model required only analyzed humidities before forecasting could commence. For our test purposes, we did not analyze humidity separately but used the FGGE III-A analysis of relative humidity.

Table 6 shows the rms errors of a 12 h forecast generated from our analysis at 1/15/78, 12Z. The errors are determined from the forecast values computed at FGGE II-A observation sites. They may seem rather high, but the forecast was computed from a non-initialized state and from data that had not been quality checked. The forecast also represents the AFGI six-level model which tends to do poorly in the upper levels because of poor resolution. It would be advantageous to compare these errors with forecast errors, where the forecast had been computed from FGGE III-A analysis. Future experiments will be geared to reducing these errors.

One possible method for reducing analysis errors that is being studied currently involves an iterative procedure. Because of poor data resolution over the oceans and many portions of the Southern Hemisphere, our analysis deteriorates if too many waves are included. A similar result was seen in the study by Baer and Tribia (1967)<sup>17</sup> who found, as did Halberstam

17. Baer, F., and J.J. Tribia, 1976: Spectral fidelity of gappy data. Tellus, 28, 215-227.

TABLE 6. RMS ERRORS OF 12 h FORECAST VALID AT 1/16/78, 00Z VS. PAGE 11-A  
OBSERVATIONS

<u>Level (mb)</u>	<u>u error (m/s)</u>	<u>v error (m/s)</u>	<u>z error (m)</u>
1000	5.64	6.34	48.31
850	6.13	6.23	65.35
700	6.30	6.23	70.47
500	7.10	7.84	79.98
400	7.66	8.26	95.52
300	10.04	9.77	126.58
250	10.01	10.54	151.04
200	13.56	12.44	160.73
150	15.69	13.52	177.02
100	15.02	12.04	214.68
70	17.11	12.49	290.66
50	19.86	12.35	333.91

(1981),<sup>18</sup> that it is not the amount of data that dictates the accuracy of an expansion but their distribution. In fact, it is useless to attempt resolving data with wave lengths smaller than the largest spatial gaps in the data. The iterative method thus begins with a very coarse resolution. Once an analysis is performed with the low resolution, extra "data points" from the analysis are added in areas of the largest gaps. A higher resolved analysis is then performed and so on until the desired resolution is achieved. The method was tested for a one-dimensional function but no conclusions can yet be made concerning its effectiveness.

18. Halberstam, I., 1981: An Investigation of Three Methods for Spectral Representation of Randomly Distributed Data. AFGL-TR-81-0.34, Contract F19628-81-C-0039, Systems and Applied Sciences Corporation.

## II. MESOSCALE FORECASTING

## A. An Error Analysis of LFM-II Forecasts During the 1982-83 Winter Season

### 1. Introduction

Studies designed to identify systematic errors in circulation variables (sea-level pressure, geopotential height, thickness, etc.) in National Meteorological Center (NMC) models have become fairly widespread. Fawcett (1969)<sup>1</sup> examined forecast errors of a baroclinic model while Leary (1971),<sup>2</sup> Colucci and Bosart (1979),<sup>3</sup> and Wallace and Woessner (1981)<sup>4</sup> looked at errors in primitive equation models. Silberberg and Bosart (1982)<sup>5</sup> have described systematic errors in forecasts of sea-level pressure and 1000-500 mb thickness for cyclones in the currently operational Limited-Area Fine Mesh Model (LFM-II). The present study differs from most others in that forecasts of mean relative humidity, boundary layer relative humidity, boundary layer temperature, and boundary layer winds were also examined.

The purpose of this analysis was to identify systematic errors in the LFM-II to investigate what might best be termed model bias. Chief motivation for the study came from the need to evaluate short-range forecast guidance in the preparation of terminal forecasts of wind, precipitation, ceiling height, and total cloud amount (see C. Mesoscale Forecasting Experiment IV, below). Products derived from LFM-II model forecasts were included in this evaluation (Chisholm et al., 1983<sup>6</sup>). In broader terms the results

1. Fawcett, E.B., 1969: Systematic errors in operational baroclinic progress at the National Meteorological Center. Mon. Wea. Rev., 97, 670-682.
2. Leary, C., 1971: Systematic errors in operational National Meteorological Center primitive equation surface prognoses. Mon. Wea. Rev., 99, 409-413.
3. Colucci, S.J., and L.F. Bosart, 1979: Surface anticyclone behavior in NMC prediction models. Mon. Wea. Rev., 107, 377-394.
4. Wallace, J.M., and J.K. Woessner, 1981: An analysis of forecast error in the NMC hemispheric primitive equation model. Mon. Wea. Rev., 109, 2444-2449.
5. Silberberg, S.R., and L.F. Bosart, 1982: An analysis of systematic cyclone errors in the NMC LFM-II model during the 1978-79 cool season. Mon. Wea. Rev., 110, 254-271.
6. Chisholm, D.A., A.J. Jackson, M.E. Niedzielski, R. Schechter, and C.F. Ivaldi, Jr., 1983: The Use of Interactive Graphics Processing in Short-Range Terminal Weather Forecasting: An Initial Assessment. AFGL-TR-83-0093, Air Force Geophysics Laboratory, Hanscom AFB, MA.

could serve the dual purpose of alerting forecasters to consistent and systematic model errors and at the same time give some clues as to areas where model physics and procedures might be improved. In addition, it is planned to incorporate the error statistics that have been generated into some type of forecast update procedure in an attempt at reducing model forecast errors. Work on this endeavor by SASC scientists has already begun.

## 2. Data Collection Procedures

The Man-computer Interactive Data Access System (McIDAS) at AFGL is used to support forecasting research studies. Software routines have been written to decode, store, display and contour the FOUS 60-78 bulletins (see Gerlach, 1982<sup>7</sup>) received over the FAA WB-604 line.<sup>8</sup> These bulletins are received twice daily approximately three to four hours after 00 GMT and 12 GMT and are stored on disk in McIDAS. All forecasts used in this study were written to magnetic tape to facilitate further analysis.

Fig. 1 is an example of a FOUS 60-78 bulletin. The values listed for each station are obtained by interpolation from the LFM-II model grid points to the FOUS station locations. Since the vertical structure of the LFM is in sigma ( $\sigma$ ) coordinates, the pressure levels represented by the model layers are approximate rather than exact. The first line of the bulletin represents initial conditions (at 00 GMT or 12 GMT). The remaining lines are forecasts at six-hour intervals from initial time out to 48 hours. In all, 89 FOUS 60-78 stations are received over the WB-604 line including six model grid points located in close proximity to the United States east coast. Fig. 2 shows the geographic locations of the available stations.

The parameters chosen for error analysis in this study include mean relative humidity, sea-level pressure, 1000-500 mb thickness, boundary layer relative humidity, boundary layer temperature, and the boundary layer wind in vector form. In the LFM-II the boundary layer is defined as the bottom 50 mb of the model atmosphere. Forecast values for each parameter at each

7. Gerlach, A.M., ed., 1982: Objective Analysis and Production Techniques. AFGL-TR-82-0394, Contract F19628-82-C-0023, Systems and Applied Sciences Corporation, pp. 99-113.

8. National Weather Service, 1981: FOUS 60-78 Bulletins. NOAA NWS Tech. Proc. Bull. No. 294, 11 pp.

OUTPUT FROM LFM 12Z 10/05/83

STA	RH	R1R2R3	VVLI	HHDDFF	TBPSPTT
ALB	78	998760	///06	642115	9109///
06	81	977682	01603	651916	9307000
12	91	969091	01602	632018	9304035
18	73	938834	00403	612518	9206008
24	52	877002	-0806	562818	9011002
30	43	815403	-1407	542817	9012000
36	43	805702	-1407	522818	8916000
42	36	724802	-1409	502918	8718000
48	26	623202	-1912	483017	8621000

Fig. 1. FOUS G1 Guidance Forecast Bulletin for Albany, NY (ALB)

RH is the mean relative humidity of the lowest three layers of the model (1000 to 490 mb) in percent. R1 is the mean relative humidity of the 50 mb thick boundary layer (1000 to 950 mb) in percent. R2 is the mean relative humidity of the lowest tropospheric layer (950 to 720 mb) in percent. R3 is the mean relative humidity of the middle tropospheric layer (720 to 490 mb) in percent. VV is vertical velocity at 700 mb in tenths of a microbar per second, with negative values representing downward motion. LI is lifted index in °C, with negative values designated by subtracting from 100 (-4 is represented by 96, for example). HH is 1000-500 mb thickness in decameters with the first digit omitted. DDFF is the direction and speed of the boundary layer wind in tens of degrees and knots. TB is the mean potential temperature of the boundary layer in °K, with the first digit omitted. PS is sea-level pressure in mb with the hundreds position omitted. PTT is the six-hour accumulated precipitation in hundredths of an inch.

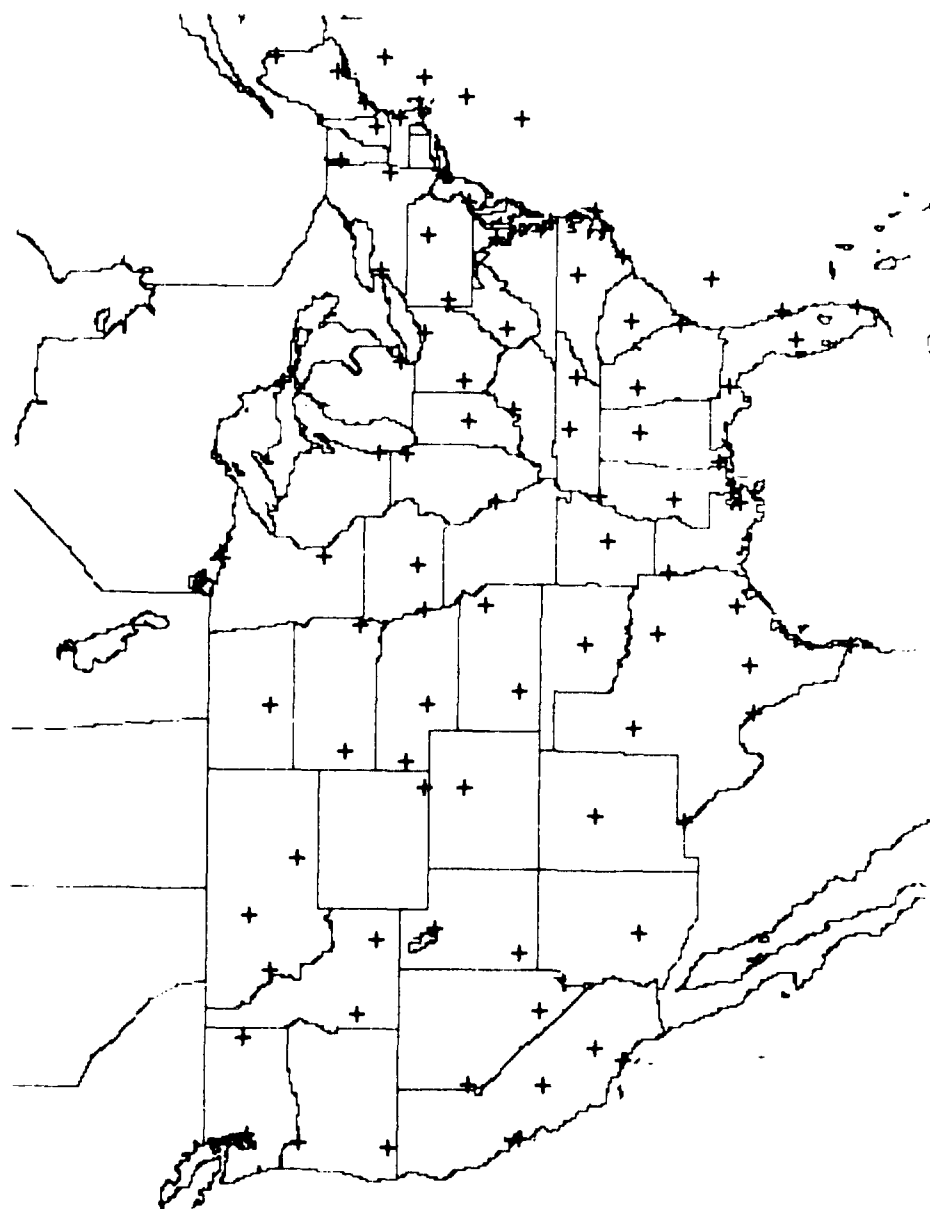


Fig. 2. Geographic Locations of the 89 FOCUS 60-78 Stations

station were obtained from the appropriate location in the FOUS 60-78 bulletin. Verification values were obtained from the initial conditions in the relevant successive FOUS 60-78 bulletins that were derived from the model initial analysis. Parameter forecast errors were accumulated at each of the 89 locations by simply subtracting the verification value from a parameter forecast value. Since verification values are available on a 12-hour cycle, only the 12, 24, 36, and 48 hour forecast errors have been analyzed. This procedure was applied to forecasts beginning on November 23, 1982 through March 31, 1983. By calculating the forecast errors in the manner described, a positive number indicates an overforecast (i.e., the forecast was too warm, too wet, pressure too high, etc.) while a negative number indicates an underforecast (too cold, too dry, pressure too low).

Errors at each of the 89 stations were further stratified into error sums from 00 GMT model runs and 12 GMT model runs. This stratification was selected to examine the sensitivity of the error fields to diurnal variations. For computational simplicity we required the availability of all variables being evaluated at all 89 FOUS stations for the pertinent forecast and verification periods. If any data were missing from a forecast or verification series the whole set was deleted. Loss of FOUS data was often caused by local computer failure or by incomplete transmission of the FOUS bulletins from Kansas City. Since these occurrences were more or less random it is believed that little or no sampling bias was introduced. Table 1 shows the total number of forecast sets that made up the final sums of the forecast errors. The average errors (bias) for each FOUS station, for each forecast parameter at the specified forecast periods were then determined. For 00 GMT

TABLE 1. NUMBER OF FORECASTS USED TO CALCULATE AVERAGE LFM-II PARAMETER FORECAST ERRORS

	00 GMT	12 GMT
12-hour forecasts	73 (57%)	76 (59%)
24-hour forecasts	72 (56%)	69 (53%)
36-hour forecasts	65 (51%)	65 (50%)
48-hour forecasts	61 (48%)	60 (47%)

there were 128 possible forecasts while at 12 GMT there were 129 possible

forecasts. Numbers in parentheses indicate the percentage of possible forecasts used to calculate the forecast errors.

In order to determine the errors for the boundary layer winds, forecast and verification winds were first separated into zonal and meridional components. The errors were then computed for each wind component individually prior to recombining them to arrive at the final average vector error.

### 3. Results

Graphic displays (contoured maps) of the average error fields were generated using the McIDAS upper air objective analysis routine. McIDAS uses a two-pass technique based on Barnes (1964)<sup>9</sup> to generate regularly spaced grid point values from which isolines (contours) are determined. Details on the capabilities of McIDAS can be found in Suomi et al. (1983),<sup>10</sup> Smith (1975),<sup>11</sup> and Wash and Whittaker (1980).<sup>12</sup>

Figs. 3 and 4 show the results of the error analysis for sea-level pressure in tenths of millibars for 00 GMT and 12 GMT, respectively. Consistent with the earlier study of Silberberg and Bosart (1982),<sup>5</sup> the model tends to underforecast sea-level pressure east of the Rockies and overforecast sea-level pressure west of the Rockies with the sharpest error gradient along or just east of the mountain chain. The largest negative errors occur slightly east of the upper Mississippi Valley while the largest positive errors occur in the southwestern United States. The maximum absolute error grows with forecast length, being everywhere less than two mb at 12 hours, but exceeding five mb by 48 hours. While the general features of the average error fields for both model initial times are the same, the 12 GMT model runs tend to underforecast pressure relative to 00 GMT model runs nationwide for all forecast periods.

9. Barnes, S.L., 1964: A technique for maximizing details in numerical weather map analysis. J. Appl. Meteor., 3, 396-409.

10. Suomi, V.E., R. Fox, S.S. Limaye, and W.L. Smith, 1983: McIDAS III: A modern interactive data access and analysis system. J. Appl. Meteor., 22, 766-778.

11. Smith, E.A., 1975: The McIDAS system. IEEE Trans. Geosci. Elec., GE-13, 123-136.

12. Wash, C.H., and T.M. Whittaker, 1980: Subsynchronous analysis and forecasting with an interactive computer system. Bull. Am. Meteorol. Soc., 61, 1584-1591.

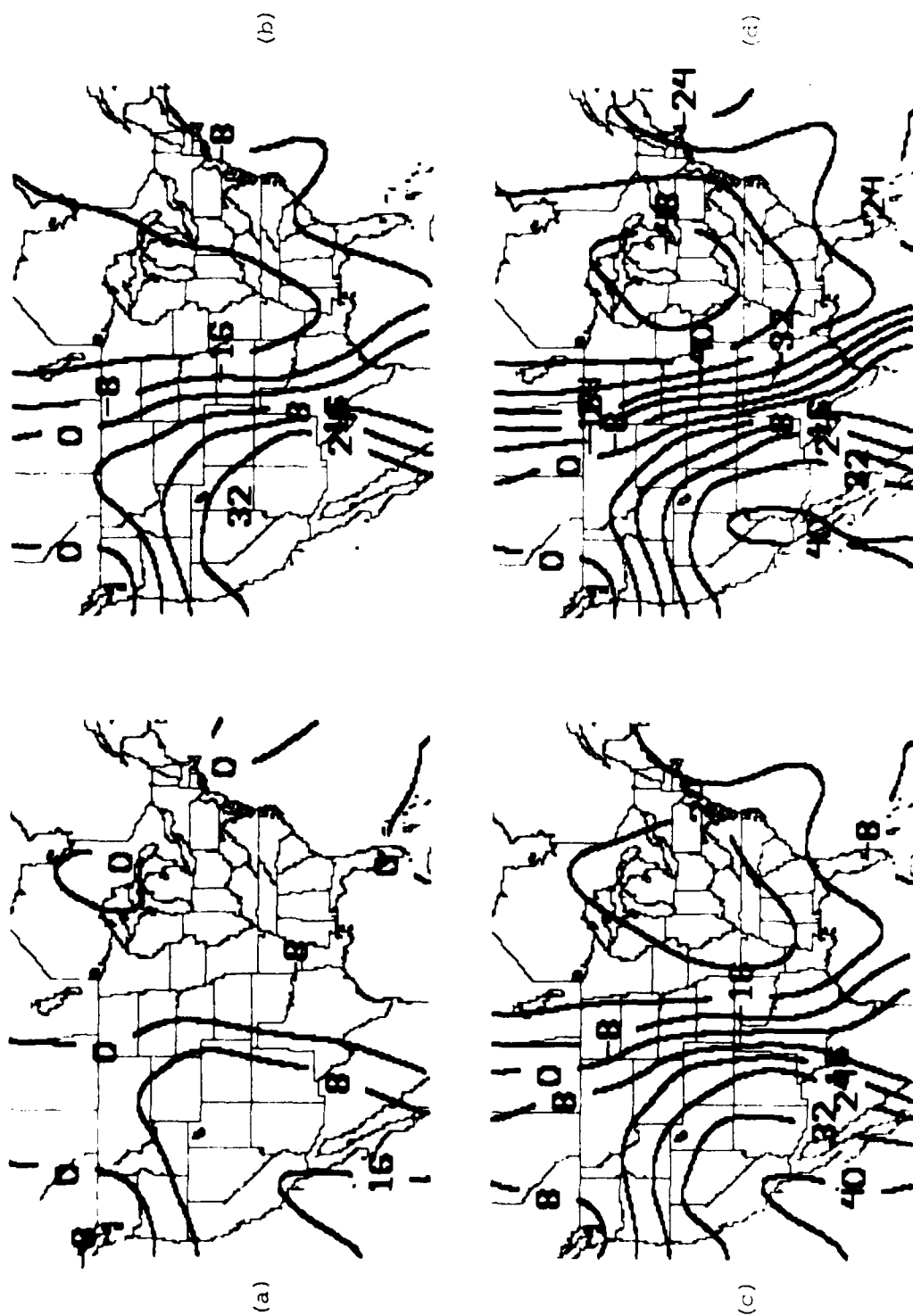


Fig. 3. Average LFM Forecast Sea-Level Pressure Error (tenths of mb) from 00 GMT Model Runs: (a) 12-hour forecasts; (b) 24-hour forecasts; (c) 36-hour forecasts; (d) 48-hour forecasts

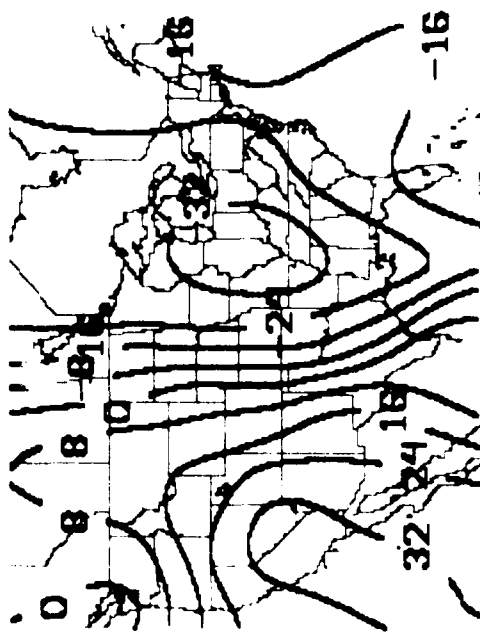
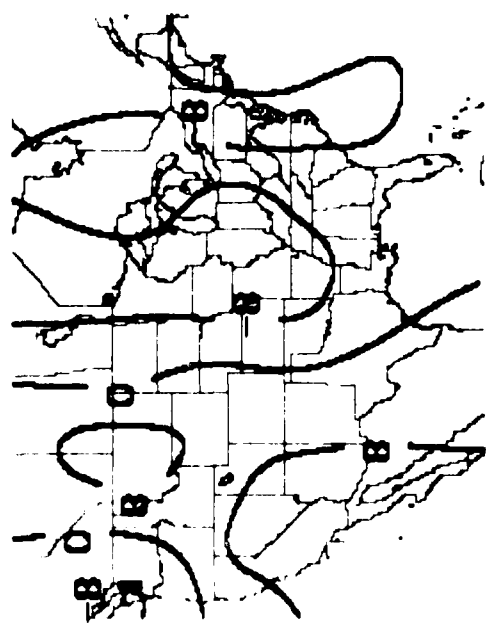
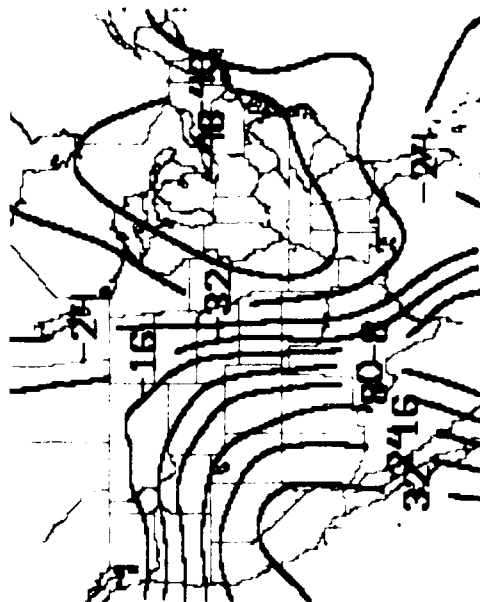
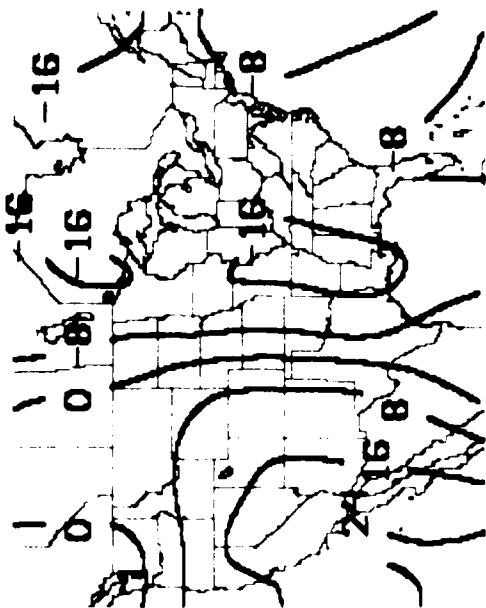


Fig. 4. Same as Fig. 3 Except 12 GMT Model Runs

Average 1000-500 mb thickness errors (in meters) are shown in Figs. 5 and 6. Thickness errors are generally small but show less coherence than the sea-level pressure error fields. For most periods thickness is forecast too high (warm) over the central United States and too low (cold) over most of the Rocky Mountain states. Distinct differences can be seen in the mean error patterns for the 00 and 12 GMT model runs particularly at 12 and 36 hours. For example, the relatively large positive 1000-500 mb thickness errors apparent in the 36-hour forecasts at 00 GMT over Texas are replaced by small negative thickness errors at 12 GMT. The magnitudes of the thickness errors are not well correlated with length of forecast since the largest positive errors occur in the 00 GMT 36-hour forecast while the largest negative errors occur in the 12 GMT 36-hour forecast. The maximum 1000-500 mb thickness errors are 20-25 meters, which are equivalent to mean virtual temperature errors of approximately  $1^{\circ}\text{C}$ .

Errors in the mean relative humidity of the model (approximately 1000-490 mb) are depicted in Figs. 7 and 8. Positive errors (i.e., too wet) that have a tendency to grow with forecast length are apparent in the northern Rocky Mountain states, particularly Montana and Idaho, for both model initial times. A dry bias that grows with forecast length at both 00 GMT and 12 GMT is noted around the Texas and Florida Gulf Coast. The dry tongue shows a tendency to expand northward from Texas, especially at the 36- and 48-hour forecast periods.

The most striking feature of the error fields for boundary layer relative humidity (Figs. 9 and 10) is a widespread tendency for the model to be too dry. A notable exception is the 12-hour forecast from the 12 GMT run in which all of the U.S. except the extreme southeast is forecast too wet. There is also a tendency for the Rocky Mountain region to be too wet at some forecast intervals for each of the model initial times. However, the magnitude of those errors is smaller than those found in the same periods for mean relative humidity (see Figs. 7 and 8). Around the Texas Gulf Coast the model underforecasts the boundary layer relative humidity by as much as 20-30 percent at the 12 GMT verification times. There also seems to be an overall diurnal effect since the largest errors in the 00 GMT runs are at 12 and 36 hours, while the largest errors in the 12 GMT runs are at 24 and 48 hours.

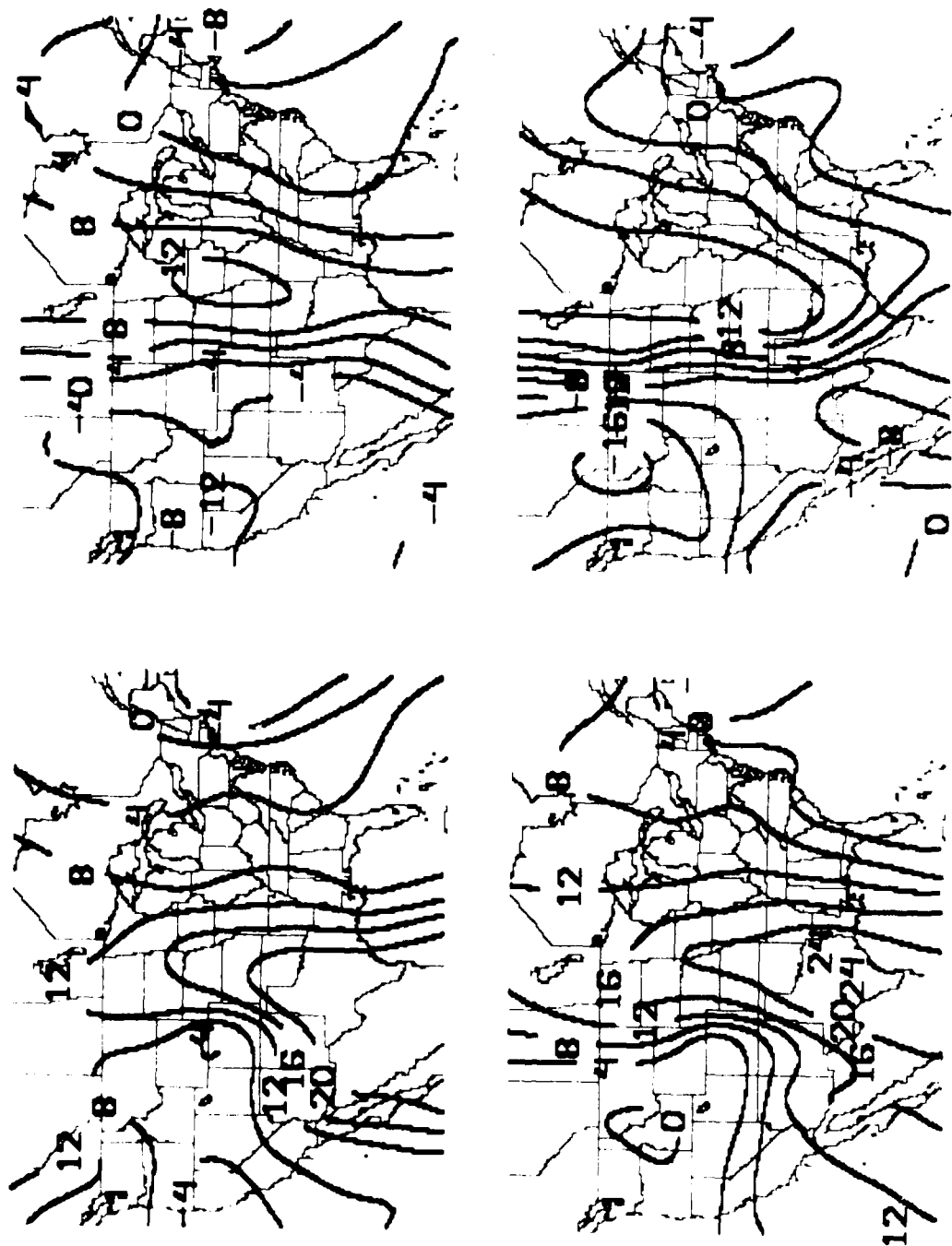


Fig. 5. Average LFM Forecast 1000-500 mb Thickness Error (meters) from 00 GMT Model Runs: (a) 12-hour forecasts; (b) 24-hour forecasts; (c) 36-hour forecasts; (d) 48-hour forecasts

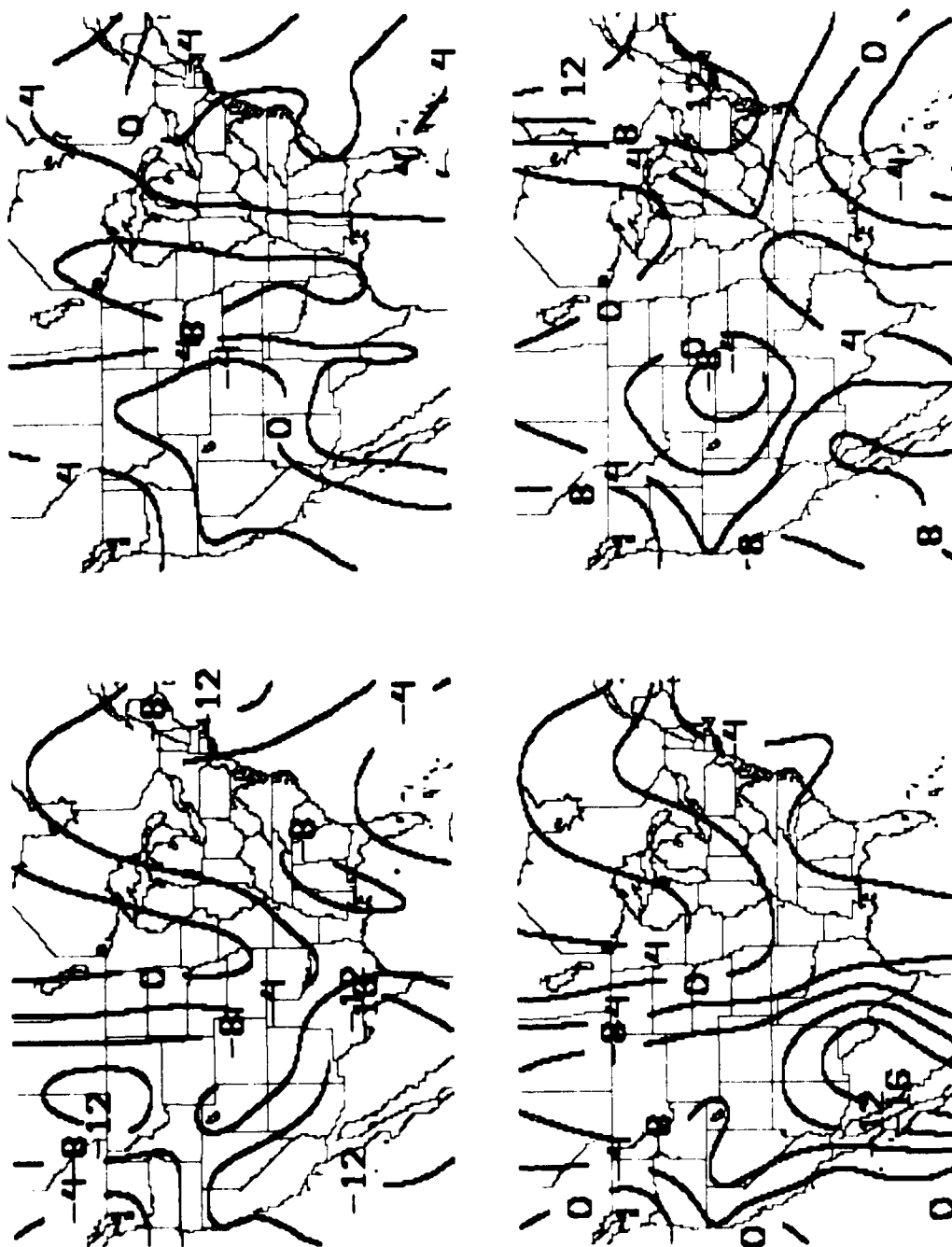


Fig. 6. Same as Fig. 5 Except 12 GMT Model Runs

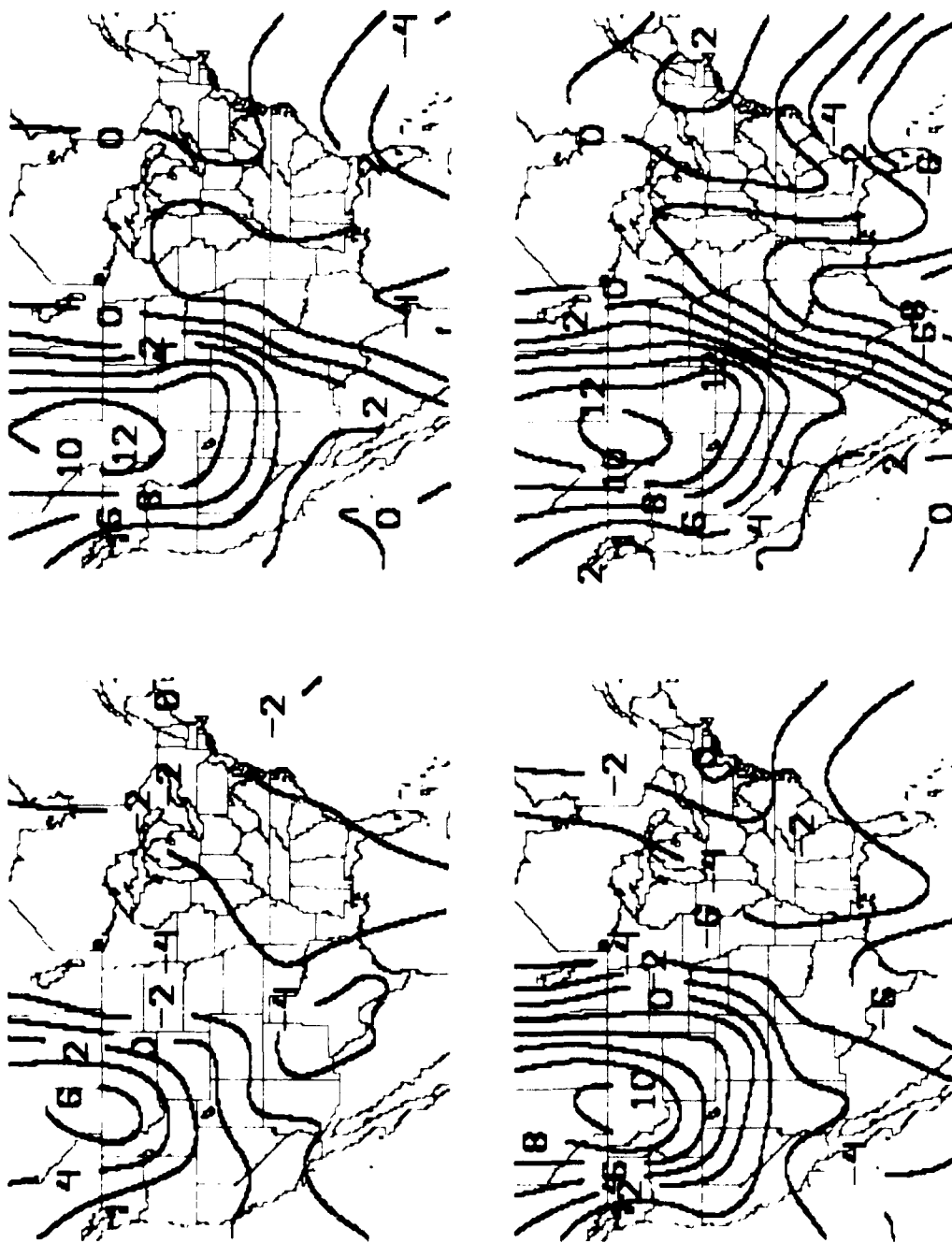


Fig. 7. Average LFM Forecast Mean Relative Humidity Error (whole percent) from 00 GMT Model Runs: (a) 12-hour forecasts; (b) 24-hour forecasts; (c) 36-hour forecasts; (d) 48-hour forecasts

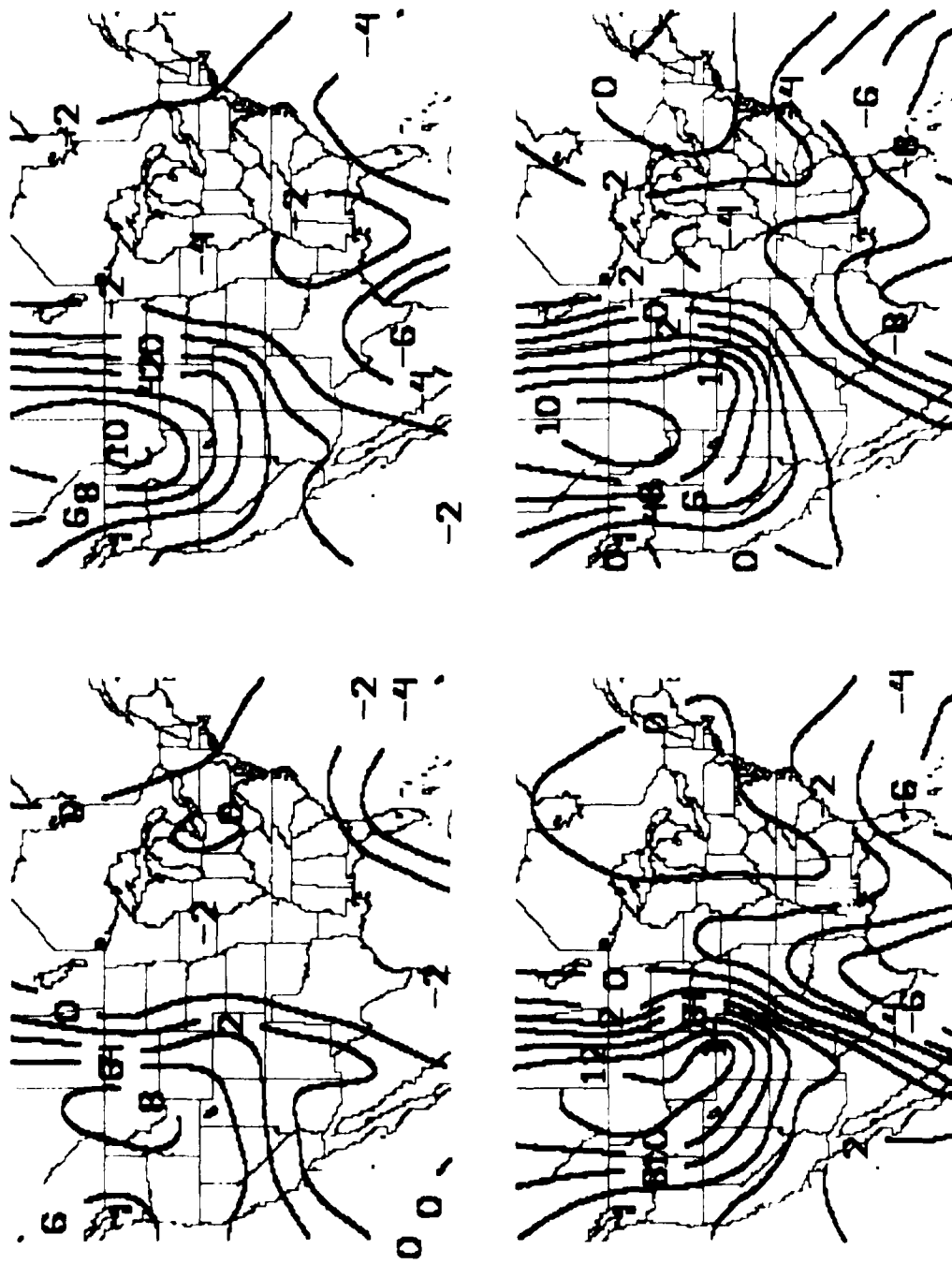


Fig. 8. Same as Fig. 7 Except 12 GMT Model Runs

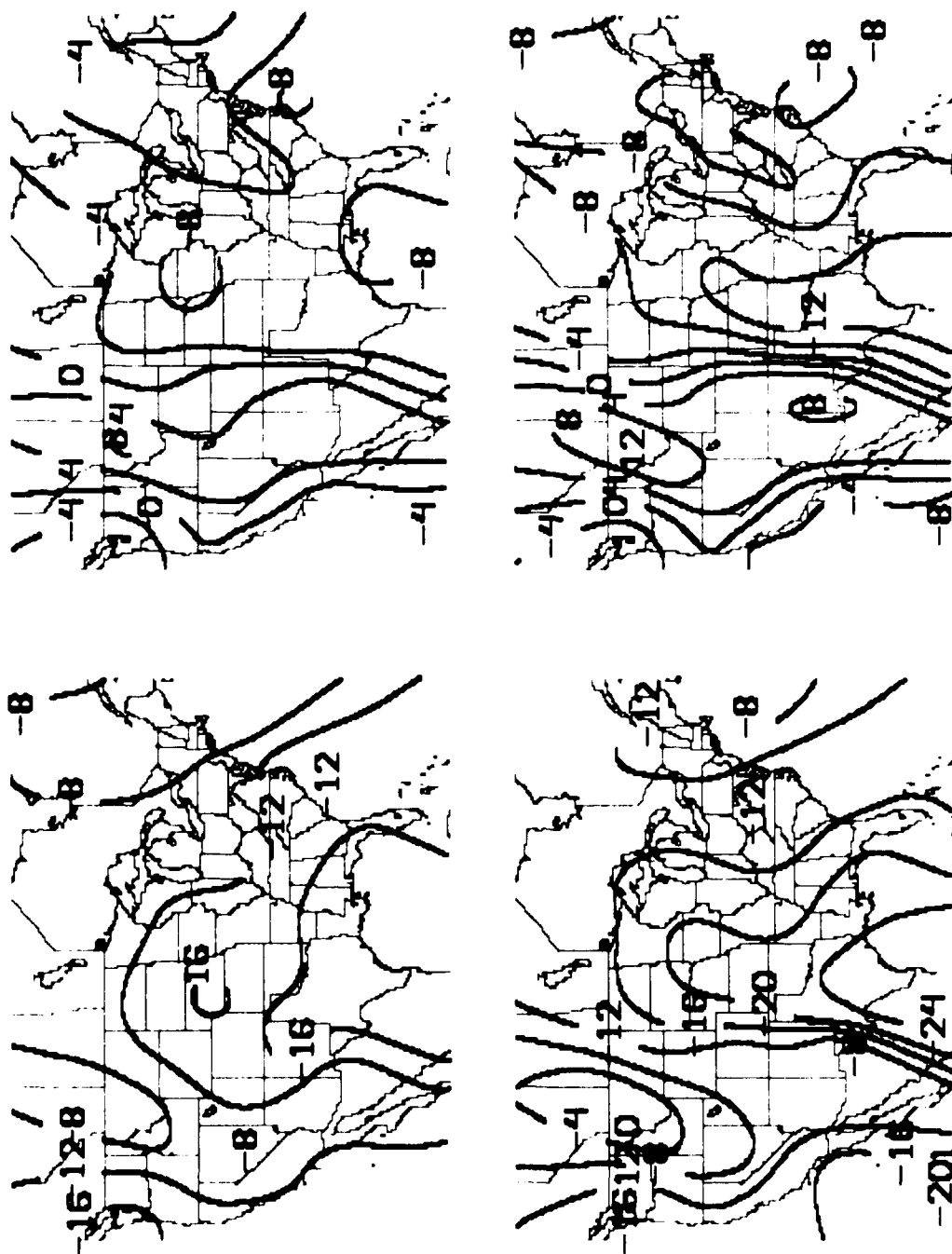


Fig. 9. Average LFM Forecast Boundary Layer Relative Humidity Error (whole percent) from 00 GMT Model Runs: (a) 12-hour forecasts; (b) 24-hour forecasts; (c) 36-hour forecasts; (d) 48-hour forecasts

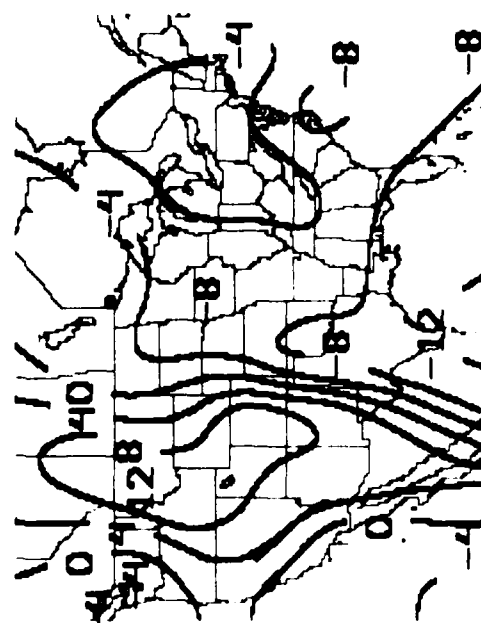
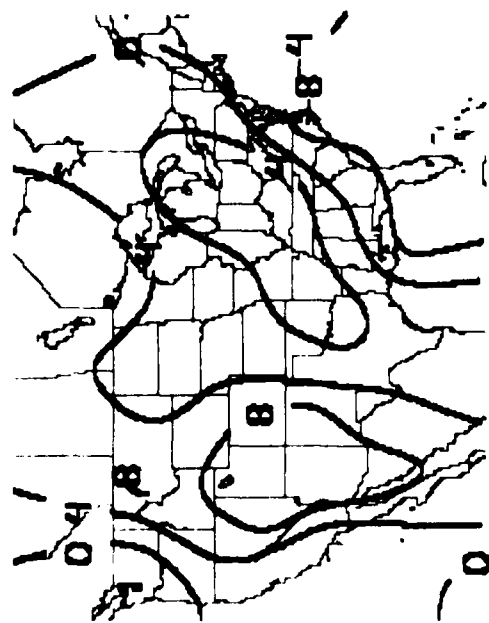
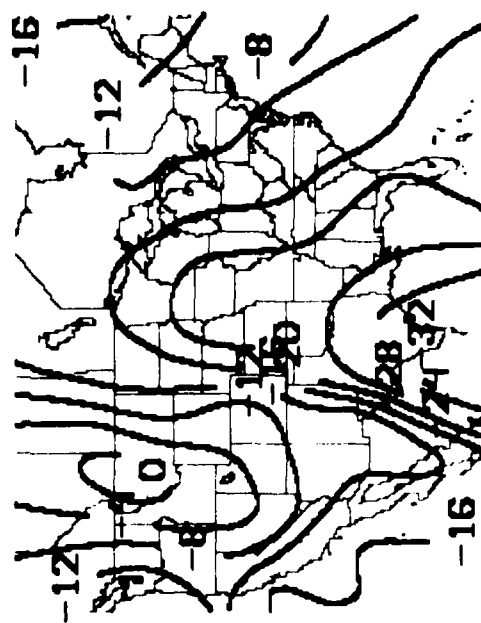
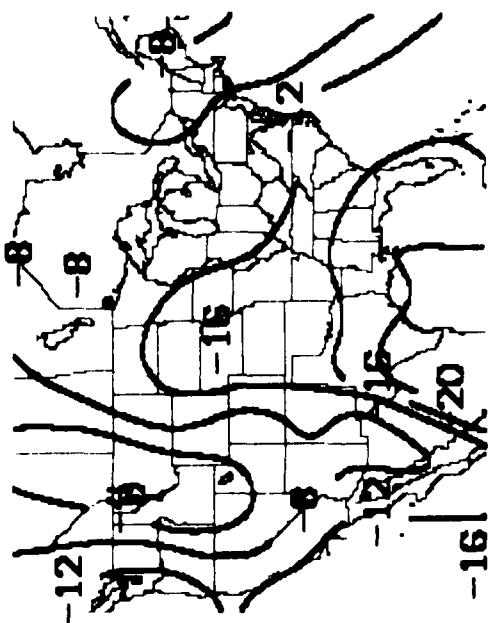


Fig. 10. Same as Fig. 9 Except 12 GMT Model Runs

The results of the error analysis for the mean potential temperature of the boundary layer are depicted in Figs. 11 and 12. A warm (over-forecasting) bias is apparent at Florida in all cases and along the rest of the Gulf coast except for the 12- and 36-hour forecasts of the 12 GMT run. A cold bias is generally found in the Great Lakes region wedging southward along the Ohio and Mississippi River axes. This feature is absent only in the 12-hour forecast from the 00 GMT model run. While the general features of the error fields for boundary layer temperature are the same for both 00 GMT and 12 GMT, the times when the maximum absolute errors occur are out of phase by 12 hours. This can be attributed to the influence of diurnal variations. The greatest positive errors for boundary layer temperature occur at the 12 GMT verification times. These time periods also show the development of a warm bias in the Rockies while at 00 GMT validation time the bias is around zero or even slightly negative.

Wind flags representing the vector errors in boundary layer wind forecasts are shown in Figs. 13 and 14. A wind error flag from the north-east indicates that the model forecast for that station had too much of a northerly and easterly component. These results suggest a broad cyclonic "error" circulation in the eastern half of the U.S. and too much of a northerly component in the nation's mid-section. Along the west and Gulf Coasts there is a tendency for too much offshore flow. No striking differences are apparent between the 00 GMT and 12 GMT runs. In general, the calculated average vector wind errors are small, although some stations (Portland and Medford, Oregon, for example) show vector wind errors of up to 8-11 knots.

#### 4. Discussion

Figs. 3-14 clearly demonstrate that the LFM-II has some bias in all of the forecast parameters examined in this study. The rather large magnitude of the errors in sea-level pressure forecasts (Figs. 3 and 4), especially at 36 and 48 hours, was rather surprising since essentially continuous sampling was used to build the data sample, rather than focusing on cyclonic or anticyclonic regimes. The large negative errors around the Great Lakes are consistent with the results of Silberberg and Bosart (1982)<sup>5</sup> who found that the LFM-II systematically overpredicted the intensity (i.e.,

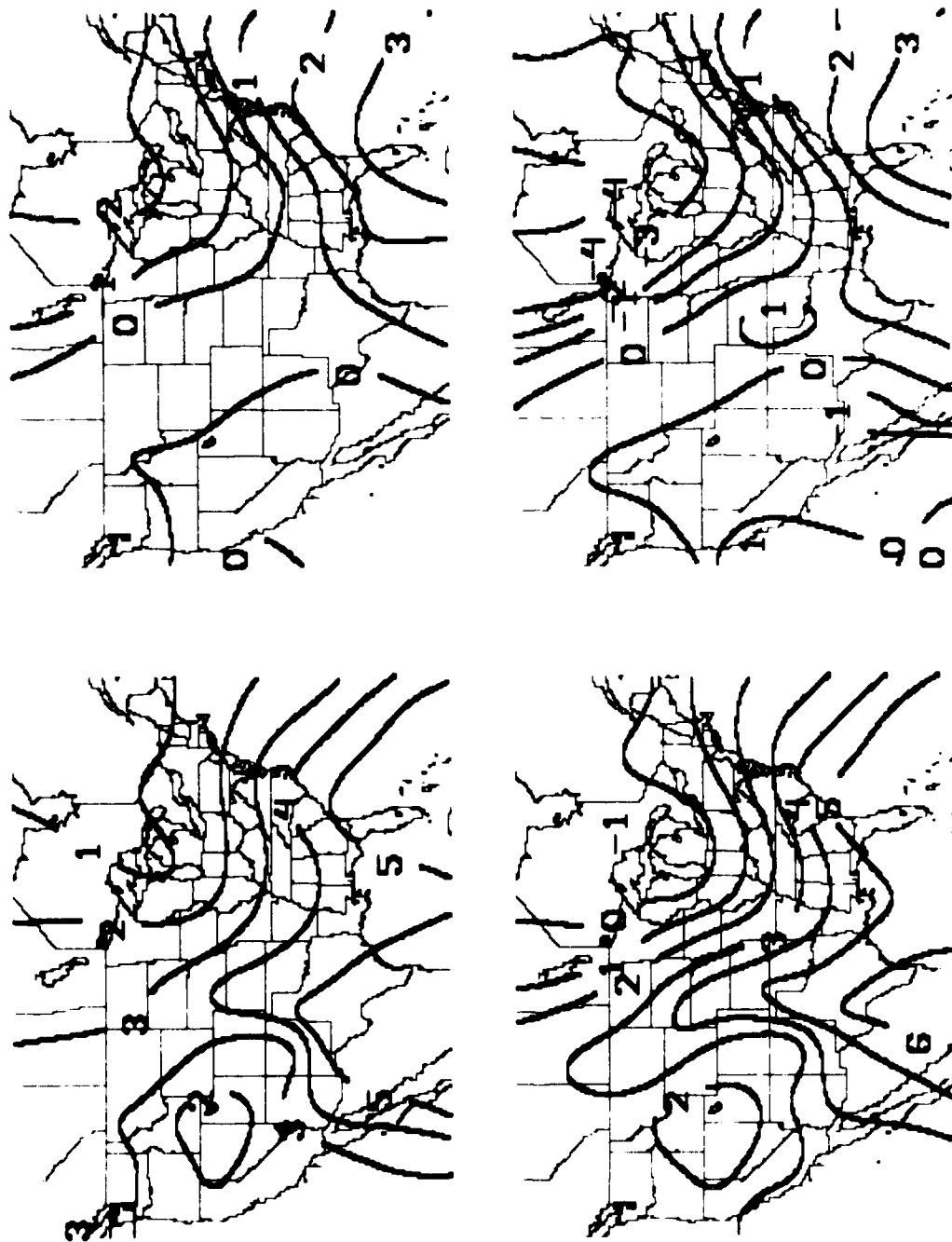


Fig. 11. Average LFM Forecast Boundary Layer Temperature Error ( $^{\circ}\text{K}$ ) from 00 GMT Model Runs: (a) 12-hour forecasts; (b) 24-hour forecasts; (c) 36-hour forecasts; (d) 48-hour forecasts

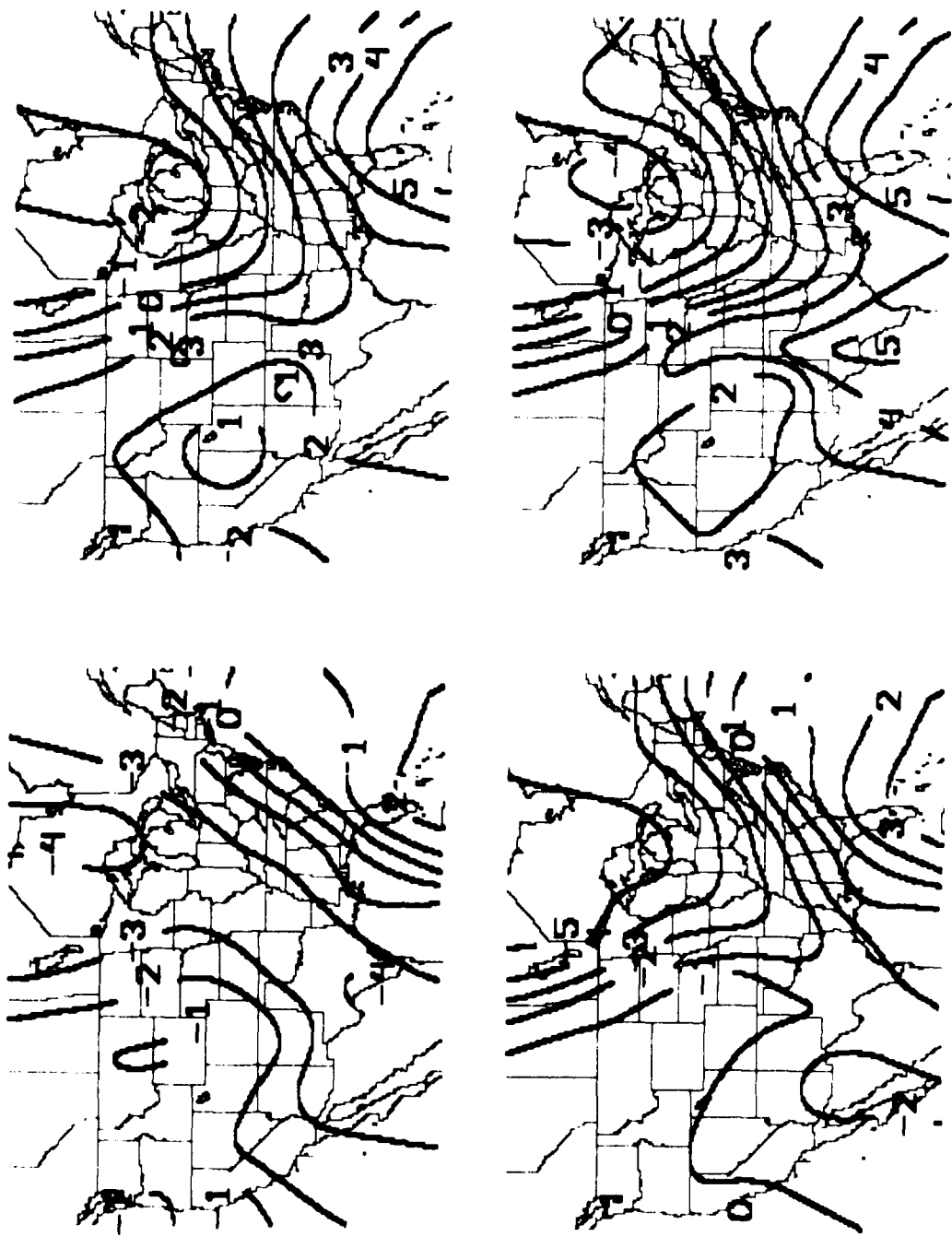


Fig. 12. Same as Fig. 11 Except 12 GMT Model Runs

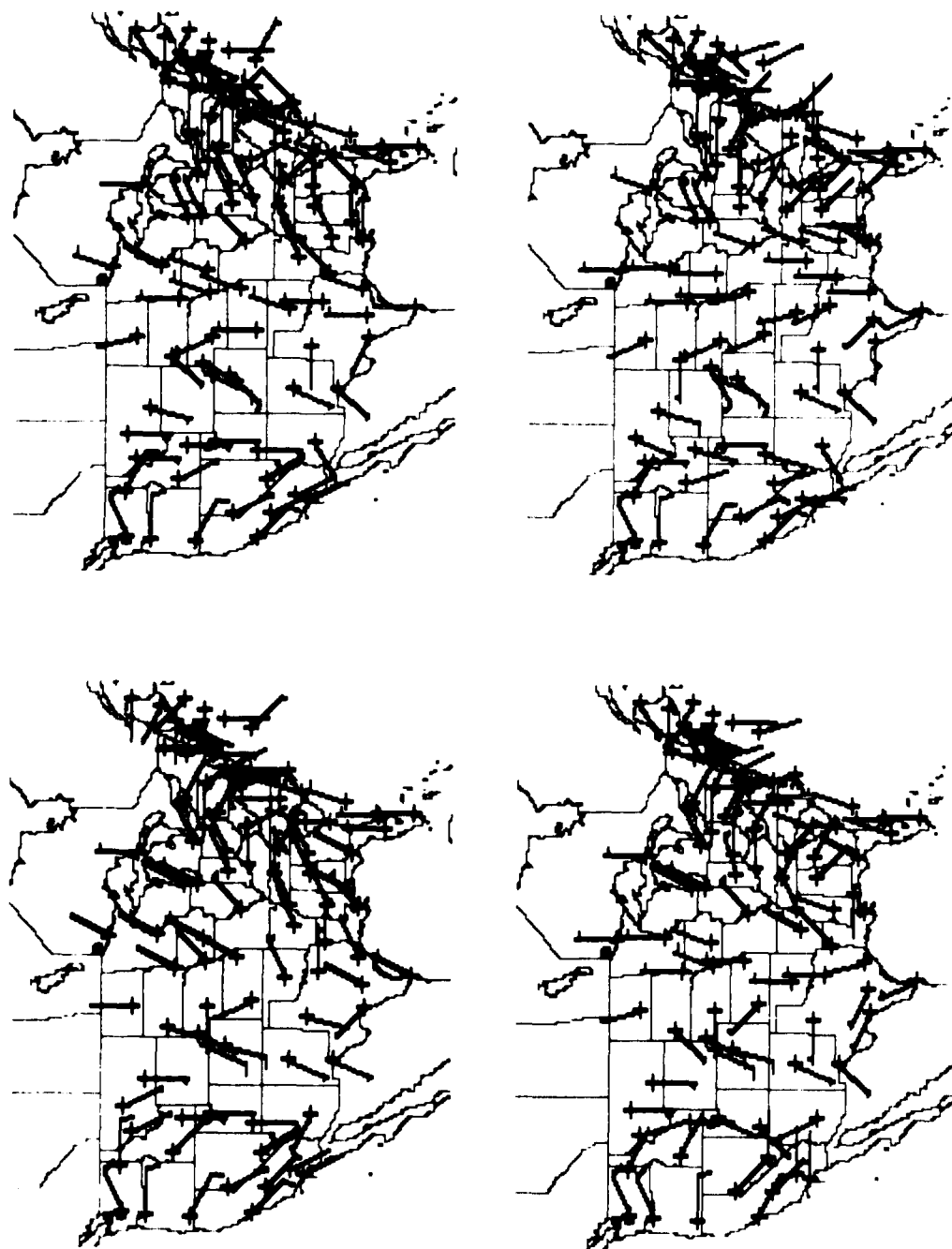


Fig. 13. Wind Flags Representing Average LFM Forecast Boundary Layer Wind Vector Error (speed in knots) from 00 GMT Model Runs: (a) 12-hour forecasts; (b) 24-hour forecasts; (c) 36-hour forecasts; (d) 48-hour forecasts

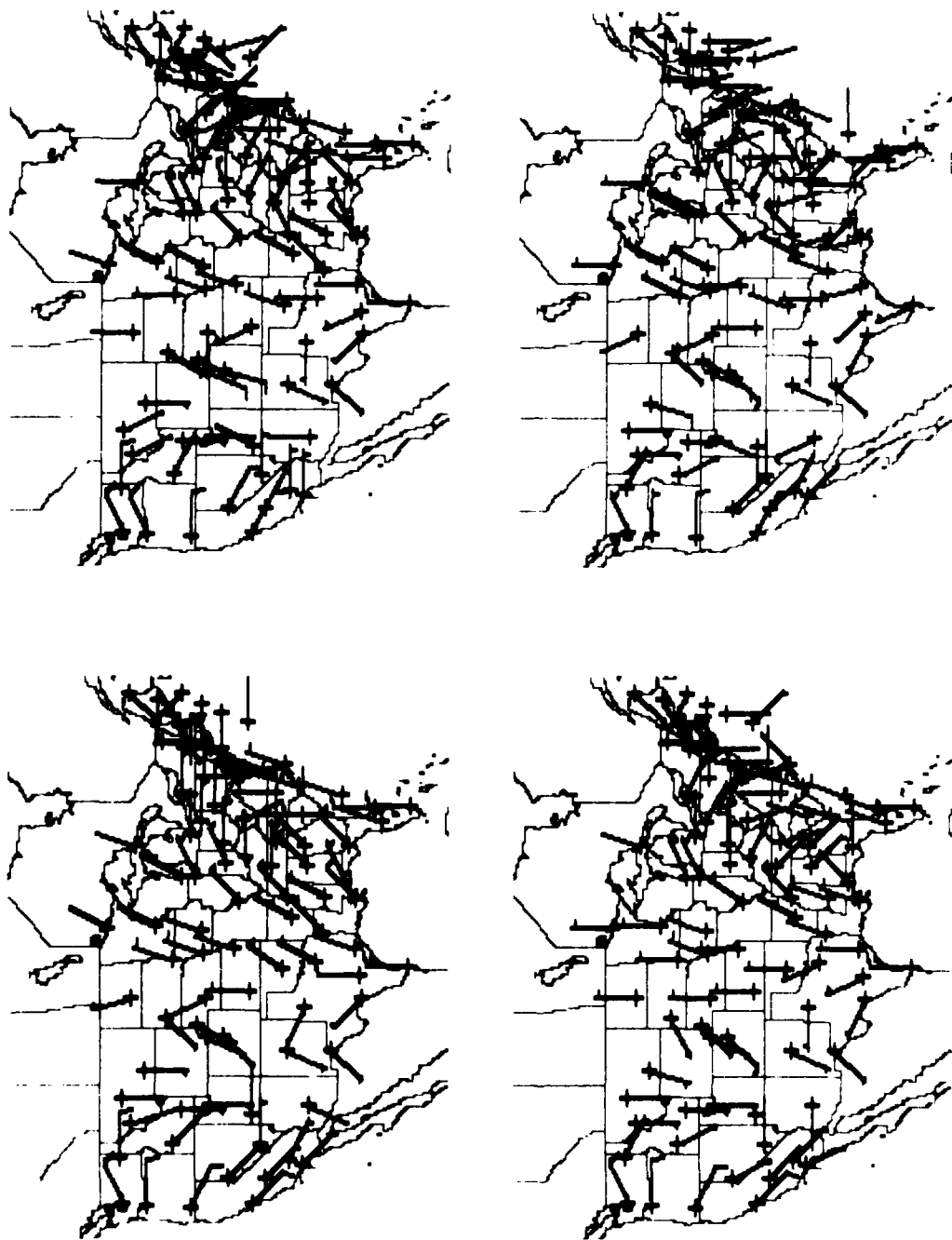


Fig. 14. Same as Fig. 13 Except 12 GMT Model Runs

predicted surface pressures too low) of cyclones east of the Continental Divide during the 1978-79 winter season. It seems the LFM-II frequently overdeepened and/or was too fast with cyclones moving out from the lee of the Rockies towards the Great Lakes. Errors of -20 to -25 mb in 48-hour forecasts were recorded at stations in Ohio, Kentucky, Wisconsin, and Illinois during the period under study. The large positive errors in the West may be partially attributable to inadequate initial analysis of storm intensity in the eastern Pacific. Maximum positive errors are found in California and Nevada, reflecting the large southern displacement of the jet stream during much of the 1982-83 winter. Cyclones frequently moved onshore from the Pacific deeper and farther south than usual. Fig. 15 shows the 36-hour sea-level pressure forecast and verifying analysis valid at 00 GMT on December 23, 1982, for the southwestern United States as contoured by McIDAS. Note the strong cyclone in the analysis centered near Reno (indicated by small square) and the poor forecast produced by the LFM. Positive errors of 24 mb and 21 mb were recorded at Reno and San Francisco, respectively. Even the 24-hour forecast at 00 GMT on December 22, 1983 (not shown) resulted in a large area of positive 8-10 mb errors in this region.

Looking at Figs. 3 and 4, and 13 and 14, the sea-level pressure error analyses are clearly reflected in the vector errors for boundary layer wind. An overprediction of a northerly component in the nation's mid-section is consistent with the geostrophic flow defined by low pressure in the East and high pressure in the West. The weak cyclonic circulation in the east-central United States seen in Figs. 13 and 14 is also consistent with the analyzed errors in the pressure forecasts. It can be inferred from the depicted winds that the LFM-II is overforecasting low level convergence in the lower Mississippi Valley, which in turn would adversely impact vertical velocity computations.

In order to make some inferences about geopotential height forecast bias at 500 mb, it is necessary to combine the results of the error analysis for sea-level pressure with those for 1000-500 mb thickness. Judging from Figs. 3 and 4, and 5 and 6, it would seem that the error patterns are fairly consistent; i.e., areas that show that the surface pressure is forecast too low also show that 1000-500 mb thickness is forecast slightly

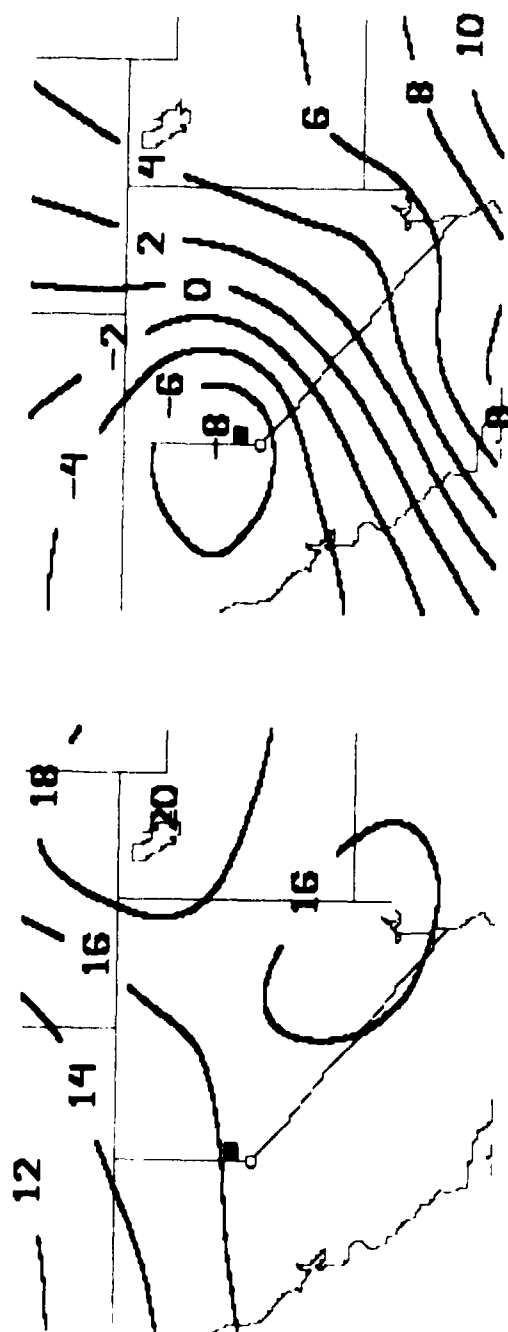


Fig. 15. (a) 36-hour LFM Forecast of Sea-Level Pressure from 12 GMT December 21, 1982, Valid 00 GMT December 23, 1982; (b) LFM Sea-Level Pressure Analysis Valid 00 GMT December 23, 1982  
(Isobars are labelled as deviations from 1000 mb)

too high. This is qualitatively what one would expect if the 500 mb height forecasts were accurate. However, if we use the approximation that a one mb pressure change represents a height change of eight meters (a reasonable estimate at least in the lower troposphere (Wallace and Hobbs, 1977<sup>13</sup>)), it soon becomes apparent quantitatively that the surface pressure errors are larger than the corresponding 1000-500 mb thickness errors. For example, a four mb pressure error would translate into a thickness error of about 32 meters. As can be seen in Figs. 5 and 6 the thickness errors never get that large. This then implies that the 500 mb height errors would have a similar pattern (although the errors are probably smaller in magnitude) to the sea-level pressure errors. As a result the 500 mb height forecast would on average be too low in the East and too high in the West.

Comparing Figs. 5 and 6 to Figs. 11 and 12 reveals several areas in which the error fields for boundary layer temperature and 1000-500 mb thickness, when viewed together, imply a temperature bias above the boundary layer. That is, there are geographic areas in which a warm (cold) bias for boundary layer temperature coincides with regions with negative (positive) bias in thickness. When this occurs it can be inferred that the model's mean temperature above the boundary layer must be compensating in order to maintain hydrostatic balance. However, the compensation will profoundly affect the model's estimation of static stability. As an example, the 36- and 48-hour forecasts from 12 GMT overforecast boundary layer temperature but underforecast 1000-500 mb thickness over Florida. The resulting "cold pool" above the boundary layer near Florida would lead to an underestimation of the static stability in this region and perhaps an overprediction of convective activity. The net effect on static stability when the sign of the bias is the same in the boundary layer and between 1000 and 500 mb is more difficult to determine because the effect of one bias opposes the other.

Errors in boundary layer temperature do not grow monotonically with forecast length as they do for sea-level pressure. Rather, the largest positive errors occur at 12 GMT verification time while the largest negative

13. Wallace, J.M., and P.V. Hobbs, 1977: Atmospheric Science: An Introductory Survey. Academic Press Inc., New York, NY, 467 pp.

errors occur at 00 GMT. An explanation for this apparent anomaly is put forth by Gerrity (1977):<sup>14</sup> the diurnal variations of errors in boundary layer temperature are attributed basically to the fact that long wave radiative cooling within the model is treated as proceeding at a constant rate, which is evidently too simple a parameterization. In any event, the strongest warm (cold) bias for boundary layer temperature in the model tends to occur at the time of day when the real atmosphere is diurnally coldest (warmest).

The strong influence of the diurnal cycle on bias of boundary layer temperature within the LFM-II in turn profoundly affects the boundary layer relative humidity error fields. When the real atmosphere reaches its diurnal minimum temperature (around 12 GMT) the model errors in boundary layer relative humidity show the largest dry bias. This is illustrated by the error results for forecasts verifying at 12 GMT (see Figs. 9 and 10). Since the relative humidity is strongly dependent on temperature, the over-forecasting of boundary layer temperature contributes directly to the enhanced dry bias. Further evidence to support this hypothesis can be found when it is noted that at the 12 GMT verification time (when the real atmosphere will be about the coldest) the moist bias usually found over the Rockies almost completely disappears. This disappearance of the moist bias can probably be attributed to the enhancement of the dry bias which is caused by the enhancement of the warm bias.

Even neglecting the diurnal temperature effect, the model appears to be too dry in the boundary layer for the major portion of the nation with the exception of the Rockies. The problem is most acute around the Gulf Coast but extends up into the central plains. Although an analysis of the boundary layer water budget has not been conducted, it would appear that the treatment of moisture flux parameterization from water bodies to the model atmosphere may need refinement. Caution must be exercised, however, since absorption of short wave radiation in the model is parameterized by the amount of water vapor present (Gerrity, 1977),<sup>14</sup> and the addition of more moisture could serve to heat up the already "hot" boundary layer.

14. Gerrity, J.F., 1977: The LFM Model-1976: A Documentation. NOAA Tech. Memo NWS NMC-60.

No diurnal variations are apparent in the error fields for mean relative humidity (Figs. 7 and 8). While the gross features are essentially the same as for boundary layer relative humidity, the magnitude of the errors in the mean relative humidity are much smaller in areas in which the model shows a dry bias and about the same or even slightly larger in regions where the model shows a moist bias in the boundary layer. This implies that, if anything, the layers above the boundary layer tend to have a more widespread moist bias across most of the U.S.

A puzzling aspect of the results for relative humidity in both the boundary layer and for the model as a whole is the pervasive moist bias located over the northern Rockies. Whether this feature is attributable to poor analysis of storms arriving from the Pacific, is in some way caused by terrain considerations, or results from some combination of both factors is unclear.

## 5. Conclusions

An analysis has been conducted to identify persistent error (bias) in certain forecasts produced by the NMC LFM-II model. Forecasts were examined from November 23, 1982 to March 31, 1983 and were stratified into 00 and 12 GMT initial analysis times. The results show obvious systematic errors in the model, some of which are startling in their magnitude. Sea-level pressure is, on average, forecast too low east of the Rockies and too high west of the Rockies and the magnitude of the errors increases with forecast length. The pressure error is reflected in the boundary layer wind forecasts with the tendency to forecast too much of a northerly component in the nation's mid-section. Bias for forecasts of 1000-500 mb thickness is relatively small, with the model too high (warm) in the central U.S. and too low (cold) in the Rockies. Model relative humidity is generally forecast higher than observed in the northern Rockies and drier than observed along the Gulf Coast.

The bias for boundary layer temperature and for boundary layer relative humidity (and to a lesser degree 1000-500 mb thickness) seems to be composed of two components, one of which varies diurnally, the other which grows with forecast length. This is demonstrated by the differences in the

error fields at 00 and 12 GMT. In general the model is too warm along the Gulf Coast and too cold in the vicinity of the Great Lakes, but the positive errors (warm bias) are enhanced at times when the real atmosphere is diurnally cold (i.e., forecasts that verify at 12 GMT) while the negative errors (cold bias) are enhanced at times when the real atmosphere is diurnally warm (00 GMT). This effect, which is probably caused by inadequate parameterization of long wave cooling within the model, is enough to mask the expected increase of forecast error with forecast length that is seen in some of the other parameters examined. The boundary layer temperature errors also contribute to enhanced errors in the boundary layer relative humidity. The Gulf Coast, which is forecast too warm, is also forecast too dry. The real problem may not be one of predicting too little moisture (although this is true as well) as much as the fact that given the same amount of water vapor, the calculated relative humidity will be lower at higher temperatures. Due to the complexity of the atmosphere and its models, the physical links between various error fields are difficult to isolate, but tracking them down may help in developing future models. It is hoped that identification of errors such as these may give clues as to ways to improve the physics on which the model is based and thereby reduce systematic forecast bias.

## B. Forecast Guidance Displays

To help evaluate the utility of a technique designed to monitor the short range (6-12 hour) accuracy of LFM model output, a modified version of the Estimated Satellite Product (ESP)<sup>15</sup> was added to the MOS module of programs FOSCON and FAMAIN.<sup>7</sup> The ESP value is based on MOS forecasts of total cloud amount and ceiling height and can be used in conjunction with satellite data to monitor the forecast movement of cloud fields. Table 2 shows the ESP values assigned for the given combination of cloud and ceiling categories.

TABLE 2. SPECIFICATION OF ESP VALUES

Cloud Category	Ceiling Category (ft)	ESP
CLEAR/SCT	-	0
BKN	>7500	1
OVC	>7500	2
BKN	3000 - 7500	3
OVC	3000 - 7500	4
BKN	1000 - 2900	5
OVC	1000 - 2900	6
BKN	500 - 900	7
OVC	500 - 900	8
BKN/OVC	≤400	9

CLEAR = 0 - 1 tenths

SCATTERED (SCT) = 2 - 5 tenths

BROKEN (BKN) = 6 - 9 tenths

OVERCAST (OVC) = 10 tenths

15. Reinke, D.L., and T.H. Vonder Haar, 1982. Weather Forecast Adjustment Using Model Output Cloud Fields and Digital Satellite Data. AFGL-TR-82-0168, Contract F19628-78-C-0207, Colorado State University, Ft. Collins, CO, pp. 9-28.

A typical plot of ESP values for MOS stations in the northeastern U.S. is depicted in Fig. 16 while Fig. 17 shows an ESP analysis for the entire U.S. By laying an ESP forecast analysis over a coincident satellite photograph it is possible to make some inferences about the performance of the MOS (and therefore the LFM model) forecasts. For example, in Fig. 18 the 6-hour ESP forecast verifying at 18 GMT on November 21, 1983, is superimposed over the 18 GMT visible satellite imagery. In most areas where the visible imagery shows essentially cloud free conditions the ESP value is less than one, which indicates clear or scattered clouds and hence a good forecast. When the ESP values and cloud imagery are not consistent, model error is indicated and it would be possible to adjust the model cloud (and other related parameter) forecasts to take into account the detected error. Such errors could arise, for example, from the model being too fast (or slow) in moving a cyclone or failing to predict properly the intensity of the system. Caution must be exercised, however, to insure that the discrepancy uncovered is caused by model error and is not due to a poor ESP analysis resulting from lack of MOS data. Once a particular model error is uncovered the longer range forecasts (12-48 hours) can be adjusted in an effort to improve these subsequent forecasts.



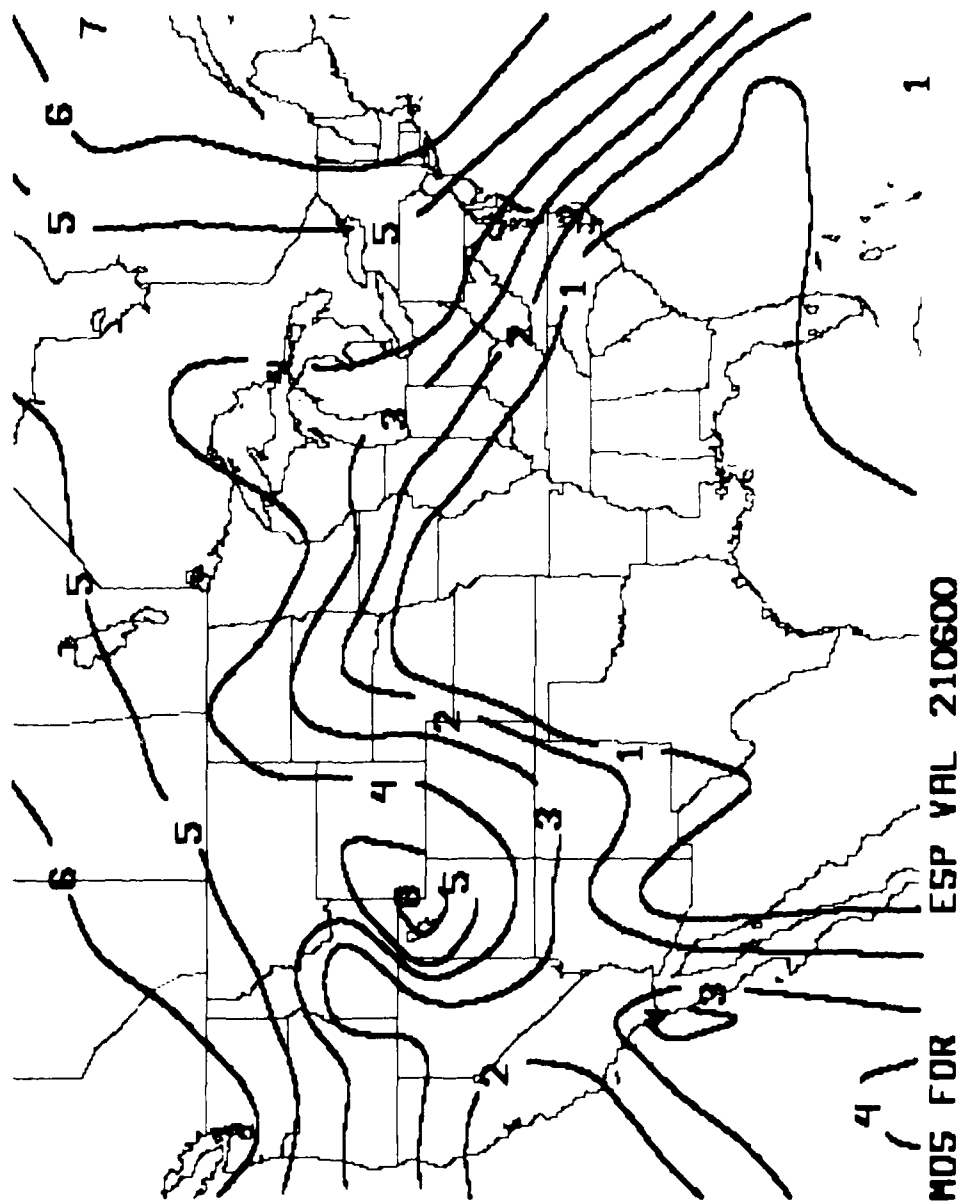


Fig. 17. Example of 6-hour Forecast Analysis of ESP Values



Fig. 18. Six-hour Forecast Analysis of ESP Values Superimposed over Coincident Visible Satellite Imagery

### C. Mesoscale Forecasting Experiment IV

Evaluation of the utility of McIDAS nowcasting and forecasting routines developed by SASC scientists was continued during Phase IV of the AFGL Mesoscale Forecast Experiment (MFE IV).

Six forecasters were involved in day-long forecast sessions using the McIDAS. One-, two-, four- and n-hour forecasts of winds, cloud amount, ceiling height, and six-hour cumulative precipitation amounts were prepared hourly for two forecast sites in the Northeastern U.S. Both sites selected were MOS forecast stations; Boston was always the primary station and the second station was chosen according to the characteristics of the particular forecast episode. The n-hour forecast length was chosen at each forecast time to verify at the nearest MOS valid forecast period. This was done so that MOS could be used as an additional forecaster for certain parameters and as guidance for the other participating forecasters.

Emphasis was placed on using McIDAS routines developed during the past year. These included routines to list, plot, and contour FOUS guidance and MOS and trajectory forecasts; improved Manually Digitized Radar (MDR) displays; implementation of a time-space weight analysis; and display of forecast guidance based on the Generalized Equivalent Markov (GEM) technique. After a forecast episode each forecaster completed a questionnaire designed to evaluate the usefulness of the products used during the case. This survey elicited comments and criticisms from the forecasters in order to help provide a basis for improvement and future development of McIDAS nowcasting routines.

Episodes used in the forecast sessions were chosen from cyclogenesis and cold front cases archived during Winter 1982 and Spring 1983 that displayed rapidly changing weather conditions, such as significant wind shifts, changes in cloud cover, alterations in ceiling height, and onset (or termination) of precipitation. Such short time scale variabilities constituted the most rigorous bases for evaluating the sensitivity and practicality of the newly developed mesoscale forecasting methods.

The format of MFE IV was identical to that of MFE III<sup>16</sup> with the

16. Gerlach, A.M., ed., 1982: Objective Analysis and Prediction Techniques. AFGL-TR-82-0394, Contract F19628-82-C-0023, Systems and Applied Sciences Corporation, pp. 129-130.

exception that ceiling height forecasts were limited to only probability forecasts. This change necessitated that modifications be made to the forecast entry, verification, and interrogation modules of the Mesoscale Forecast Facility (MFF)<sup>17</sup> program.

On a weekly basis archived data were re-stored for access through McIDAS as required in MFE IV. For each episode the required satellite visible - infrared imagery was re-stored from magnetic tape to satellite digital areas. Images were then loaded on consecutive TV frames to facilitate the use of time sequence loops. Surface, upper air, MDR, and FOCUS guidance data were loaded into history disk files from magnetic tape. Missing observations provided by AFGL were entered manually through the terminal keyboard.

MFF forecast verification procedures included saving the verification statistics on magnetic tape, documenting statistics for the forecasts entered, and generating listings of past forecasts and skill scores upon request.

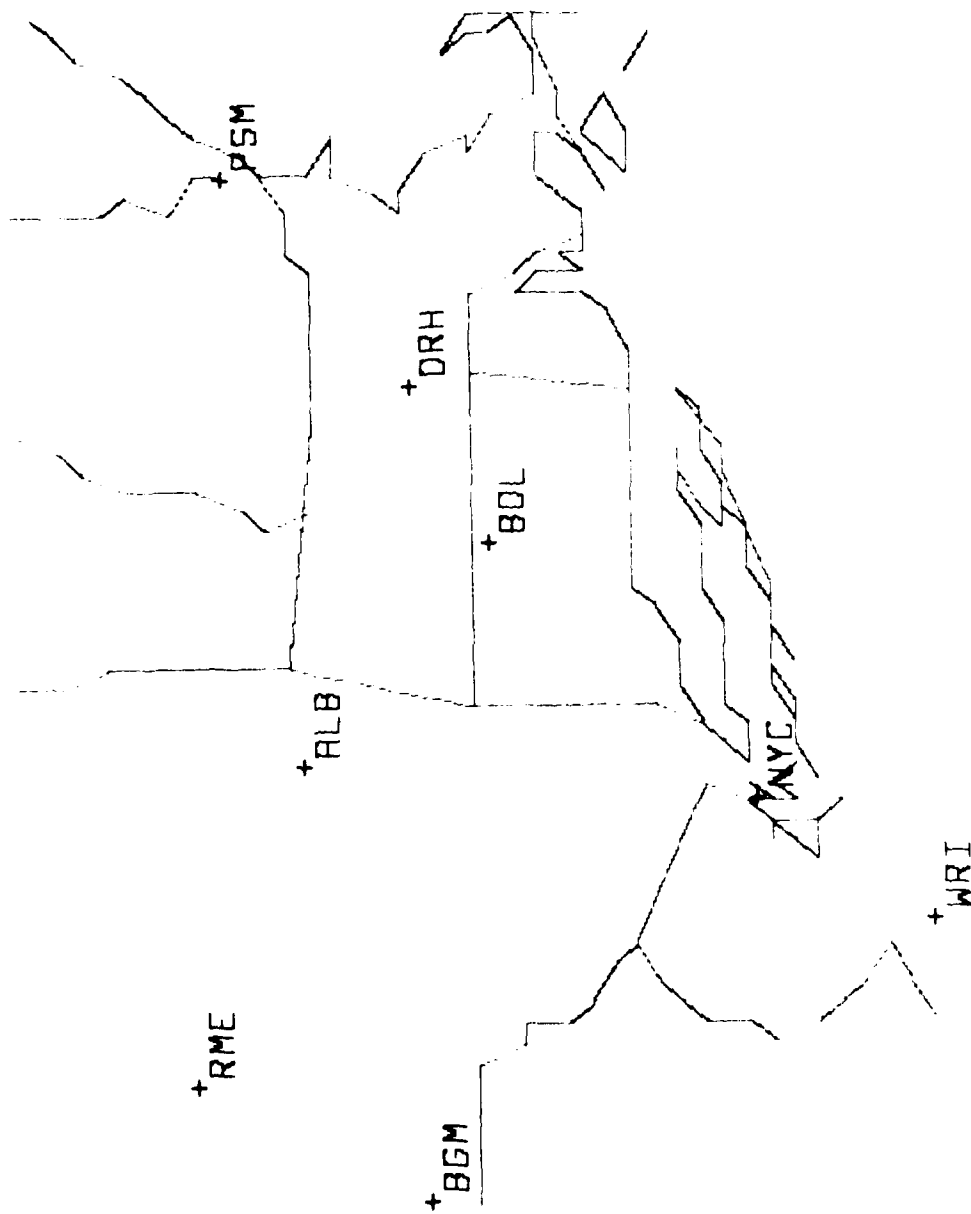
17. Gerlach, A.M., ed., 1980: Computer-Based Weather Research. AFGL-TR-80-0069, Contract F19628-79-C-0033, Systems and Applied Sciences Corporation, pp. 49-73.

#### D. Enhanced Manually Digitized Radar (MDR) Displays

Participants in the 1983 Mesoscale Forecasting Experiment (MFE IV) identified a need for additional display information in conjunction with the MDR grid and contouring routine GRDMDR. They suggested: 1) identifying and geographically locating National Weather Service network radars, FAA/ARTC radars, U.S. Air Force radars, and NWS local warning radars; and 2) providing a first approximation to the spatial distribution of MDR data input for transformation to a longitude-latitude based grid. In order to integrate this background information with displays of contoured MDR D/VIP fields, further software development was required. To this end, two subroutines were developed and implemented as command options to GRDMDR. Documentation of the command options in GRDMDR is provided in Appendix II-A.

Identifying and locating the MDR radar sites is accomplished with the SID option of GRDMDR. As input to the display, this subroutine extracts routinely stored latitude-longitude values and the three- or five-letter station identifiers from a disk file on McIDAS containing the most recent 100 national MDR grids (100 hours). Fig. 19 is a plot of the station identifiers produced by the SID option. The map area of  $4^{\circ}$  latitude by  $6^{\circ}$  longitude is generated by specifying a three-letter station identifier from the surface observations network rather than a station from the MDR network. This provides maximum flexibility in choosing a map area of this size because the spatial distribution of surface stations is far greater than that of MDR radar sites. This feature is incorporated in all other display options as well as in the SID option.

In addition to storing station identifiers and latitude-longitude information, McIDAS also decodes and saves the operational status contractions encoded by radar sites when appropriate. These contractions indicate whether stations are reporting echoes, and therefore to a limited degree provide data coverage on a station by station basis. Table 3 shows the three-bit values used in McIDAS to represent the status contractions and the interpretation of each. The actual contraction as it appears when encoded is indicated in parentheses. Status values 0 and 5-7 are software induced; they do not represent transmitted status information.



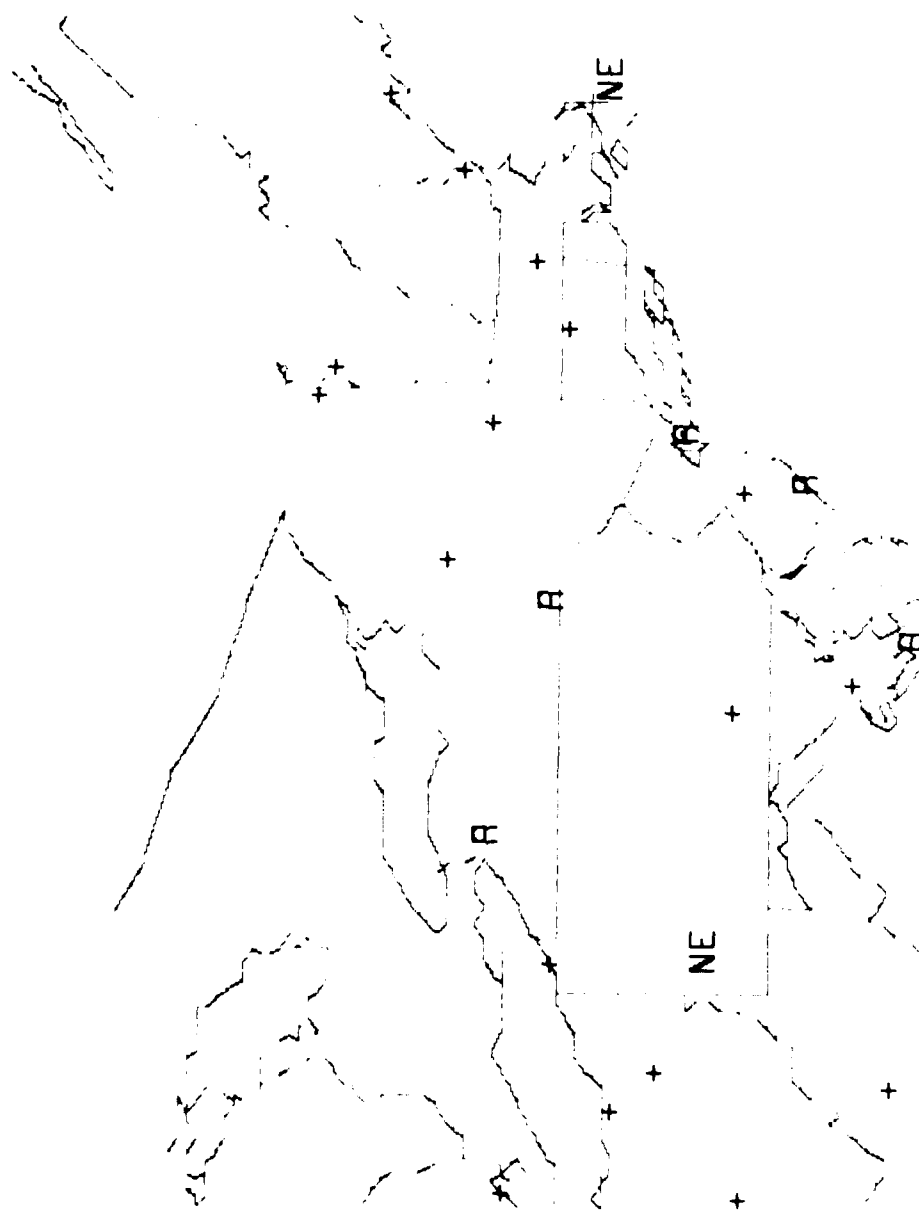
510 710530

Fig. 19. Plot of MDR Station Identifiers Produced by SID Option of GRDMR  
 (Map defines 4° latitude x 6° longitude centered on surface observation station PSF - Pittsfield, MA)

TABLE 3. MDR STATUS INFORMATION

<u>Three-bit Status</u> <u>Value</u>	<u>Interpretation</u>	<u>Plotting</u> <u>Symbol</u>
0	No data received for station or could not be deciphered by decoder	+
1	Equipment performance normal in PPI mode; no precipitation echoes observed; surveillance continuing (PPINE)	NE
2	Equipment inoperative or out of service for preventive maintenance (PPIOM)	OM
3	Observation omitted or not available for reasons other than those above (PPINA)	NA
4	Radar operating below performance standards (ROBEPS)	ROBEPS
5	MDR data available for station; MDR decoder found valid MDR data	A
6-7	Invalid status bits	None

Status reports are plotted with the PST option of GRDMR. The last column of Table 3 shows the symbols used by this option to display graphically the status of reporting radar sites. Fig. 20 is a plot of the status reports produced by the PST option. Stations that do not report on a regularly scheduled basis (i.e., NWS local warning radars) are shown as plus sign (+) to avoid cluttering the display. When combined with a contoured analysis of MDR D/VIP levels to form a composite display (see Fig. 21), the PST option shows the distribution of stations that contribute to the analysis.



PST 210630

Fig. 20. Plot of Status Reports Produced by PST Option of GRDMR  
(Symbols as in Table 3)



MDR SIC 210630  
PST 210630

Fig. 21 Contoured Analysis of MDR D/VIP Levels with Status Report Overlay  
Produced by NEC and PST Editions of GRDMR

Appendix II-A. Documentation of Command option in GRDMR

COMMAND: DISPLAY MDR DATA

KEY IN: MR

PROGRAMMER: FOURNIER, IVALLI

DATE: JUNE 83

CHAPTER: 9

COMMAND FORMATS:

1) MR LIS

This option lists the 100 entries in the MDR hours table and indicates the most recent day/time available.

2) MR ECO

This option lists the meanings of each of the six reported D VIF levels.

3) MR SID DTG MAP COLOR PAN N

This option plots the MDR site ID's on a mercator map background or over a satellite image (GOES).

DTG - any valid day/time group (DDHH30) for MDR

MAP - USA, NEW, EST, WST, -, two-letter P.O. state abbreviation or three-letter surface station ID for mercator map background, SAT for satellite overlay

COLOR - 1 for RED, 2 for GREEN, 3 for YELLOW

PAN - For specifying panel or quadrant display

N - Panel number (1-4)

4) MR PST DTG MAP COLOR PAN N

This option plots the report status of MDR sites on a mercator map or over a GOES satellite image.

ALL PARAMETERS ARE THE SAME AS IN OPTION 3.

NE = No Echo, A = Data Available, + = No Data Received

OM = Out for Maintenance, NA = Data Not Available

ROBEPS = Radar Operating Below Performance Standard

AD-A142 441

OBJECTIVE ANALYSIS AND PREDICTION TECHNIQUES - 1983(U)

2/3

SYSTEMS AND APPLIED SCIENCES CORP VIENNA VA

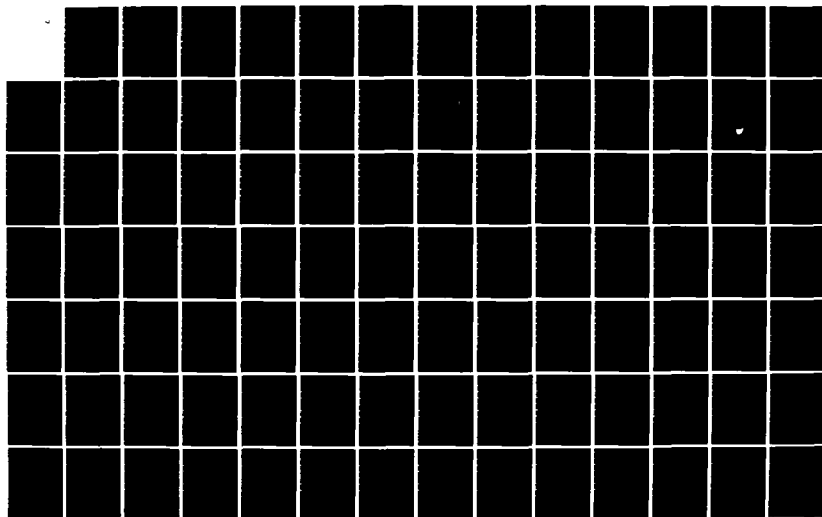
A M GERLACH 30 NOV 83 SCIENTIFIC-3 AFGL-TR-83-0333

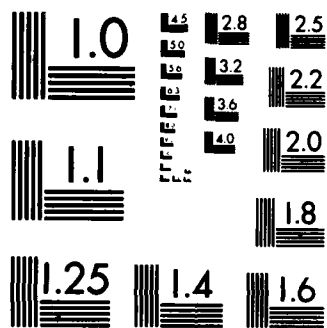
UNCLASSIFIED

F19628-82-C-0023

F/G 4/2

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

5) MR STA DTG ID<sub>1</sub>...ID<sub>i</sub>

This option lists: 1) the report status for the specified MDR sites and/or 2) the number of reports and echoes decoded.

DTG - Any valid MDR day/time (DDHH30) to list (2) above or DDHH30 to DDHH30 to list several consecutive hours

If ID's follow the DTG then the status of those sites (1) and of all accumulated (2) are listed.

6) MR MRC DTG MAP TEM FUN COLOR PAN N

This option remaps the MDR national grid to a mercator grid, saves that grid in a terminal gridfile, and initiates ZC (which sec.) to draw the analysis.

DTG - AS IN OPTION 3

MAP - AS IN OPTION 3 or a three-letter surface data station ID to produce a 4° latitude by 6° longitude map centered approximately on that station's latitude/longitude, or NW and SE corners (in LLWWW format)

TEM - Search area for remapping grid points:

PNT = MDR (i, j) point nearest lat/lon point

LAP = Nearest MDR (i, j) point and points (i+1, j), (i-1, j), (i, j+1), (i, j-1)

BOX = Nine i, j points centered on PNT

FUN - How to use search area points in conversion:

AVE - Take average of points

MAX - Take maximum value in search area and assign it to the lat/lon grid point

→NOTE: Default values are TEM = LAP, FUN = AVE

COLOR - AS IN OPTION 3, or NON to generate grid but draw no analysis

PAN, N - AS IN OPTION 3

7) MR SAT DTG MAP TEM FUN COLOR

This option remaps the MDR grid to a mercator grid whose bounds are defined by MAP, saves the grid, and initiates ZC to overlay the analysis contours on the currently displayed satellite image.

ALL PARAMETERS ARE AS IN OPTION G.

NOTES: MDR data are reported on a polar-stereographic national grid. MR remaps these data to lat/lon grids which are compatible with all other nowcasting grids and software.

### III. BOUNDARY LAYER METEOROLOGY

## A. Wind Shear Research

### 1. Shear Specification Study

SASC's involvement in the study of slant wind shear (SWS) predictions over a runway was to devise and implement programming and data processing techniques for manipulating large data bases of wind and temperature measurements. The objective of the study was to investigate the accuracy of offset remote tower systems for predicting SWS. Two statistically-based approaches using offset tower data from the AFGL Weather Test Facility (WTF) at Otis AFB were pursued: extrapolation techniques and screening regression determined equations. Processing methods devised for the former are described in Gerlach (1982).<sup>1</sup> The discussion that follows is devoted generally to the screening regression system used to determine equations useful for SWS predictions. The discussion is fairly detailed in order to provide useful information on the statistically-based screening regression procedure because of its potential value as a tool in many other applications.

Several computerized multiple screening regression systems developed by other groups were scrutinized for possible use in statistically deriving equations for predicting SWS. Of the systems reviewed, software was chosen that is similar to that described in the IBM SST package<sup>2</sup> which provides stepwise multiple screening regression (SMSR) statistical techniques for analyzing the relationship between a dependent variable and a set of independent variables in the order of their importance. The system as modified by SASC for SWS prediction requires four separate steps of data processing (see Fig.1) to produce coefficients for equations that ultimately predict dependent variables given independent variables.

Program EXTRAT in step (a) inputs from Tape A a set of one-minute average wind and temperature measurements from the AFGL WTF remote sensing tower instruments. Covering the period March 28 through May 4, 1981, these

1. Gerlach, A.M., ed., 1982: Objective Analysis and Prediction Techniques. AFGL-TR-82-0394, Contract F19628-82-C-0023, Systems and Applied Sciences Corporation.

2. IBM, 1968: System/360 Scientific Subroutine Package (360-CM-03X) Version III Programmers Manual. IBM Technical Publications Department, White Plains, NY.

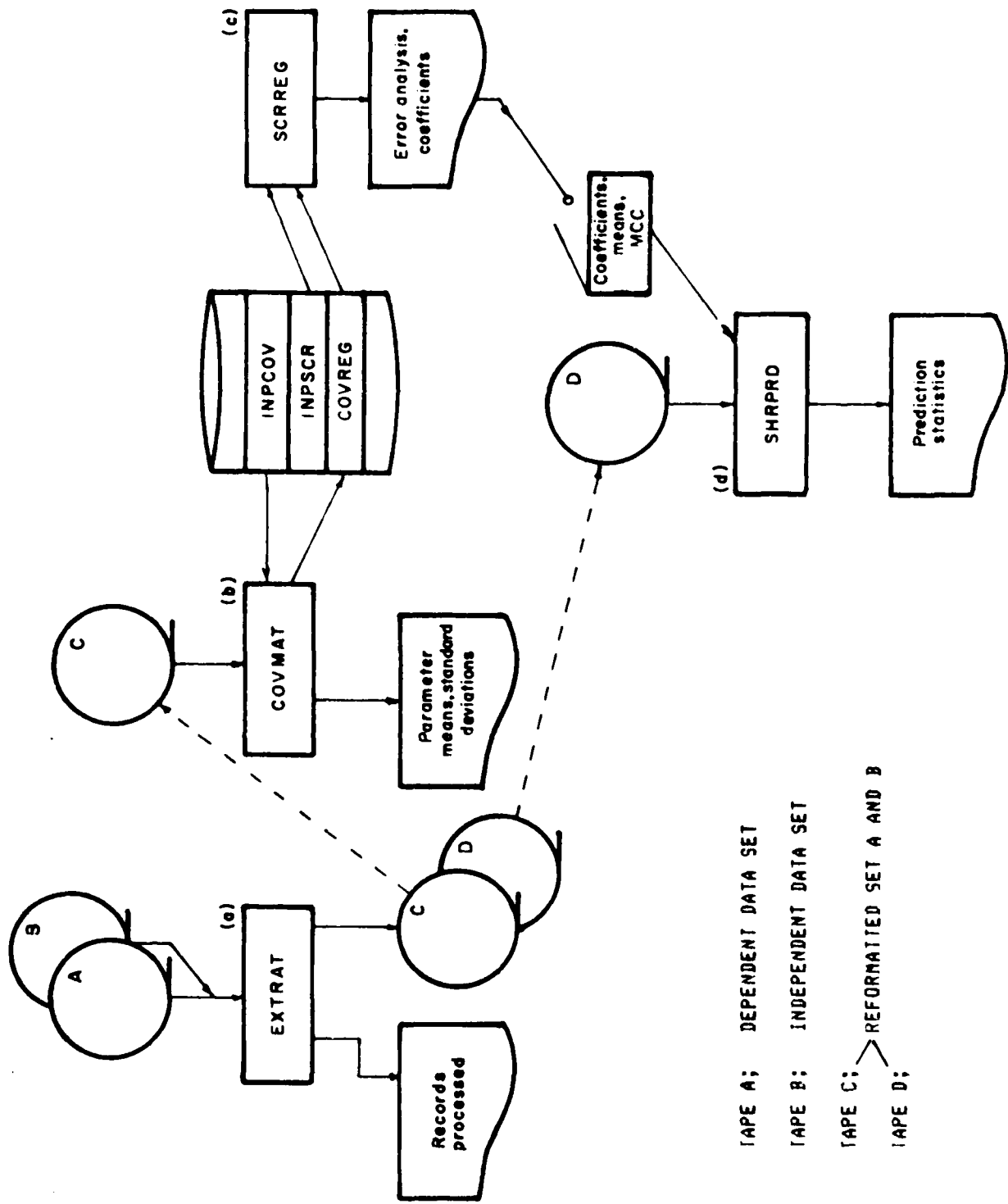


Fig. 1. Stepwise Multiple Screening Regression (SMSA) System Configuration

measurements were originally used in the previous study. Wind shear speeds, temperature differences, sines and cosines of the time of day to reflect diurnal variability for each observation, together with wind shear speeds between the surface and the highest positioned instruments at intervals of 0, 5, 10, 15, 30, and 60 minutes after initial observation, are computed and written to Tape C in a format suitable for step (b). These variables are listed in Table 1 in order of their position in a C tape record.

Program COVMAT in step (b) computes the terms of a covariance matrix from the set of data contained on Tape C. The resulting covariance matrix COVREG is stored on permanent disk file as a necessary product for step (c). Means and standard deviations of pertinent variables are computed and printed for use further on in the SMSR sequence of events. Prior to step (b) execution, control information to the COVMAT program is keyed into the computer and stored on permanent disk file called INPCOV. Organization of input parameters and their definitions for this file are pointed out in Fig. 2. These inputs control which variables in a C tape record to use for computing the terms of a covariance matrix. In this case, 25 out of a possible 38 variables were desired; their names are shown in Fig. 2 together with an array of integer 1 and 3 values stipulating variables to use or ignore respectively within a C tape record. The order of variables associated with the resulting covariance matrix is, in this case, redefined from a number of possible desired variables of 38 to a sequence of 25.

Having generated a covariance matrix, step (c) is initiated. Program SCRREG utilizes COVREG to analyze the relationship between dependent variables and a set of independent variables and for selecting independent variables in the order of their magnitude of contribution toward forecasting other dependent variables. (Any variable used in generating the covariance matrix can be named as a dependent variable.) Definitions of independent and dependent variables are input to program SCRREG via a set of control parameters outlined in Fig. 3 and stored on permanent disk file named INPSER. The set of input parameters shown defines nine chosen independent variables that will be used to compute coefficients for equations to predict, in this case, six dependent variables of wind speed shears 0, 5, 10, 15, 30, and 60 minutes later. Concluding each run, error analyses, resulting equations, and multiple correlation coefficients (MCC) of the predicted SWS speeds are printed for investigation.

TABLE 1. C TAPE RECORD ORGANIZATION AND PARAMETER DEFINITIONS

RECORD POSITION	MNEMONIC	DEFINITION
1	DATE	YEAR, MONTH, DAY, AND MINUTE OF OBSERVATION
2	A100	WIND SPEED TOWER A, 100FT
3	P100	" " " P, 100FT
4	Q100	" " " Q, 100FT
5	A 50	" " " A, 50FT
6	P 50	" " " P, 50FT
7	Q 50	" " " Q, 50FT
8	X	" " " X, SURFACE
9	Q 10	" " " Q, 10FT
10	A100-X	WIND SPEED SHEAR BETWEEN A100-X
11	P100-X	" " " " P100-X
12	Q100-X	" " " " Q100-X
13	A 50-X	" " " " A 50-X
14	P 50-X	" " " " P 50-X
15	Q 50-X	" " " " Q 50-X
16	Q 10-X	" " " " Q 10-X
17	A100-A 50	" " " " A100-A 50
18	P100-P 50	" " " " P100-P 50
19	Q100-Q 50	" " " " Q100-Q 50
20	Q100-Q 10	" " " " Q100-Q 10
21	Q 50-Q 10	" " " " Q 50-Q 10
22	T100-TX	TEMPERATURE DIFFERENCE=(TEMP,100'-TEMP,X)
23	T 50-TX	" " " =(TEMP, 50'-TEMP,X)
24	SIN(TM)	SIN(TIME OF DAY)=SIN((60HOUR+MINUTE)/1440*360)
25	COS(TM)	COS(TIME OF DAY)=COS((60HOUR+MINUTE)/1440*360)
26	GILL-X00	WIND SPEED SHEAR BETWEEN GILL-X AT TIME 0
27	GILL-X05	" " " " GILL-X 5 MIN. LATER
28	GILL-X10	" " " " " 10 MIN. LATER
29	GILL-X15	" " " " " 15 " "
30	GILL-X30	" " " " " 30 " "
31	GILL-X60	" " " " " 60 " "
32-38		UNUSED



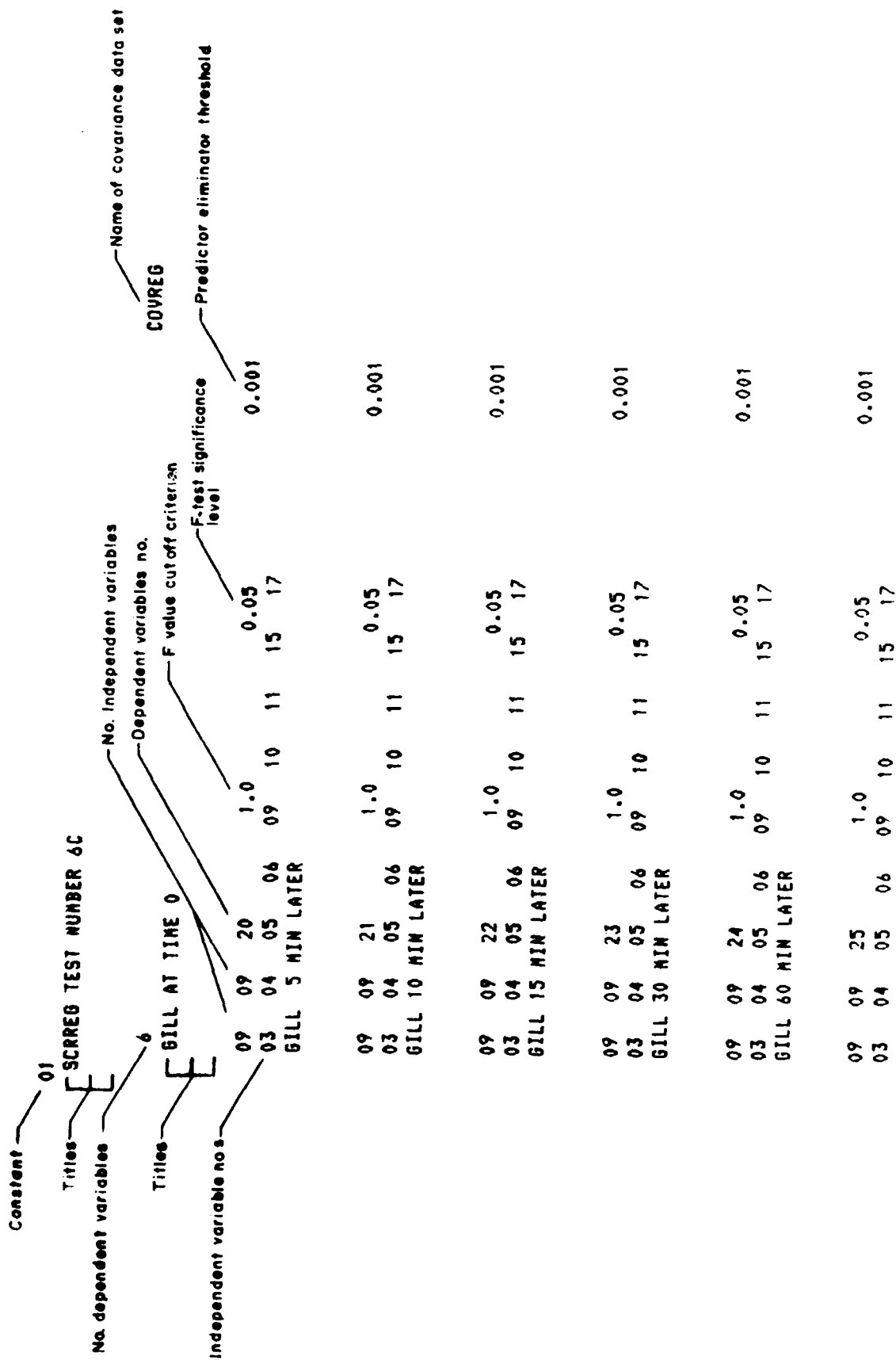


Fig. 3. Organization of Input Parameters and Definitions for INPSR

Program SHRPRD was developed in step (d) to determine how well JMB-generated equations predict actual temporal SWS. To facilitate this investigation, an ensemble of wind and temperature data from the same test site for a period of record from March 16 through March 19, 1981, was made available on Tape B, and for convenience, step (a) was utilized again to provide the data on an additional Tape D in a format identical to Tape C, Table 1, for input to SHRPRD. Equations containing coefficients and MCC values from step (c) together with a priori variable means from step (b) are made available to program SHRPRD via punched card input. Tape D records are then read in, whereupon pertinent variables in each record are used in equations to compute SWS at 0, 5, 10, 15, 30, and 60 minutes later and compared to the actual SWS occurring at the same time lapse intervals. Statistics of these comparisons are distributed in array form suitable for comparing with SWS predictions derived using extrapolation techniques in the previous study.

As a result of step (d) it was discovered that too few severe SWS predictions were being computed. The program in step (d) was therefore further modified to include a prediction enhancement algorithm defined by Eq. (1):

$$\Delta V_E = \frac{\Delta V - \overline{\Delta V}}{r} + \Delta V \quad (1)$$

where

$\Delta V$  = originally computed SWS speed

$\overline{\Delta V}$  = a priori mean SWS speed

$r$  = multiple correlation coefficients (MCC)

$\Delta V_E$  = enhanced predicted SWS speed.

Step (d) was again implemented to reprocess all requested test cases using the prediction enhancement algorithm. Final results of the study are shown in Brown (1983).<sup>3</sup>

## 2. Acoustic Doppler Sounder Data Base

An acoustic Doppler sounder located at the AFGL Weather Test Facility at Otis AFB routinely collected u and v components of vertical wind profiles at 13 levels from the surface to 170 meters on a daily basis from May 1981 thru January 1983. The data were processed by SASC as described in Gerlach (1982),<sup>1</sup> in order to create an archived acoustic Doppler sounder data base.

3. Brown, H.A., 1983: Prediction of Slant Wind Shear with an Offset Tower Observation System. AFGL-TR-83-0308, Air Force Geophysics Laboratory, Hanscom AFB, MA.

useful for future boundary layer wind studies.

Data are stored on magnetic tapes that have assigned tape numbers and corresponding periods of record as outlined in Table 2.

TABLE 2. ACOUSTIC DOPPLER SOUNDER DATA BASE TAPE NUMBERS AND PERIODS OF RECORD

<u>Archive</u> <u>Tape No.</u>	<u>Start Date</u>	<u>End Date</u>	<u>Backup Tape</u> <u>No.</u>
DOP001	5/26/81	9/29/81	DPB001
DOP002	9/29/81	11/28/81	DPB002
DOP003	11/30/81	7/22/82	DPB003
DOP004	9/27/82	1/12/83	DPB004

To facilitate use of the acoustic Doppler sounder data base, the contents of the data base tapes are here described.

The acoustic Doppler sounder data are written on seven-track tapes at a recording density of 800 BPI by a computer having 60-bit word architecture. Each file of data on tape contains a header record for denoting the data time information relevant to the data records that follow in the same file. The header record contains seven 60-bit right justified binary words signifying the following:

- Word No. 1 = Run number
- 2 = Time average (nine minutes)
- 3 = Year (1981-1983)
- 4 = Month (1-12)
- 5 = Day (1-31)
- 6 = Hour (0-23)
- 7 = Minute (0-59)

Each successive record contains 507 60-bit words of nine-minute averages of u,v wind components at thirteen levels in the vertical atmosphere over a period of two days or less. Each word contains four packed 15-bit integer quantities of u followed by v averaged wind components for level one, followed by u,v components for level two, etc., for thirteen levels; followed by the same sequence of events for the next nine-minute time average and so on. The integer u,v components are multiplied by ten; negative quantities are noted by a one-bit present in the high order 15th bit. Missing data are filled by ones.

The acoustic Doppler sounder data base is further accessible through coherently formatted printouts of the observed sounder wind components created during the data base development.

## B. Atmospheric Attenuation Effects on Remote Sensing

Transmission data from ground based remote sensing instruments in the infrared (8-12 $\mu$ ), visible (.55 $\mu$ ), 3-5 $\mu$ , and laser (1.06 $\mu$ ) frequency spectra collected in December 1981 through February 1982 were made available in digital form for investigating their usefulness for tactical decision aids during snowfall conditions.

Software previously developed by SASC provided several standard statistical summaries of the transmission data for data quality review and subsequent measurement variability studies for three types of snow storms; namely, heavy snowfall, light snowfall, and storms accompanied by high winds. The standard statistical summaries consisted of data means, variance, bias, correlation coefficients, and graphic displays in the form of scatter diagrams.

Transmission measurements in the spectral regions mentioned were recorded every minute during snow storms. Based on this high sampling rate and the fact that the magnitude of transmission received in the range from 0 to 100 percent correlates well from Eq. (2) with the intensity of snowfall rate, led to the conclusion that by providing duration statistics of transmissivities in dB/km during snowfall conditions, useful decision aids could be provided.

$$T = 10^{-.0064(PT)} \quad (2)$$

where

PT = percent transmission

T = transmission (dB/km).

Software was developed to categorize T computations from the digital transmission data and to distribute the categorized values into cumulative probability distributions of the frequencies of their durations in time intervals of one minute through 45 minutes or more. Included in the software development by SASC was a specialized graphics plotting routine for displaying cumulative probabilities as a function of time duration. The ordinate of such a graph is more complex than the ordinate of a conventional log scale in that

four differently scaled log cycles increase from the bottom of the graph to the halfway point and thereafter decrease in scale to the top of the graph. Probabilities of events that are normally distributed will appear on the graph as a straight line. This routine was extensively used to produce graphs of cumulative probabilities of categorized transmissivities in dB/km as a function of time durations. Fig. 4 displays such a graph for the .55 $\mu$  channel in a type C snowstorm.

Program SPECTR was obtained to provide power spectrum analyses and autocorrelations of given data quantities over sequential time periods. Two basic modifications to this program were made to (a) plot all results via computer graphics, and (b) incorporate a red-noise filter to be applied to the remote sensing measurements during the autocorrelation analysis.

Figs. 5 and 6 show typical products displayed by the modified SPECTR program for a type A storm using raw transmission data (0 to 100 percent transmission values) for the IR 8-12 micrometer band.

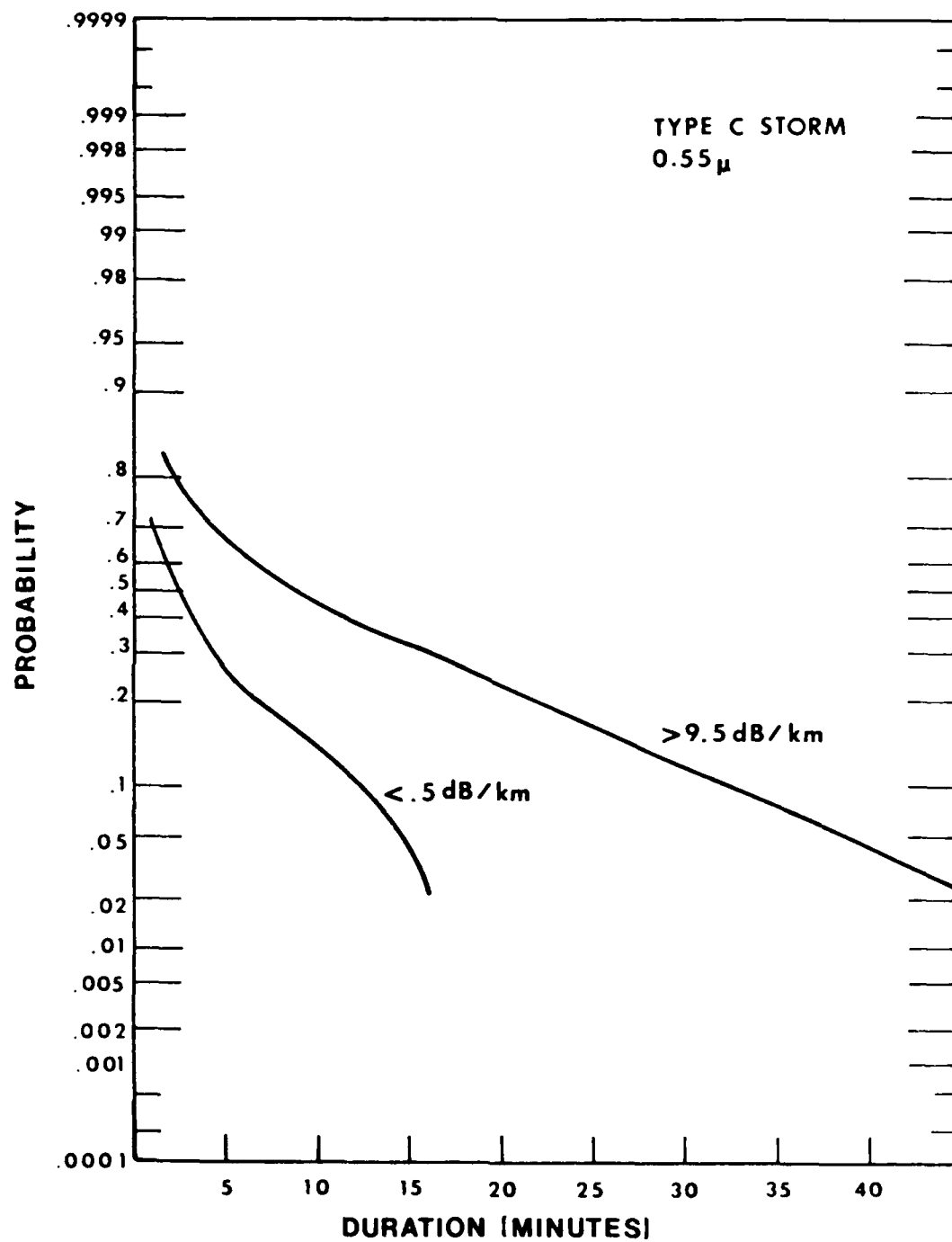


Fig. 4. Typical Duration Statistics Display

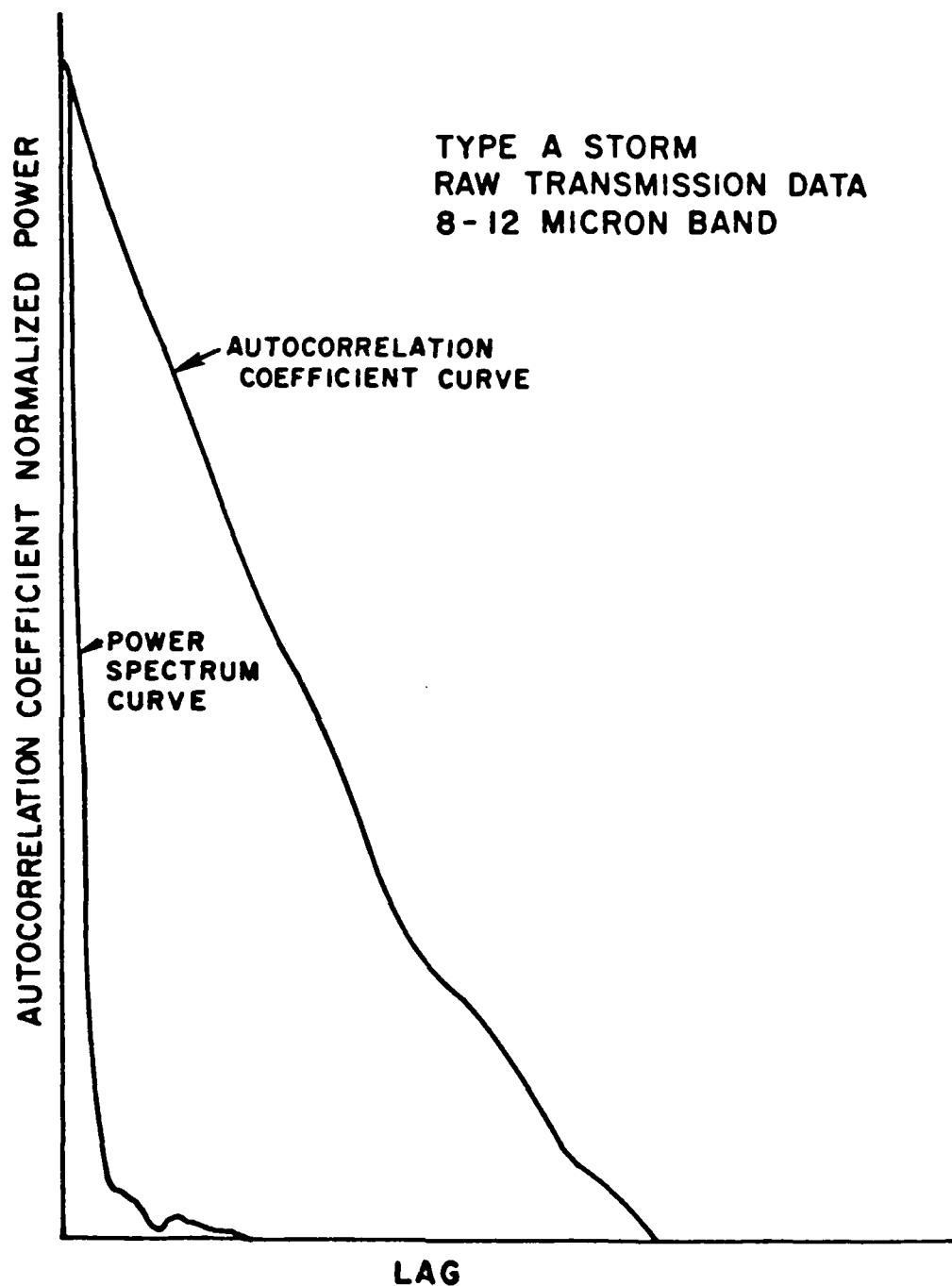


Fig. 5. Typical Autocorrelation Display

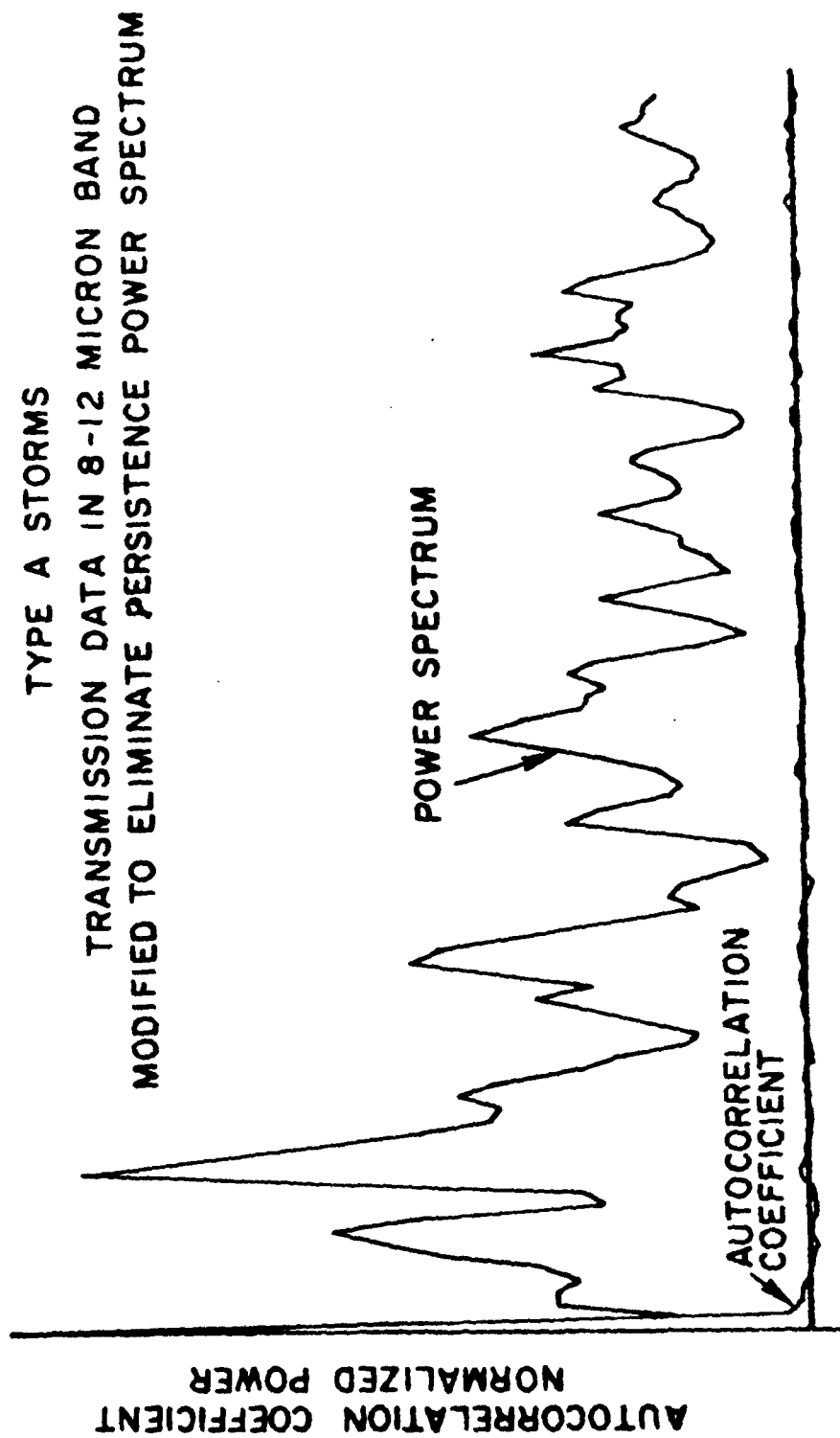


Fig. 6. Typical "Red Noise" Power Spectra Display

### C. Atmospheric Chemical Dispersion Modeling

AFGL is conducting research seeking to improve the mathematical models used to predict dispersion of chemicals in the atmosphere following space shuttle launches or possible chemical spills. SASC has adapted several existing models, making additions and modifications as required to enhance their usefulness for Air Force purposes.

In 1983 AFGL acquired a chemical spill model produced by the Shell Oil Development Company. One of the most complete Gaussian plume dispersion models in existence, SPILL is an unsteady state model representing the evaporation of a chemical spill and the atmospheric dispersion of the vapors. The model estimates concentrations of the vapors of neutral density and lighter than air gases as a function of time and distance downwind from a spill and outputs amount and location of maximum concentration and concentration at specific points. The computer program for this model, incompatible in part with the AFGL CDC 6600 computer, was revised and modified. A major component of the program listed by Shell was missing and was reconstructed by SASC.

SPILL will be used as a standard against which simpler models can be tested. It also may be expanded to include such factors as heavy gas effects and concentration fluctuations.

A second model, the NASA REEDM, was acquired by AFGL and translated by SASC for the CYBER 750 computer. It is used by NASA to assess the environmental impact of space shuttle launches at Cape Kennedy. The REEDM computer program predicts the dispersion of exhaust effluents under various meteorological conditions. Air Force Space Division plans to use this code, with appropriate changes to reflect near-field fallout and terrain differences, to predict the consequences of space shuttle launches at Vandenberg AFB, CA. AFGL acquired the model to subject it to critical analysis in determining potential deficiencies that could be corrected by software changes before it is used by the Air Force.

Just before and immediately after a shuttle launch, rocket engine exhaust causes formation of a large cloud of hot buoyant effluents near the ground that subsequently rises and entrains ambient air until the temperature and density of this ground cloud reach equilibrium with ambient conditions. The rocket engines also leave an exhaust trail that extends throughout the depth of the troposphere.

The REEDM model calculates peak concentration, dosage, and deposition downwind of the launch site to provide pre-launch forecasts of the environmental effects of the exhaust diffusion and post-launch environmental analysis.

The code received from NASA was designed to operate on a CDC Cyber 750 using Network Operating System (NOS) transaction processing. AFGL CDC equipment operates under NOS/BE (Batch Environment) processing. NOS system control language formats, and loader, product set, and system utility control statement formats, in many instances proved to be significantly different from NOS/BE. The REEDM code contains multiple options for plotting, printing, or choosing concentration/dosage model calculations, washout model conditions, or gravitational deposition results. These options are under control statement operating formats known as procedures. Procedures are groups of control statements that perform a single function within the job control stream and can be executed repeatedly. Most of the procedures to run the REEDM model had to be revised or eliminated because the flexibility and allocation of memory or file definition allowable under NOS is not available in NOS/BE.

The plotting procedures in the REEDM model required a computer graphics package, Graphics Compatibility System (GCS), developed by the U.S. Army Engineer Waterways Experiment Station, Vicksburg, MI. Since this package was not available at AFGL, plotting had to be completely revamped for in-house systems. Successful operation of four test runs of the REEDM was achieved.

The REEDM and Shell models will be run comparatively using varying parameters to explore their respective efficiencies and deficiencies relating to the environmental consequences of launches.

#### D. Fog Prediction Modeling

AFGL is developing short range fog prediction techniques. As part of this project a two-dimensional fog prediction computer model was obtained from the Naval Environmental Prediction Research Facility (NEPRF), Monterey, CA. Various revisions and modifications were made to create an AFGL version. Both the original and the AFGL model were tested on six observed advective fog case studies provided by the Calspan Corporation. Nearly 40 plots were developed to illustrate the changes made to the NEPRF model to enhance its capability.

Two displays are shown in Figs. 7 and 8 which are samples of the types of plots provided by SASC. Fig. 7 shows the horizontal temperature structure of warm moist air being advected over cold water at horizontal distances of 190 km and 90 km. Fig. 8 portrays the liquid water content of fog formed by warm moist air advected over cold water, distributed by distance and height. The results of the AFGL version of the NEPRF model are labeled "errors corrected."

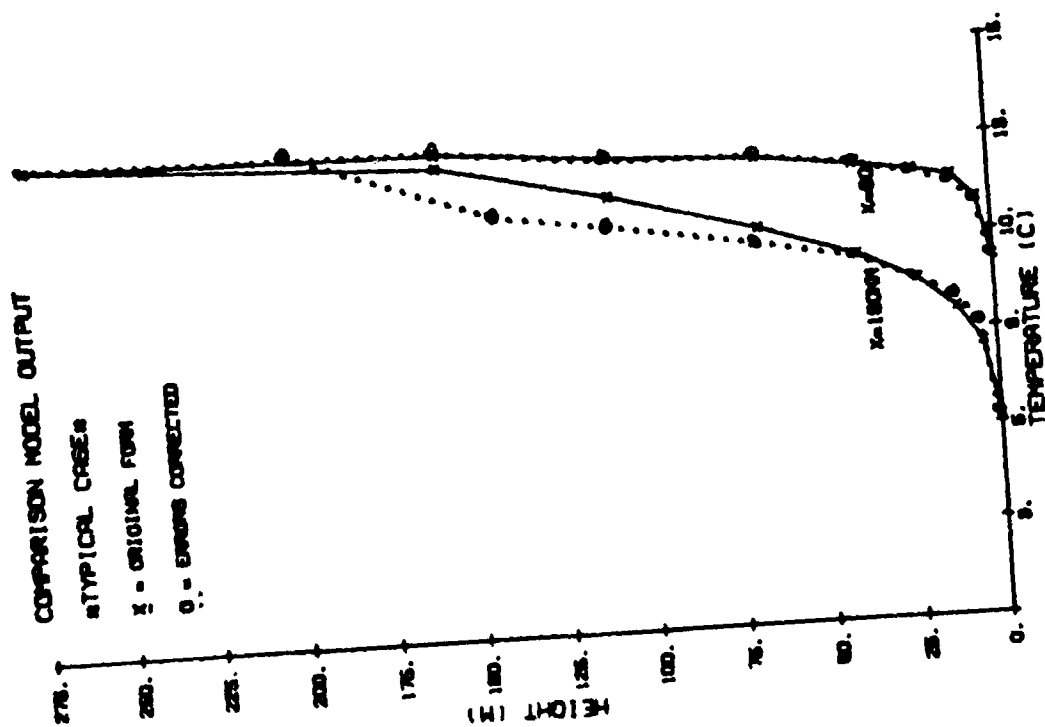


Fig. 7. Horizontal Temperature Structure of Warm Moist Air Advected over Cold Water at Distances of 190 km and 90 km

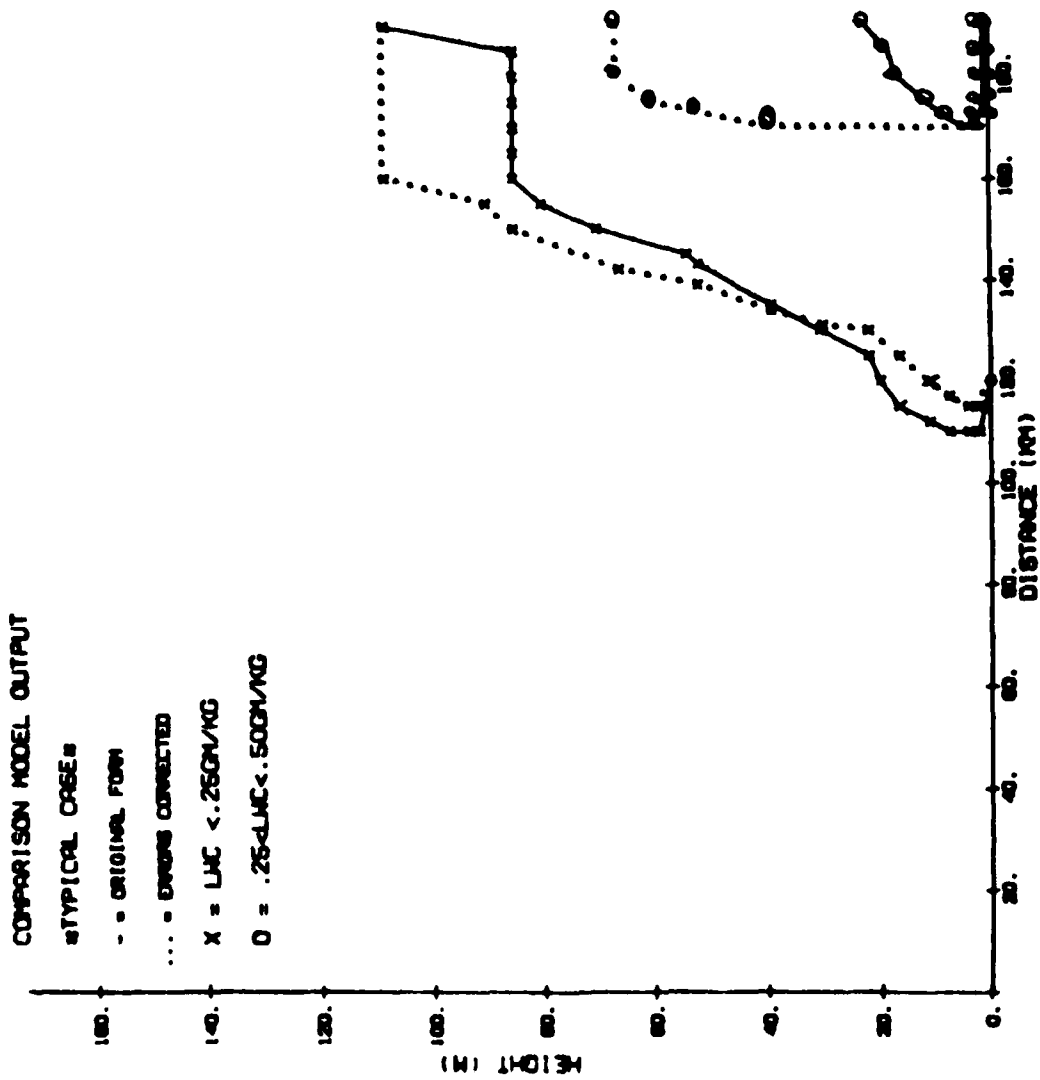


Fig. 8. Liquid Water Content of Fog Formed by Warm Moist Air Advected over Cold Water

#### E. Weather Instrumentation Evaluation

Expansion of the MAWSE computer program, which is the processing vehicle for translating sensor outputs from weather instruments at the AFGL Weather Test Facility, continued in 1983. Visibility instruments for the Battlefield Weather Observation and Forecast System (BWDFS) program were introduced periodically and evaluated by comparison with known sensors. Up to 19 sensor raw voltage samplings taken every six seconds can be interspersed with the normal one-minute mean data on MAWS magnetic tapes produced at WTF. Operators at the MAWS console there or at AFGL can select the instruments to be polled from the more than 70 in place on the ground or on towers at WTF. Sensor selection may be changed or dropped at any time during a tape recording. The MAWSE computer program senses a raw voltage record, identifies the selected sensors, and prints out the sensor name and voltage.

#### F. WTF Maintenance and Operation

SASC is responsible for the maintenance and operation of the WTF. In addition to daily site and instrument maintenance and repair, the following characteristic activities were conducted.

Stability problems in the GMQ-32 Transmissometer pulse amplifiers were corrected, based on information supplied by Tasker Systems.

An interface device was received from Ertek Co., Newton Center, MA. Consisting of an analog to frequency transmitter, optical isolating link, and a frequency to analog receiver, the equipment was installed and placed in operation.

Upon receipt of the necessary replacement parts, the GMQ-13 Cloud Height Set projector was overhauled.

A GMQ-32 Transmissometer was modified for use on a 40 ft baseline. Foundations were poured, tower supports erected, and the transmissometer installed. A "light trap" was constructed in the form of a louvered back-stop painted black, installed immediately behind the projector of the 500 ft/1,000 ft dual baseline transmissometer. Purpose was to reduce background count.

A joint AFGL/DOT-FAA field program involving visibility sensors supplied by various manufacturers is in progress. Sources are HSS, Inc.; Wright and Wright, Inc.; Impulse Physics, USA; Qualimetrics; Belfort Instrument Company; Marconi; and Meteorological Research Institute. Installation foundations were constructed, and power cables and cables for data collection were installed at the facility "X" location to provide high quality comparison data. Manufacturers' technicians are installing their own equipment with technical assistance provided by AFGL and SASC. The AFGL MAWS is being used to record data.

#### G. Microwave Refractive Structure of the Lower Atmosphere

The RAYTRACE program is a geometric optics model of wave propagation through an inhomogeneous atmosphere having a vertically stratified index of refraction. The program calculates the direction of wavefront propagation by solving a parabolic trajectory approximation.

Program RAYPROG was formulated specifically for analyzing atmospherically induced anomalous propagation on Line-of-Sight (LOS) communication links. Data collected during the meteorological measurement program conducted on two Digital European Backbone Line-of-Sight Links in Germany in October - November 1981 were used with the RAYPROG1 and RAYPROG4 programs (modifications of RAYPROG) as the basis of the AFGL report Atmospheric Effects on DEB Communication Network.<sup>4</sup> The programs were also used to process data obtained on the Monadnock-Prospect Hill LOS link in 1982, for the technical memorandum Meteorological Measurements on the Monadnock-Prospect Hill Communication Link.<sup>5</sup>

RAYPROG1 and RAYPROG4 calculated and plotted:

1. Cross section of the terrain along the propagation path.
2. Paths of the wave trajectories relative to the terrain.
3. Paths of rays reflected by the earth.
4. Vertical profile of the modified refractive index as a function of height.

Throughout the year several modifications of the RAYPROG1 program were formulated. In REFLECTRAYPROG1 a subroutine was added that allows a combination of refraction and reflection to be considered in the plotting of the raytrace. Reflection caused by the ground configuration had been incorporated in the raytrace, but in checking the time of arrival of the rays at the receiver, it was noted that the delay of some rays was too lengthy to be caused by either atmospheric refraction or surface reflection. The modification allows reflection from an arbitrary atmospheric layer. The atmospheric

4. Morrissey, J.F., O.R. Coté, and Y. Izumi, 1983: Atmospheric Effects on DEB Communication Network. AFGL-TR-83-0116, Air Force Geophysics Laboratory, Hanscom AFB, MA.

5. Izumi, Y., and J.F. Morrissey, 1983: Meteorological Measurements on the Monadnock-Prospect Hill Communication Link. AFGL Tech Memo No. 84, Air Force Geophysics Laboratory, Hanscom AFB, MA.

reflection subroutine will be used in an attempt to correlate the longer time delays with atmospheric turbulence measurements and predictions (see Fig. 9).

RAYPROGSUB underwent several improvements as follows:

1. RAYPROGSUB is able to graph an earth's surface outline from a random number of points. It will dispense with the preliminary step of processing points read from a topographic map into the MAPJIM program to produce a prefixed number of smoothed points. It will produce a more precise representation of the earth's surface.
2. The program is more flexible in the way it allows the microwaves to reflect from the earth's surface. There are three options for reflection:
  - a. No reflection (ray is terminated if it hits the earth).
  - b. Specular reflection.
  - c. Reflection directed toward the receiving antenna.
3. The program was further refined into subroutines for flexibility.

RAYPROGCROSS was developed for use on troposcatter links. It is a two-dimensional program that determines the intersection of the transmitting and receiving beams given the mean refractive index profile. This program will be the basis for a three-dimensional program in the future. It is now being used to analyze data gathered to demonstrate the performance of the TRC70 troposcatter radio in several meteorologically diverse areas of the world. Turbulent refractive data will also be included in the model.

Programs to evaluate refractometer data gathered on aircraft have been partially developed. These require decommutation of high speed (200 sample/sec) refractive index data and subsequent analysis using Fast Fourier Transform (FFT) techniques, smoothing, and plotting schemes. These routines have been used on data gathered in Florida in September 1983. Although only partially developed these programs were able to demonstrate certain shortcomings in the refractometer which led to an instrument redesign.

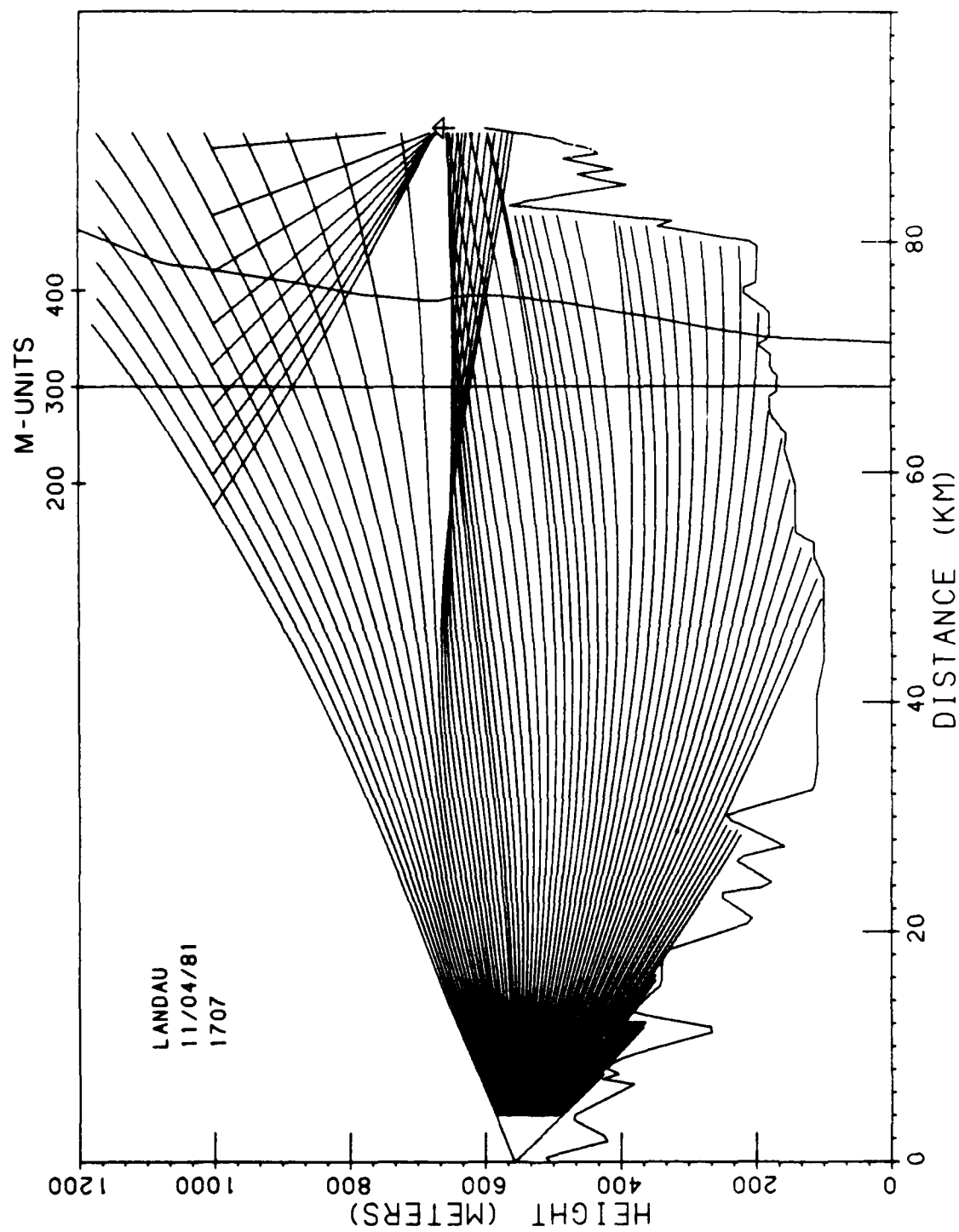


Fig. 9a. Raytrace for Refractive Conditions Measured at 1707 CET, November 4, 1981, at Landau, with Rays Reflected from an Arbitrary Atmospheric Layer above and Directed toward the Receiver

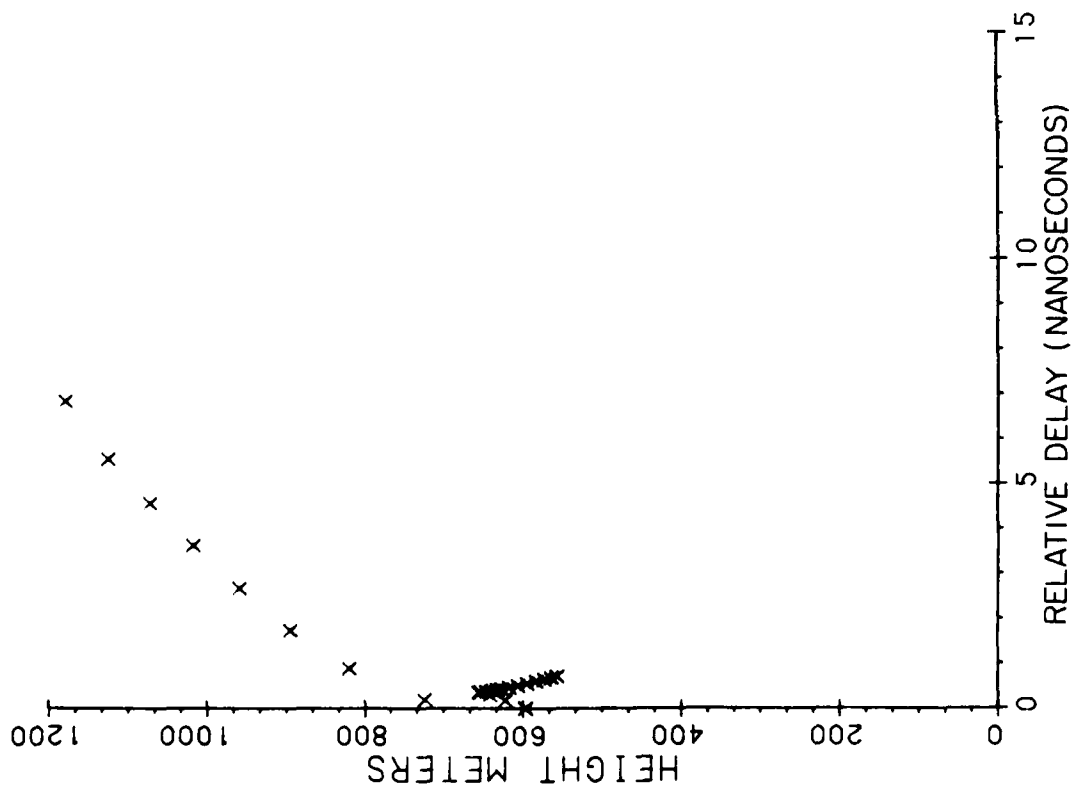


Fig. 9b. Relative Time of Arrival of the Reflected Rays Directed toward the Receiver  
(Circles represent the reflected rays)

#### IV. RADAR METEOROLOGY

#### A. Single-Doppler Radar Detection of Synoptic-Scale Wind Anomalies

Wind field analysis through use of a single Doppler radar was suggested by Probert-Jones (1960).<sup>1</sup> Antenna scanning strategy and techniques for interpretation of the resultant velocity patterns in widespread precipitation were successively developed and refined by Lhermitte and Atlas (1961),<sup>2</sup> Caton (1963),<sup>3</sup> Browning and Wexler (1968),<sup>4</sup> and most generally by Waldteufel and Corbin (1979).<sup>5</sup>

Harmonic analysis of the Doppler velocities measured throughout a complete circle of azimuth angles, at constant elevation and range, yields divergence, wind speed and direction, and deformation as functions of the zeroth, first, and second harmonic, respectively. These properties of the wind field are averages along the scanning circle; therefore the wind field around the radar, within the radius of the scanning circle, is necessarily characterized as horizontally homogeneous.

Natural wind fields, however, are variable, not homogeneous. A wealth of interesting and possibly useful information is concealed in velocity anomalies remaining after removal of the zeroth, first, and second harmonics of the Doppler velocity as a function of azimuth. For example, velocity perturbations encompassing a scale much less than the radar scanning circle have been successfully identified as mesocyclones (Donaldson, 1970<sup>6</sup>) or

1. Probert-Jones, J.R., 1960: Meteorological use of pulse Doppler radar. Nature, 186, 271-273.

2. Lhermitte, R.M., and D. Atlas, 1961: Precipitation motion by pulse Doppler. Proceedings, Ninth Weather Radar Conference, AMS, Boston, 218-223.

3. Caton, P.A.F., 1963: Wind measurement by Doppler radar. Meteor. Mag., 92, 213-222.

4. Browning, K.A., and R. Wexler, 1968: The determination of kinematic properties of a wind field using Doppler radar. J. Appl. Meteor., 7, 105-113.

5. Waldteufel, P., and H. Corbin, 1979: On the analysis of single-Doppler radar data. J. Appl. Meteor., 18, 532-542.

6. Donaldson, R.J., Jr., 1970: Vortex signature recognition by a Doppler radar. J. Appl. Meteor., 9, 661-670.

tornado circulations (Brown and Lemon, 1976<sup>7</sup>) if strong shear is oriented azimuthally with respect to the scanning radar, or as divergence capping convective updrafts (Lemon and Burgess, 1980<sup>8</sup>) if the strong shear is oriented radially. These mesoscale features, inferred from perturbations in the Doppler velocity field, occupy a very small portion of the total area scanned by the radar. They range in size from one km<sup>2</sup> or less for a tornadic vortex signature to several hundred km<sup>2</sup> for a large supercell divergence signature.

The current investigation is directed toward retrieval of a wind field that varies in a smooth manner throughout the scale of the radar scanning circle. The initial approach is through analysis of several model wind fields. The basic tools of this analysis are variations in the radius of the radar scanning circle, and bisection of the Doppler velocity pattern both along and across the direction of mean flow.

The first model for analysis is a curved wind field with constant speed. For simplicity, it is assumed that elevation angles are sufficiently small so that the Doppler radar detects only the components of horizontal wind along its beam. The complexities of appreciable elevation angles and vertical motion components may be introduced later without difficulty.

Geometrical relationships are depicted in Fig. 1. The radius of the circle scanned by the radar is  $r$ , and the distance of the radar from the center of curvature of the wind field is  $R$ . It is assumed that  $R > r$ , an assumption easily granted because  $r$  is controllable. As the radar scans, its azimuth angle from a reference position is given by  $\alpha$ . In the interest of simplicity, Fig. 1 is set up so that  $\alpha = 180^\circ$  is directed toward the center of curvature of the wind field. The angle  $\beta$ , with the same reference position as  $\alpha$ , indicates direction from wind field center of curvature to any particular location scanned by the radar. The wind speed  $V$  is assumed constant in this model.

In Cartesian coordinates, the wind field can be expressed as the sum of components  $u$  and  $v$ , where  $u$  is directed along  $\beta = 90^\circ$  and  $v$  along

7. Brown, R.A., and L.R. Lemon, 1976: Single Doppler radar vortex recognition: Part II - tornadic vortex signatures. Preprints, 17th Conference on Radar Meteorology, AMS, Boston, 104-109.

8. Lemon, L.R., and D.W. Burgess, 1980: Magnitude and implications of high speed outflow at severe storm summits. Preprints, 19th Conference on Radar Meteorology, AMS, Boston, 364-368.

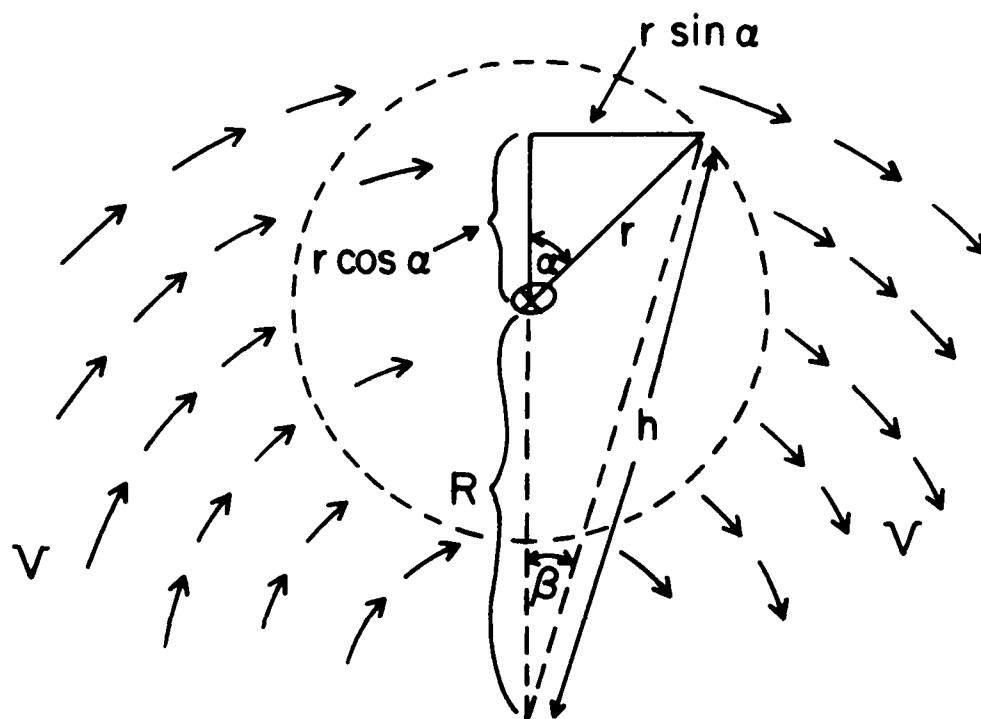


Fig. 1. Geometry of Radar Scan of a Circular Wind Field with Constant Speed  $V$

(Radius of scan is  $r$  and its azimuth angle is  $\alpha$ ; distance of radar from circulation center is  $R$ ;  $h$  is distance and  $\beta$  is azimuth angle of any point on the scan circle with respect to circulation center.)

$\beta = 0^\circ$ . Inspection of Fig. 1 readily shows that  $u = V \cos \beta$  and  $v = -V \sin \beta$ . The radar can detect motion only along its beam. At any moment, then, the Doppler velocity  $V_D$  is given by wind components along the direction  $\alpha$ . Therefore

$$V_D = u \sin \alpha + v \cos \alpha = V(\sin \alpha \cos \beta - \cos \alpha \sin \beta). \quad (1)$$

It is useful to express  $\beta$  in terms of other variables, and this can be done through solution of the right triangle in Fig. 1. Accordingly,

$$\sin \beta = (r \sin \alpha) / h \quad (2)$$

$$\cos \beta = (R + r \cos \alpha) / h \quad (3)$$

and after substitution of Eqs. (2) and (3) into Eq. (1),

$$V_D = V(\sin \alpha)R/h. \quad (4)$$

It is also useful to express  $h$  in terms of the controllable variables  $r$  and  $\alpha$  and the distance  $R$  from radar to center of curvature of the wind field. After considerable manipulation of the right-triangle relationships of Fig. 1, the final result is

$$V_D = \underbrace{V \sin \alpha}_{\text{VAD}} \underbrace{\left[ 1 + r^2/R^2 + 2(r/R) \cos \alpha \right]^{-1/2}}_{\text{CWC = circular wind correction}}. \quad (5)$$

Note that  $V \sin \alpha$  is the velocity-azimuth display, or VAD, of a constant linear wind moving toward  $\alpha = 90^\circ$ , and the term in brackets with its exponent is the modification of the VAD effected by curvature of the wind field.

For all values of  $r/R < 1$ , the wind field crosses the radar scanning circle at right angles at two points. At these points, the wind vector is directed along the radar beam, and the maximum speed  $V$  and minimum speed  $-V$  are observed. The angular locations of the maximum and minimum in the Doppler velocity pattern are obtained by setting  $V_D = V$  in Eq. (5). It all reduces to

$$\cos \hat{\alpha} = -r/R \quad (6)$$

where the  $\hat{\alpha}$  refers to both maximum and minimum Doppler velocity.

If the speed of the circular wind field is not constant, but is proportional to distance from the center of the circulation, the wind field is in solid rotation and its speed is

$$V = V_{\max} h/h_{\max} \quad (7)$$

where  $V_{\max}$  is the speed attained at  $h_{\max}$ , the radius of maximum wind. Let  $V = V_R$  at  $h = R$ , the distance of the radar from circulation center. As long as the scanning circle is entirely within a semicircle of the solidly rotating wind field, the radar encounters a variable wind speed given by Eq. (7). Substitution of this speed function, and its value at  $h = R$ , into Eq. (4), which is the Doppler velocity in a circular wind field, yields

$$V_D = V_{\max} (h/h_{\max}) \sin \alpha (V_R/h) (h_{\max}/V_{\max})$$

which neatly simplifies to

$$V_D = V_R \sin \alpha \quad (8)$$

which is identical to Doppler velocity in a linear wind field of constant speed, directed toward  $\alpha = 90^\circ$ .

Preliminary studies are under way of the Doppler velocity patterns in linear wind fields with constant speed gradients both along the flow and across the flow. Numerical treatment indicates a superficial resemblance of the latter condition to the circular flow with constant speed. These may be distinguished, however, through control of  $r$  and through comparison of opposite halves of the harmonic analysis. These studies will be pursued in 1984 with emphasis on the specification of the appropriate wind model from observed Doppler velocity patterns, which, it is hoped, will lead to at least rough estimates of synoptic-scale vorticity from single-Doppler observations.

## B. Study of Relative Storm Motions

The first hour in the lifetime of a convective storm that can be revealed by radar is usually characterized by the emergence of many small and distinct echoes. Although some of these emerging echoes soon decay or merge with another, others remain as distinct entities for periods of ten minutes or more. Relative motions among surviving nearest-neighbor echo pairs may be measured because radars employed for study or operational analysis of convective storms are generally programmed to repeat the same antenna scanning procedure every five to seven minutes.

In the few convective storms that produce severe weather, radar-detectable predictors of the damaging events typically appear during the second hour of storm lifetime. The most reliable severe-storm predictor is the mesocyclone, a persistently rotating portion of a storm, averaging five km in diameter, which can be detected by means of Doppler radar. Nearly all mesocyclones are associated with some form of severe weather, and about half of them spawn one or more tornadoes. Furthermore, the warning time between verification of a mesocyclone and tornado touchdown averages 20 minutes.

The processes leading toward establishment of a mesocyclone have not been clearly defined, except for the recognition that thermodynamic conditions favorable to well-organized, energetic convection are required. Certainly the mismatch in scale between synoptic observations and convective processes inhibits understanding of the antecedent conditions for mesocyclones.

However, it seems reasonable to postulate that one of the pre-conditions for mesocyclone development is an environmental wind field with convergence and cyclonic vorticity on an inter-storm scale of 50 km or so. It also seems reasonable to use emerging storm echoes as imperfect but perhaps revealing tracers of the environmental wind field prior to mesocyclone development. Any predictive value of the early inter-storm echo motions would be useful in identifying storms with enhanced probability of severe weather production, but not so distinctively that an immediate warning would be justified. Such storms should command greater attention than others and perhaps be subjected to specialized surveillance techniques,

because some of them are likely to develop destructive power at a later time.

An excellent case for testing the value of relative storm motions was found in the observations of the Joint Doppler Operational Project in Oklahoma. Two storm groups, separated by a little less than two hours in start-up time and 100 km in location, produced strongly contrasting weather events. The first one of these failed to establish a mesocyclone and produced no severe weather. Near the town of Rush Springs, several funnel clouds were observed, but they did not extend to the ground. The second, later storm developed an intense and long-lived mesocyclone, several small tornadoes, and finally an intense maxi-tornado that destroyed a corner of the town of Piedmont and adjacent rural areas.

There was a striking contrast between the two storms in the kinematic wind field properties inferred by study of relative echo motions during the earliest 40 minutes in which echo pairs appeared at low levels in each storm. Inferred values of convergence in the tornadic Piedmont storm averaging a value of  $6 \times 10^{-4} \text{ s}^{-1}$  exceeded those in the non-severe Rush Springs storm by more than an order of magnitude. Also, the average vorticity inferred in the Rush Springs storm, only  $10^{-5} \text{ s}^{-1}$ , was very weakly anticyclonic; the corresponding vorticity value of  $4 \times 10^{-4} \text{ s}^{-1}$  in the Piedmont storm was strongly cyclonic. A more detailed discussion of this case and rationale for its analysis were given by Donaldson and Snapp (1983)<sup>9</sup> in a paper prepared and presented during the reporting period.

Two other cases of relative storm motion during the initial hour were also examined. On a day with no severe storms, the relative echo motions were slight, indicating negligible convergence and vorticity. On a day with a severe hailstorm accompanying a mesocyclone, the mean inferred vorticity during the earliest half hour was cyclonic and about the same value,  $4 \times 10^{-4} \text{ s}^{-1}$ , as in the tornadic Piedmont storm, but the inter-storm motions were also very weakly divergent, with an average value less than  $0.4 \times 10^{-4} \text{ s}^{-1}$ .

9. Donaldson, R.J., Jr., and M.R. Snapp, 1983: Interstorm motion as a mesocyclone precursor. Preprints, 21st Conference on Radar Meteorology, AMS, Boston, 7-10.

These investigations, conducted manually and to a certain degree subjectively, are time-consuming and obviously have no immediate operational application. They do, however, indicate that relative echo motions during the first hour of a storm's lifetime may have some predictive value for assessment of the probability of subsequent development of a mesocyclone and severe weather events. Accordingly, a search for automated techniques appears to be justified.

An initial attempt to model inferred divergence by measurement of temporal variability in echo reflectivity distribution in a horizontal plane showed some promise, but was thwarted by strong differences in the growth rate of the various echoes. This led to the conceptualization of a technique independent of echo growth rate. The basic premise of this technique is that one can identify cells by locating peaks in the reflectivity pattern. This premise was the basis of a tracking algorithm by Crane (1979)<sup>10</sup> in which he identified peaks only three dB above background. Wieler, Harris and Snapp (1982)<sup>11</sup> showed that Crane's three dB peaks are not always significant and can not always be tracked. SASC has since developed a fast algorithm that allows one to detect peaks at any arbitrary level and to correlate them in two dimensions. To date the technique has been applied to data from the Piedmont, OK, tornadic storm of April 30, 1978, and has been successful in detecting ten dB peaks and distinguishing between weather targets and ground echoes. The next step is to examine the temporal changes in the distribution of those peaks and to determine if reasonable correlations can be made using the nearest-neighbor assumption. That is, it will be determined if peaks observed at successive times that are closest spatially can be reliably identified as the same physical feature. If this identification process is successful, a tracking algorithm will be developed. Otherwise an examination of the statistics of the distributions of echoes and how they change with time will be pursued.

10. Crane, R.K., 1979: Automatic cell detection and tracking. IEEE Trans. Geoscience Elect., GE-17, 250-262.

11. Wieler, J.G., F.I. Harris, and M.R. Snapp, 1982: An Evaluation of An Automatic Cell Detection and Tracking Algorithm. AFGL-TR-82-0368, Air Force Geophysics Laboratory, Hanscom AFB, MA.

### C. Radar Detection of Gust Fronts

The gust front is usually defined as the shallow boundary between cold air flowing out of a storm at low levels and warm air flowing in toward the storm. Abrupt changes in wind speed, direction, or both (commonly referred to as shear) along these fronts near the ground pose hazards to aircraft take-offs and landings. With a Doppler radar scanning in the lower planetary boundary layer there is potential for identifying these strong shears since the Doppler radar can detect particulate motion directed along its beam. These motions have been shown to approximate closely the component of the actual local airflow. With this information locations of strong shears can be pinpointed. This section summarizes preliminary SASC results in developing an automated shear detection algorithm and suggests future investigations and improvements.

In the AFGL data archives there are two gust front cases which occurred on April 30 and May 2, 1978 in Oklahoma. Both are described by Donaldson and Bjerkaas (1980).<sup>12</sup> The April 30 gust front exhibits a bow-like configuration with both a long-beam and a cross-beam orientation. The cross-beam segment does not show a change in sign in the radial velocities, as one often expects with gust fronts. Instead, rather strong radial velocities ( $20\text{--}30\text{ ms}^{-1}$ ) are observed on either side of a line of weak velocities ( $2\text{--}6\text{ ms}^{-1}$ ), all with components toward the radar. In contrast, the May 2 front was oriented nearly parallel to the beam and had weak radial velocities. However, a change in sign of the radial velocities across the front is evident. These are two quite different cases and should provide a good test for any gust front detection algorithm.

Several differencing and averaging schemes were employed in an effort to find the optimum technique for gust front (shear line) detection. First efforts involved computation of gate-to-gate (radial and tangential) shears. Software for averaging velocities and shears radially in variable length blocks (one to nine gates) was written. Velocity averages could be

12. Donaldson, R.J., Jr., and C.L. Bjerkaas (1980): Gust front structure observed by Doppler radar. Preprints, 19th Conference on Radar Meteorology, AMS, Boston, 157-161.

computed first and then the shears, or shears computed first and then averaged. There was little difference between the results of the velocity averaging and the shear averaging. Of course, the nine-gate averaged data yielded much smoother results than did the unaveraged data.

In the April 30 case high computed shears associated with apparent noise in gates ahead of the storm are greater than the shears along and in the immediate vicinity of the front. This noise could be removed by imposing a threshold of 15 dBZ. However, this value is much too high to use in all cases, as gust fronts can occur in regions of weak returns with reflectivities less than ten dBZ. This is often seen in the mature stages of the storm, when the front is several tens of kilometers ahead of the leading edge of precipitation.

It seems likely that this inability to identify the gust front uniquely may be eliminated by taking differences in radial velocity values (RVD's) between gates spaced several kilometers apart rather than between consecutive gates. For both April 30 and May 2, lines of the highest difference values from each radial coincide with the front. However, where the front seems to be oriented along the beam or possibly washed out, these differences are not significant and the front cannot be found by simply looking for highest shears or differences.

In an effort to deal with this problem of the front being parallel to the beam, one might consider computing tangential (azimuthal) shear. However, tangential shears in the May 2 case are weak. One possible alternative not yet investigated would be to compute azimuthal RVD's.

Because only two gust front cases exist in AFGL's ten cm Doppler data, additional gust front data were selected and acquired from the library of the National Severe Storms Laboratory. In all, eight cases were obtained. Since software has been written or modified to accommodate NSSL data, the schemes mentioned earlier in this section, and others, will be applied to these data. The problems of forecasting movement of the front and estimating peak gusts at the ground during frontal passage will be addressed.

## D. Mesocyclone Detection and Classification Algorithm

### 1. Introduction

Mesocyclones, the vorticity parents of tornadoes, are rotating regions located within some intense and well-organized thunderstorms. Mesocyclones have been successfully detected for many years through visual inspection of the velocity field supplied by a Doppler radar. Visual analysis of Doppler data in real time, however, demands the constant and alert attention of a meteorologist. Economical and reliable employment of human resources in an operational scenario requires the aid of an automated detection scheme. SASC scientists have developed such a scheme in supporting AFGL in the Next Generation Weather Radar (NEXRAD) program.

### 2. The Detection Algorithm

Criteria for the detection of mesocyclones were set forth by Donaldson (1970),<sup>13</sup> and later by Burgess (1976).<sup>14</sup> Three basic requirements are:

- \* Significant azimuthal shear must exist between closed velocity contours of opposite sign (provided storm motion has been removed).
- \* Shear pattern and closed isodops must extend vertically for a height interval comparable to the horizontal diameter.
- \* Shear pattern must persist for a time interval greater than one half of the revolution period of the feature.

This algorithm has evolved from a two-dimensional mesocyclonic shear algorithm proposed by Hennington and Burgess (1981),<sup>15</sup> further developed by Zrnic et al. (1982),<sup>16</sup> and a three-dimensional mesocyclone tornado vortex

13. Donaldson, R.J., Jr., 1970: Vortex signature recognition by Doppler radar. J. Appl. Meteor., 9, 661-670.

14. Burgess, D.W., 1976: Single Doppler radar vortex recognition: Part I - mesocyclone signatures. Preprints, 17th Conference on Radar Meteorology, AMS, Boston, 97-103.

15. Hennington, L.D., and D.W. Burgess, 1981: Automatic recognition of mesocyclones from single Doppler radar data. Preprints, 20th Conference on Radar Meteorology, AMS, Boston, 704-706.

16. Zrnic, D.S., L.D. Hennington, and J. Skelton, 1982: Automatic Recognition of Mesocyclones from Single Doppler Radar Data. AFGL-TR-82-0291, Contract ESD 2-0949, National Severe Storms Laboratory, NOAA.

signature (TVS) algorithm developed by Forsyth et al. (1981).<sup>17</sup> We have taken some of the attributes of these algorithms and developed an algorithm that runs in real time and automatically archives the mesocyclone attributes for later analysis.

The mesocyclone detection algorithm is a four-dimensional resolution-dependent detection algorithm. It was initially designed to run in a stand-alone fashion on data supplied from the AFGL five cm Doppler radar. Because our initial case studies gave promising results the algorithm was modified to run concurrently with other special analysis modules (SAMS) within the real-time Modular Radar Analysis Software System (MRASS).

The algorithm is capable of detecting cyclonic and anticyclonic circulations over a wide range of scales. It computes the horizontal and vertical extent, average shear, momentum, rotational kinetic energy, and temporal persistence of these circulations. We have included the detection of tornado vortex signatures (TVS's) in our algorithm, since the same assumptions can be made about TVS's and mesocyclones in deriving our resolution dependency.

Hennington and Burgess (1981)<sup>15</sup> and Zrnic et al. (1982)<sup>16</sup> used a combination of momentum and shear thresholds to determine if a pattern vector is saved for further processing. In order to compensate for the cross-beam degradation of resolution with range, our algorithm applies resolution-dependent thresholds for shear and peak velocity difference as criteria for retaining pattern vectors. The momentum of a pattern vector is defined as the product of the measured velocity difference ( $V_e - V_b$ ) and the azimuthal distance  $R(O_b - O_e)$ ; shear is the quotient of these two quantities.

The velocity pattern around a non-divergent mesocyclone is generally modelled by a Rankine combined vortex (RCV). This means that the core of the mesocyclone, defined as the region between velocity peaks of opposite sign, is assumed to rotate as a solid body. As the mesocyclone is moved out in range the peak velocities decrease significantly as more and more of the mesocyclone is encompassed by the radar pulse volume. At the same time the diameter of the feature appears to increase.

17. Forsyth, D.E., C.L. Bjerkaas, and P. Petrocchi, 1981: Modular radar analysis software system (MRASS). Preprints, 20th Conference on Radar Meteorology, AMS, Boston, 696-699.

We have derived our resolution dependence by considering the effect of increasing range on the resolution of a RCV. Brown and Lemon (1976)<sup>18</sup> present some enlightening results of a RCV model developed by Zrnic and Doviak (1975).<sup>19</sup> This paper addresses the problem of identifying a TVS described by a RCV as it is moved out in range.

In this analysis we have assumed data are collected at increments equal to the antenna beamwidth. Thus the basic indicator of the resolution dependence of a feature is the ratio of the radar beamwidth (BW) to core radius (CR). If this ratio is much less than unity the radar can resolve the meso-circulation fairly well. As this ratio increases there is a degradation in the observed feature. At beamwidth to core ratios greater than two the true attributes of the feature are difficult if not impossible to resolve.

A minimum azimuthal shear of  $5 \times 10^{-3} \text{ s}^{-1}$  for mesocyclone detection was first proposed by Donaldson (1970)<sup>13</sup> and later corroborated by Burgess (1976)<sup>14</sup> from a study of several years of mesocyclone data. Brown's and Lemon's data show that a beamwidth twice the core radius would smooth the peak velocities to approximately 62 percent of their original magnitude, and the core radius would appear to be slightly larger than its true size. Consequently the shear reported for the feature would be somewhat less than 62 percent of the 'true' shear. If the ratio of the beamwidth to the core radius of a tornadic feature is large the apparent core radius can be much larger than the actual size. For example, when the beamwidth is three times the core radius the peak velocities detected by the radar may be 1.5 core radii from the center of the rotation. Likewise for beamwidth to core radius ratios of five and ten, the peak velocities will be located at 2.5 and five core radii from the center of the rotation. Since no objective criteria for the detection of TVS's have been established we based our thresholds on the

18. Brown, R.A., and L.R. Lemon, 1976: Single Doppler radar vortex recognition: Part II - Tornadic vortex signatures. Preprints, 17th Conference on Radar Meteorology, AMS, Boston, 104-109.

19. Zrnic, D.S., and R.J. Doviak, 1975: Velocity spectra of vortices scanned with a pulse-Doppler radar. J. Appl. Meteor., 14, 1531-1539.

guidelines adopted during the Joint Doppler Operational Project (JDOP) (Burgess et al., 1979<sup>20</sup>). These are the presence of shear greater than  $5 \times 10^{-2} \text{ s}^{-1}$  and a rotational velocity greater than  $30 \text{ ms}^{-1}$ .

An illustration of the decrease of velocity and shear across a RCV or mesocyclone as the BW/CR ratio is increased is shown in Fig. 2. Since we can only resolve features with a BW/CR ratio of less than two, we have assumed a linear drop off from our thresholds of  $15 \text{ ms}^{-1}$  and  $5 \times 10^{-3} \text{ s}^{-1}$ .

### 3. Preliminary Results

As a test of the algorithm we have processed several volume scans of data collected during JDOP. Our test data are from April 30, 1978, starting at 1810 CST, just prior to the Piedmont tornado touchdown. In the first volume scan the algorithm identifies a well-developed mesocyclone. Output from the second volume scan reveals that the mesocyclone is moving to the east and that there is a two-dimensional TVS shear at elevation angles of 5.0 and 6.9 degrees, the peak velocities at these levels being 33 and  $-36 \text{ ms}^{-1}$ . When the antenna drops down from the next volume scan (1823 CDT) we have a two-dimensional TVS shear with peak velocities of 35 and  $-36 \text{ ms}^{-1}$  at 100 m above ground level. On the completion of the third volume scan the algorithm identifies a three-dimensional TVS. The Piedmont tornado touched down at 1820 CDT, at which time radar was scanning at high elevation angles. The development of this tornado was rapid at low levels; the tornado was on the ground by the time the antenna dropped to 0.3 degrees. This feature is identified by the algorithm in two more volume scans and appears to weaken slightly although maintaining very strong shear.

### 4. Conclusions

A resolution-dependent pattern recognition algorithm for detecting mesocyclones and tornado vortex signatures has been developed. The algorithm identifies three-dimensional cyclonic and anticyclonic circulations and computes their average shear, angular momentum, and rotational kinetic energy.

20. Burgess, D.W., R.J. Donaldson, Jr., T. Sieland, and J. Hinkelman, 1979: Final Report on the Joint Doppler Operational Project (JDOP) 1976-1978: Part I - Meteorology Applications. NOAA Tech. Memo ERL NSSL-86, 84 pp.

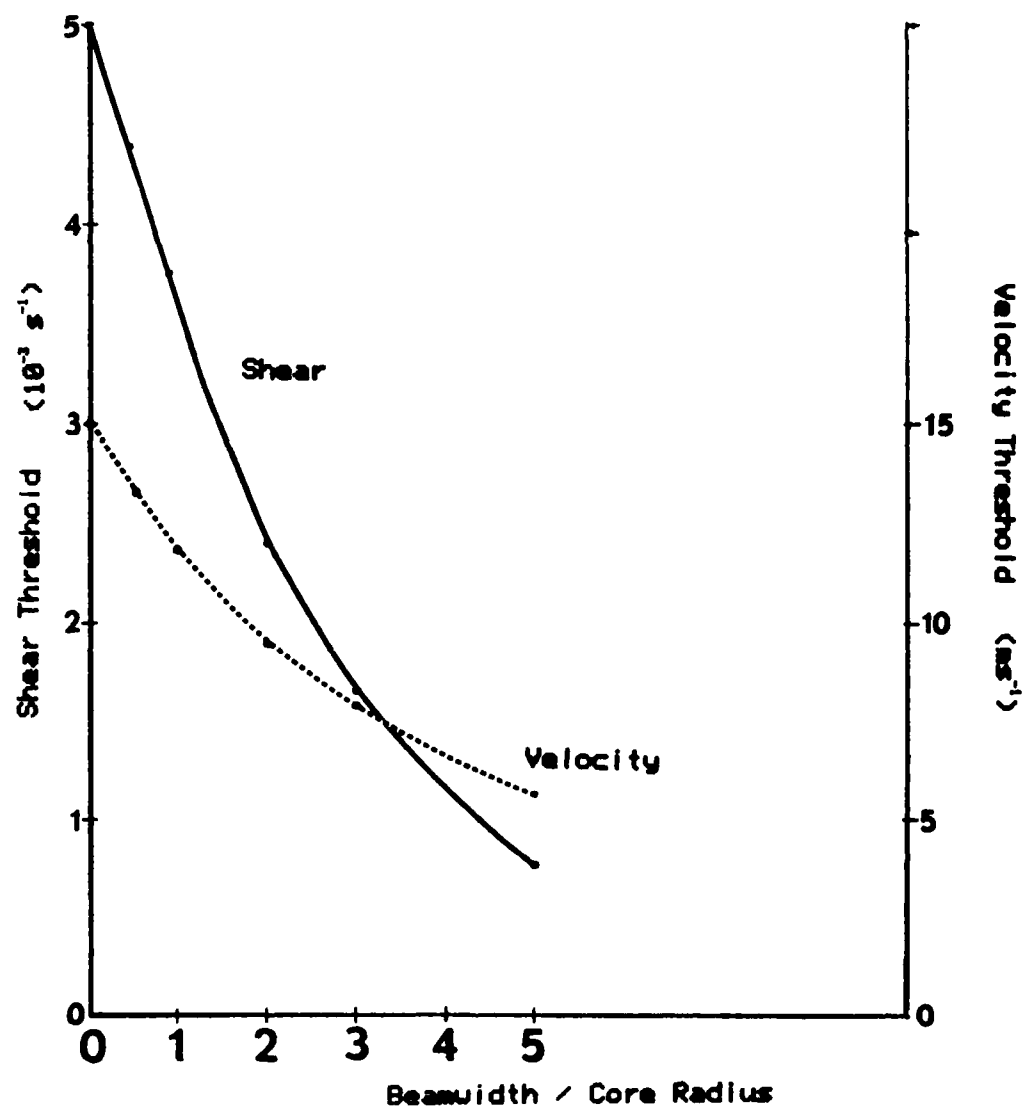


Fig. 2. Resolution-Dependent Thresholds for Shear and  $\Delta$ Velocity, as a Function of the Ratio of Radar Beamwidth (BW) and Vortex Core Radius (CR)

The time continuity of the detected features is confirmed by a persistence criterion, and the peak velocities at every elevation angle are recorded. Initial test of the algorithm appears promising. More studies must be done in this area to improve our understanding of mesocyclones and the mechanisms involved as they spawn tornadoes. The interactions among a mesocyclone's momentum, shear, and kinetic energy throughout its life cycle should yield some valuable statistics regarding the potential destructive power of the circulation.

Further refinement of the algorithm will take place as more data are processed. In particular the minimum feature size, the velocity difference and shear thresholds, and the beamwidth to core radius ratio cut-offs can only be evaluated after extensive testing. Probability of detection and false alarm statistics need to be compiled over a wide range of cases in order to give us a better understanding of the algorithm's performance. For an in-depth discussion of the algorithm's processing scheme the reader is referred to Wieler and Donaldson (1983).<sup>21</sup>

21. Wieler, J.G., and R.J. Donaldson, Jr., 1983: Mesocyclone detection and classification algorithm. Preprints, 13th Conference on Severe Local Storms, AMS, Boston, 58-61.

#### E. Sub-Storm Scale Motions from Single Radar

SASC scientists investigated the feasibility of extracting sub-storm scale airflow ( $\sim 5$ -15 km) from the data of one radar using correlation techniques similar to those of Rinehart (1979)<sup>22</sup> and Smythe and Zrnic (1983).<sup>23</sup> Details of this investigation are available in Smythe (1983)<sup>24</sup> and Smythe and Harris (1983).<sup>25</sup> There are two important conclusions from this effort:

1. The techniques do not detect airflow but apparently track precipitation generators that exist at heights above the analysis levels.
2. Spatial and temporal resolution available with the typical weather radar results in substantial errors in the precipitation pattern tracks.

The basic recommendation in Smythe and Harris (1983) is that such correlation techniques cannot be used reliably as wind detection schemes in an automated processing and analysis environment. Therefore no further effort will be expended on this task.

22. Rinehart, R.E., 1979: Internal Storm Motions from a Single Non-Doppler Weather Radar. NCAR/TR-146+STR, 262 pp.

23. Smythe, G.R., and D.S. Zrnic, 1983: Correlation analysis of Doppler radar data and retrieval of the horizontal wind. J. Clim. Appl. Meteor., **22**, 297-311.

24. Smythe, G.R., (1983): Sub-storm scale motions as determined from single-Doppler and non-Doppler radar data. Preprints, 13th Conference on Severe Local Storms, AMS, Boston, 109-112.

25. Smythe, G.R., and F.I. Harris (1983): Sub-cloud layer motions from radar data using correlation techniques. Scientific Report (in preparation), Contract F19628-82-C-0023, Systems and Applied Sciences Corporation.

## F. Systems Support

There were two main efforts during 1983: one involved development of software to use the Chromatics CGC 7900 color display as an integral part of the AFGL MRASS real time analysis and display system; the second involved development of software to facilitate evaluation of radar data.

### 1. Chromatics-Related Software Development

At present the Perkin-Elmer 3242 minicomputer not only analyzes the radar data but also produces all of the graphical output and sends this output to the Chromatics 7900 in picture form; i.e., the intelligent Chromatics is being used simply as a dumb terminal. This unnecessarily ties up the host computer, since the Chromatics could be generating the graphics. During the past year considerable effort has been expended in attempting to transfer the display generation task to the Chromatics. First, it has been essential to establish communication between the operating systems of the two machines. When the P-E 3242 uses the CGC 7900 as a dumb terminal, data can be transferred at 9600 baud. But when communications between the operating systems were attempted, transfers at any rate were initially found to be impossible. After considerable experimentation, however, SASC succeeded in achieving a data transfer rate between the two operating systems at 4800 baud. This was accomplished by means of a routine written on the P-E 3242 that changes the attributes (baud rate, parity, stop bits, data bits) of the CGC 7900. This routine is general enough to change the attributes of any interactive terminal.

A parallel effort to solving the communication problems has been the conversion over to the Chromatics of the MRASS routines for graphics generation that are resident on the P-E 3242. All of the software has been modified, with about 25 percent having been converted to the language C. The only routines that have not been totally debugged at this time are REFPL (the reflectivity plotting routine) and TRKPL (the map plotting routine). The REFPL routine as it is configured on the P-E 3242 uses a rather large data file (~ 0.5 Mbytes in 50 Kbyte records). Attempts to transfer this file to the Chromatics have been unsuccessful, since the Chromatics cannot handle such large files, and even if it could, the time required to transfer

the file would be prohibitive for real time operations. Currently several alternatives are being explored. These data transfers will be further facilitated by development of a monitor program that will control the software acquisition of data.

The TRKPL routine is currently configured to plot the map of Oklahoma, where the routine was first used. It will be modified to draw a Massachusetts map.

## 2. Radar Evaluation Software

During the past year, several software programs were developed or modified to facilitate the evaluation of radar data collected by the AFGL radar. These programs involve calibration and display of the data, determination and display of the radar beam illumination pattern, and evaluation of the individual data bits going from the radar processor to the computer used for analysis.

Calibration software development involved taking existing software that directed the calibration process and expanding it to allow greater versatility, establishing an archive of all calibrations, and plotting the calibration data on the TRILOG printer plotter, with an option to plot onto the Chromatics 7900. Radar operators are now for the first time able to preserve a calibration for later evaluation and to obtain a hard copy plot of any calibration, current or past. Fig. 3 is a sample plot of the output.

Before one can evaluate the radar data obtained from weather targets, knowledge of the illumination pattern of the transmitted beam is essential. Software was developed to utilize AFGL hardware to measure, display, and analyze the beam patterns for the AFGL radar. The signal from a transmitter on Nobscot Hill, eight km from the AFGL radar, is measured as a function of azimuth and/or elevation angle. The data are plotted unsmoothed on the TRILOG printer plotter along with a log of all ancillary data. In addition, the data are analyzed to yield the intensity, location, and width of the main lobe and side lobe. A sample output is shown in Fig. 4. This beam pattern is an intermediate pattern obtained while AFGL personnel were making adjustments to the antenna feed. Current one-way side lobes are approximately 18 dB below the peak. Output from this program

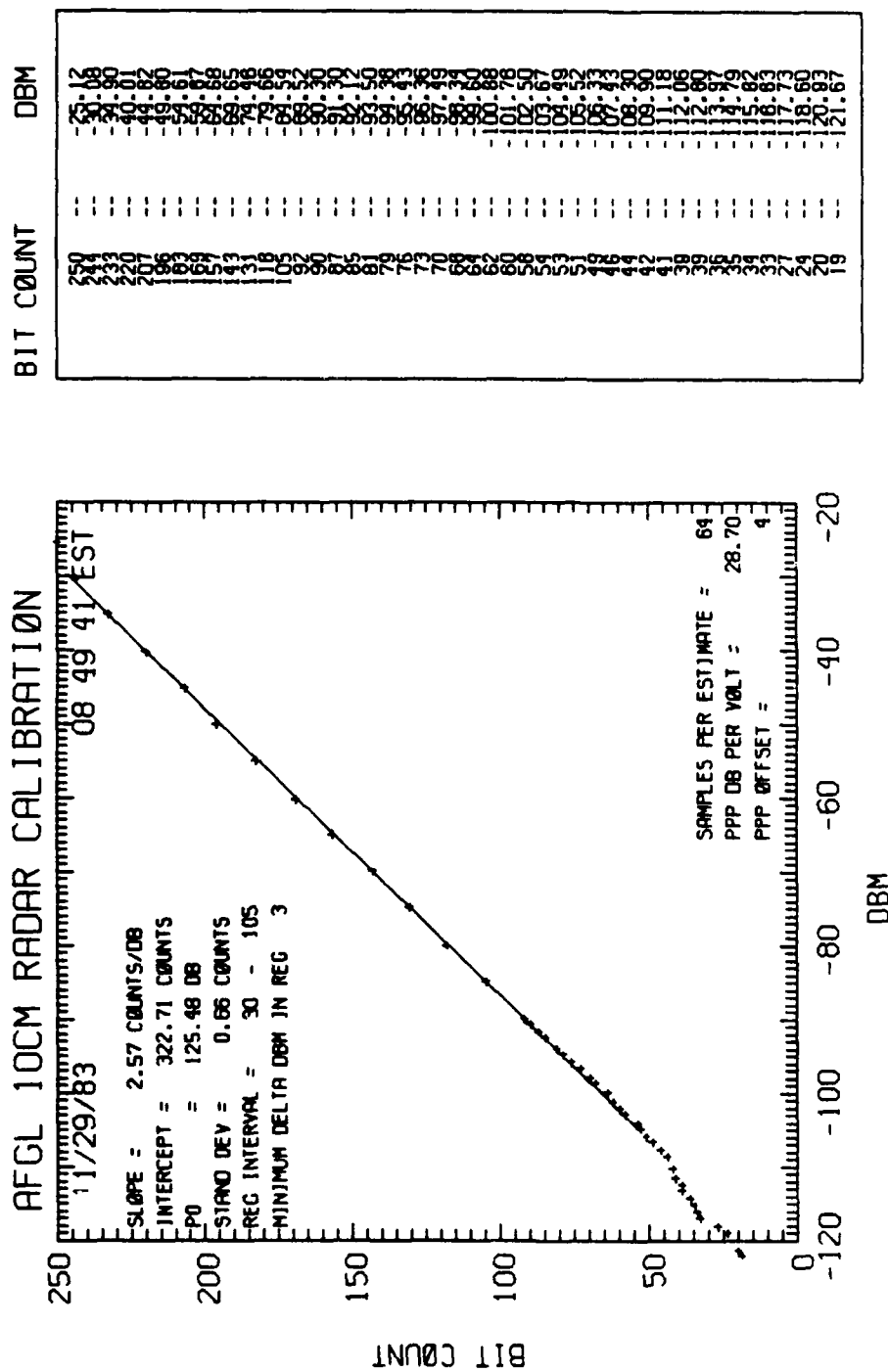


Fig. 3. Plot of Radar Calibration for the AFGL 10 cm Radar Located in Sudbury, MA

(Bit count data from the processor are plotted against their corresponding power levels in dB below a milliwatt. A least-squares fit has been plotted for the range of -30 to -100 dBm. Characteristics of this curve are printed in the upper left of the figure. Data values used to plot these points are included in the table on the right.)

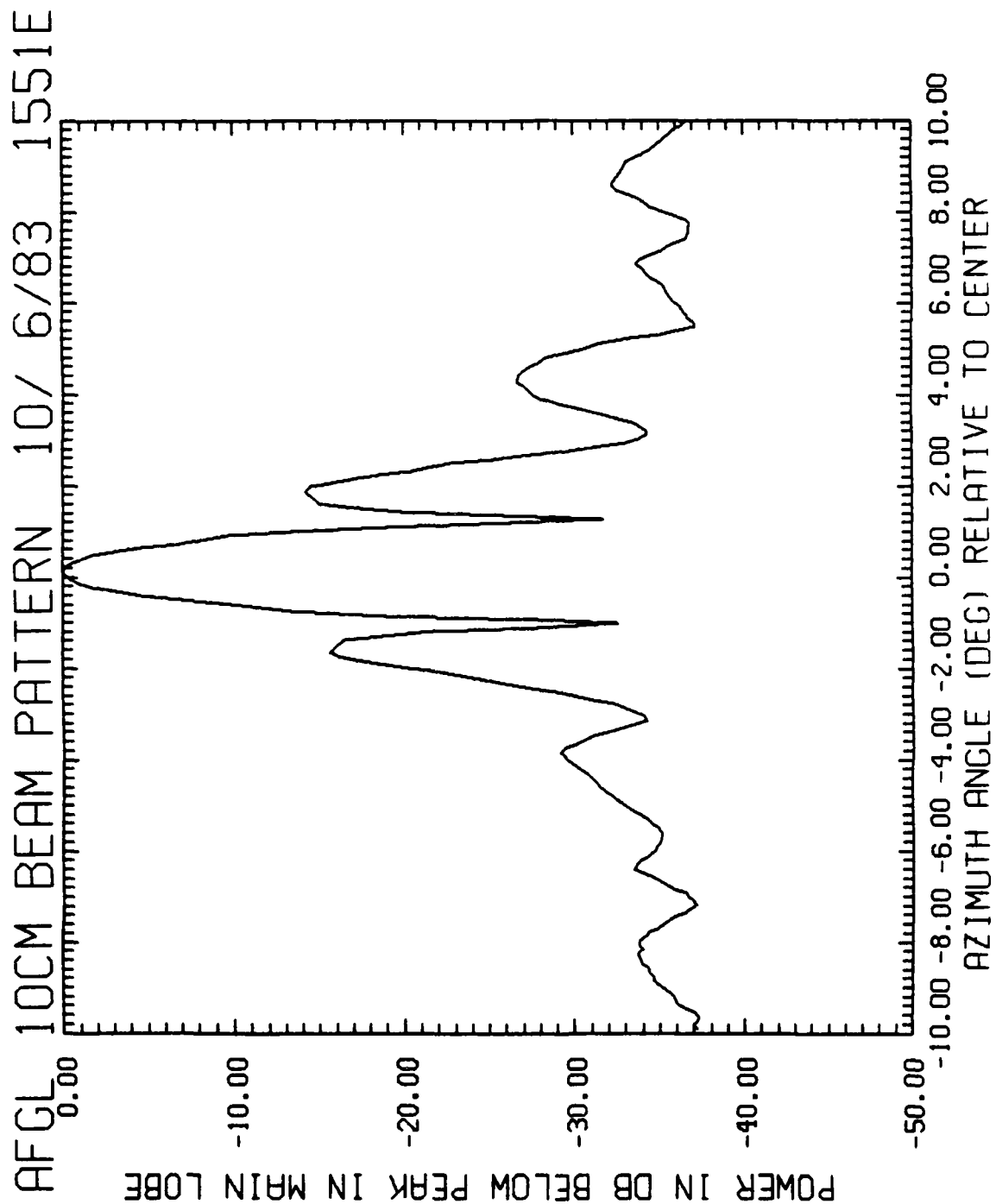


Fig. 4. Plot of the One-Way Beam Pattern for the AFGL 10 cm Radar  
(Power is plotted against azimuth angle. Power is measured relative to peak value at the center; azimuth is measured relative to location of transmitter on Nobscot Hill.)

has been compared with output of an antenna pattern recorder and agrees perfectly. This software capability provides AFGL with the unique capability of obtaining beam pattern determinations at any time and under any weather conditions.

A problem with all radars in the past has been the improper processing or recording of the radar data. To reduce the possibility of this occurrence with the AFGL radar, a program was written to monitor the data entering the P-E 3242 minicomputer from the processor. This program checks the bits in each word and determines whether each is set or not and accumulates the number of occurrences of each event for each bit. This exercise allows one to check on whether certain bits are being "hung up" or not.

Another function of this program is to check the output of the velocity, variance, and reflectivity channels of the processor. A histogram of the number of occurrences of each bitcount is constructed for each channel, with the added capability of thresholding the velocity and variance data with respect to received power. This program has already been used to find a bias in the velocity data at very low power levels. Also, it confirmed the erratic nature of the variance estimates and a bit problem in velocity estimated at low power levels. The intent is to run this program periodically as a quick quality control test for the processor.

### 3. Computer Operations

SASC software and hardware support to AFGL/LYR included acquisition of service contracts with Perkin-Elmer for maintenance of the P-E 7/32, P-E 6/16 and P-E 3242 computers and from Scopus for preventive maintenance of the removable disk packs for these machines. In addition, Perkin-Elmer, Bartlett Associates, and Chromatics, Inc. handled hardware problems not covered by the maintenance agreements.

SASC subscribes to the Perkin-Elmer user subscription service which automatically provides updates to all system software. In the past year SASC received through this service the latest operating system (Rev. 6.2) and associated software.

## G. CO<sub>2</sub> Doppler Lidar

### 1. Introduction

AFGL and SASC are developing a CO<sub>2</sub> laser wind system in response to an Air Force need to determine the existence of wind shear in the vicinity of airfields and to evaluate aircraft wake vortices. General Electric Co. (Syracuse) designed a high energy, high pulse repetition frequency, pulsed heterodyne CO<sub>2</sub> laser/scanner system for the above application. However G.E. was not able to assemble and test the lidar system within the contractual funding. The project is being completed by AFGL and SASC personnel.

A reorganization of AFGL resulted in transfer of the project from the Atmospheric Sciences Division to the Optical Physics Division. A two-month delay in the project was incurred.

### 2. Trailer Preparation

Design changes in the trailer interior modification plans were submitted to correct previous oversights and to accommodate changes in the lidar system configuration.

Final installation and testing of the trailer's environmental control system was performed; it was found to be operating properly.

Work was begun on the RFI shielding of the control area in the forward section of the trailer. The shielding is needed to prevent the high levels of RFI emanating from the laser system from interfering with the operation of the data processing and storage systems.

Source of toxic fumes that were accumulating within the trailer was determined. Imported plywood used to panel the trailer walls contained a volatile and extremely irritating bonding agent. Removal of the plywood and replacement with domestic plywood eliminated the problem.

### 3. CO<sub>2</sub> Laser System

Working with the AFGL engineering group SASC designed an elevator system to support and raise and lower the laser scanner through the roof of the trailer.

The entire laser system was moved to a laboratory area where it was assembled as a mock-up, to verify the compatibility of the assembled system's dimensions with the modifications being made to the trailer's interior.

Work was begun on assembly of the low pressure CO<sub>2</sub> local oscillator

and injection lasers. When assembly is completed they will be operationally tested on the system mock-up.

An assessment of the potential safety hazards associated with the laser system was performed. As a result of that assessment, procurement of the required safety equipment and design of a safety interlock and failsafe system were begun.

A Technical Memorandum on the Doppler lidar system was published by AFGL.<sup>26</sup> The memorandum describes each of the major system components and gives their then-current assembly or operational status. The general assembly and test plan for the system is also outlined.

#### 4. CO<sub>2</sub> Lidar Milestones

##### Completion dates

- |                                                                       |               |
|-----------------------------------------------------------------------|---------------|
| a. Lidar trailer modifications                                        | December 1983 |
| b. Preparation of environmental assesment                             | December 1983 |
| c. Assembly and testing of low pressure lasers                        | February 1984 |
| d. Installation of laser system into trailer                          | March 1984    |
| e. Preliminary testing of lidar system in a non-Doppler configuration | January 1985  |

#### 5. Lidar Beam Modeling

##### a. Introduction

To evaluate the effects of the beam pattern of the new 10.6  $\mu$ m AFGL lidar system, a modelling effort was initiated that simulates the beams. The transmit and receive apertures on this lidar are co-axial annuli, with the transmit aperture nested within the receive aperture (see Fig. 16 of Jacobs, 1982<sup>27</sup>). The transmit annulus has an inside diameter (ID) of 120 mm and an outside diameter (OD) of 160 mm while the ID and OD for the receive aperture are 200 mm and 300 mm, respectively. The annular shape and the non-coincidence of the two apertures must be taken into account in determining the appropriate constants for the lidar equation and the expected noise level for the Doppler estimations.

26. Moroz, E.Y., S.B. Alejandro, and F.I. Harris, 1983: System Integration and Test Plan for the AFGL CO<sub>2</sub> Lidar Doppler Wind Measuring System. AFGL Tech. Memo No. 85, Air Force Geophysics Laboratory, Hanscom AFB, MA.

27. Jacobs, G.B., 1982: Heterodyne Detector/Electronics Design and Alignment of a Pulsed CO<sub>2</sub> Laser Wind System. AFGL-TR-82-0348, Contract F19628-80-C-0184, General Electric Co., Syracuse, NY.

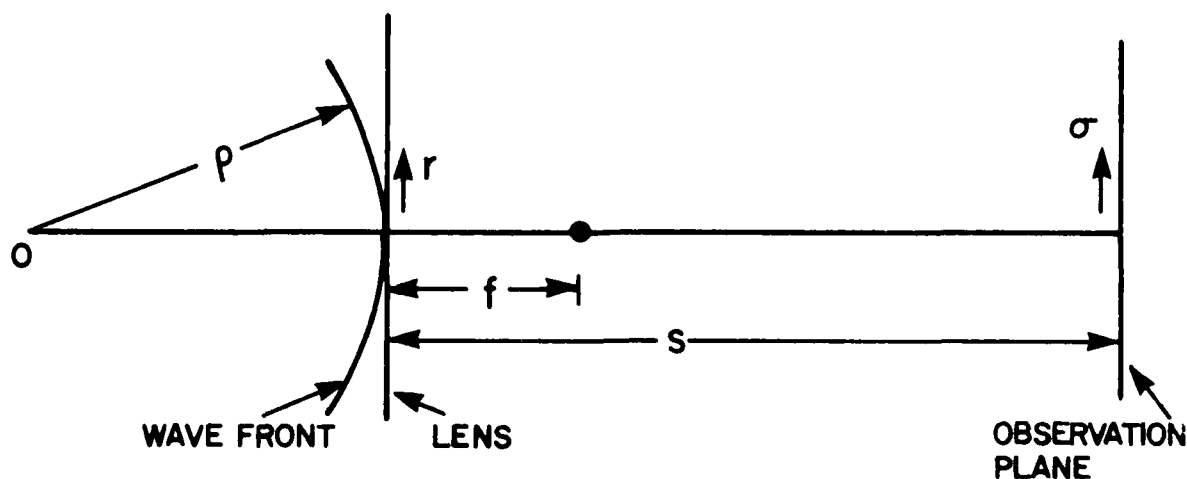
The purpose of the current research effort is to shed light on expectations for the new system, and to provide a basis for the determination of a beam correction factor for the  $\beta$  measurements. Since not all of the details of the transmitted beam and telescope, for example, the amplitude and phase distribution across the transmit aperture, nor all the details of the focusing capabilities of the telescope are known, the theory and resultant software being used allow for a high degree of flexibility with regard to beam and aperture parameters.

#### b. Theory

To assess the propagation effects of an electromagnetic beam the following simplifying but realistic assumptions are involved:

- \* The beam is circularly symmetric.
- \* All considerations are to be in the Fraunhofer or far-zone for the aperture.

The approach used here is to follow Dickson (1970)<sup>28</sup> by solving the Kirchhoff diffraction integral in the reduced form derived by Boivin (1964)<sup>29</sup> for circularly symmetric pupil functions.



28. Dickson, L.D., 1970: Characteristics of a propagating Gaussian beam. Appl. Opt., 9, 1854-1861.

29. Boivin, A., 1964: Théorie et Calcul des Figures de Diffraction de Révolution. Les Presses de l'Université Laval, Québec.

We assume a point source at 0, resulting in a beam that has a radius of curvature at the lens of  $\rho$  and a coherently illuminated complex amplitude pupil function  $F(r^2)$  where  $F(r^2)$  is circularly symmetric about  $r = 0$  at the lens. The lens has a focal length of  $f$  and is confined to a radius of  $0 \leq r \leq a$ . The amplitude distribution  $A$  is then examined at some distance  $s$  from the lens as a function of  $\sigma$ , the radial distance from the beam axis at  $s$ . The equation to be solved then is

$$A(z, s, f, \rho) = \frac{C_0}{s} \int_0^a F(r^2) \exp \left[ \frac{i\pi}{\lambda} \left( \frac{1}{s} - \frac{1}{f} + \frac{1}{\rho} \right) r^2 \right] \cdot J_0(zr) r \, dr \quad (9)$$

where  $\lambda$  is the transmitted wavelength,  $z = \frac{2\pi}{\lambda} \cdot \frac{\sigma}{s}$ , and  $J_0$  is the zero-order Bessel function. This equation is solved numerically for various values of  $\lambda$ ,  $s$ ,  $f$ , and  $F(r^2)$ . The source is assumed to be sufficiently distant from the mirror for  $1/\rho$  to be considered small.

#### c. Computational Method

Eq. (9) has been put into the following finite difference form:

$$A_j = \frac{C_0}{s} \sum_{i=0}^n F(r_i^2) \left[ \cos \left\{ \frac{2\pi}{\lambda} r_i^2 \left( \frac{1}{s} - \frac{1}{f} \right) \right\} + i \sin \left\{ \frac{2\pi}{\lambda} r_i^2 \left( \frac{1}{s} - \frac{1}{f} \right) \right\} \right] \cdot J_0(z_j r_i) r_i \Delta r \quad (10)$$

where  $n\Delta r = a$ . This expression is then evaluated for particular values of  $s$ ,  $f$ , and  $\lambda$  and for sufficient  $z$ 's ( $\sigma$ 's) so that beam shape can be adequately defined. From the real and imaginary parts of  $A_j$  we then compute the intensity ( $I$ ) and phase ( $\phi$ ) distributions as a function of  $\sigma$  where

$$I = \text{Re}(A)^2 + \text{Im}(A)^2$$

$$\phi = \tan^{-1} \left[ \frac{\text{Im}(A)}{\text{Re}(A)} \right]$$

#### d. Computational Sensitivity

It is important to ensure that the results obtained through numerical methods are not artifacts of the numerical method itself. A number of tests were performed with particular emphasis on sensitivity to resolution.

#### 1) Resolution in $r$ for Integration

Here the steps in  $r$  must be sufficiently small for the three functions containing  $r$  (Eq. (10)) to be well-defined. Two of these functions are periodic, the trigonometric and the Bessel. In each of these, the interval in  $r$  is confined to be at most ten percent of the wavelength. For the configuration described in the introduction it has been found that a resolution of 0.4 mm will produce results without apparent artifacts. However, in order to be conservative a resolution of 1/5 of that, namely 0.08 mm, has been used for most calculations.

#### 2) Resolution in $\sigma$

For this test, the resolution was related to the number of points as a function of beamwidth. Beamwidth is proportional to the wavelength of the transmitted radiation divided by the diameter of the transit aperture ( $D$ ). The resolution in  $\sigma$  is therefore specifiable through

$$\Delta\sigma = \frac{\lambda}{D \cdot N}$$

where  $N$  is any specified number. Through tests it was found that  $N = 100$  is sufficient to produce unaliased results. Here again, a conservative approach is used and  $N$  is taken to be 200 for most computations.

#### e. Results

The software that has been developed is highly interactive, allowing a number of parameters to be varied easily. These variables include:

- \* Beam specification

- shape
- size
- amplitude distribution (Gaussian or top hat, truncated or infinite)
- location of focus
- wavelength

- \* Computation specification

- resolution in  $r$  and  $\sigma$

The output is in the form of cross-sections transverse to the beam axis of the beam intensity and phase.

#### f. Circular Versus Annular

A series of cases was run where a circular Gaussian beam 20 mm across evolved into an annulus with a Gaussian amplitude distribution. The amplitude at the edge of the beam is ten percent of the peak amplitude. Figs. 5a,b through 8a,b show the results of these calculations for a ten km range and a focus at 99999 km. Fig. 5a, the circular Gaussian result, has a main lobe with a half-power half beam width (HPHW) of about 3.47 m and a series of side lobes almost 40 dB below the peak. The phase (Fig. 5b) is nearly zero to about nine m, beyond which there are sharp deviations occurring on the edges of the lobes. For Fig. 6 the peak amplitude is shifted ten mm off the beam axis, resulting in a double peaked transmitted beam at the aperture, with 20 mm separating the two peaks. The HPHW is now 1.58 m and the first lobe is only -15.75 dB below the peak. As with the circular Gaussian, the phase remains nearly zero across most of the main lobe. Further separation of the peaks results in a narrowing of the main lobe (Figs. 6a-8a) and a decreasing of the difference between the peaks in the main and side lobes, to the point where at a separation of 140 mm the main lobe is 0.54 m across (two times the HPHW) but the first side lobe is only 8.13 dB below the peak. The phase (Figs. 7b-8b) remains flat across the beam but becomes shifted some  $10^\circ$  from the optimum of  $0^\circ$ .

One can conclude the following: First, to keep the side lobes small and the phase minimally affected, the circular Gaussian is optimal; second, the only real benefit of increasing an annulus size is to decrease the HPHW. There is a point of diminishing returns here, however, because of the accompanying side lobe amplification and, perhaps, the shift in the phase. However, the effects of the latter are not known at this time.

#### g. Truncation Effects

For a simple Gaussian beam extending to infinity the resultant beam pattern at any range is of course another Gaussian, with no side lobes. It is the truncation of the pattern that yields the lobes. For a fixed transmitter configuration, i.e., an annulus of 70 mm radius and 20 mm width, the level of truncation was varied. The amplitude at the beam edge compared to

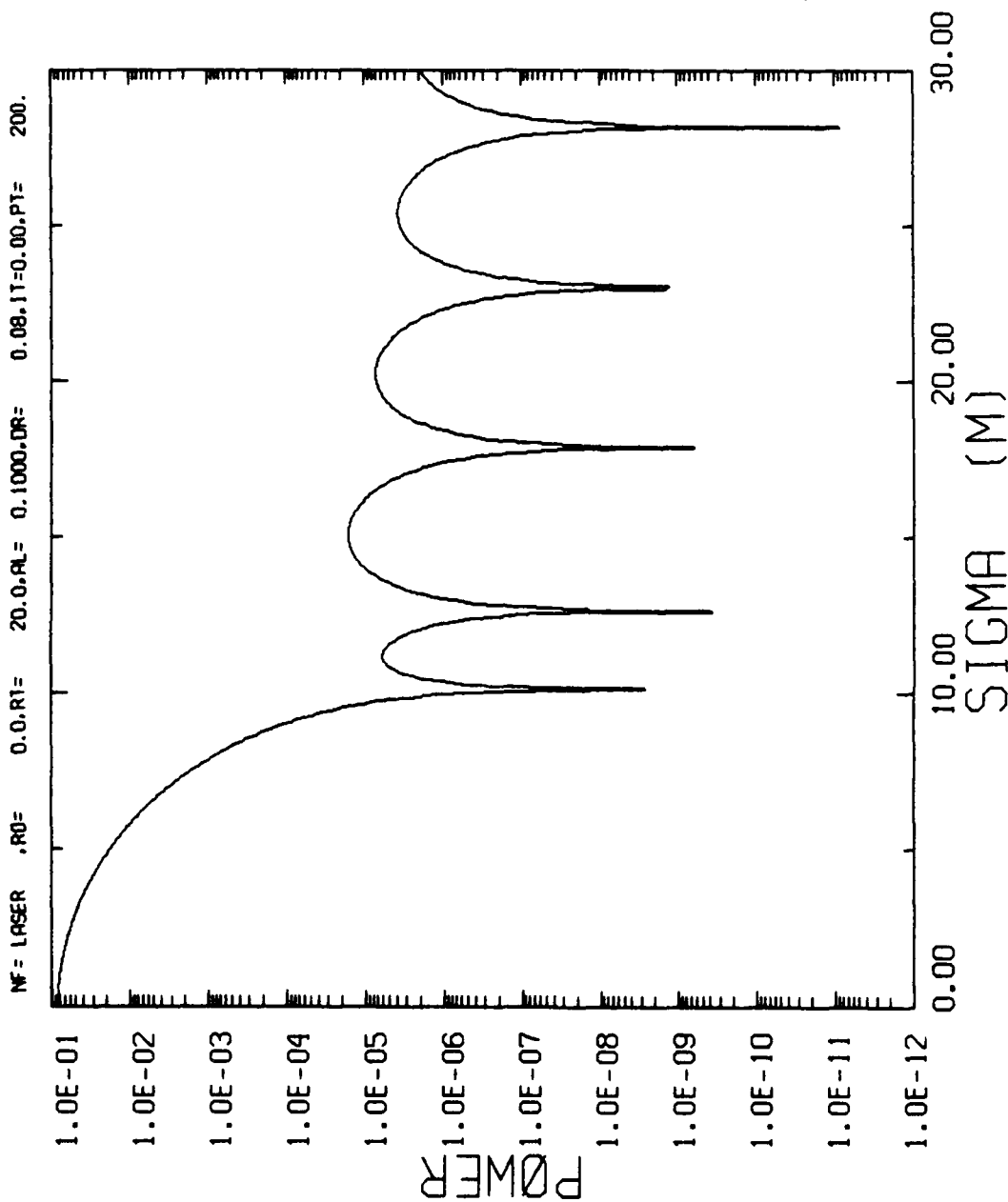


Fig. 5a. Plot of One-Way Power (arbitrary units) vs. Distance from Beam Axis of a Lidar Beam at a Range of 10 km from the Telescope  
(Beam at telescope is circular Gaussian of 20 mm radius with resolution of 0.08 mm and focus at infinity. Beam is truncated at 20 dB below peak power.)

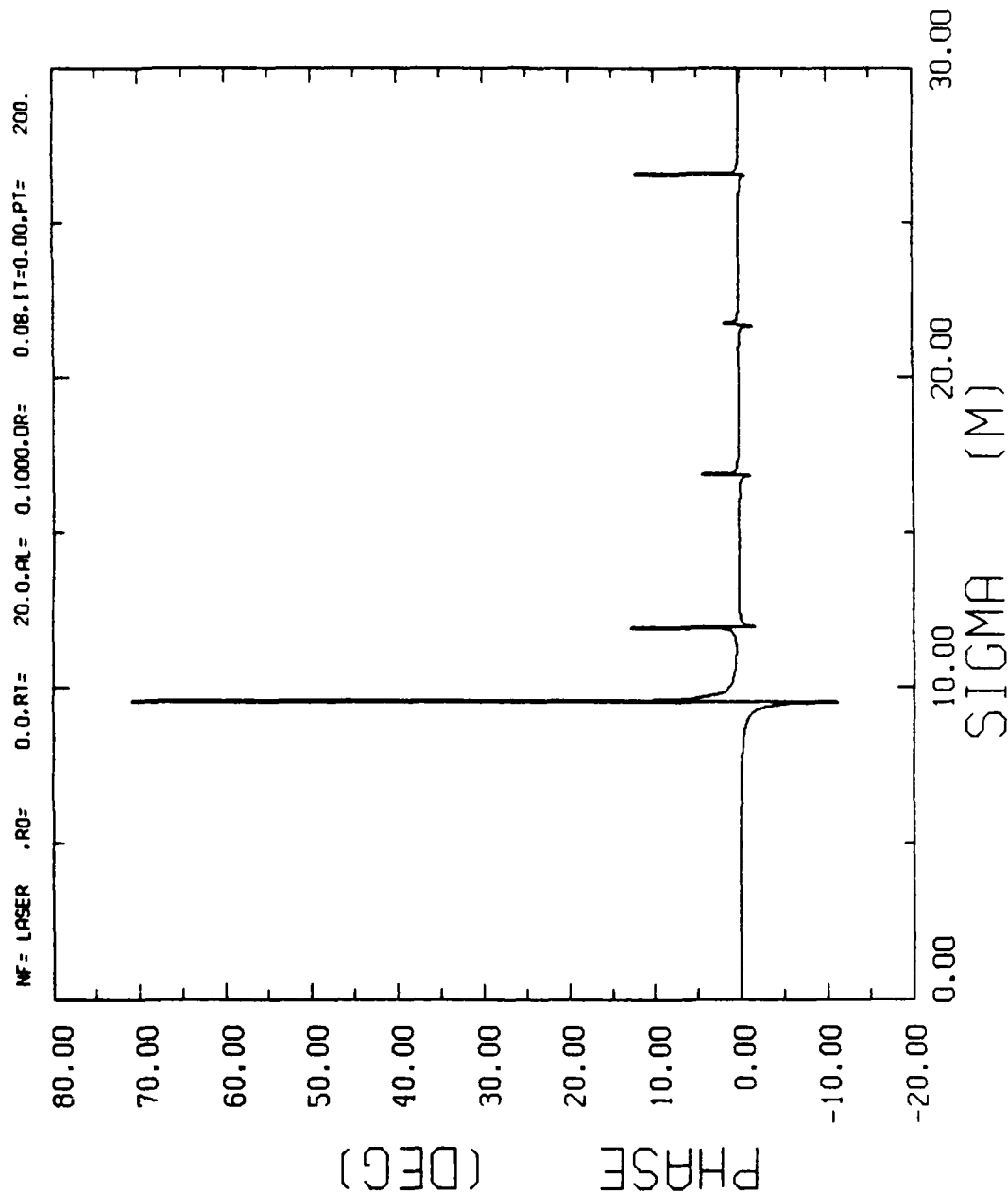


Fig. 5b. Plot of Phase (deg) of the Transmitted Beam vs. Distance from the Beam Axis for Conditions Noted in Fig. 5a

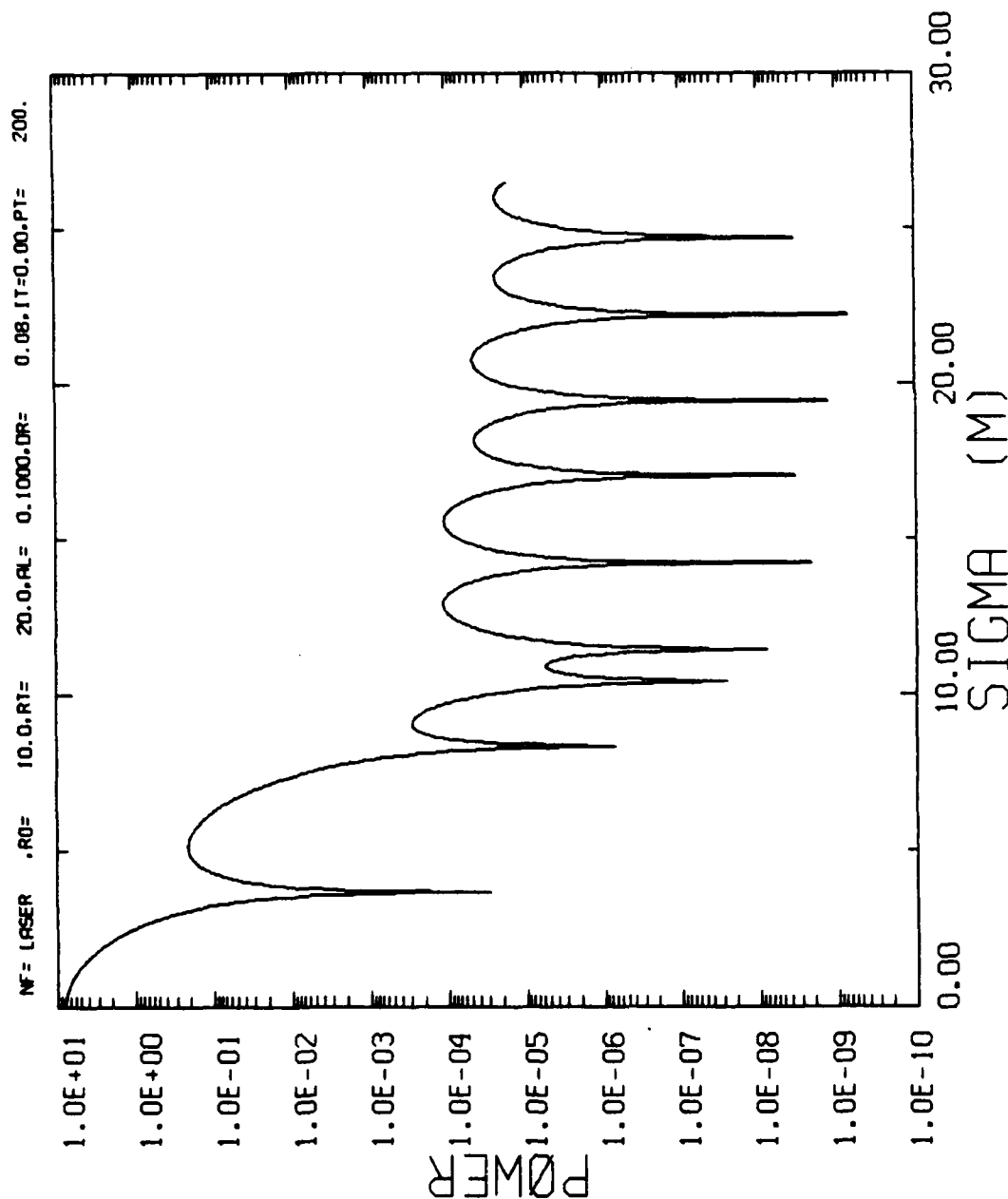


Fig. 6a. As in Fig. 5a but for Annular Beam with a 20 mm Wide Annulus Centered 10 mm from Beam Center

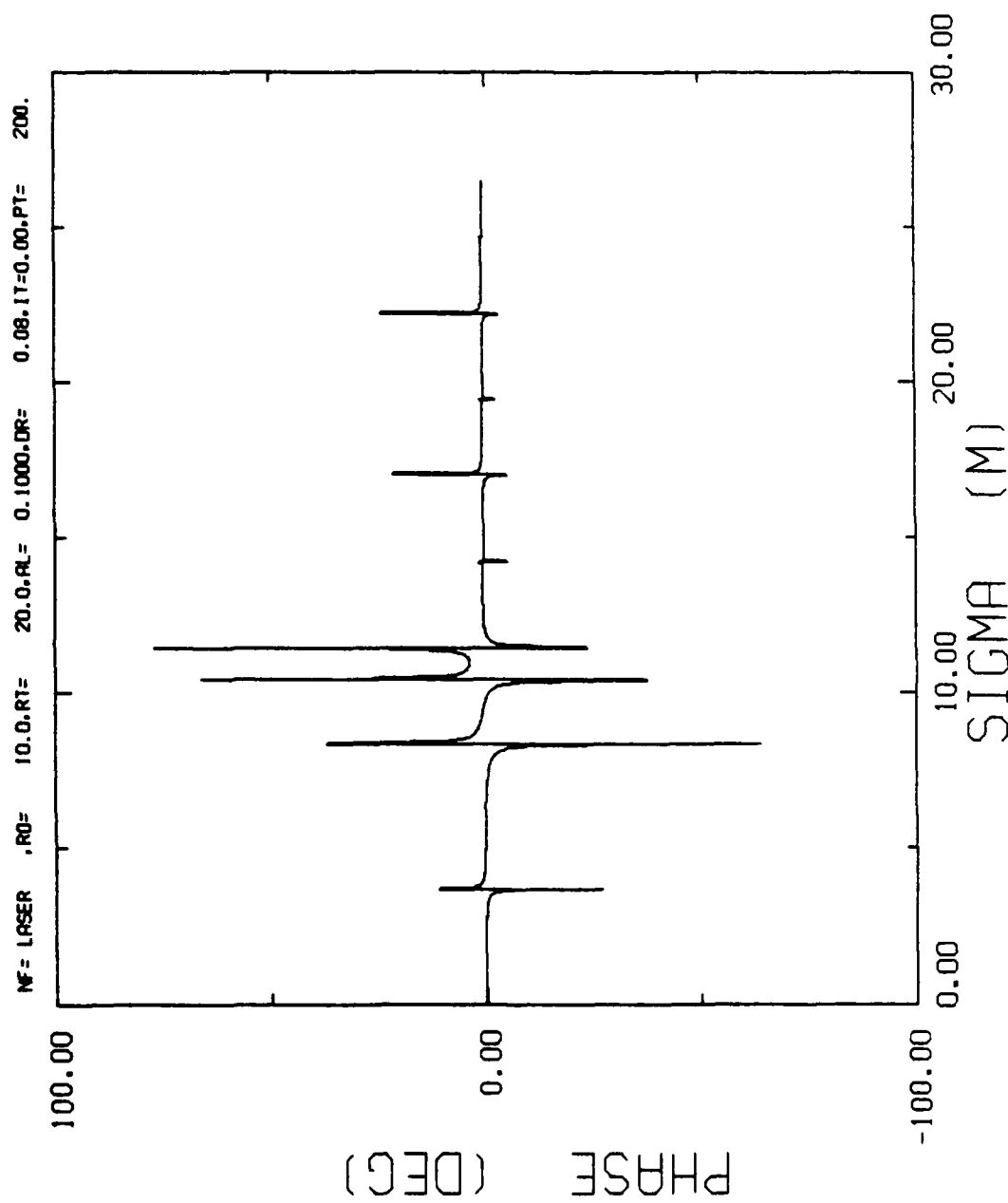


Fig. 6b. As in Fig. 5b but for Annular Beam with a 20 mm Wide Annulus Centered 10 mm from Beam Center

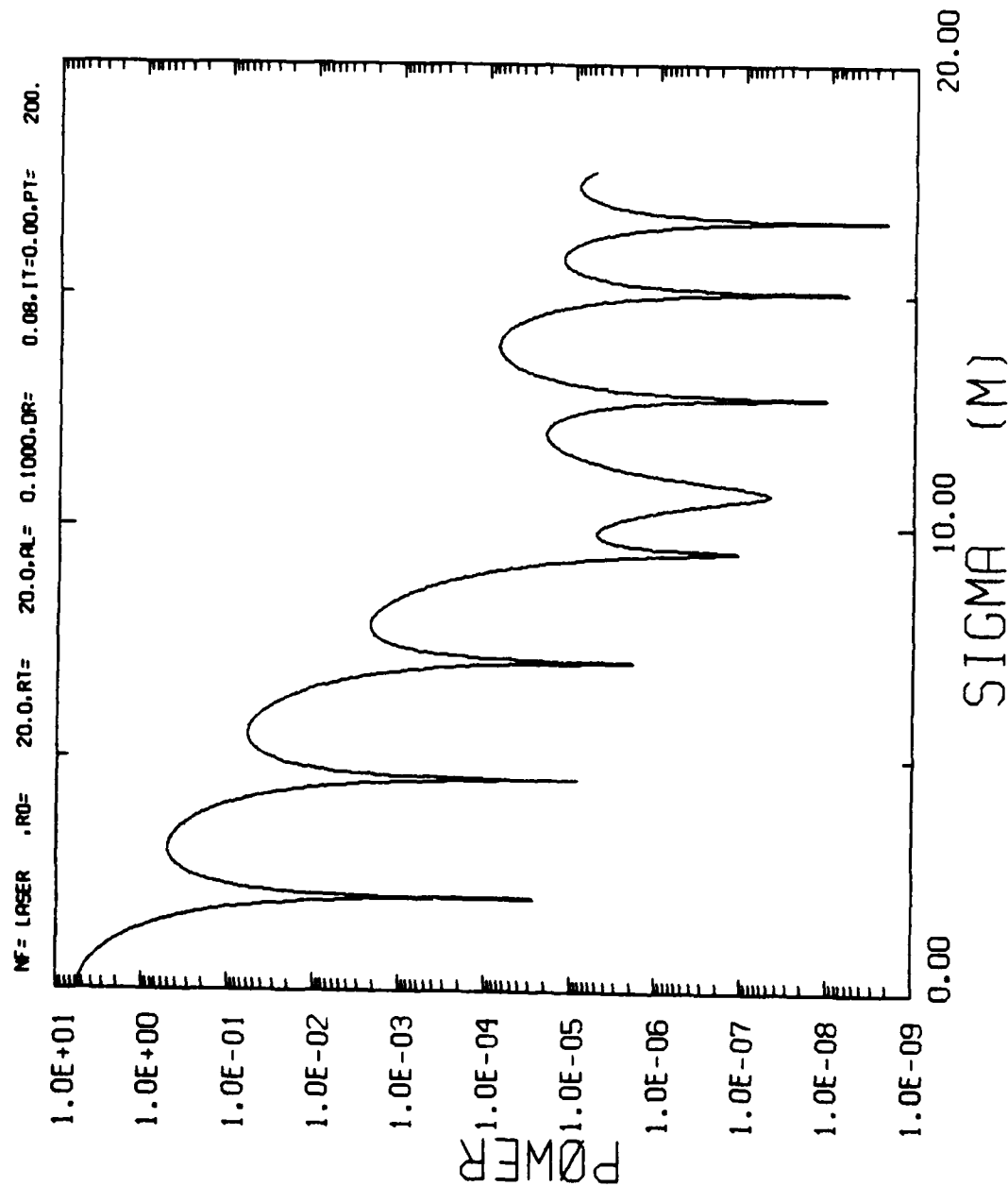


Fig. 7a. As in Fig. 5a but for Annular Beam with a 20 mm Wide Annulus Centered 20 mm from Beam Center

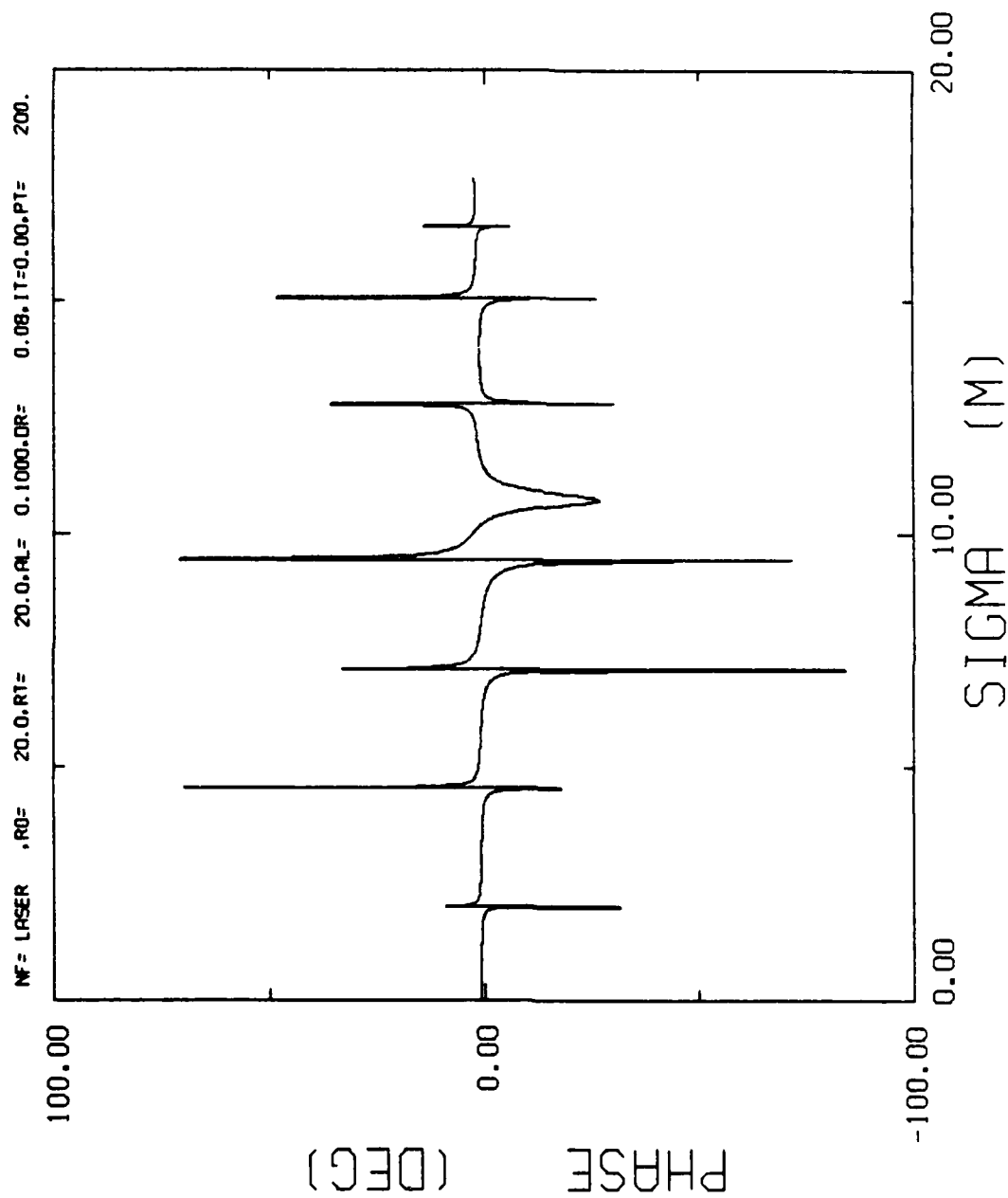


Fig. 7b. As in Fig. 5b but for Annular Beam with a 20 mm Wide Annulus Centered 20 mm from Beam Center

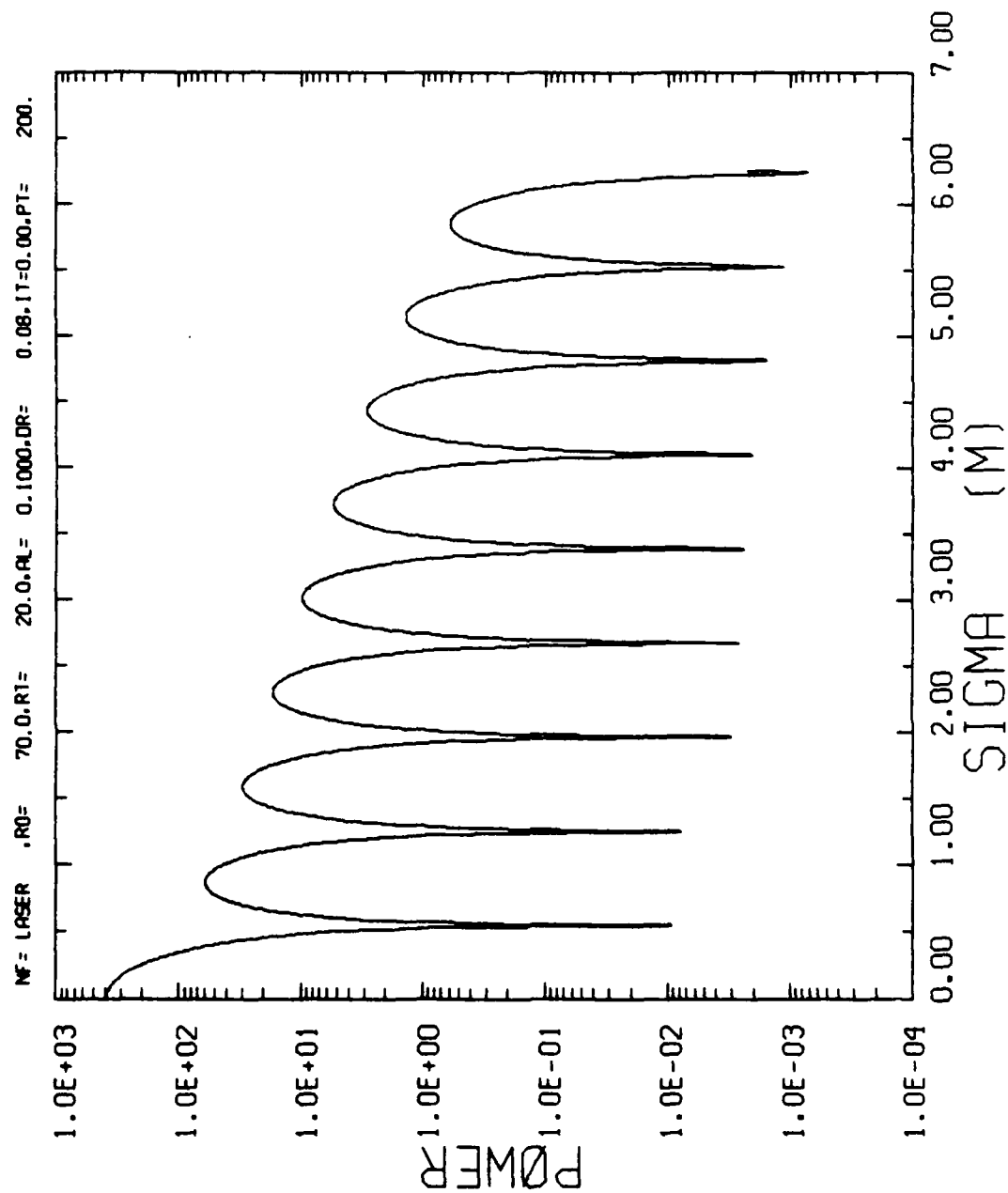


Fig. 8a. As in Fig. 7a but with Annulus Centered 70 mm from Beam Axis

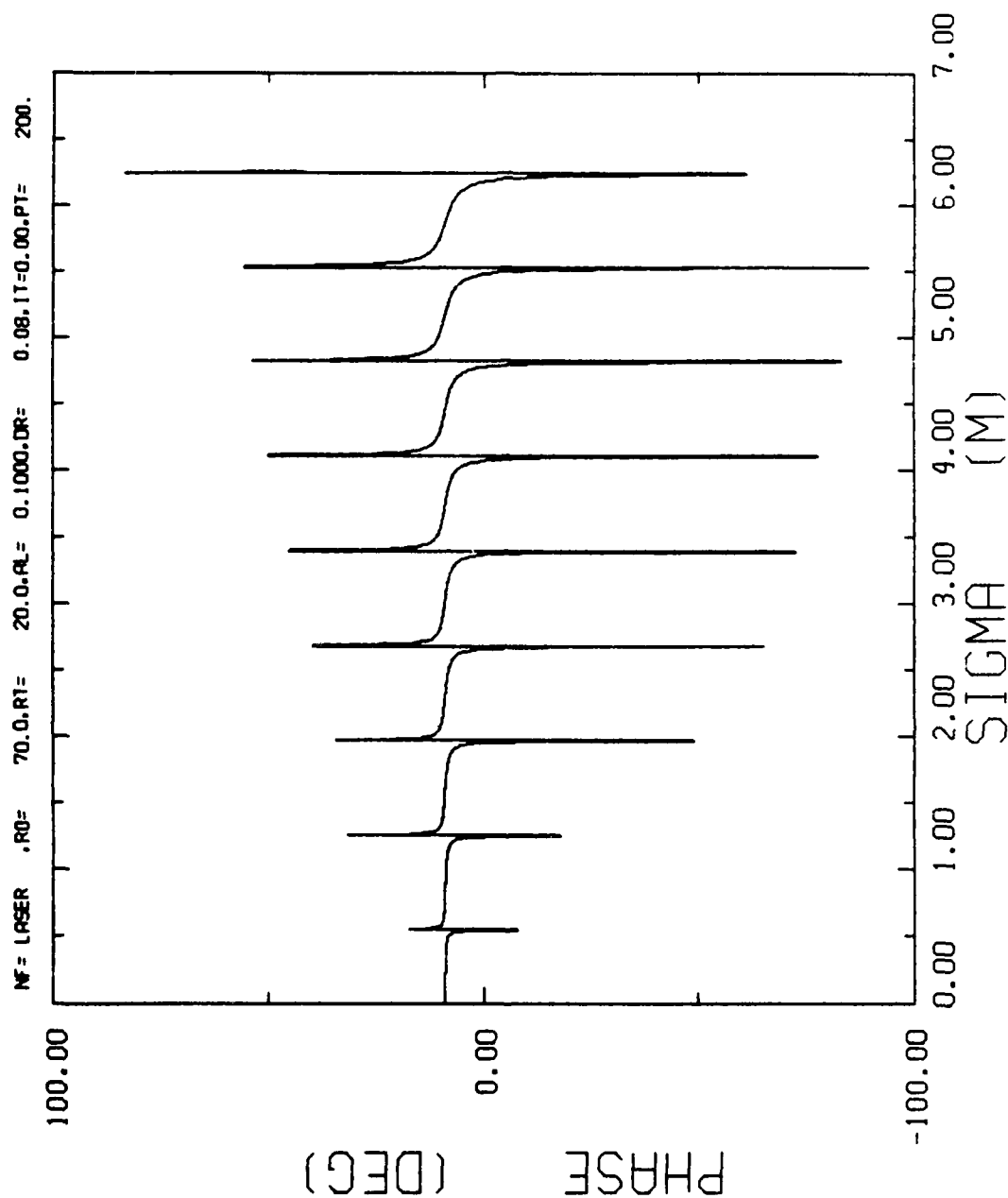


Fig. 8b. As in Fig. 7b but with Annulus Centered 70 mm from Beam Axis

that at the center was varied from 0.7 down to 0.1 (see Figs. 9-11). While the net effect on the beamwidth is minimal, the side lobe structure is affected quite significantly. The best behaved structure in terms of lobe peaks decreasing monotonically occurs with the lower truncation values (Figs. 9-11). Interestingly enough, the fall-off is actually worst for the minimal truncation.

We can conclude from this and the previous section that the important factor in determining beam quality (of those examined here) is the geometry of the transmitting aperture. The distribution of amplitude across the beam is not at all important for most practical considerations.

#### h. Summary

Through a model of the transmission of a coherent electromagnetic beam, a better understanding of the structure of a lidar has been obtained. It is now understood that an annular beam like that of the AFGL lidar will transform in the far-field to a main central lobe with side lobes, similar in structure to a circular beam. The main lobe beamwidth is significantly smaller for the annular beam but the side lobes are more than 30 dB higher than the side lobes due to a circular aperture. It is also shown that the effects of amplitude distribution across the beam are not an important factor.

The software that has been developed by SASC will be released to AFGL personnel for further expansion and development.

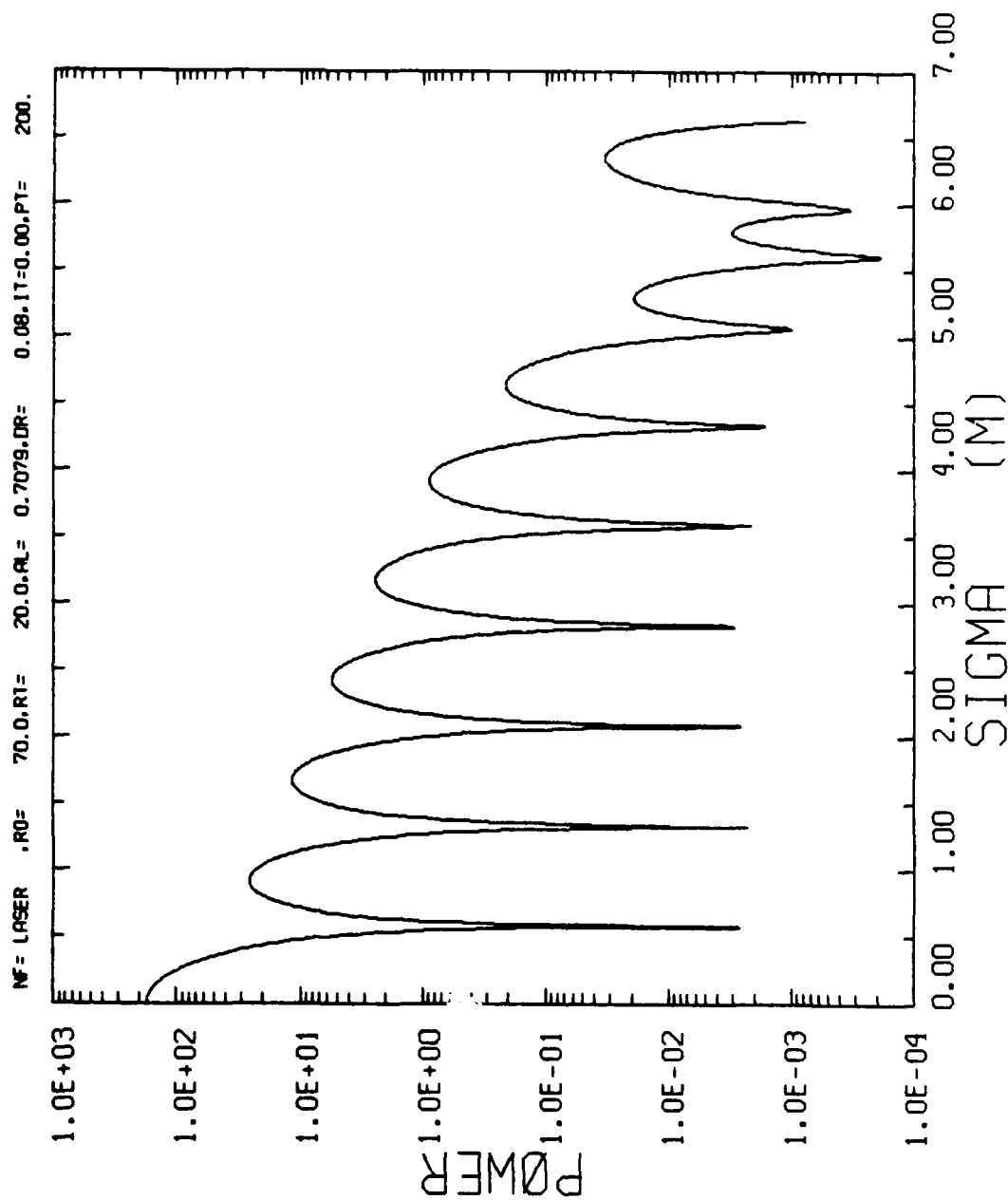


Fig. 9a. As in Fig. 8a but with Beam Truncated 3 dB Below Peak

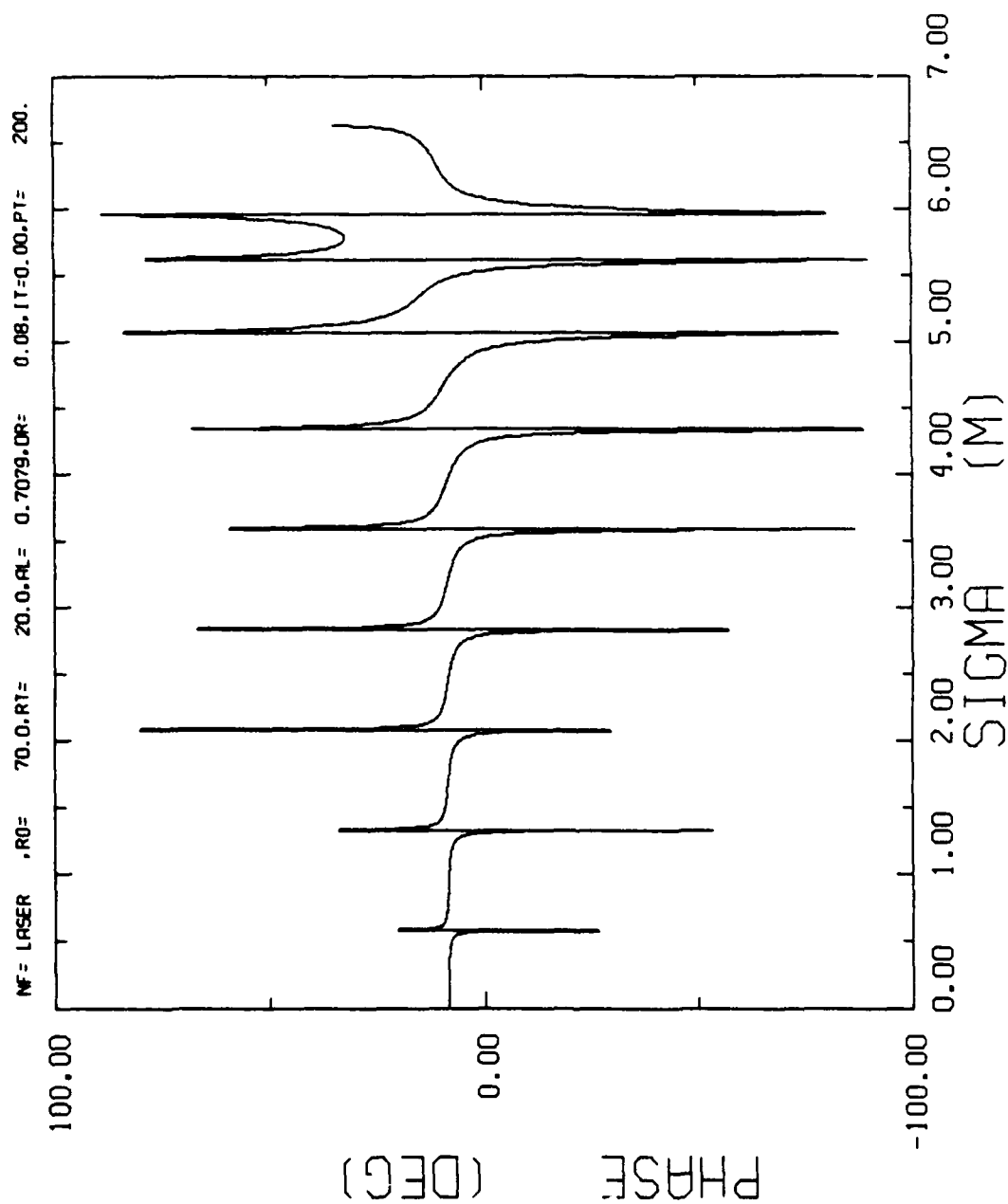


Fig. 9b. As in Fig. 8b but with Beam Truncated 3 dB Below Peak

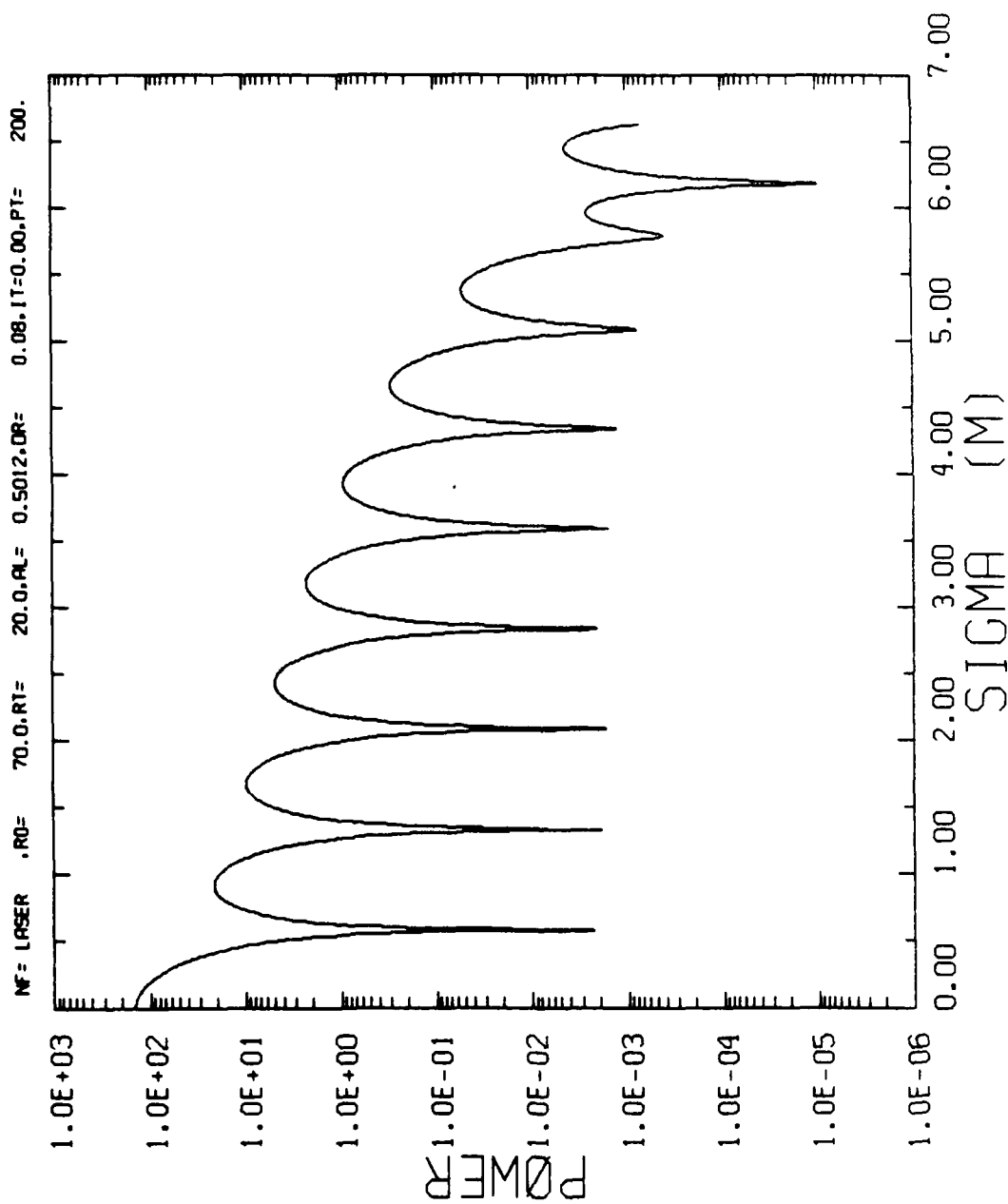


Fig. 10a. As in Fig. 8a but with Beam Truncated 6 dB Below Peak

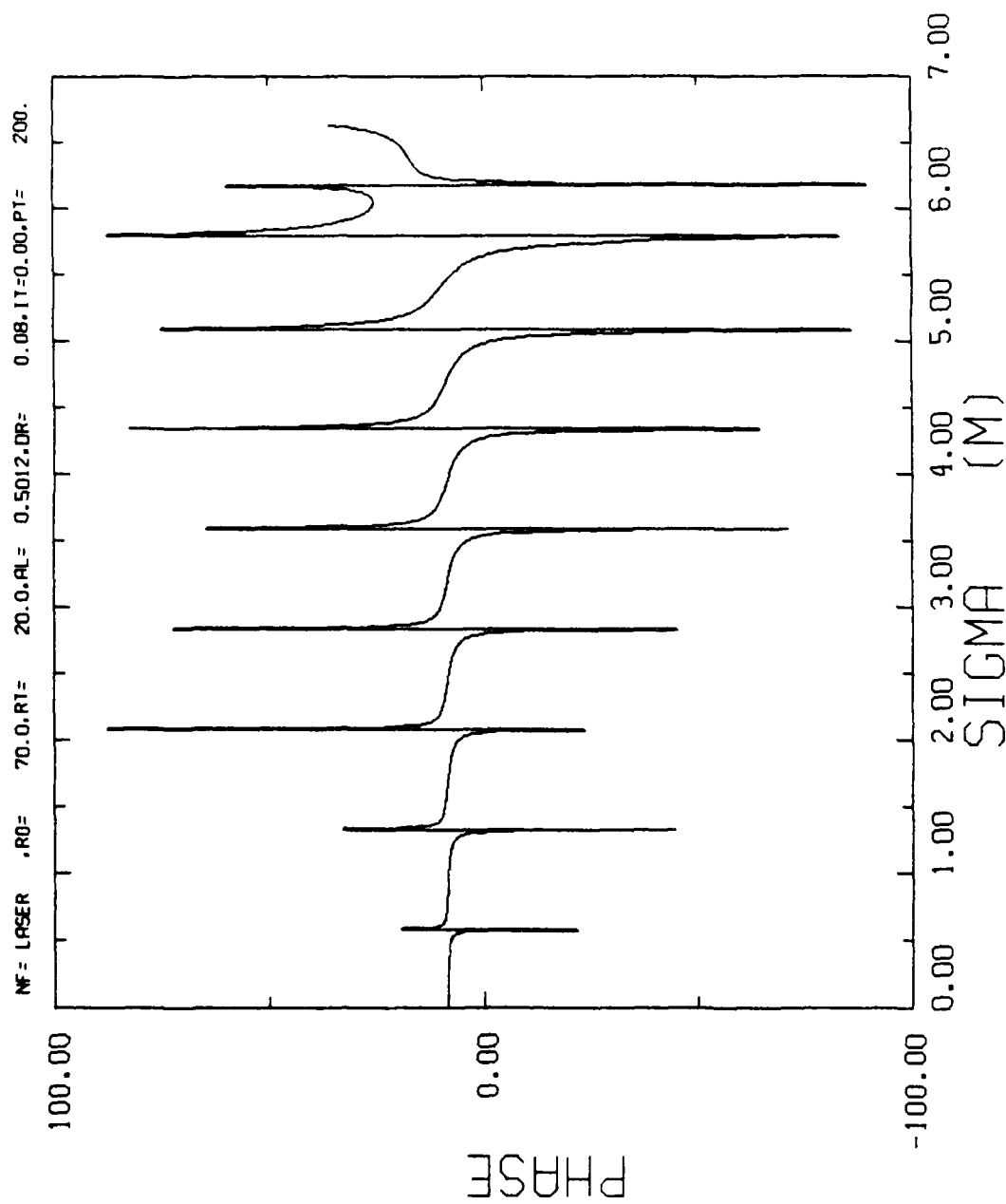


Fig. 10b. As in Fig. 8b but with Beam Truncated 6 dB Below Peak

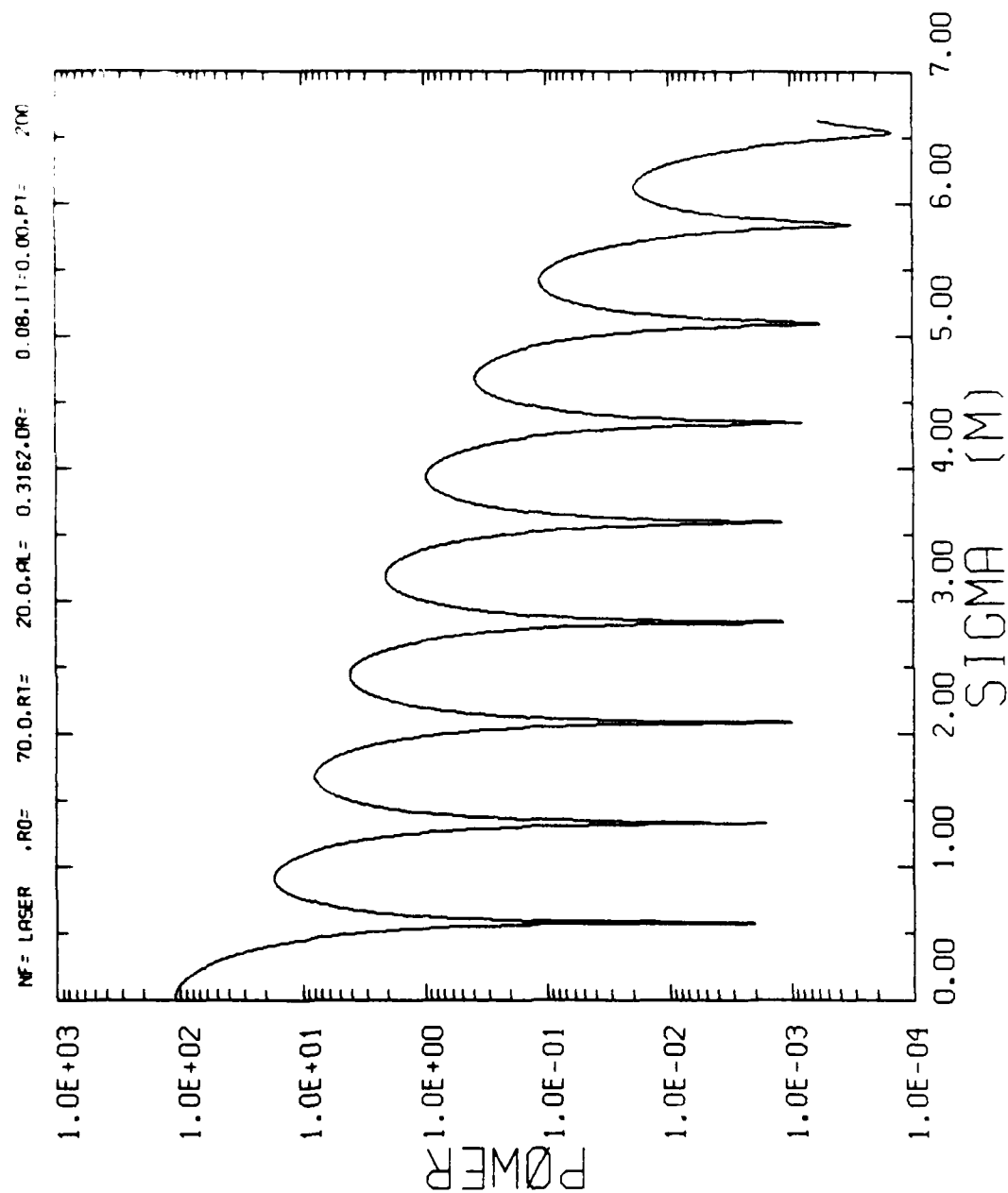


Fig. 11a. As in Fig. 8a but with Beam Truncated 10 dB Below Peak

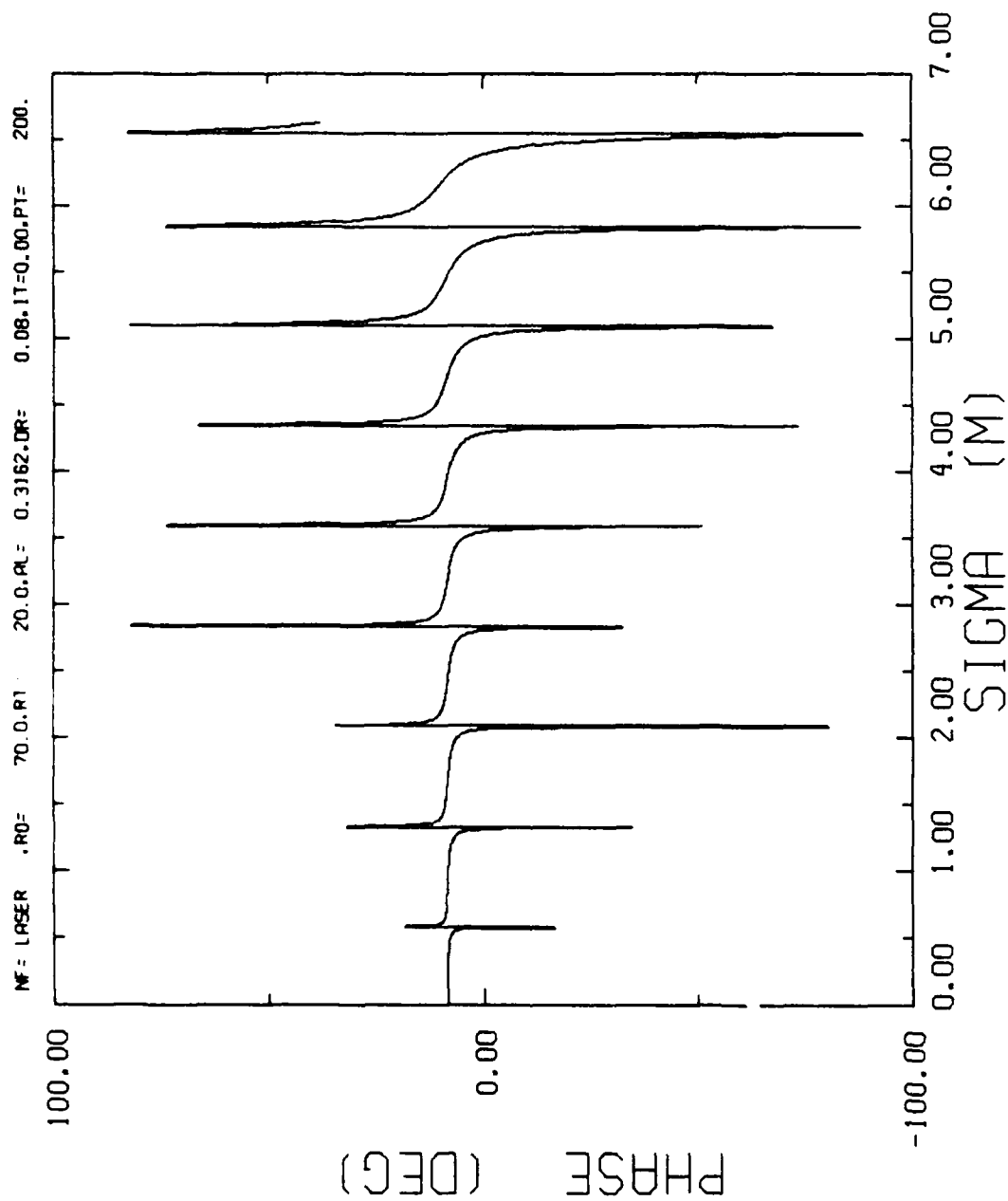


Fig. 11b. As in Fig. 8b but with Beam Truncated 10 dB Below Peak

V. SATELLITE METEOROLOGY

#### A. VAS Data Processing Software

SASC software engineers were responsible for data processing support of ongoing research into quantitative evaluation of multispectral data from GOES-VAS. Emphasis was placed on development of a simple statistical analysis routine and modifications to and extensions of existing data retrieval routines.<sup>1</sup>

Comparisons of radiance values measured by the VAS instrument and radiances calculated using radiosonde data are not entirely straightforward because of the different characteristics of the two detectors. The VAS instrument provides a near-instantaneous measurement of radiant energy from an atmospheric column whose horizontal cross-sectional area is equal to the spatial resolution of the sensor, approximately 14 km at subpoint. The radiosonde measures point temperature and dew point temperature at various altitudes through the atmosphere as it rises and is displaced horizontally by prevailing currents. To reduce the disparity between the two measuring techniques it is advisable to calculate a horizontal average of VAS measurements to account for the drift of the radiosonde and to reduce sensor noise. A statistical analysis option was added to the data retrieval routine VASVAL to perform this data averaging on the retrieved data. Statistics calculated include sample mean, standard deviation, maximum value, minimum value, range of values, sum of the values, sum of the squares of values, mean plus standard deviation, and mean minus standard deviation. Also listed are the number of values used in the calculations and the number of bad, i.e., missing, data points encountered in the sample. Statistics can be generated for radiance values and radiant temperatures. Sample size and location are determined by the electronic cursor and can be specified via direct computer commands or manually set via the cursor joysticks. A typical listing is shown in Fig. 1.

Preliminary analysis of radiance values measured by GOES-VAS and those calculated from radiosonde data input to algorithms developed at

1. Gerlach, A.M., ed., 1982: Objective Analysis and Prediction Techniques. AFGL-TR-82-0394, Scientific Report No. 2, Contract F19628-82-C-0023, Systems and Applied Sciences Corporation, pp. 213-214.

VAS BAND 10 RADIANCE

DATE: 2982066

TIME: 23000

5.9	5.7	5.9	5.9	5.5	5.1	4.7	4.7	4.6
6.0	6.1	5.9	5.5	5.0	4.5	4.1	4.0	3.6
5.6	5.5	5.2	4.9	4.4	4.0	3.9	3.9	3.4
5.4	5.2	5.2	5.2	5.2	4.7	4.5	4.2	4.0
5.1	4.9	4.6	4.5	4.7	4.7	4.5	4.2	4.2
5.2	5.0	4.6	4.5	4.2	3.9	3.5	3.5	3.6
5.1	4.9	4.6	4.2	3.9	3.9	3.9	3.9	4.2
4.5	4.4	4.4	4.5	4.4	4.1	4.1	3.9	3.7
4.0	4.0	3.7	3.6	3.6	3.7	3.6	3.5	3.1

N	=	81	MAX	=	6.10	X+SD	=	5.23
MEAN	=	4.52	MIN	=	3.10	X-SD	=	3.80
SD	=	.72	RANGE	=	3.00			
SUMX	=	365.900	SUMXSQ	=	1694.530			
THERE WERE 0 BAD POINTS WITHIN THE CURSOR								

Fig. 1. Sample Output Statistics Generated by Program VASVAL

AFGL<sup>2</sup> indicated a significant positive bias of 10 to 30 percent in the VAS measurements for all spectral bands. In an effort to locate the source of this bias, test routine PIXDAT was written to retrieve and list the 12-bit IR count value (X), the space-look or calibration value (Y), and the line scale factor (Z) for each spectral band at a single picture element location. X, Y, and Z are data values contained within a sounder file. These values were then converted to radiance values ( $\text{mW/M}^2\text{sr cm}^{-1}$ ) by the relationship

$$\text{RADIANCE} = (X - Y/16) * 2 \text{ EXP}(F + Z - 12) \quad (1)$$

where F is a scaling factor dependent on spectral band or channel number. Radiant temperature values were also calculated using the relationship

$$T = \{ \{ \text{FK2}/\text{ALOG}(\text{FK1}/\text{RADIANCE} + 1) \} - \text{TC1} \} / \text{TC2} \quad (2)$$

where FK2, FK1, TC1, and TC2 are constants that vary with channel number. Table 1A lists those values measured from space background away from the earth's edge. As can be seen, the calibration values (labelled Y SUB Z) are approximately one-half the measured values (labelled RAW DATA). While some variability of IR counts sensed when looking into space can be attributed to sensor noise, a difference of this magnitude indicates an undocumented error in the calibration value. It was determined that doubling the calibration value and not halving the measured value produced a more likely solution, so Eq. (1) was modified to

$$\text{RADIANCE} = (X - Y/8) * 2 \text{ EXP} (F + Z - 12). \quad (3)$$

Communication with the Space Science and Engineering Center at the University of Wisconsin, which provided the data and algorithms, established that an updated set of constants used in Eq. (2) had been calculated. These new constants and Eq. (3) were included in PIXDAT and the sample data from Table 1A were recalculated. The updated results are shown in Table 1B. The appropriate corrections were also made in the data display routine

2. Valovcin, F.R., 1981: A Ground Truth Analysis of DMSP Water Vapor Radiances. AFGL-TR-81-0323, Air Force Geophysics Laboratory, Hanscom AFB, MA, pp. 13-14.

TABLE 1A. UNCORRECTED RAW AND CALCULATED VAS DATA

CHANNEL	RAW DATA	Y SUB Z	DELTA F	RADIANCE	TEMP
13	100	50	2	12.500	171.32
14	73	35	3	19.000	186.17
15	103	52	2	12.750	174.55
16	95	46	2	12.250	175.07
17	74	38	2	9.000	170.73
18	96	48	0	.187	236.40
19	132	65	1	8.375	173.45
20	84	42	1	5.250	174.16
21	87	44	1	2.687	211.51
22	120	58	0	1.937	215.28
23	105	51	0	.211	242.19
24	136	64	0	.070	246.37

TABLE 1B. CORRECTED RAW AND CALCULATED VAS DATA

CHANNEL	RAW DATA	Y SUB Z	DELTA F	RADIANCE	TEMP
13	100	100	2	0.000	0.00
14	73	70	3	1.500	126.18
15	103	104	2	0.000	0.00
16	95	93	2	.500	113.39
17	74	77	2	0.000	0.00
18	96	97	0	0.000	0.00
19	132	130	1	.250	112.95
20	84	84	1	0.000	0.00
21	87	89	1	0.000	0.00
22	120	116	0	.125	168.47
23	105	103	0	.008	194.30
24	136	126	0	.008	214.54

SMOVER and the data retrieval routine VASVAL. A cursory check of the new results indicates that the biases in the VAS radiances have been greatly reduced.

Since the test routine PIXDAT provides a useful method of quickly listing point data for all 12 spectral bands, it was modified for use as an application program called VASDAT. Header information listing the date and time of the data along with the data location in satellite coordinates (line, element) and earth coordinates (latitude, longitude) was added to the data table. Sensor angles relative to the earth surface location of the data, both zenith and elevation, were calculated as were the solar zenith and elevation angles (in degrees-minutes-seconds format) and included in the output. A sample listing generated by VASDAT appears as Fig. 2.

```

***      A F G L   M C I D A S      ***      PROGRAM:  VASDAT
VAS DATA FOR 2982066 AT 23000 GMT
SATELLITE COORDS = 3609, 5137, EARTH = 31.04, -95.56
SAT ZENITH ANGLE = 423052, ELEVATION ANGLE = 472908
SOLAR ZENITH ANGLE = 1180646, ELEVATION ANGLE = -280646
CHANNEL  RAW DATA  Y SUB Z  DELTA F  RADIANCE  TEMP
13         272         84         2     47.000    222.71
14         162         70         3     46.000    222.83
15         298        104         2     48.500    226.64
16         322         91         2     57.750    237.27
17         398         79         2     79.750    259.61
18         205         97         0      .422    251.59
19         842        128         1     89.250    270.56
20         684         84         1     75.000    271.74
21         322         83         1     14.937    259.02
22         332        120         0      6.625    245.89
23         133        102         0      .121    232.54
24         376        124         0      .246    269.12

```

Fig. 2. Sample Output of VAS Data Generated by Program VASDAT

## B. Satellite-Manually Digitized Radar Feasibility Study

A combined SASC/AFGL study was undertaken to determine the feasibility of establishing a quantitative procedure for inferring equivalent Manually Digitized Radar<sup>3</sup> (MDR) Digital Video Integrator and Processor (D/VIP) levels from visible-infrared satellite imagery for cases of moderate to deep convection (M/DC).

The approach to data handling is viewed as a two-phase process. Phase one, described in Gerlach (1982),<sup>4</sup> was to identify and archive the MDR, visible, and infrared data sets for days and regions of M/DC. The stretched Visible and Infrared Spin-Scan Radiometer (VISSR) digital data from the GOES-5 satellite were obtained from the McIDAS Sony Cassette Archive System (MSCAS). The MDR data, extracted from the high speed FAA WB-604 line, decoded, and stored by McIDAS,<sup>5</sup> were also accessed through and processed with this facility. Phase two is to process cases selected from the archive. This report covers the initial processing procedure under phase two, which is ongoing.

The data stored for each case are the case number, latitude and longitude of the MDR box, MDR box D/VIP level, satellite number-year-date identifier, normalized brightness values, and infrared counts. As a first step in processing cases, visible and infrared arrays are reduced to single values such that each case is defined as an MDR, visible, and infrared value collectively. The criterion first proposed for the reduction was to choose the brightest visible and coldest infrared pixels as imagery representative of each MDR box. The basis for this method was the finding in earlier studies of tropical regions that bright clouds are positively correlated with

3. (1979) National Weather Service Radar Code User's Guide, NOAA.

4. Gerlach, A.M., ed., 1982: Objective Analysis and Prediction Techniques. AFGL-TR-82-0394, Scientific Report No. 2, Contract F19628-82-C-0023, Systems and Applied Sciences Corporation, pp. 217-226.

5. Gerlach, A.M., ed., 1981: Technique Development for Weather Forecasting. AFGL-TR-82-0020, Final Report, Contract F19628-81-C-0039, Systems and Applied Sciences Corporation, pp. 124-137.

areas of convective rainfall as viewed from visible satellite imagery (Martin and Suomi (1971),<sup>6</sup> Woodley and Sancho (1971),<sup>7</sup> and Woodley et al. (1972)<sup>8</sup>). This finding led others to propose that the brightness of these clouds is an indication of the probability of precipitation. Griffith and Woodley (1973)<sup>9</sup> and Reynolds and Vonder Haar (1973)<sup>10</sup> found that visible brightness is directly related to the height of tropical convective clouds. Processing the imagery arrays in this manner has been named maximum point processing to indicate that the largest or maximum visible and infrared digital values are extracted. This method was later expanded upon to explore processing on average value computed from the top five, ten, fifteen, twenty, and twenty-five percent of the brightest and coldest pixels. This is known as average percent processing.

Actual case processing was achieved with the development and implementation of program FRQDIS. Modular in program design, FRQDIS consists entirely of subroutines that perform specific tasks. These tasks are outlined for the general user as command options. Appendix V-A provides documentation of the command options in FRQDIS. For this application, FRQDIS manages reduction-type identifiable named disk files containing processed case save file data. File structure is based on a three-dimensional frequency distribution where occurrence values are stored under and accessed through array coordinates defined by the MDR, visible, and infrared intersection. Table 2 shows the range of data values for MDR, visible, and infrared data and the corresponding coordinates used within the frequency

6. Martin, D.W., and V.E. Suomi, 1971: A Satellite Study of Cloud Clusters Over the Tropical Atlantic Ocean. Final Report, Tropical Cloud Organization Studies, Task Order No. 3 to STAG, Contract E-127-69-(N), Space Science and Engineering Center, Madison, WI.

7. Woodley, W.L., and B. Sancho, 1971: A first step toward rainfall estimation from satellite cloud photographs. Weather, 26, 279-289.

8. Woodley, W.L., B. Sancho, and A.H. Miller, 1972: Rainfall Estimation from Satellite Cloud Photographs. NOAA Tech. Memo. ERL GD-11, Boulder, CO.

9. Griffith, C.G., and W.L. Woodley, 1973: On the variation with height of the top brightness of precipitating convective clouds. J. Appl. Meteor., 12, 1086-1089.

10. Reynolds, D., and T.H. Vonder Haar, 1973: A comparison of radar-determined cloud height and reflected solar radiance measured from the geosynchronous satellite ATS-3. J. Appl. Meteor., 12, 1082-1085.

distribution. Fig. 3 illustrates the three-dimensional frequency distribution, visualized as a volume whose dimensions are defined by the visible, infrared, and MDR upper-bound coordinates.

TABLE 2. DATA RANGE AND CORRESPONDING 3-D FREQUENCY DISTRIBUTION COORDINATES FOR MDR AND SATELLITE IMAGERY

	<u>DATA RANGE</u>	<u>DISTRIBUTION COORDINATES</u>
MDR D/VIP level	0 - 6	1 - 7
Normalized brightness	< .06	1
	.06 - .10	2
	.11 - .15	3
	.16 - .20	4
	.21 - .25	5
	.26 - .30	6
	.31 - .35	7
	.36 - .40	8
	.41 - .45	9
	.46 - .50	10
	.51 - .55	11
	.56 - .60	12
	.61 - .65	13
	.66 - .70	14
	.71 - .75	15
	.76 - .80	16
	.81 - .85	17
	.86 - .90	18
	.91 - .95	19
	> .95	20
Infrared counts	0 - 255	1 - 256

With the development of FRQDIS to satisfy the initial requirements of phase two and the continuing effort to collect cases of M/DC, an expanded step-through procedure was established to process cases as they were collected. This procedure consists of executing command options from both SATMDR<sup>4</sup> and FRQDIS in an orderly sequence.

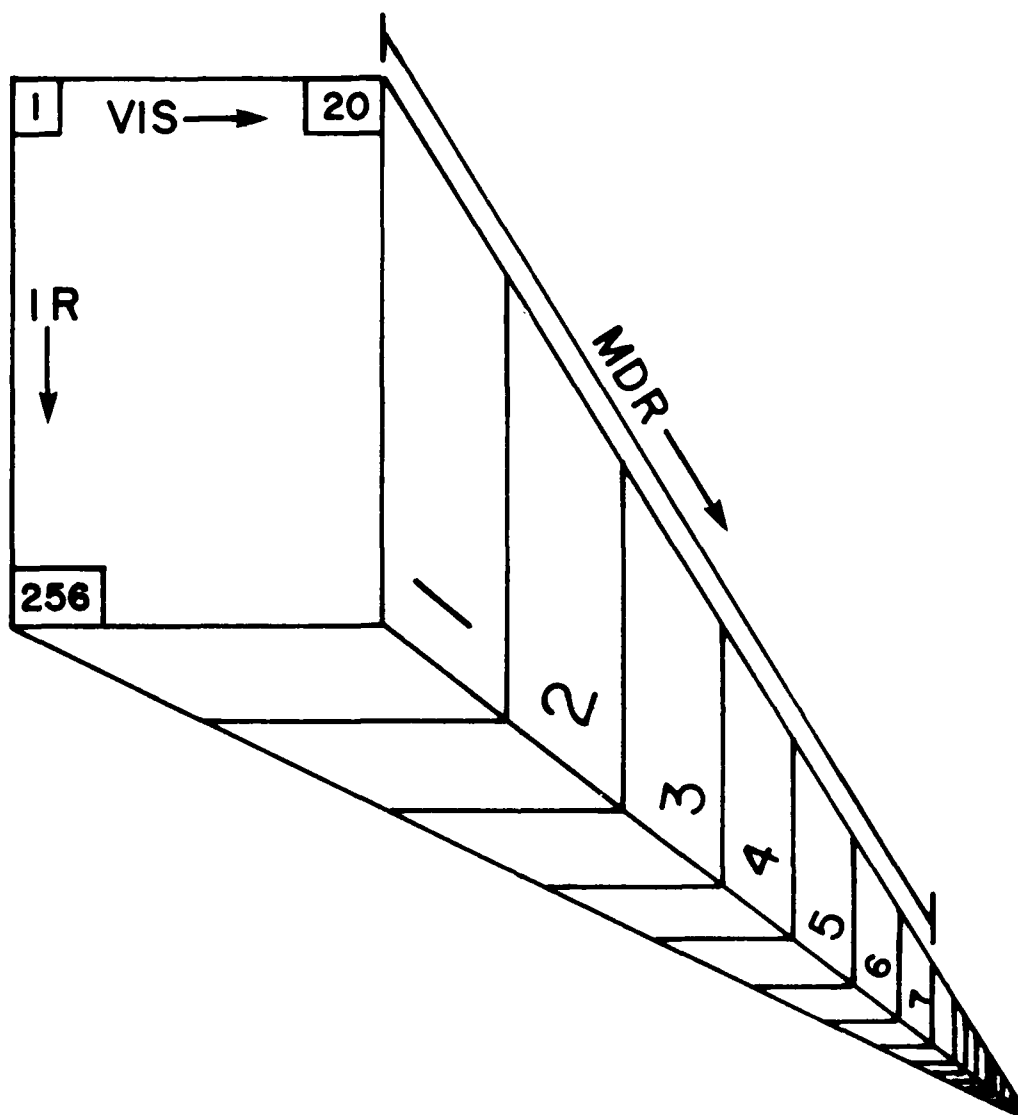


Fig. 3. Graphical Representation of 3-D Frequency Distribution  
(Volume dimensions correspond to visible, infrared, and MDR upper bound coordinates)

Six disk files were created and initialized with the INI option of FRQDIS. This option defines the size or dimensions of the volume to contain occurrence values and then sets all values to zero. After a case save file had been filled using the CAS option of SATMDR, cases were processed using the MAX option of FRQDIS for maximum point processing, and the FVE, TEN, PTN, TWE, and TFV options for average percent processing. Totals for the number of cases processed to date were monitored with the STA option of FRQDIS and any summation errors were rectified on an element basis with the ADD and SUB options. Listings of the distributions for any or all MDR planes could be viewed at any time with the PRI option of FRQDIS. Files were saved on magnetic tape as backup using the SAV option.

The total number of cases processed for each of the six files was 4094. Table 3 shows the breakdown of this total according to number of cases per D/VIP level. In order to determine whether this distribution represents a sample of some population distribution, it was proposed to collect a larger sample of the population and then compare the two distributions using the chi-square test for independence. A low probability would indicate that our sample distribution was not biased from collecting areas encompassing M/DC, since the probability would be high that we were dealing with two samples of the same population rather than samples from two different populations.

TABLE 3. DISTRIBUTION OF PROCESSED CASES BY D/VIP LEVEL

<u>D/VIP Level</u>	<u>No. of Cases</u>
0	956
1	1054
2	1259
3	428
4	272
5	101
6	24

The larger sample was defined as reported D/VIP levels covering the national MDR grid for the days and times used in case collection. To extract the reported levels, a temporary option was added to program GRIMDR

to allow the user to simply specify the valid time(s) from the case collection archive. Tabulated values for each level were stored in the recovery sector of file MAXFRQ and updated each time a national grid was processed. Table 4 shows the distribution of reported D/VIP levels for this larger sample.

TABLE 4. DISTRIBUTION OF REPORTED D/VIP LEVELS FROM MDR NATIONAL GRIDS FOR DAYS AND TIMES USED IN CASE COLLECTION

<u>D/VIP Level</u>	<u>Frequency</u>
1	6003
2	6850
3	2172
4	1222
5	374
6	60

Table 5 is the row (r) by column (c) contingency table for the two samples. Values are percent relative frequency and expected values are shown in parentheses along with row-column marginal frequencies. The value of chi-square computed from this table is 0.55. In determining degrees of freedom we note that row and column totals are fixed because of the requirement that each sample distribution sum to 100. This would imply  $c + r$  conditions, but the last condition is determined from all others. Thus, we have  $c + r - 1 = 7$  independent conditions applied to  $c \cdot r = 12$  variables or five degrees of freedom. The five percent limit for chi-square with five degrees of freedom is 11.07, obtained from Table VI of Probability and Statistics for Engineers and Scientists (Walpole and Myers, 1972<sup>11</sup>). The observed value of chi-square is considerably smaller than this value; hence the probability is high (> 90 percent) that for the two samples collected, sample size has minimal bearing on the distribution of reported D/VIP levels.

To satisfy the need for graphical representation of the three-dimensional frequency distribution for data analysis, an interactive routine was developed to display frequency distribution data in histogram form. Program FRQHIS is utilized to plot frequency of occurrence values vs. D/VIP

<sup>11</sup> Walpole, R.E., and R.H. Myers, 1972: Probability and Statistics for Engineers and Scientists. The MacMillan Company, 506 pp.

TABLE 5. OBSERVED AND EXPECTED PERCENT RELATIVE FREQUENCIES

D/VIP LEVEL SAMPLE	D/VIP LEVEL						TOTALS
	1	2	3	4	5	6	
M/DC CASES ONLY	33.6 (34.8)	40.1 (40.6)	13.6 (13.3)	8.7 (8.0)	3.2 (2.7)	0.8 (0.6)	100
NATIONAL GRIDS	36.0 (34.8)	41.1 (40.6)	13.0 (13.3)	7.3 (8.0)	2.2 (2.7)	0.4 (0.6)	100
TOTALS	69.6	81.2	26.6	16.0	5.4	1.2	200

level for a user-specified visible-infrared combination. The imagery pair can be defined as single values, a range of values, or a combination of the two over the visible-infrared data domain. Fig. 4 is an example of a histogram produced by FRQHIS. The legend at the top right corner indicates the visible-infrared pair selected. Appendix V-B provides documentation of the command format in FRQHIS.

Two steps remained to complete the initial processing procedure. The first was to compress the frequency distribution in the infrared to equivalent infrared temperature bins of  $2^{\circ}\text{C}$  each, in order to generate a relatively large number of points in a sizable number of the resulting  $20 \times 84 \times 7$  distributions. The last step was to generate rain/no-rain exceedence probability distributions for the range of reported D/VIP levels. Both steps were accomplished through further software development. The first was written as new options to FRQDIS and the second as a new foreground program.

File compression was achieved with the COM option of FRQDIS. This command initializes the compressed file with specified dimensions, converts non-linear infrared counts to  $2^{\circ}\text{C}$  bins and sums the corresponding frequencies, names the new file with the old file name, and finally deletes the old file. The PRT option was then used to generate hard copy listings of individual MDR planes. Summed frequencies were verified by adding occurrence values obtained from a listing of the non-compressed distribution and comparing the value with the PRT listed value. Compression procedures were eventually applied to all six disk files. Fig. 5 is a listing of the compressed frequency distribution for MDR plane coordinate 2 (D/VIP level 1) and maximum point processed data.

As described earlier the MDR, visible, and infrared data are stored as a three-dimensional frequency distribution containing visible-infrared arrays for each MDR plane. The first plane (D/VIP level 0) contains all the non-raining  $R_{NR}$  points, each point identified by the set  $(V_{NR}, IR_{NR})$ . The remaining six MDR planes (levels 1-6) represent the rain matrices in which the pair  $(V_R, IR_R)$  provides the coordinates for reported MDR level points  $R_R$ . The rain distribution(s) and no-rain matrix are combined to yield a probability of rain distribution as a function of visible and infrared values, expressed as

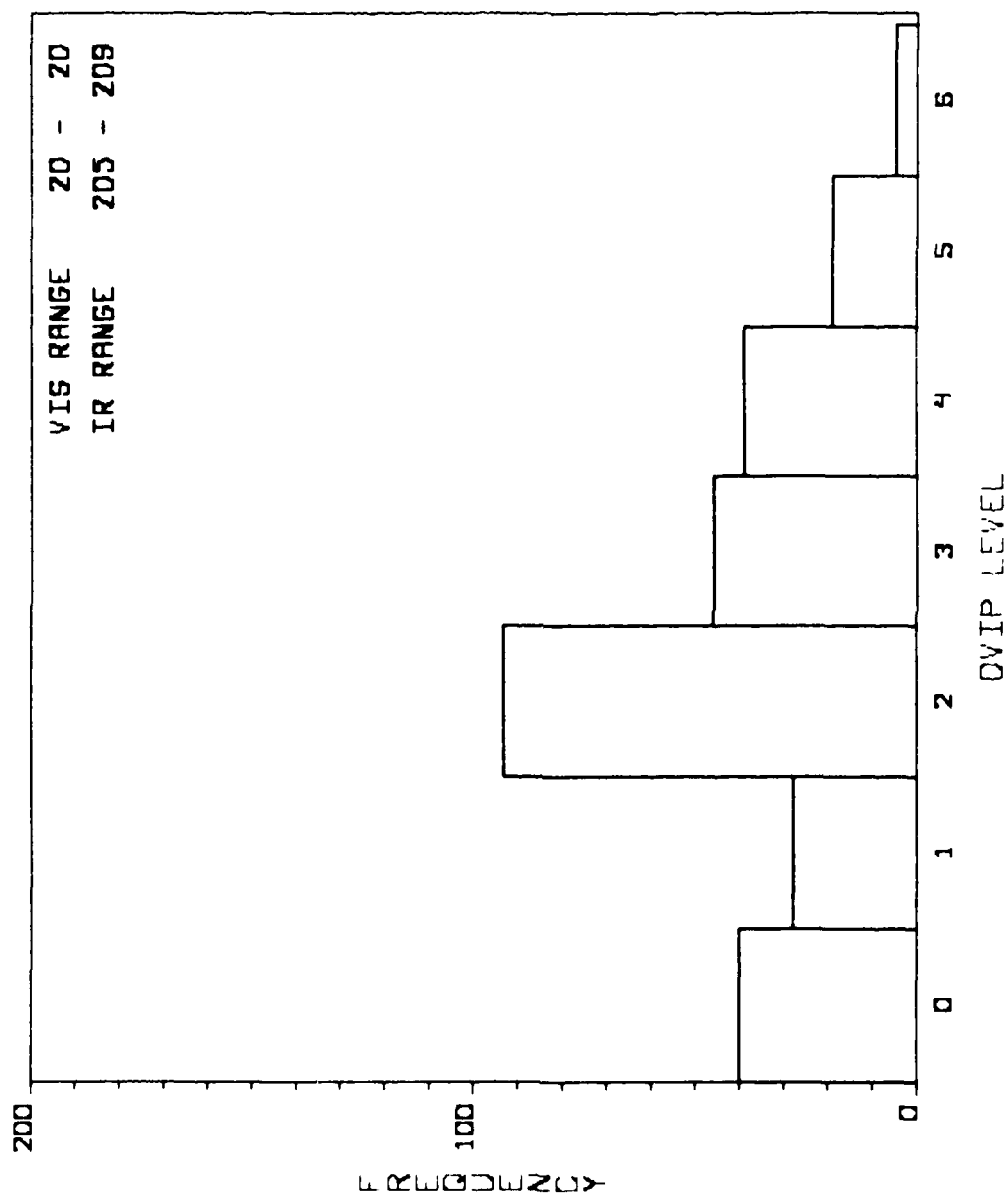


Fig. 4. Histogram Plot of MDR D/VIP Levels Produced by FRQHS

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
37.0 13	35.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35.0 13	33.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33.0 13	31.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31.0 13	29.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29.0 13	27.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27.0 13	25.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25.0 13	23.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23.0 13	21.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21.0 13	19.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19.0 13	17.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17.0 13	15.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15.0 13	13.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13.0 13	11.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11.0 13	9.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9.0 13	7.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7.0 13	5.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5.0 13	3.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3.0 13	1.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1.0 13	-1.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-1.0 13	-3.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-3.0 13	-5.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-5.0 13	-7.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-7.0 13	-9.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-9.0 13	-11.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-11.0 13	-13.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-13.0 13	-15.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-15.0 13	-17.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-17.0 13	-19.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-19.0 13	-21.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-21.0 13	-23.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-23.0 13	-25.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-25.0 13	-27.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-27.0 13	-29.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-29.0 13	-31.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-31.0 13	-33.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-33.0 13	-35.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-35.0 13	-37.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-37.0 13	-39.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-39.0 13	-41.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-41.0 13	-43.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-43.0 13	-45.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-45.0 13	-47.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-47.0 13	-49.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-49.0 13	-51.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-51.0 13	-53.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-53.0 13	-55.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-55.0 13	-57.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-57.0 13	-59.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-59.0 13	-61.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-61.0 13	-63.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-63.0 13	-65.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-65.0 13	-67.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-67.0 13	-69.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-69.0 13	-71.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-71.0 13	-73.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-73.0 13	-75.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-75.0 13	-77.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-77.0 13	-79.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-79.0 13	-81.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-81.0 13	-83.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-83.0 13	-85.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-85.0 13	-87.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-87.0 13	-89.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-89.0 13	-91.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-91.0 13	-93.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-93.0 13	-95.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-95.0 13	-97.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-97.0 13	-99.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-99.0 13	-101.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-101.0 13	-103.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-103.0 13	-105.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-105.0 13	-107.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-107.0 13	-109.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-109.0 13	-111.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Fig. 5. Listing of Compressed Frequency Distribution Produced by PRT Option of FRQDIS

(Equivalent infrared temperature intervals in °C and visible brightness categories as in Table 2; data from file MAXFRQ)

$$PF_{RS}(V, IR) = \frac{\sum_{i=m}^n R_R(V, IR)}{\sum_{i=m}^n R_R(V, IR) + R_{NR}(V, IR)} \quad (4)$$

$R_R(V, IR)$  is the frequency of occurrence of rain with associated  $(V, IR)$  magnitudes summed over the range of D/VIP levels  $m$  to  $n$ .  $R_{NR}$  is the corresponding frequency of no-rain.

Using this empirical relationship, exceedence probability distributions for  $n = 6$  and  $m = 1, 2, 3, 4$ , and  $5$  were computed and listings generated with the development and implementation of program PRBDIS. Documentation of the command format of PRBDIS is provided in Appendix V-C.

To verify PRBDIS output, listings of the compressed distribution for each MDR plane were generated with the PRT option of FRQDIS. Visible-infrared pairs were selected at random and probabilities were calculated from Eq. (4). Results were then compared with PRT compatible listings generated with the PRI option of PRBDIS. Exceedence probability distributions were eventually generated for all six disk files. Fig. 6 is a listing of the rain vs. no-rain probability distribution for maximum point processed data.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
27.0 12	55.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
28.0 12	56.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
29.0 12	57.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
30.0 12	58.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
31.0 12	59.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
32.0 12	60.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
33.0 12	61.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
34.0 12	62.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
35.0 12	63.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
36.0 12	64.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
37.0 12	65.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
38.0 12	66.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
39.0 12	67.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
40.0 12	68.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
41.0 12	69.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
42.0 12	70.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
43.0 12	71.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
44.0 12	72.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
45.0 12	73.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
46.0 12	74.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
47.0 12	75.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
48.0 12	76.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
49.0 12	77.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
50.0 12	78.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
51.0 12	79.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
52.0 12	80.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
53.0 12	81.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
54.0 12	82.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
55.0 12	83.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
56.0 12	84.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
57.0 12	85.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
58.0 12	86.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
59.0 12	87.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
60.0 12	88.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
61.0 12	89.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
62.0 12	90.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
63.0 12	91.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
64.0 12	92.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
65.0 12	93.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
66.0 12	94.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
67.0 12	95.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
68.0 12	96.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
69.0 12	97.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
70.0 12	98.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
71.0 12	99.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
72.0 12	100.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
73.0 12	101.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
74.0 12	102.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
75.0 12	103.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
76.0 12	104.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
77.0 12	105.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
78.0 12	106.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
79.0 12	107.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
80.0 12	108.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
81.0 12	109.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
82.0 12	110.1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

Fig. 6. Listing of Rain/No Rain Percent Probability Distribution Produced by PRI Option of PRBDIS

(Temperature intervals and brightness categories as in Fig. 5; data from file MAXFRQ; -1 indicates no-rain data not available)

# Appendix V-A. Command Documentation for Program FRQDIS

COMMAND: FRQDIS

KEY IN: FD

PROGRAMMERS: R. FOURNIER  
C. IVALDI

DATE: NOV 82

## COMMAND FORMAT:

FD FUN PARM1 - PARMN

## COMMAND DEFAULT PARAMETERS:

NONE

## PARAMETER DESCRIPTION:

FUN Function requested

### 1. FD INI

Create and initialize the 3-D frequency distribution file. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. PARM2 to PARM4 are the number of frequency intervals in the frequency distribution file. If PARM2 is negative, then the file is deleted from the system.

### 2. FD STA

List on terminal CRT or print status of 3-D frequency distribution file. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. If a hard copy of the status is desired, set PARM2 to "PRI."

### 3. FD ADD

Add constant to specified frequency distribution. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. PARM2 is the constant to add to the frequency count. If PARM2 is zero, the frequency count is incremented by one. PARM3-4 are the VIS, IR, and MDR coordinates.

### 4. FD SUB

Subtract constant from specified frequency distribution. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. PARM2 is the constant to subtract from the frequency count. If PARM2 is zero, the frequency count is decremented by one. PARM3-4 are the VIS, IR, and MDR coordinates.

### 5. FD PRI

Print VIS, IR frequency plane for specified frequency distribution. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. PARM2 is the MDR coordinate plane to print (1-7). This is not a generalized function. It is very specific to the MDR satellite study.

6. FD TAB

List or print VIS interval table. PARM1 is either "LIS" to list intervals on CRT or "PRI" to print intervals on line printer.

7. FD SAV

Save specified frequency distribution file on tape. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. PARM2 is the tape number. PARM3 (optional) is the file number to locate the specified frequency distribution file on tape.

8. FD RES

Restore specified frequency distribution file from tape. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. PARM2 is the tape number. PARM3 is the file number to locate the specified frequency distribution file on tape.

9. FD FRQ

Computes variable IR, fixed VIS class with relative frequency distribution from specified 3-D frequency distribution file. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. PARM2 is "IR" for MDR vs. IR or "VIS" for MDR vs. VIS relative frequency distribution computation. PARM3 is the number of IR class intervals desired. The choices for PARM3 are 17, 15, 5, or 1. If a single or selected range of MDR plane(s) is/are required, set PARM4 to "MDR" with PARM5 and PARM6 as the start and ending range of the MDR plane coordinates. If PARM2 is "VIS," then PARM3 through PARM5 are identical to PARM4 through PARM6 for the "IR" option above. This is not a generalized function. It is very specific to the MDR-satellite study.

10. FD MAX

Maximum point processing of case(s) from case save file. This function will extract the brightest and coldest pixels and update the frequency distribution if desired. PARM1 is the first case number in the case save file. PARM2 is "TO" for processing more than one case. PARM3 is the last case number in the case save file to process. To list or print results without updating the frequency distribution, PARM1 is either "LIS" or "PRI." PARM2 is the first case number. PARM3 is "TO" for more than a single case and PARM4 is the last case number.

11. FD FVE ( 5%)  
TEN (10%)  
FTN (15%)  
TWE (20%)  
TFV (25%)

Average percent processing of case(s) from case save file. These functions will compute the average of the top five, ten, fifteen, twenty, or twenty-five percent brightest and coldest pixels and update the respective frequency distribution if desired. PARM1 is the first case number in the case save file. PARM2 is "TO" for processing more than one case. PARM3 is the last case number in the case save file to process. To list or print results without updating the frequency distribution, PARM1 is either "LIS" or "PRI." PARM2 is the first case number. PARM3 is "TO" for more than a single case and PARM4 is the last case number.

12. FD COM

Compress non-linear IR digital counts to  $2^{\circ}\text{C}$  increments. PARM1 is frequency distribution to compress. In essence, PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name.

13. FD PRT

Prints VIS, compressed IR frequency plane for specified frequency distribution. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. PARM2 is the MDR coordinate plane to print (1-7). This is not a generalized function. It is very specific to the MDR-satellite study.

Appendix V-B. Command Documentation for Program FRQDIS

COMMAND: FRQDIS

KEY-IN: FH

PROGRAMMER: C. IVALDI

DATE: MAY 1983

COMMAND FORMAT:

FN FUN PARM1-PARMN

COMMAND DEFAULT PARAMETERS:

NONE

PARAMETER DESCRIPTION:

FUN Function requested

1. Z

Generate histogram plot of z-axis coordinates vs. frequency of occurrence from FRQDIS compatible 3-D frequency distribution file. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. PARM2 and PARM3 are the start and ending y-axis coordinates. PARM4 and PARM5 are the start and ending x-axis coordinates. PARM2-5 must lie within respective coordinate bounds. If PARM3 < PARM2, PARM3 is set to PARM2. This rule also applies to PARM4 and PARM5.

Appendix V-C. Command Documentation for Program PRBDIS

COMMAND: PRBDIS

KEY IN: PD

PROGRAMMER: C. IVALDI

DATE: OCT 83

COMMAND FORMAT:

PD FUN PARM1 - PARMN

COMMAND DEFAULT PARAMETERS:

NONE

PARAMETER DESCRIPTION:

FUN Function requested

1. PRI

Compute and print rain/no rain probability distribution for specified frequency distribution. PARM1 is a 3-character string that will be used as the first three characters of the 3-D frequency distribution file name. If a single or selected range of MDR plane(s) is/are required, set PARM2 to "MDR" with PARM3 and PARM4 as the start and ending range of the MDR plane coordinates.

## C. Experimental Cloud Analysis Program

### 1. Introduction

SASC is supporting the AFGL effort to improve the Air Force automated cloud analysis model known as RTNEPH, as specified under AWS Geophysical Requirement 2-80; Improved Cloud Data Handling and Analysis. The RTNEPH program is an updated version of the 3DNEPH program described by Fye (1978).<sup>12</sup> RTNEPH replaced 3DNEPH on August 1, 1983, as the operational cloud model run at the Air Force Global Weather Central. RTNEPH operationally provides a hemispheric analysis of total cloud and cloud layer amount every three hours plus special analysis for local areas of interest on demand. Cloud analysis is performed by analyzing satellite imagery and merging the results with observational reports.

An experimental version of the RTNEPH computer program has been implemented on the AFGL McIDAS. The AFGL program, known as RDNEPH (for research and development nephanalysis), was designed using the RTNEPH satellite analysis algorithm as a baseline.

### 2. RDNEPH

The satellite analysis algorithm is a statistical thresholding technique. A grayshade histogram of thermal IR counts is generated from satellite imagery over a 25 nm (1/8 mesh) resolution grid box (Fig. 7). A clustering algorithm combines peaks in the IR histogram into thermally distinct grayshade clusters. Each cluster is assigned a single grayshade that is representative of the entire range of grayshades composing the cluster. The single representative grayshade (RGS) is used to determine if the cluster exceeds a cloud/no-cloud threshold. The RGS value is converted to an equivalent radiance temperature through a table lookup procedure; several corrections are then applied to compensate for atmospheric attenuation and scan geometry bias. The final, corrected temperature is compared to the background temperature of that same 1/8 mesh grid box. Clusters are determined to be cloud when the background temperature exceeds the computed cluster temperature by an amount greater than an empirical cloud threshold. A more complete

12. Fye, F.K., 1978: The AFGWC Automated Cloud Analysis Model. AFGWC-TM-78-002, Air Weather Service, Scott AFB, IL.

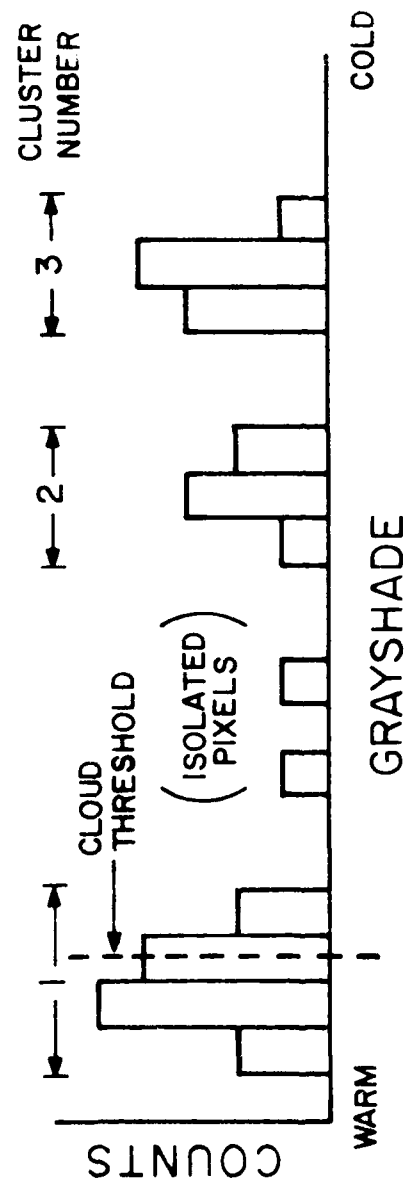


Fig. 7. IR Grayscale Histogram

AD-A142 441

OBJECTIVE ANALYSIS AND PREDICTION TECHNIQUES - 1983(U)  
SYSTEMS AND APPLIED SCIENCES CORP VIENNA VA

3/3

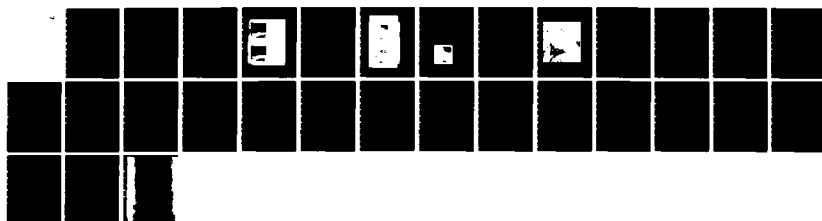
A M GERLACH 30 NOV 83 SCIENTIFIC-3 AFGL-TR-83-0333

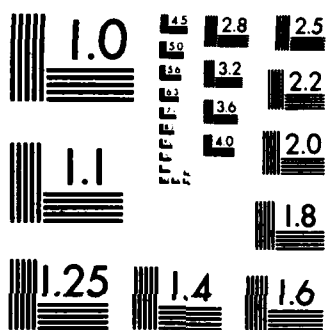
UNCLASSIFIED

F19628-82-C-0023

F/G 4/2

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

description of the satellite processor is given by Bunting et al. (1983)<sup>13</sup> and Gerlach (1982).<sup>14</sup>

To aid evaluation of the analysis algorithm, the RDNEPH program provides for user interaction at various levels of the algorithm through the use of runtime option flags. To date three flag combinations that control clustering, RGS determination, and isolated pixel processing have been investigated.

The process of combining histogram peaks into clusters is controlled by a grayshade scan radius. Moving from the cold to warm end of the grayshade histogram, peaks are combined if they are within a set number of grayshade intervals, the scan radius. After each complete pass across the histogram, the scan radius is incremented and clustering continues out to a user defined scan maximum. The magnitude of the scan maximum influences the number of cloud layers that can be detected by the model. In general the greater the scan radius maximum, the poorer the cloud layer discrimination. The clustering process is further controlled by a cluster significance criterion. After each pass of a single scan radius the resulting clusters are subjected to a peak size threshold. When all clusters exceed this threshold the clustering process terminates and clusters are passed directly to the cloud decision procedure. The significance criterion is also controlled by option flags. The interaction between scan radius and cluster significance provides fine tuning in situations of multiple cloud layers and partially cloud filled scenes.

As described above, cloud thresholding of a cluster is performed on a single grayshade level that is determined to be representative of that cluster. The selection of that grayshade, the RGS, can be critical when the threshold falls within cluster bounds as illustrated by cluster 1 of Fig. 7. There are four possible grayshade values that can be used to represent a cluster: the warmest, the mean, the most frequently occurring, and the coldest. The value that is used is controlled by a processing flag. Note that for cluster 1 in the figure, only the coldest grayshade option would result in the cluster being declared cloud (i.e., colder than the

13. Bunting, J.T., R.S. Hawkins, R.P. d'Entremont, and G.B. Gustafson, 1983: R&D nephanalysis at the Air Force Geophysics Laboratory. Preprints, 5th Conference on Atmospheric Radiation, AMS, Boston, 272-275.

14. Gerlach, A.M., ed., 1982: Objective Analysis and Prediction Techniques. AFGL-TR-82-0394, Scientific Report No. 2, Contract F19628-82-C-0023, Systems and Applied Sciences Corporation, pp. 227-241.

threshold); all others would result in threshold rejection and the cluster being interpreted as background (i.e., clear). Clusters 2 and 3 clearly exceed the threshold level regardless of the RGS option selected and would therefore be unconditionally interpreted as two separate cloud layers.

The final set of run options to be discussed controls the processing of isolated pixels that remain unassigned after clustering (Fig. 7). In early versions of the program these pixels were unconditionally assigned to the next warmer cluster for cloud threshold processing (pixels assigned to cluster 1 in Fig. 7). This resulted in the isolated pixels often being interpreted as background (i.e., clear). However, investigation of these pixels showed that they are usually isolated high cloud tops embedded in a surrounding stratiform deck or isolated sub-resolution cumulus and that they occur much more frequently than previously assumed. A second option was initiated that forces the isolated pixels to be assigned to the next colder cluster (pixels assigned to cluster 2 in Fig. 7). This procedure greatly reduced the underanalysis of isolated cloud tops but resulted in some cloud free areas being interpreted as cloud covered. A third option retains the assignment to colder clusters for all groups of isolated pixels except the warmest; these pixels are subjected to the cloud thresholding procedure as if they composed a separate cluster. The results of this third option are still under investigation.

### 3. Algorithm Evaluation

To evaluate the results of various combinations of runtime options, a new type of display was generated from the model output. As described above, each cluster within an  $1/8$  mesh grid box is represented by a single grayshade selected to characterize the entire cluster for the cloud threshold procedure. After thresholding, resulting cloud layers can be represented as images through the RGS. The procedure is to replace each pixel in the original IR images with the RGS of the cloud layer (cluster) of which that pixel is a member. The computer generated pseudo-image that is produced, referred to as an RGS display, can be displayed on a McIDAS video monitor. Fig. 8 is a side-by-side display of an original IR image and the RGS pseudo-image for a single whole mesh grid box (100 nm) over a land background within neph grid box 45. The image is from June 11, 1982, and is part of



Fig. 8. IR Image (left) and RGS Display (right) for a Single Whole Mesh Box with a Land Background - June 11, 1982

a data save made at GWC for the RDNEPH effort at AFGL. The IR image is on the left and the RGS is on the right. If the model is performing correctly the cloudy areas of the RGS display should resemble the cloudy areas in the image, both in shape and in grayshade level. Of interest in this scene are the clear area running top to bottom on the left half of the box (clear areas are represented as black in the RGS display), the multiple cloud layers in the right half of the box, and the broken clouds in the lower left corner. Over this whole mesh box the model has done a good job detecting the cloud boundaries due to the large thermal contrast between the (warm) land and the cloud areas. The cloud model has more trouble over a variable background. Fig. 9 shows another whole mesh box in neph box 45. In contrast to the background in Fig. 8, this box contains a coastline boundary with a relatively cool ocean surface and a warm land surface. This figure shows the IR image (9a) and the RGS display (9b) as before, plus the corresponding visible image (9c). Over the ocean a low stratus cloud deck is clearly discernible in the visible image but is very difficult to detect in the IR. This is due to the cool ocean background ( $\sim 10^{\circ}\text{C}$ ) underlying the low stratus clouds at about the same temperature. Low thermal contrast between the cloud and background has caused the cloud detection algorithm to misinterpret some of the cloud as clear background (illustrated by the black area in the center region of the RGS display). Other areas of interest in Fig. 9 are the patches of cloud (gray) in the upper center of the RGS display. By examining the corresponding IR and visible images it can be seen that these are actually cloud-free areas along the ocean side of the coastline. The ragged edges on the land side of the RGS cloud areas exactly match the coastline contours recognizable in the imagery. This overanalysis along coastal boundaries is caused by the coarseness of the background temperature data base. Refer to Gerlach (1982)<sup>14</sup> for a description of the nephanalysis data base structure. Background temperatures used in the cloud threshold procedure are maintained at a 1/8 mesh (25 nm) grid resolution. In locations where a 1/8 mesh box contains a coastline boundary, the land surface temperature is used to represent the entire box regardless of the relative proportions of land and ocean surfaces.<sup>15</sup> Under conditions of a large thermal gradient

15. Personal communication from Maj. J. Koenig (1983).

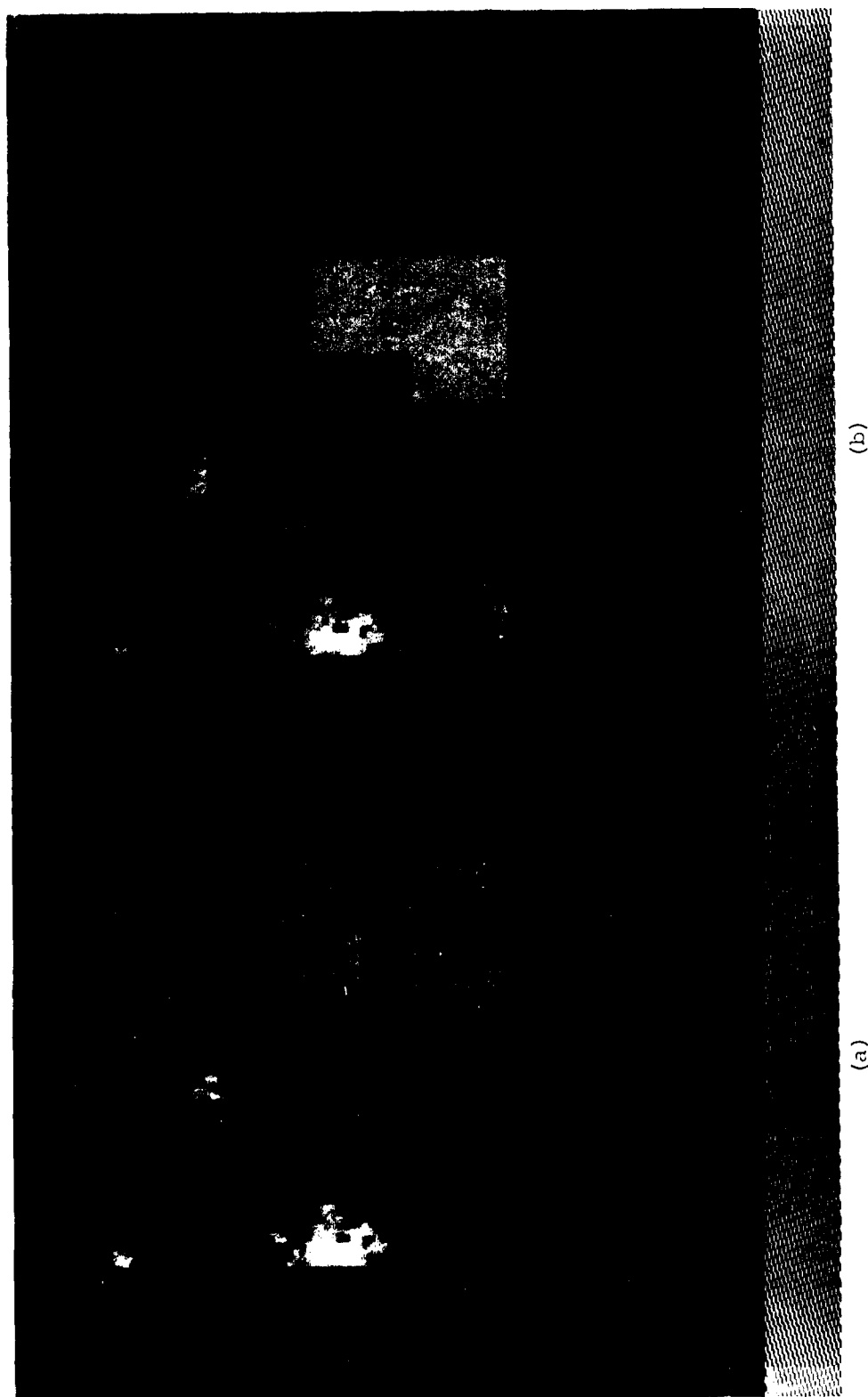
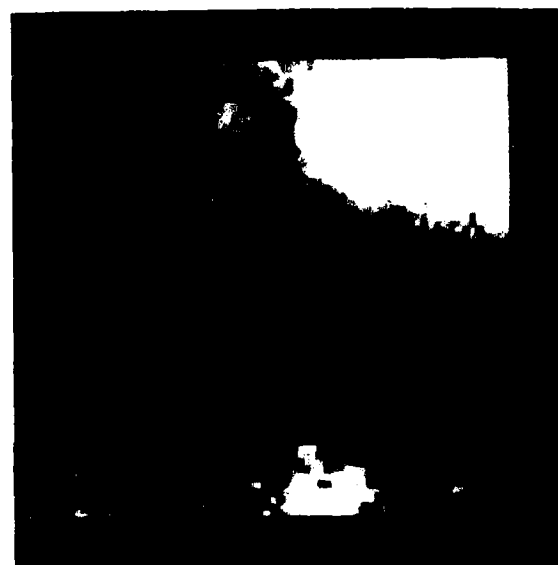
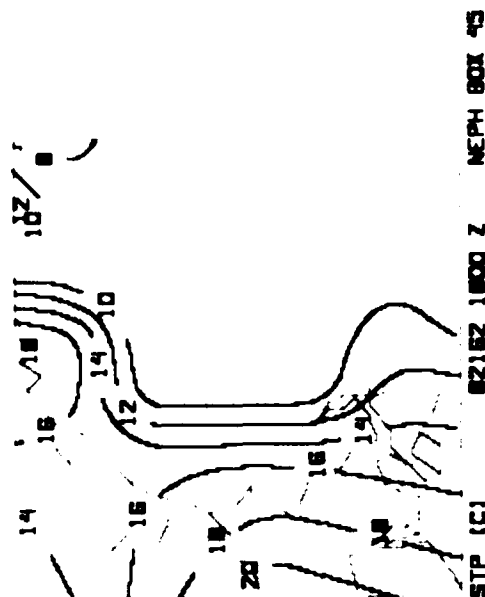


Fig. 9. IR (a) and RGS (b) Images for a Single Whole Mesh Box with a Coastline Boundary -  
June 11, 1982



(c)



(d)

Fig. 9. Visible (c) Image Plus Surface Temperature Analysis (d) for a Single Whole Mesh Box with a Coastline Boundary - June 11, 1982

between a cool water surface and an adjacent warm land surface as in this example (refer to Fig. 9d), the ocean background is misrepresented as too warm. Thus when the analysis program thresholds the satellite-sensed cold water against the too warm background temperature, the scene is misinterpreted as cloud.

#### 4. Sharpness

In addition to evaluating the cloud analysis accuracy, a second priority at AFGL is to produce a more binary cloud analysis; that is, an analysis that is either completely cloud-free or completely cloud-covered at the analysis resolution. A measure of this quantity is sharpness, generally reported as a percent score. For example, a 20-20 score would give the fractional number of grid boxes with a total cloud amount within 20 percent of total clear or total overcast at each analysis grid point.

Sharpness scores have been computed for cloud analysis over five neph grid boxes in the northern hemisphere. Analysis valid time was 20Z on June 11, 1982. Although analysis accuracy is a separate problem, run-time options were selected that produced the most accurate analysis when compared to manual interpretation of the imagery. The results show that the cloud analysis program produces a very sharp analysis with an overall 20-20 score of 0.901. Sharpness scores were also computed as a function of background type. Ocean backgrounds produced a sharper analysis at 0.913 for 15,858 grid points than land background regions at 0.809 for 2,106 points; coastline points were excluded. No general conclusions can be drawn from these results due to the large disparity between the number of land vs. ocean points; however, it does draw attention to unexpectedly sharp analysis over ocean. Neph box 45 is composed of approximately 30 percent land points (1,125) and 60 percent water points (2,374), a proportionately larger number of land points than the five box study above. A sharpness computation of the cloud analysis for box 45 alone produced a 20-20 score of 0.861 over ocean and 0.738 over land, again showing that ocean areas contribute more to overall sharpness. The RGS display for the entire neph grid box 45 is given in Fig. 10. Cloud boundaries over ocean tend to follow the 1/8 mesh grid box pattern (appear geometric) much more than clouds over land. The 1/8 mesh grid is the resolution of the background temperature data base.



Fig. 10. RGS Display for Neph Grid Box 45

(Analysis valid time 20Z, June 11, 1982)

Examination of low cloud over ocean situations has revealed that the cloud/no-cloud decision is often made on a temperature difference of less than  $1.0^{\circ}\text{K}$ ; consequently a small change in background temperature across an  $1/8$  mesh boundary can cause an opposite cloud determination over a thermally homogeneous cloud field. This would produce an  $1/8$  mesh grid pattern to appear as cloud boundaries as in Fig. 10. This effect is enhanced by the difference between the temperature resolution of the background data base at  $0.1^{\circ}\text{K}$  and the satellite radiance temperatures at  $1.56^{\circ}\text{K}$ . Thus for low cloud situations, a very high sharpness score for an analysis over ocean is at least partially an artifact of the analysis program and not representative of nature.

#### 5. Accuracy

While a quantitative measure of sharpness, such as the 20-20 score, can be made directly from the nephanalysis results, no direct measure of accuracy is available. In the past attempts to measure accuracy have required an investigator sitting at an image display screen attempting to estimate subjectively a percent cloud cover at some fixed resolution by examining a visible or IR image. This type of study was performed to compare the 3DNEPH IR processing algorithm against an AFGL designed algorithm (d'Entremont et al., 1982<sup>16</sup>). In that study 350 cases were selected for investigation. The entire set of 350 covered an area only 34 percent as large as a single neph grid box and yet required several man-months of effort to collect and analyze the data. As a consequence of the time required to process data the 350 selected cases were culled from a much larger set in order to provide a manageable number. The possibility of introducing a bias into the data set always exists when this type of selection is made.

Considerable effort has gone into developing a scheme of measuring cloud analysis accuracy that will minimize these limitations. The resulting method is to compare cloud model results against a standard cloud truth data base. This data base is to be composed of 6 to 8 neph grid boxes that

16. d'Entremont, R.P., R.S. Hawkins, and J.T. Bunting, 1982: Evaluation of Automated Imagery Analysis Algorithms For Use in the Three-Dimensional Nephanalysis Model at AFGWC. AFGL-TR-82-0397, Air Force Geophysics Laboratory, Hanscom AFB, MA.

are representative of a range of geographic and climatic conditions. The cloud truth data base will be generated using a background blanking technique. This technique requires several trained observers to make independent, subjective cloud/no-cloud decisions over a small area by analyzing data from as many sources as possible. Once cloud boundaries have been established, background blanking will be done at an interactive image processing terminal. Working from single, multiple, or composite images the observer will interactively raise or lower a neutral background color across the image gray-scale until only cloud areas remain visible above the background threshold. The processor can then automatically store the cloud boundary information in the cloud truth data base and from that compute percent cloud amount at any resolution down to the individual pixel size.

This technique does not eliminate the need for a subjective determination of cloud boundaries but it does provide several advantages. Data from a variety of sources are accommodated, including multispectral IR and microwave imagery. Precise cloud boundaries are accurately delineated and stored as opposed to a subjective cloud amount estimate. Also it becomes possible to analyze a large amount of data quickly allowing for the size of the projected cloud truth data base. Plans are under development to implement a threshold blanking capability on the second generation AFGL McIDAS currently in procurement.

## D. Microprocessor-Based Satellite Data Ingest System

### 1. Introduction

Under a contractual agreement between the National Oceanic and Atmospheric Administration and AFGL, a redundant data path in the AFGL GOES mode A ground station is being made available to a consortium of users known as the Northeast Area Remote Sensing System (NEARSS). Included in the agreement is a requirement to develop a central processing facility to ingest, store, and transmit over a dedicated communication network subsets of the GOES imagery transmission. To satisfy this requirement an INTEL 8086-based microprocessor with fixed disk mass storage was purchased and assembled. Development of software to handle the task and configuration of the central processor is being accomplished by AFGL personnel. SASC software engineers are participating in various phases of the software system specification and development including operating system configuration, data base design and management, automated scheduling of the ingest/store/transmit cycle, and terminal command processing.

### 2. Software System Specifications

Before software development can begin, a more detailed specification of the required tasks must be developed. To this end the following task outline has been identified:

#### a. Data Ingest

- 1). Automated scheduling of real-time ingest sequence
- 2). Operator override of automated sequence
- 3). Image subset specification
- 4). Ingest of navigation parameters
- 5). Ingest of imagery data

#### b. Data Storage

- 1). Data base design
- 2). Data base manager

#### c. Data Transmission

- 1). Communication network manager
- 2). Data transmission routine

With the division of the process into identifiable tasks, decisions regarding an operating system can be made. First, it must be a real-time operating system equipped to handle the asynchronous collection and dissemination of data and operator commands. Second, it must provide for multitasking to encourage the simpler modular approach to program design and to allow concurrent processing of overlapping tasks. Third, it must allow synchronized task processing and intertask communication through system service requests, since the various tasks composing an ingest/transmit sequence must execute in an orderly fashion and require the sharing of common data. Finally, the operating system must be flexible enough to accommodate unanticipated system design changes or expansion. The iRMX86 real-time multitasking executive designed by INTEL meets all these requirements and was therefore chosen.

### 3. Data Base Management

The singular nature of this system, i.e., to manage GOES mode A data, simplifies the requirements of its data manager considerably. The only information items requiring consistent handling and storage are parameters identifying an image and defining its areal extent, resolution, and location; orbital parameters defining the satellite's location and motion; and the visible and IR data points comprising the image. These three groups of data — directory data, navigation data, and imagery data — are the fundamental units in the system data manager.

Each data type was assigned a data structure and a file type to systematize its handling and storage. Directory data and navigation data have a fixed format and therefore could be treated as logical records within a fixed length file. Imagery data have a more variable nature due to the flexibility in specifying the size of the image subset to be ingested and so were assigned to individual files whose size can be varied. A further consideration was the interrelated nature of the three data bases: each directory data entry has exactly one navigation data entry and one imagery file that correspond to it. Taking advantage of this interrelationship, entries in the directory and navigation data files corresponding to a single image have been assigned the same record number. The name of the corresponding imagery data file is also chosen to include that number as an integral

part. Using this technique only the directory data file must be searched to locate any information regarding a particular image, since that directory entry location is also the key to the navigation data and imagery data file. Disk input/output requirements are also substantially reduced because of the smaller size of the directory file. A detailed description of each data base file type and structure appears in Appendix V-D.

A rotating entry scheme was utilized for directory entries in an effort to reduce operator intervention and provide for limited archive capability. The rotating scheme automatically overwrites the oldest data set with the newest, eliminating the need for operator specified assignment. A temporally continuous set of imagery data is left intact by this method should the need for old data arise.

With a data base structure defined, the software modules that make up the data base manager could be designed and written. The primary subset of modules is composed of an unbuffered, record-directed input/output subsystem. These modules provide FORTRAN-callable links to the low overhead iRMK86 Basic Input Output System (BIOS) while simplifying the activation call by reducing the number of calling arguments. The modules designed were: DYNASG, a routine to dynamically assign a logical unit number to a hierarchically named file; SCRA, a routine to position the file pointer to the beginning of a specified record; READW, a synchronous, unbuffered data read routine; and WRITW, a synchronous, unbuffered data write routine.

Using the I/O subsystem modules as a foundation, data specific handling routines were developed. ENTDIR and ENTNAV were designed to update a complete entry in the directory data file and navigation data file respectively and LISDIR and GETNAV provide single entry retrieval capabilities. Imagery file management routines developed include: ARANAM, to build an imagery data file name based on a directory entry number; GENAR, to dynamically generate an imagery data file; ZEROAR, to initialize the file to all zeros; and AREMOV, to delete an imagery file.

#### 4. Scheduler Routine

A scheduler routine is the main job of the automated ingest system. It activates the various tasks that initialize the data files, ingest data, and transmit the imagery. An integral part of the scheduler is a clock

task maintaining the current year, day of year, and number of seconds elapsed since 00 GMT. It utilizes the programmable interval timer available on the 8086 CPU and is set at system initialization by the time information contained in the GOES IR documentation stream. A separate task within the scheduler monitors the clock and initiates a main or controlling task at 25 and 55 minutes past each hour. The main task reads from disk storage an event file containing the current sequence of commands identifying the sector of imagery to be ingested for that period. If no operator-supplied command string exists for the current time, a default sequence is used. Each command is a set of instructions for a particular task associated with the ingest/transmit sequence. The main task initiates each task in turn, suspending itself while each executes. A test is made upon receipt of each task's completion message for error conditions to determine if the next task is to be initiated or if the entire chain of events is to be aborted. After successful processing of the event file or an unscheduled abort the main task suspends itself to await reinitialization by the monitor task.

#### 5. Terminal Communications

Although the ingest system is being designed to run without operator intervention, that capability is being included in the software to meet unanticipated requirements. Routines to handle command requests and inter-user communications of text messages were developed to handle terminal communications.

The command string interpreter TTYTTY was designed to validate and decode commands of ASCII text received from the command console. TTYTTY verifies that the requested task is valid by comparing the six-letter task request to entries in a file containing all currently defined task names. If the request is invalid an error message is transmitted and processing terminates. Valid requests are further processed to decode any command parameters. Alpha text is packed two characters per word and numeric text is converted to its integer equivalent while simultaneously checking for word limit overflow. The requested task is then initiated and parameters are passed to it via the iRMX86 intertask communication procedure.

The terminal output routine, module THOUPUT, was designed to validate and reformat ASCII text strings prior to terminal transfer. THOUPUT checks that the message length does not exceed terminal output specifications and inserts carriage control characters. The complete message is copied

to a memory segment compatible with the requirements of the terminal handler, an operating system layer supporting asynchronous I/O between the user console and THOUTPUT. Transfer of control is made to the terminal handler task to complete the output procedure.

#### 6. Summary

The specifications for and development of a microprocessor-based GOES mode A data ingest system have begun. The definition of the required tasks and selection of an appropriate operating system have been completed. A data base management scheme has been defined and developed. Work is continuing on defining and coding the individual tasks.

## Appendix V-D. Data File Structures Description

File Name: DATDIR

Size: 7 Sectors (896 16-bit words)

Type: All integer data

Purpose: Contains directory entries for digital areas (satellite data files)

### Structure:

WORD	PURPOSE
1,2	Double precision integer representing maximum number of bytes on disk allocated for digital areas
3,4	Double precision integer representing number of bytes currently being used for areas
5	Oldest area number
6	Next area number to be generated in sequence
7-16	Directory entry for area 1
17-26	Directory entry for area 2
887-896	Directory entry for area 89

A maximum of 89 digital areas can exist at one time. Each area cannot exceed 32750 sectors (8,384,000 bytes) in length.

### Directory Entry Structure:

WORD	PURPOSE
1	SSYY - Satellite ID number and year
2	DDD - Julian day
3	HHMM - Greenwich time of data (hours and minutes)
4	LCOR - Starting line number
5	ECOR - Starting element number
6	LRES - Line count increment, represents resolution approximately
7	ERES - Element count increment
8	LSIZE - Number of lines of data
9	ESIZE - Number of elements per line
10	STATUS - -1 = area non-exist, 0 = no data, 1 = area OK, 2+ = area in use

File Name: LNDFIL

Size: 90 sectors (11524 16-bit words)

Type: Mixed, some 2-byte integer, some 4-byte integer

Purpose: Contains navigation information for all digital areas and  
most recent navigation

Structure (Sector 0 through 89):

WORD*	PURPOSE
1	SSYY - Satellite - year identifier
2	DDD - Julian day
3	HHMM - hour-minute of image start
4	zero-filled
5,6	ORBIT TYPE - 1 = classical, 2 = Phillips. Set to 1 for NEARSS
7,8	EPOCH DAY (YYMMDD)
9,10	EPOCH TIME (HHMMTT)
11,12	SEMI-MAJOR AXIS
13,14	ECCENTRICITY
15,16	INCLINATION
17,18	MEAN ANOMALY
19,20	ARGUMENT OF PERIFOCUS
21,22	RIGHT ASCENSION OF ASCENDING NODE
23,24	ATTITUDE - DECLINATION
25,26	ATTITUDE - RIGHT ASCENSION
27,28	ATTITUDE - PICTURE CENTER LINE
29,30	SPIN RATE
31,32	FRAME GEOMETRY - DEGREES OF SWEEP IN LINE DIRECTION
33,34	FRAME GEOMETRY - TOTAL NUMBER OF LINES
35,36	FRAME GEOMETRY - DEGREES OF SWEEP IN ELEMENTS
37,38	FRAME GEOMETRY - TOTAL NUMBER OF ELEMENTS
39,40	CAMERA GEOMETRY - PITCH
41,42	CAMERA GEOMETRY - YAW
43,44	CAMERA GEOMETRY - ROLL
45,46	(RESERVED)
47,48	BETA ADJUSTMENT
49,50	TIME FOR WHICH BETA ADJUSTMENT COMPUTED
51,52	(RESERVED)
53,54	SUN SENSOR MOUNTING ANGLE
55,56	SKEW
57,58	PRECESSION RATE
59,60	PRECESSION DIRECTION
61,62	BETA CODE ONE
63,64	
65,66	
67,68	
69,70	
71,72	

\*A WORD is a 16-bit integer

## WORD

## PURPOSE

73,74	BETA CODE TWO
75,76	
77,78	
79,80	
81,82	
83,84	
85-128	NOT CURRENTLY USED

## Notes:

Sector 0 contains the most recent navigation available. Data in Sectors 1 to 89 correspond to data in digital areas 1 to 89.

File Name: X00XNN (where NN is the area number 01 to 89) )

Size: Variable (always a multiple of 256 data elements (bytes) up to a maximum of 8,384,000 bytes)

Type: Unsigned byte (8-bit) data

Purpose: Hold satellite digital data

This file contains satellite digital data. All relevant information pertaining to these data is contained in the area directory file DATDIR.

## E. McIDAS Software Enhancement

SASC McIDAS software efforts consisted of three types of activities: debugging existing software, expanding the capabilities of operational programs through addition of new options, and development of new programs. The following sub-sections detail the efforts made in these areas.

### 1. Three-Dimensional Graphics

Investigation into graphics text annotation uncovered the existence of three-dimensional drawing that had previously remained unused because of lack of documentation. The graphics driver ANNOT contains options that allow the user to control both the depth (thickness) and number of times an alphanumeric character or string of characters is repeated with varying internal levels. Documentation of the options was appended to the original write-up of ANNOT and appears in Appendix V-E.

### 2. Image Frame Identification

To reduce user search time in finding TV-loaded satellite imagery for a particular day, a general driver was developed and implemented to monitor the 400 frame image field. Program LISENT lists all unique satellite-year-day identifiers that exist in the field for the requesting terminal. When used in conjunction with the program that lists all frames and times for a specific identifier, the user can pinpoint the frame(s) of interest in a minimal amount of time. Appendix V-F provides documentation of the command format of LISENT.

### 3. Graphics Drawing

Graphic displays generated on the Plot Display File (PDF) by programs accessing modules in the McIDAS graphics library GRALIB revealed that vertical line width was variable, apparently determined according to the internal level (color) chosen. An investigation into the library module INITPL, a subroutine that executes plot initialization procedures, confirmed the observation. Elements of a 15 word line width common block, corresponding to the 15 internal levels available for plotting on the PDF, were indexed with and set to the value(s) of the chosen internal level(s).

To rectify the situation, INITPL was modified to define a fixed line width for all internal levels, and all drivers that access GRALIB modules for plotting were then recompiled so that the revised version would be used.

#### 4. Software Documentation

In anticipation of the extended use of nowcasting library modules in software development, documentation of the library SVCLIB was undertaken and completed by SASC. The documentation procedure also included cleaning the library of modules no longer accessed by nowcasting drivers and other programs, and modules not compatible with AFGL McIDAS. For each module retained, the subprogram type, calling sequence, and description of the arguments were recorded. Any additional information describing the use of the routine, such as files accessed, variable type, calling routine (if any), and other helpful comments, were also indicated.

Subroutine DECODE, an SVCLIB module that unpacks hourly surface data, was replaced by a version that returns unpacked expansion data as well as basic data. To provide more cloud cover and height information to nowcasting plot and analysis contouring routines, this version was later modified to assign valid ceiling heights of 0 to stations reporting the lowest cloud level as overcast with no cloud height information and to assign a value of 999 (unlimited ceiling) for any station that does not report any level as either broken or overcast. In association with DECODE, a new SVCLIB module was written to determine composite cloud information. Subroutine CMPCLD is given the word containing the packed cloud cover category numbers of all three levels as generated in DECODE, and returns the category number associated with the lowest cloud level reporting the greatest amount of cloud cover.

#### 5. Map Specification Compatibility in Nowcasting

Generating an integrated display, such as one produced by an objective analysis and contouring routine and a plotting routine to locate geographically the raw data input to the analysis, requires a map projection common to the routines that produce it. McIDAS automated objective analysis and contouring routines maintain considerable flexibility in providing the user a number of ways to specify a map, while nowcasting plot routines remained restrictive in this area.

Consequently, the surface and upper-air plotting routines IASVCA and IARAOB were modified to include command options for specifying a map with earth coordinates (latitude - longitude), based on either a center-point or NW and SE corners. Changes were also made to include software checks to adjust the map size if necessary, and eliminate boundary inconsistencies when superimposing gridded data. In this way, increased map specification compatibility among nowcasting graphic routines has increased integrated display capabilities on McIDAS.

#### 6. Data Plotter Enhancement

To complete incorporation of a ship and buoy data base into McIDAS nowcasting routines, modifications were made to the data plotting routine WISPLT. Ship and buoy data management routines were incorporated into the data collection module, permitting display of temperature, dew point temperature, pressure, and winds from these data sources when appropriate. Additional changes were also made to adhere to a more modular programming method and more closely follow the programming logic used in the graphics data plotter IASVCA.

#### 7. Input Parameter Processing

The McIDAS command line interpreter TTPROC contains a macro command language enabling a single key-in to generate a command file composed of several command strings to be executed sequentially. In this way long repetitive sequences can be executed with a minimal amount of effort. However, if any routine initiated by a macro command file requests further user input, the requests are routed to the command file and not the user console, thus eliminating interactive capability.

Subsequent research indicated that the input queue query routine included in every application program, subroutine IQ, automatically routed all requests for input to the source of the job initiation, either the user console or the macro command file. It was necessary to modify IQ to force routing of all requests for input after the first to the user console. This was accomplished by addition of a flagword that is set by the initial input request. When set, all subsequent entries to module IQ bypass the input selection code and use the default source, the user console. In this way

a program can be used within a macro command file and still retain its interactive capability.

Several McIDAS routines require as input argument parameters that are a combination of alphabetic and numeric characters, such as 4DG or HT1. The existing command line interpreter TTPROC was designed to separate alphabetic and numeric values, passing alpha text as ASCII bit patterns and numeric text as integer values, preventing mixed character input. To rectify this situation the line scanner module in TTPROC was modified to recognize a new control character, the comma. A comma in the input command string directs the scanner to pass all subsequent characters as ASCII bit patterns, packed in 24-bit words, until a second comma is detected. At this point normal input string decoding resumes. This addition to TTPROC removes the need to repack mixed text and numeric values within application routines, thus increasing their flexibility and eliminating unnecessary input data processing.

Appendix V-E. Documentation of Three-Dimensional Drawing for Program ANNOT

COMMAND: TEXT ANNOTATION, PAGE 2

THREE - DIMENSIONAL TEXT ANNOTATION

ZA WWDDHH COLORS \*TEXT ..... CR

PARAMETER DESCRIPTION:

DD - TEXT DEPTH OR THICKNESS (90 - 99)  
WHERE UNITS PLACE INDICATES THICKNESS.

COLORS - PACKED WORD CONTAINING INTERNAL LEVELS (0 - 9).  
LEVELS ARE EXTRACTED FROM RIGHT TO LEFT.

EXAMPLE:

ZA 9430 123 \* HELLO

THE WORD "HELLO" IS DISPLAYED IN YELLOW WITH A HEIGHT OF 30  
PIXELS AND A DEPTH OF 4. THE DISPLAY IS THEN REPEATED FOR THE  
COLORS GREEN AND RED FOR A COMBINED DEPTH OF 12 PIXELS.

# Appendix V-F. Command Documentation for Program LISENT

COMMAND: LIST ALL UNIQUE SSYYDDD IDENTIFIERS

KEY IN: LE

PROGRAMMER: C. IVALDI

DATE: DEC 82

CHAPTER: 4

## COMMAND FORMAT:

LE OUTDEV

## COMMAND DEFAULT PARAMETERS:

LE LIS

## PARAMETER DESCRIPTION:

OUTDEV = Output device for listing  
LIS for terminal CRT  
PRI for line printer

## EXPLANATION:

The LE command is used to list all unique satellite-year-day identifiers that exist in the image field for the requesting terminal. Identifiers are ordered by image day in rows.

## F. McIDAS Operations and Enhancement

### 1. Northeast Area Remote Sensing System (NEARSS) Development

The following equipments were acquired for construction of the microprocessor-based satellite ingest system:

- a. Hazeltine Executive 80 Terminal, manufactured by Hazeltine Corporation, Greenlawn, NY.
- b. ZX-86 16-bit single board computer, manufactured by Zendix Corporation, Dublin, CA.
- c. Micro 4000 Data Modem, manufactured by Micom Systems Inc., Chatsworth, CA.
- d. Model F880 Magnetic Tape Transport, manufactured by Cipher Data Products, San Diego, CA.
- e. SA 810/860 Diskette Storage Drives, manufactured by Shugart, Sunnyvale, CA.
- f. Comark's Diskstor System, manufactured by Comark Corporation, Waltham, MA.
- g. Rimfire 38A Winchester Disk Controller, Manufactured by Computer Product Corporation, Plymouth, MA.
- h. Priam 70m/byte Winchester Disk, manufactured by Priam Corporation, San Jose, CA.

Integration work was begun on cables, interfaces, and housing for these equipments. Software for the operating and receiving system is a joint effort of SASC, AFGL, and Massachusetts Institute of Technology. The AFGL satellite ground station will be the data source. Satellite images will be processed by the system and sent over high speed data lines to user stations throughout the New England area.

### 2. McIDAS Enhancement

Two electronic circuits were fabricated to enhance McIDAS operational capabilities. The first is a Sony remote playback control box, which saves time and effort when many hours of satellite imagery have to be played back into the system. Previously, the operator had to leave the terminal area and

go to another room to advance, play, rewind, or stop player action for each satellite picture being processed into the system. This required some ten trips for each cassette. Now the only time the operator leaves the terminal area is to change cassettes. Features of the remote control box are a fast forward, rewind, playback, and stop switch. All switches have indicator lamps that indicate to the operator what mode of operation the Sony playback unit is in. The control box consists of four 2SC634A transistors with associated resistors, switches, and light emitting diodes. An additional remote jack was installed in the Sony player to handle the remote control box.

The second circuit fabricated for the system operator is a remote indicator box for the Laserfax 850. The Laserfax 850 is integrated into the McIDAS for making hard copies of presentations on the system's TV monitor. The remote indicator, installed in the Terminal 21 area, indicates to the operator what mode of operation the Laserfax 850 is in. The different modes of operation are: power on or off, rapid advance, recording, process, and standby. Prior to installation of the remote indicator the operator on Terminal 21 had to leave the terminal area and go to the Terminal 22 area, located in another room, to determine if the picture making process had been completed for each hard copy being made. Once completed, another hard copy command could then be entered into the system. If, however, another hard copy command had been entered into the system before completion of the previous hard copy, the first hard copy process would have been terminated, producing an unfinished image.

### 3. McIDAS Operation and Maintenance

McIDAS downtime during the year consisted of a two-day period when a new ceiling and area lighting were installed; a one-day period when a new motor and spindle for the system's 80m/byte disk were installed; and a one-day period when a new polarization motor was installed in the antenna system. There was restricted operational capability for a four-day period due to the loss of the FAA line, disconnected accidentally by the telephone company.

The following major items required maintenance during the year. The end item is listed, followed by the components that were replaced to bring the equipment back to operating condition.

#### a. Harris 80m/byte Disk

5 amp fuse, motor and spindle assembly

- b. Digital Video Storage System (DVSS)  
RT1 and R37 located in the 5v power supply
- c. Off Line Ingest System (ODIS)  
Q4 and CR12 located in the Ann Arbor power supply
- d. Terminal 21 Red Enhancement Drawer  
Coax line driver LH033
- e. Terminal 21 Ann Arbor Terminal  
Q1, Q2, and C26 in the power supply
- f. Digital Interface No. 3  
Word address latch chip "A" (8202)
- g. Microprocessor Development System (MDS 800)  
Read/write/record buttons on diskette storage drives
- h. Tektronix Display Formatter DF2  
U318 ram memory chip
- i. VAS Preprocessor  
E66 (74LS169) located on the averaging board
- j. Sony Archive System Recorder 1  
Fast forward and reverse idler assembly
- k. McIDAS Antenna System  
Polarization motor
- l. Hazeltine Executive 80 Terminal  
Monitor board

LEND

FILMED

8