GL-TR-89-0099

AD-A210 108

SATELLITE ENHANCED NUMERICAL WEATHER PREDICTION

R. N. Hoffman, R. Isaacs, J.-F. Louis, T. Nehrkorn and M. Mickelson

Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139-3794

11 April 1989

Final Report
July 1986 - February 1989

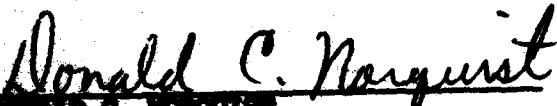Approved for public release; distribution unlimited

GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AFB, MASSACHUSETTS 01731-5000

89 7 07 041

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

Donald C. Norquist   Donald A. Chisholm

DONALD C. NORQUIST   DONALD A. CHISHOLM, Chief
Contract Manager   Atmospheric Prediction Branch

FOR THE COMMANDER

Robert A. McClatchey

ROBERT A. McCLATCHEY, Director
Atmospheric Sciences Division

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

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, 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 require that it be returned.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
	GL-TR-89-0099

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Atmospheric and Environmental Research, Inc.		Geophysics Laboratory

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
840 Memorial Drive Cambridge, MA 02139	Hanscom AFB, MA 01731-5000

8a. NAME OF FUNDING / SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
Geophysics Laboratory		F19628-86-C-0141

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
Hanscom AFB, MA 01731-5000	61102F	6670	10	CB

11. TITLE (Include Security Classification)

Satellite Enhanced Numerical Weather Prediction

12. PERSONAL AUTHOR(S)
R.N. Hoffman, R. Isaacs, J.-F. Louis, T. Nehrkorn and M. Mickelson

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final Report	FROM 7/86 TO 2/89	1989/4/11	116

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Data assimilation; Numerical weather prediction; Satellite retrievals; SSM/T-1; SSM/T-2.

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

Satellite observations hold great promise for improving numerical weather prediction (NWP). During the current contract effort we have made substantial progress in several areas related to improving NWP, especially the analysis and forecast of moisture and cloudiness, through space borne observing systems. We have developed, implemented and evaluated techniques for retrieving geophysical parameters from satellite observations for use in NWP. We have tested the impact of these data and data from other sources in a series of observing system experiments (OSEs). We have developed and applied a comprehensive data simulation methodology. We have conducted a series of observing system simulation experiments (OSSEs) to evaluate the impact of new observing systems.

Our work in the area of techniques development has focused on retrieving relative humidity profiles from three distinct sources: the civilian infrared sounder (TOVS), the

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Donald Norquist		GL/LYP

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

19. ABSTRACT (continued)

operational nephanalysis (3DNEPH) and the DMSP microwave sounder (SSM/T-1,2). In summary the real data sources, TOVS and 3DNEPH, provided only marginally useful data, while in simulation, the SSM/T retrievals were significantly useful.

A series of real data OSEs were conducted to measure the impact of retrieved humidity profile data described above as well as the impact of other satellite data. In addition two data denial experiments examined the effect of using limited surface-based data or of using satellite observations only. Such data systems are clearly inferior to the full operational system. However for the first few days of data denial such a system may be adequate. Finally another OSE was used to analyze the performance and the impact of the new physics package of the GSM. No clear improvement was seen in the forecast errors of this experiment.

We have conducted a series of realistic OSSEs to assess the impact of a Doppler lidar wind (DLW) sounder and the SSM/T-1,2 temperature and humidity profiles. The addition of DLW profiles significantly improved the initial state specification, especially in the S.H. extratropics relative to our control experiment. The addition of the SSM data significantly improved the moisture analyses in the tropics and S.H. extratropics.

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

ii

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Tables

1. INTRODUCTION

This report describes the work performed by Atmospheric and Environmental Research, Inc. (AER) for the Air Force Geophysics Laboratory (AFGL) under contract number F19628-86-C-0141, entitled "Satellite Enhanced Numerical Weather Prediction" for the entire period of the contract, 3 July 1986 through 3 February 1989.

There were two main purposes to this work. The first purpose was to study the impact of existing and planned satellite observing systems on numerical weather prediction. This was done by performing a series of data assimilation experiments, with both real and simulated data. The second purpose was to develop techniques for analyzing satellite observations to retrieve geophysical parameters most useful for numerical weather prediction (NWP). This work was done in concert with the data assimilation experiments. This series of experiments made it possible to achieve the secondary goal of this work, which was to test extensively the AFGL global data assimilation system (GDAS).

The AFGL GDAS, composed of an optimal interpolation data analysis scheme, a nonlinear normal mode initialization module and a weather forecast model, is described in Section 2 and documented in Scientific Report No. 2 (Hoffman, 1986). A number of improvements and modifications to the analysis program were performed during the contract period. These modifications were described in Scientific Report No. 4 (Hoffman et al., 1988), and are summarized in Section 2.2.

The assimilation experiments that used real data (called observing system experiments, or OSE's) first dealt with the impact of currently available satellite data, and explored the possibility of using satellite data as the primary source of information. Hence the first four experiments included using no satellite data at all, using the current operational system, and using a purely space-based system, either over a portion of the globe or everywhere. Scientific Report No. 3 (Louis et al., 1987) covered these first four experiments, which are also summarized in Sections 3.2 to 3.4

A second series of assimilation experiments explored the possibility of using different retrieval techniques than those used now operationally. They include a physical retrieval of temperature, the use of cloud

data to infer moisture profile (Scientific Report No. 5, Nehrkorn and Hoffman, 1988), and a physical retrieval of moisture. Results of these experiments are presented in Sections 3.5 to 3.7.

The final real data experiment used the forecast model developed at AFGL, which includes a new set of more sophisticated parameterization schemes for the physical processes. In addition to the data assimilation experiment, a detailed evaluation of the new model was performed, including some runs with a one-column version of the model. The results of these tests and evaluations are described in Section 3.8. More details can also be found in the Appendices.

Section 4 reports the experiments that used simulated data (observing system simulation experiments, or OSSE's). Two possible sources of data were explored: temperature and moisture data from the SSM/T microwave sounder (the temperature data exist but only recently have they been used operationally) (Scientific Report No. 1, Isaacs, 1987), and wind data from a possible Doppler lidar system. The WINDSAT OSSE is described in detail by Hoffman et al. (1989) (Scientific Report No. 7) and the SSM OSSEs, including the simulation of the SSM data are described in detail by Grassotti et al. (1989) (Scientific Report No. 6). In this report in Section 4.1, general OSSE design issues are discussed. Then the nature run (Section 4.2) and simulated data (Section 4.3) are described. Section 4.4 summarizes our experimental design and Section 4.5 summarizes our results and conclusions.

Finally, conclusions and suggestions for further research are given in Section 5.

Appendices provide more details on the studies that have not been the subject of separate scientific reports, namely AER's physical retrieval of water vapor (Appendix A), the one-dimensional tests performed on the new physical package for the AFGL GSM (Appendix B), the results of the observing system experiment that used this new GSM (Appendix C), and a study on model parameter optimization (Appendix D).

To facilitate evaluation of this work, Table 1 summarizes the correspondence between the items of the contract's statement of work and the various sections of this report, as well as the scientific reports and papers published under this contract.

Table 1. Project summary and references

Line Item	Description	Reference	Report Section
0001	Improve AFGL global data assimilation system and test the impact of various types of satellite data.	Isaacs et al. (1986a)	2
	Documentation of ASAP Modifications to ASAP	Isaacs et al. (1986a) Hoffman (1986) Hoffman et al. (1988)	
0001AA	183 GHz water vapor retrievals		4
	Review study	Isaacs (1987) Isaacs et al. (1986b) Grassotti et al. (1989) Hoffman (1988) Isaacs et al. (1988)	
	SSM/T-1,2 data simulation		
	NOSAT OSSE STATSAT OSSE SSMSAT OSSE SSM+TOVS OSSE WINDSAT OSSE	Hoffman et al. (1989) Grassotti et al. (1989)	
0001AB	Satellite data OSEs		3
	NOSAT OSE STATSAT OSE GLASAT OSE	Louis et al. (1987, 1988, 1989)	
0001AC	Data denial OSEs		3.4
	NOCON OSE NOCOR OSE	Louis et al. (1987, 1988, 1989)	
0001AD	Physical satellite retrievals		3.6, App. A
	RH retrievals from TOVS development RH retrievals from TOVS evaluation	Isaacs (1988)	
0001AE	Use of cloud data		3.7
	RH retrievals from 3dNEPH development RH retrievals from 3dNEPH evaluation NEPHSAT OSE	Nehrkorn and Hoffman (1988a, 1988b)	
0001AF	Physical parameterizations		3.8
	GSM evaluation Model tuning CSMSAT OSE Including diabatic effects in initialization	Knowlton et al. (1989)	App. B App. D App. C

2. THE AFGL GLOBAL DATA ASSIMILATION SYSTEM

2.1 Description

The series of 4D data assimilation experiments described here make use of existing computer codes and data bases at AFGL. The assimilation cycle consists of a sequence of three major steps: analysis, normal mode initialization, and forward integration of the initialized state to the next analysis time. Some pre-/post-processing of the data may occur between the major steps. An entire assimilation run is a repetition of this sequence until an initialized analysis is obtained at the ending time.

Optimal interpolation (OI) provides the mechanism for obtaining regularly gridded initial conditions, i.e., analyses, from incomplete, irregularly spaced data. The AFGL Statistical Analysis Program (ASAP), is actually composed of four individual programs. The ASAP OI was developed by SASC for AFGL/LYP and is documented by Norquist (1986, 1988). The ASAP OI is based on the NMC assimilation system, as reported in the literature and in personal communications; this system was thoroughly redesigned and recoded. According to Norquist (1986), ASAP was developed originally from the 1979 multivariate OI procedure as described by Bergman (1979) and by McPherson et al. (1979), with some assistance from NMC personnel (Morone, 1983, private communication). The ASAP OI is a multivariate analysis of height and wind components and a univariate analysis of relative humidity. The equations for the weights assigned to the data, as well as the computation of the horizontal and vertical correlation functions, follow Bergman (1979). The analysis error evolves according to simple rules (Norquist, 1986). The great circle distance method for correlation functions equatorward of 70° latitude is included as described by Dey and Morone (1985) without changing the Bergman formulation (including map factor) for latitudes poleward of 70° latitude. An important characteristic of the ASAP OI is that it is done on the sigma vertical coordinate system of the forecast model.

Data used by the height-wind analysis include Type 1 observations (radiosondes, pibals, etc.), Type 2 observations (aircraft), Type 4 observations (satellite-retrieved temperatures or thicknesses) and Type 6

observations (cloud drift winds (CDWs)). The Type 3 surface observations are not used at all. Earlier versions of ASAP included the use of surface reports in a preliminary surface pressure analysis. However, it was found that the surface pressure analysis was noisy and tended to destabilize the entire assimilation procedure. Consequently, it was decided to eliminate the surface pressure analysis. In retrospect, it might have been better to retain a surface pressure analysis, but only to anchor the satellite height retrievals. Alternatively, the height-wind analysis could be modified to include surface observations.

The normal mode initialization (NMI) is a part of the forecasting system which adjusts the initial data in such a way that undesirable gravity waves do not grow in the forecast. The AFGL NMI was developed by SASC, based on the NMC NMI described by Ballish (1980). The AFGL version, much of which is taken intact from the NMC codes, is discussed by Norquist (1982a) and Tung (1983). The NMI program is actually two programs, one which calculates the normal modes (i.e., eigensolutions of the discrete linearized model equations) and one which performs two iterations of the Machenhauer initialization procedure which sets the time tendencies of the gravity modes approximately to zero. We altered the Machenhauer initialization procedure to make use of the tendencies calculated by the AFGL GSM (see Section 2.2). In addition, we collaborated with L. Knowlton and D. Norquist (Knowlton et al., 1989) to develop and evaluate the Sugi (1986) dynamic normal mode initialization scheme. This scheme should be more stable than the usual Machenhauer iteration when dynamical processes are included in the initialization and it eliminates the need to store the normal modes. Unfortunately, this work did not progress far enough to be applied in any of the experiments described here.

The AFGL global spectral model (GSM) is based on the NMC GSM designed by Sela (1980). For the version used here (except in the GSM experiment), the physics routines are taken almost intact from NMC (circa 1983). The hydrodynamics, i.e., the nonadiabatic, inviscid dynamics including vertical and horizontal advection, time stepping, and transformations between spectral and physical space, were completely redesigned, as documented by Brenner et al. (1982, 1984).

There are a number of parameters in the assimilati n cycle codes that can be adjusted. Typically, the parameters have the same values as used by Brenner et al. (1984) and Halberstam et al. (1984). Briefly, the spectral resolution of the forecast model is rhomboidal 30. The Gaussian grid of the forecast model (analysis) contains 76 x 96 (62 x 61) latitude longitude points. There are 12 layers, the first (top) 5 of which have no moisture. Except for the GSM experiment, the sigma interfaces are at 0.00, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.375, 0.50, 0.65, 0.80, 0.925, and 1.00, the time scheme used is centered semi-implicit with a time step of 17.25 minutes and a Robert time filter is applied with a constant of 0.04. A spectral diffusion coefficient of $6 \times 10^{15} \ m^{-4} s^{-1}$ is applied during the forecasts to all prognostic variables except for moisture. In the NMI, two Machenhauer iterations are applied to modes for the four largest equivalent depths which have periods less than or equal to 48 hours.

·2.2 Modifications to the ASAP

During the course of this work, modifications were made to several aspects of the ASAP system, in the preprocessing of the data, the optimal interpolation itself, and in the normal mode initialization. These modifications are summarized here.

In all experiments, except for the three-day preliminary assimilations, the CDW data (Type 6 data) were combined (i.e., locally averaged) into super-obs. There are two principal reasons for doing this: First, to limit the total number of observations, so that computer memory restrictions are not exceeded, and second, the CDW errors are strongly correlated horizontally because the main error source is due to height assignment (McPherson, 1984). Because of new quality control codes inserted by the NASA/NMC Special Effort the quality control translation tables for Type 6 data in ASAP were altered. For some of the experiments, the satellite height observational error statistics were altered, on the basis of a colocation study (Table 2).

Table 2

Observational error standard deviations (m) used for satellite and radiosonde heights. Note that we use different values for February and June. The STATSAT and NOSAT experiments were already underway before the most up-to-date values were available. For comparison, satellite height observational standard deviations equivalent to the temperature observational error statistics used by NMC in 1979 are included. The radiosonde errors are from Dey and Morone (1983) and the satellite errors are from Norquist (1986; pers. comm., 1987). The NMC-1979 satellite errors are calculated hydrostatically from the standard deviations of temperature error from McPherson et al. (1979), the vertical correlation of temperature error from Bergman (1979) and an assumed height error at 1000 mb uncorrelated with the temperature errors.

Pressure (mb)	Radiosonde	February NOSAT/STATSAT	February NOCON/NOCOR	June NOSAT/STATSAT	June NOCON/NOCOR	NMC (1979)
1000	7.0	9.75	9.75	9.75	9.75	9.75
850	8.0	23.67	14.94	25.39	12.53	12.45
700	8.6	35.19	19.56	32.92	20.87	18.66
500	12.1	49.27	27.88	31.12	26.76	28.48
400	14.9	70.07	34.88	53.61	32.74	33.51
300	18.8	97.87	43.51	68.66	40.32	39.40
250	25.4	95.09	46.59	65.47	43.60	42.14
200	27.7	98.32	47.40	57.34	44.66	46.46
150	32.4	90.82	43.80	48.44	46.51	53.37
100	39.4	75.95	56.60	54.28	54.28	58.77
70	50.3	61.32	61.30	53.61	53.61	61.76
50	59.3	67.91	67.91	64.75	64.75	64.85

Another change to the preprocessing of the data is a new, simpler and more efficient method of converting the pressure layer temperatures given by the satellite data into sigma level heights. The new method was tested on a series of radiosonde profiles, and was found to be equally or more accurate than the old method, especially in the stratosphere.

This new interpolation method, as well as the changes made to the optimal interpolation procedure are described in Hoffman et al. (1988). This report also completes the ASAP documentation that can be found in Hoffman (1987). The changes to the OI are a corrected and improved handling of the least squares equations and a new solver for these systems of equations based on standard LINPACK routines. Three inconsistencies in specifying the normal equations were corrected. First of all, we now force a strict decoupling of the wind and height analyses in the equatorial belt by setting all uz, vz, zu and zv correlations to zero in the observation-observation matrix and the observation-gridpoint right hand sides. Secondly the calculation of the right hand sides has been corrected. Finally we now do a complete (z, u, v) "underground" analysis to insure best possible agreement between surface and first level temperatures. The new procedure allows a similar data selection for the underground and first atmospheric layer.

In the normal equations solver, we have implemented improved checks for singularity or non-positive definite systems. Previously, some of the near singular cases were retained, resulting in erroneous weights for the data. The new criterion for singularity is based on the matrix condition number.

To test these changes to the OI, as well as the new derivation of heights from satellite data, we performed a simulation experiment similar to the ones described in Section 4, using the FGGE observing system (STATSAT experiment). Global rms statistics of the test results showed that the modifications improve the wind analyses slightly at all levels, while the height analyses have about the same level of error, except in the stratosphere, where the modified system is noticeably better.

In the Machenhauer procedure for the normal mode initialization, it is desirable to use the tendencies computed by the GSM. To accomplish this we altered the GSM to write out tendencies. We then developed a program to add in the linear tendency terms, not explicitly calculated by the GSM and perform a single Machenhauer projection. The GSM tendency and Machenhauer projection calculations are repeated twice to obtain a complete initialization. The NMC Machenhauer projection codes were used with no alterations. The procedure was tested by temporarily making small changes to the AFGL GSM so that its physical constants and parameterizations would agree with the NMC tendency calculation used in the previous NMI. Extremely good agreement was obtained.

3. OBSERVING SYSTEM EXPERIMENTS

3.1 Overview

Observing systems experiments (OSEs) have often been conducted to quantify the impact of various satellite observing systems (Isaacs et al., 1986a). Recent OSEs include studies of the impact of aircraft wind data (Barwell and Lorenc, 1985), tropical wind data (Paegle et al., 1986), Monex wind data (Ramamurthy and Carr, 1987), VAS temperature data (Aune et al., 1987), and VAS moisture data (Mostek et al., 1988). At times the results of such impact tests have been debated. Now there is a general agreement on the qualitative impact of satellite temperature soundings on numerical weather prediction. Gilchrist's (1982) general conclusions, which are still valid, are that satellite sounding data (1) are vital to defining the large-scale structure of the atmosphere, and (2) have a positive impact on forecast skill albeit not a very large one by conventional measures, i.e., skill scores for continental areas.

We have reexamined the impact of satellite temperature profile data and examined the impact of physical versus statistical retrieval methods and the impact of denying all conventional data either globally or in a region by conducting OSEs using the Air Force Geophysics Laboratory (AFGL) global data assimilation system (GDAS). Table 3 summarizes these experiments and the data sources they used. The control experiment (denoted STATSAT) potentially makes use of all observations in the 1979 Global

Table 3. OSE data sources summary

OSSE	Ground-based	CDW	TOVS	Remarks
STATSAT	X	X	X	Full FGGE system
NOSAT	X			No satellite data
GLASAT	X	X	X	GLA TOVS retrievals
NOCON		X	X	Data denial
NOCOR	X	X	X	Regional data denial
NEPHSAT	X	X	X	Cloud derived RH
GSMSAT	X	X	X	New physics

Weather Experiment (GWE) Level II-b data set. (In fact, the AFGL GDAS makes no use of surface observations.) Experiment NOSAT, in which all satellite data are withheld, measures the impact of satellite data. These first two experiments serve as a quantitative baseline for succeeding experiments reported here and provide a measure for calibrating the OSSE results. Additionally, comparing the results of these experiments with those of previous studies serves to validate the AFGL GDAS. In particular, although there are differences between the Level II-b data and the Level II-a data which were available to the operational centers, the results of STATSAT may be directly compared to the operational National Meteorological Center (NMC) archived analyses. Some comparisons with the III-b analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF) have also been made.

Experiment GLASAT deals with the impact of a physical instead of a statistical method of retrieving temperature profiles from satellite observed radiances. Physical retrievals are expected to outperform statistical retrievals in extreme situations, which are under-represented in the data base used to develop the statistical retrievals. On the other hand physical retrievals require accurate modeling of radiative processes in the atmosphere. In controlled tests using simulated HIRS and AMTS data the physical approach was found to be better than the statistical approach (Phillips et al., 1988). The improvement was small for the HIRS

instrument, which produced the data used in our experiments. Our experiments determine the extent to which this improvement carries over to analyses and forecasts when real rather than simulated data are used. A similar experiment has been recently conducted by Dey et al. (1988).

The physical retrievals used in GLASAT were prepared at the NASA Goddard Space Flight Center Laboratory for Atmospheres (GLA) using the methods described by Susskind et al. (1984) and Susskind and Reuter (1985). These retrievals are not the most current GLA retrievals. In particular, they do not include a moisture retrieval and they are not interactive, that is they were not produced within the cycling of the GLA data assimilation system (Reuter et al., 1988). The newer interactive retrievals would not be appropriate for testing the AFGL system, however, since the background fields used are produced by the GLA system.

Physical retrieval of moisture data was attempted, using an algorithm developed at AER. However, a study of the retrieval error statistics did not show a sufficient improvement to justify running a complete experiment.

Another source of moisture data might be found in imagery satellite data since the presence of clouds indicates, to the very least, high relative humidity. Using the US Air Force 3DNEPH cloud analysis data as a substitute for satellite imagery, we developed statistical relationships between the 3DNEPH data and relative humidity profiles. An experiment (NEPHSAT) tested this idea. Even though the results show only a marginal impact, the slight improvement seen in the moisture analysis warrants further effort in this direction.

Our next experiments, denoted NOCON and NOCOR, examine the effect of denying conventional data either globally or in a continent-sized region, specifically Europe and the USSR. Previous data denial OSEs, in attempting to determine if satellite observations might be a cost effective replacement for the radiosonde network, have generally retained the surface observations. In the data denial experiments reported here all conventional data, including surface observations, are excluded. The results of NOCON therefore give an indication as to what extent a purely space-based system can satisfy the requirements of numerical weather prediction (NWP). The results of NOCOR may be used to quantify the effects on analysis and fore-

cast skill of large voids in the conventional data coverage. In a qualitative sense these effects are seen through comparisons of Northern and Southern Hemisphere analyses and forecasts for similar seasons. By comparing NOCOR to STATSAT we can quantify these effects since the synoptic situation and data sample are held fixed.

The final observing system experiment was performed with the new version of the AFGL GSM, including the new physical parameterization package. The data used were the same as in the STATSAT experiment. This experiment, named GSMSAT, quantifies the effect of modifying the model used to provide the first guess in each assimilation period.

Each experiment consists of assimilation runs for two 7-day periods, one each during the GWE Special Observing Periods (SOPs): February 8 through 15, 1979, and June 17 through 24, 1979. The SOP I and SOP II periods are treated in the same fashion. Each assimilation run consists of a series of assimilation cycles, and each cycle in turn is made up of a 6-hour forecast that serves as a first guess for the analysis, an optimum interpolation analysis which combines the first guess fields with the observations, and a nonlinear normal mode initialization of the analysis. The initialized analysis is the starting point for the next 6-hour forecast, which is then used as the first guess of the subsequent assimilation cycle. The forecast model used for the 6-hour forecast is a complete global spectral model (GSM). This model is also used to produce forecasts out to 4 days starting from days 3, 5, and 7 of the assimilation runs.

To provide the initial conditions for the first assimilation cycle of each assimilation experiment, we started from an initialized NMC analysis 3 days prior to the beginning of the study period. We then performed a 3-day spin-up assimilation run for each SOP period using the full GWE data base. This allows the initial fields to be consistent with our assimilation system.

3.2 Baseline experiment: STATSAT

The first experiment uses all the available upper air data in the FGGE dataset. These included satellite data retrieved by a statistical method. Therefore we named this experiment STATSAT. Since the data used are

for the most part identical, the differences that can be observed between the STATSAT analyses and the NMC or ECMWF analyses are mainly due to differences in the analysis/forecasting systems, within which we include the quality control and data selection functions. We mainly compared the STATSAT results with the NMC uninitialized analyses.

The evolution of differences during the assimilation cycle and forecasts is displayed in Fig. 1. The dot-dashed line shows the evolution of the root mean square (rms) difference between the STATSAT and the NMC 500 mb height analyses. The solid lines are the rms differences between the STATSAT forecasts and the NMC analyses, which we call the forecast errors, also at 500 mb. The first and last forecasts of the series are shown. Finally the dotted line is the rms difference between the STATSAT forecast and the corresponding STATSAT analysis. Each hemisphere and season is plotted separately. The analysis differences are larger in the Southern Hemisphere than in the Northern Hemisphere, and larger in the winter than in the summer. The average difference between the analyses is about 40 to 50 m. This is larger than the differences found by Lambert (1988) between the NMC and ECMWF analyses for the 1980 to 1984 period: 24.5 m in January and 37.5 m in July at 500 mb. It should be noted, however, that both the NMC and ECMWF analysis systems are operational and subject to constant improvement and tuning, that the ECMWF analyses used in Lambert's study were initialized and thus presumably less noisy, and that one week of assimilation in each season, in our study, may not be entirely representative of the true difference.

The forecast errors grow somewhat more slowly in the summer than in the winter. After 48 hours, in the Northern Hemisphere, the rms forecast errors become nearly independent of the reference analysis. In the Southern Hemisphere, by contrast, the STATSAT forecasts remain closer to the STATSAT analyses than to the NMC analyses throughout the length of the forecasts. This shows the influence of the first guess on the analysis in the Southern Hemisphere. The 96 hour 500 mb and the 72 hour 1000 mb forecasts would generally be considered poor forecasts.
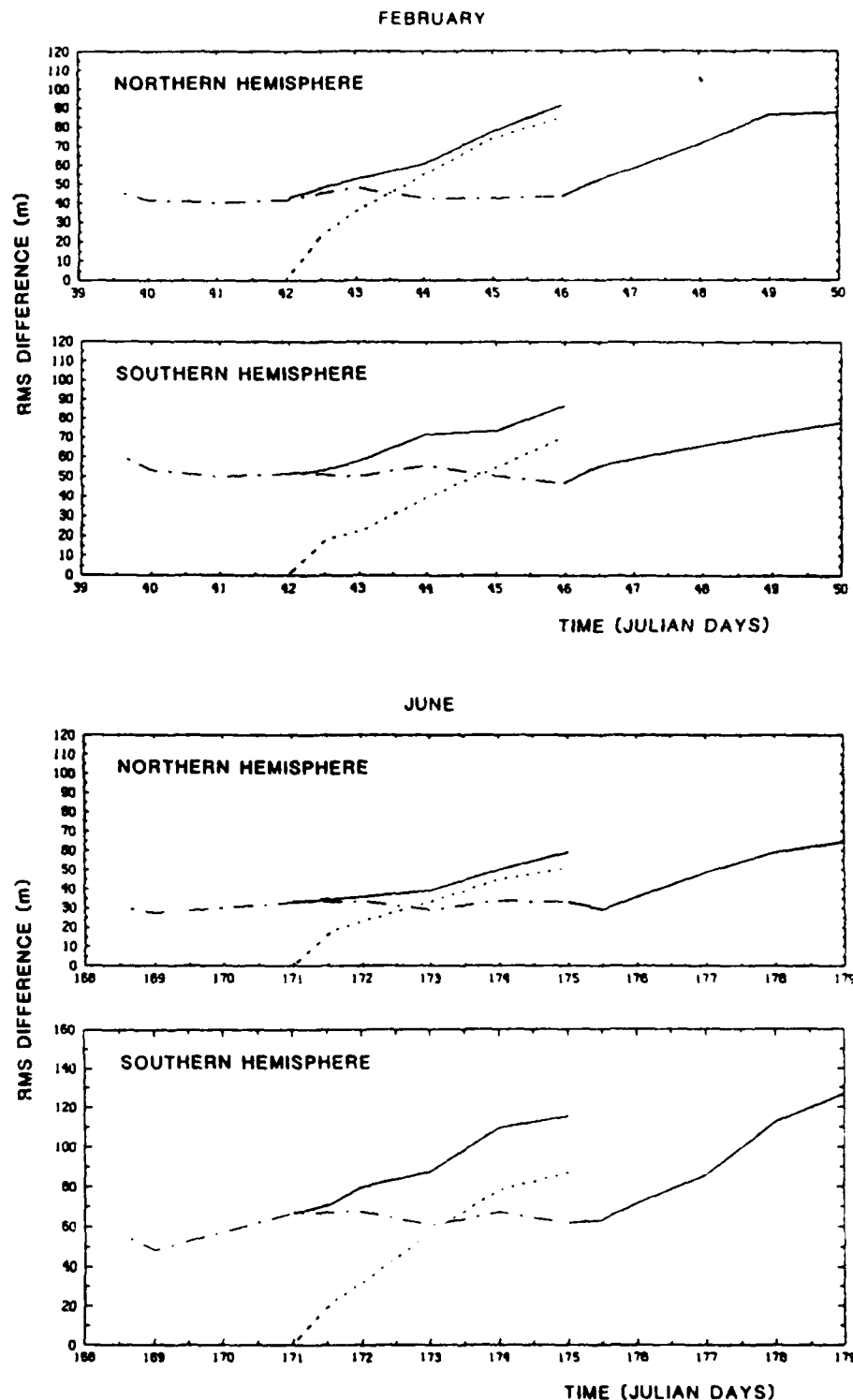
Fig. 1. *Time evolution of rms differences between the 500 mb height analyses and forecasts. The comparison shown are STATSAT analyses - NMC analyses (dot-dashed lines), STATSAT forecasts - NMC analyses (solid lines) and STATSAT forecasts - STATSAT analyses (dotted lines). Each hemisphere and season is plotted separately.*

We have also compared the STATSAT analyses to the radiosonde data. This comparison (not shown) indicates a global rms difference of 15 to 20 m at 500 mb, about half the rms difference between the STATSAT and NMC analyses.

The largest differences between the STATSAT and NMC analyses occur over Antarctica. The June STATSAT analyses are systematically higher than the NMC analyses, by up to 400 m at 500 mb. This means the polar vortex is less intense in the STATSAT analyses than in the NMC ones. In addition, large differences in the phase and amplitude of the small waves make the two Southern Hemisphere analyses noticeably different. Lambert (1988) and Trenberth and Olson (1988) have also found large differences between the NMC and ECMWF analyses over Antarctica, implying that no analysis can be considered reliable in that region.

Other fairly large differences (up to 150 m) are associated with intense oceanic systems, such as the Aleutian trough in February, where STATSAT tends to spread the baroclinic zone further South. Finally some differences that can be quite significant from the synoptic point of view exist at 1000 mb, especially over North America and Europe. In these cases the NMC analyses are generally confirmed by the ECMWF analyses. One can suspect that the AFGL system is deficient in this respect because it does not use surface data.

Considering that the AFGL GDAS is an experimental system, and has not had the benefit of the kind of optimization that comes from an operational usage, the differences that have been observed between the STATSAT and NMC analyses appear reasonable. Some improvements could be hoped for. In particular, not using surface data at all seems to be detrimental. Nevertheless, the AFGL GDAS is a properly formulated assimilation system, which can provide meaningful results in the observing system experiments that have been performed and are reported here.

3.3 No satellite data experiment: NOSAT

In the NOSAT assimilation experiment all satellite data are removed from the dataset, and only conventional data are used. Thus differences between the NOSAT and STATSAT analyses occur mainly over the oceans
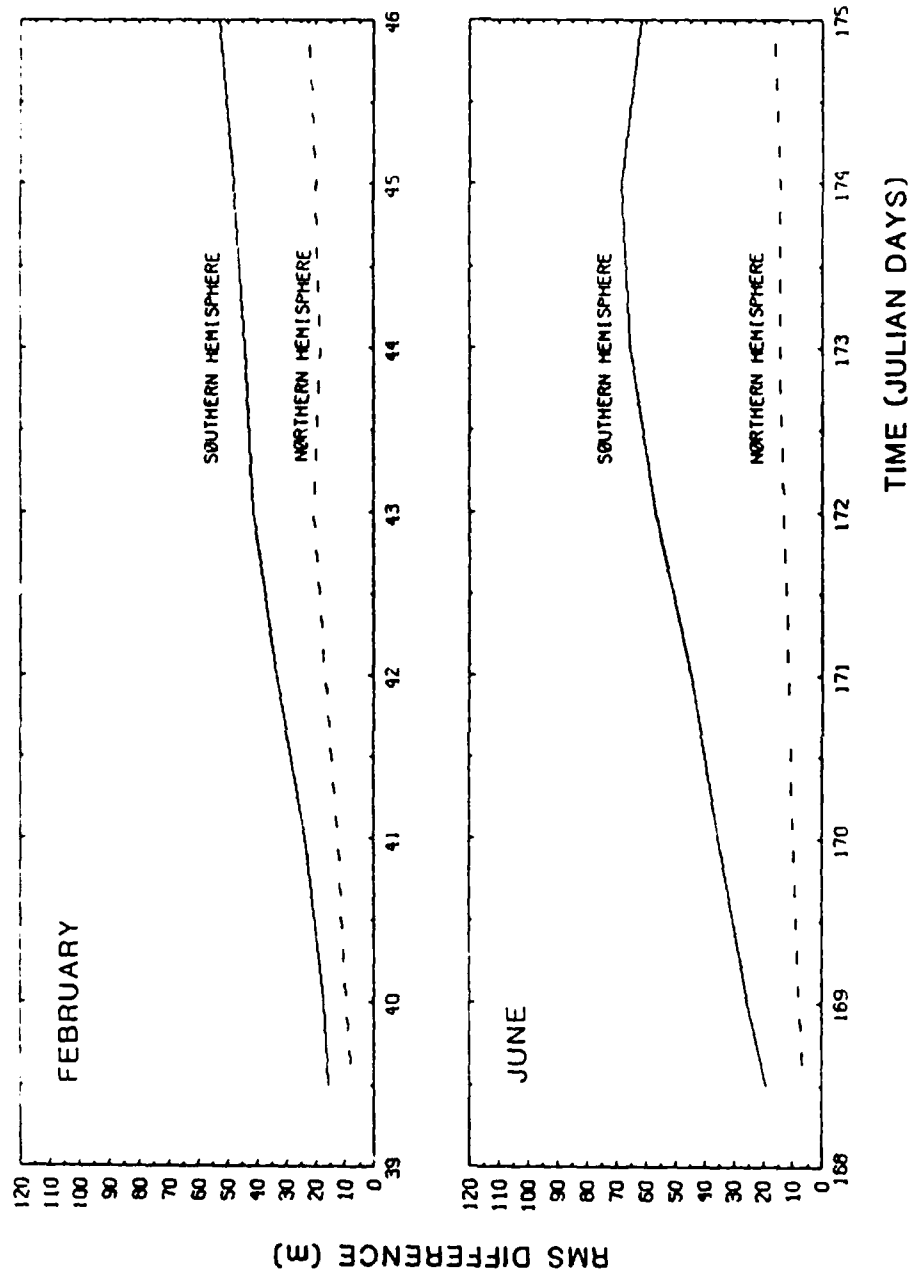
Fig. 2. Time evolution of rms differences between the NOSAT and STATSAT 500 mb height analyses, for each hemisphere and season.

and the Southern Hemisphere. When compared to the NMC analyses, the NOSAT fields exhibit similar kinds of differences as the STATSAT fields do, but usually somewhat larger. For example, there is the same tendency to spread the baroclinic zone farther south in the Pacific trough.

As can be seen in Fig. 2, the differences between the NOSAT and STATSAT analyses grow during the period. In the Southern Hemisphere some differences already exceed 100 m after the first day of assimilation, and soon reach more than 300 m. After June 20 the two Southern Hemisphere analyses have little in common.

This is especially true southwest of Australia, where on 22 June the NOSAT analysis shows a ridge where both the STATSAT and NMC analyses have a trough. Most of the differences between the two analyses, however, occur in the regions where there are no radiosonde stations (Fig. 3).

Toward the end of the assimilation period, some rather large differences begin to appear, in the Southern Hemisphere, even in regions where radiosonde data are available. An examination of the data rejection files indicated that, at the two stations where the differences are largest, observations were included in the STATSAT analyses but were rejected by the NOSAT analyses because they were too far from the first guess. This suggests that, at least in the Southern Hemisphere, satellite data are beneficial even where conventional data exist.

The different behavior in the two hemispheres is also apparent in the forecasts. While, in the Northern Hemisphere, the NOSAT forecasts are quite similar to the STATSAT forecasts, in the Southern Hemisphere, the NOSAT forecasts quickly differ from the STATSAT forecasts. As early as 72 hours into the forecasts, the Southern Hemispheric fields are so different from the NMC STATSAT analyses that they can be considered useless. Compared to the NOSAT analyses, however, these same forecasts remain quite similar to the analyses, sometimes up to 96 hours, indicating a strong influence of the first guess on the NOSAT analyses (Fig. 4). This also shows that one needs to be very careful how one defines forecast "errors".
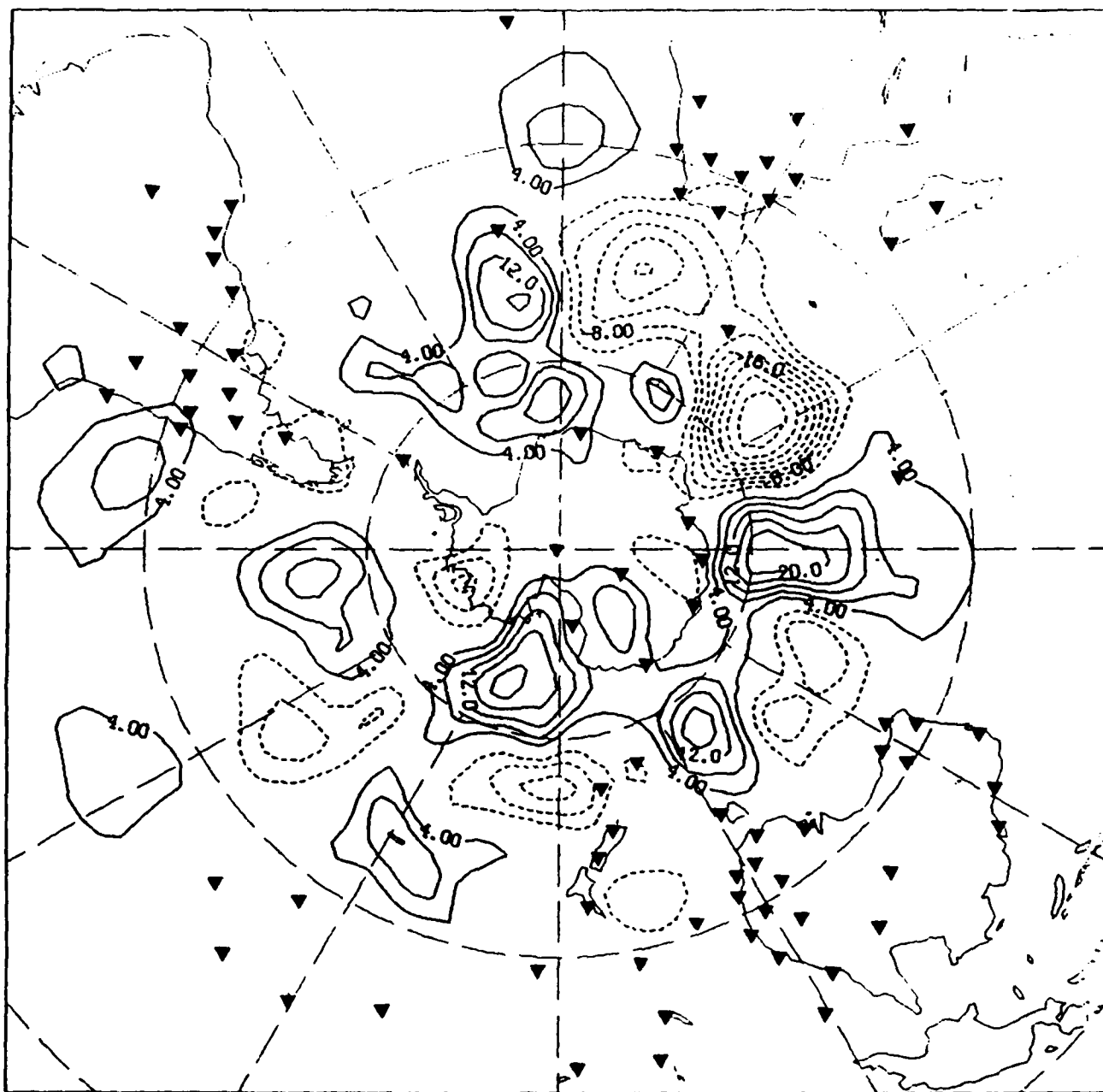
Fig. 3. The difference between the NOSAT and STATSAT 500 mb height analy-
ses (dm) for the Southern Hemisphere at 00 GMT on 21 June 1979. The
radiosonde stations are indicated as triangles. In this and all sub-
sequent difference maps the zero line has been suppressed for clarity.
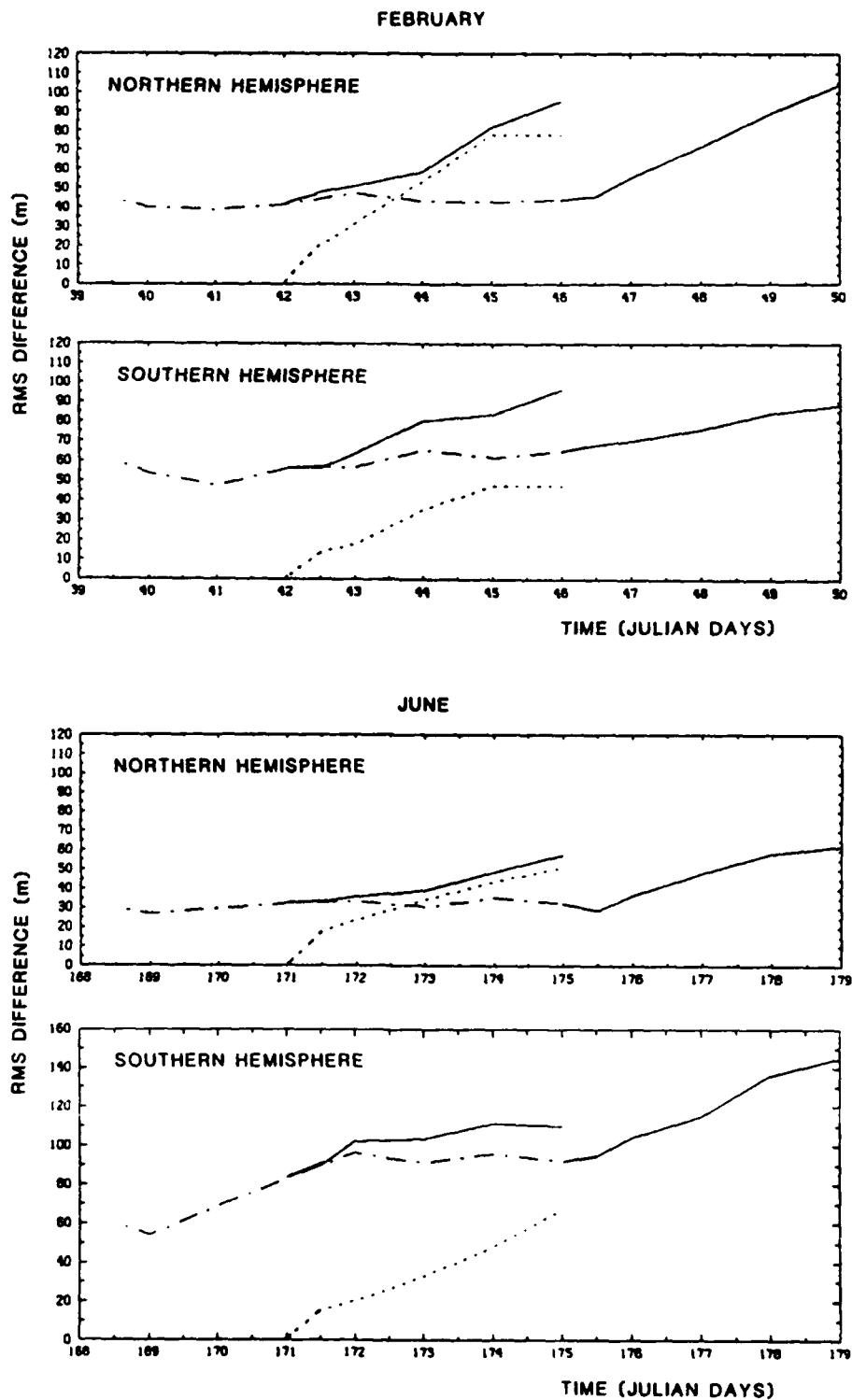
Fig. 4. As Fig. 1, but for the NOSAT analyses and forecasts.

3.4 Space based system experiment: NOCON and NOCOR

Two experiments were designed to study the effect of a completely space-based observing system. The first of these was the extreme case, in which all the conventional data are ignored in the FGGE dataset. Only satellite data are used, but these are used over the continents as well as over the oceans, in contrast to the STATSAT experiments in which satellite data are used only over the oceans. This was the NOCON experiment. In the second data denial experiment, NOCOR, conventional data were ignored only over a region of the globe, specifically Europe and USSR.

Since satellite sounders measure temperature, the absolute height of at least one level is needed as boundary condition to integrate the hydro-static equation and derive the height of all the other levels. In other data denial experiments, this reference height was improved by an analysis of surface data (e. g. Halem et al., 1982). Only the upper air data were ignored. We have noted before that, in the AFGL GDAS, the surface observa-tions are not used. Hence, in these experiments, satellite "heights" are anchored only by the 6 h forecast in regions where radiosondes are absent or denied. The errors in the reference level's height affect all the other heights and it can be expected that the whole analysis will eventually be corrupted. In addition, the moisture analysis used in the two experiments reported here uses only Type 1 data. Consequently, during the NOCON ex-periment there was no moisture analysis at all: the moisture fields com-puted by the forecast model are simply carried along.

Details of the results of the NOCON and NOCOR experiments can be found in Louis et al. (1987). They can be summarized by Fig. 5, which shows the 1000 to 500 mb thickness rms differences to the radiosonde observations, for the three experiments STATSAT, NOCON and NOCOR during the February ex-periment. As expected the NOCON and NOCOR analysis errors keep growing during the assimilation experiment period, and the NOCON errors grow faster than the NOCOR errors. In fact, the maps of the differences between NOCON or NOCOR and STATSAT (Fig. 6) show that even in the region of the NOCOR data denial, the differences are smaller for NOCOR than for NOCON, indi-cating that the model is able to carry better information into that re-gion. Nevertheless, the NOCOR forecast errors grow faster than the STATSAT forecast errors and, after about 2 days, reach the level of the NOCON fore-cast errors.
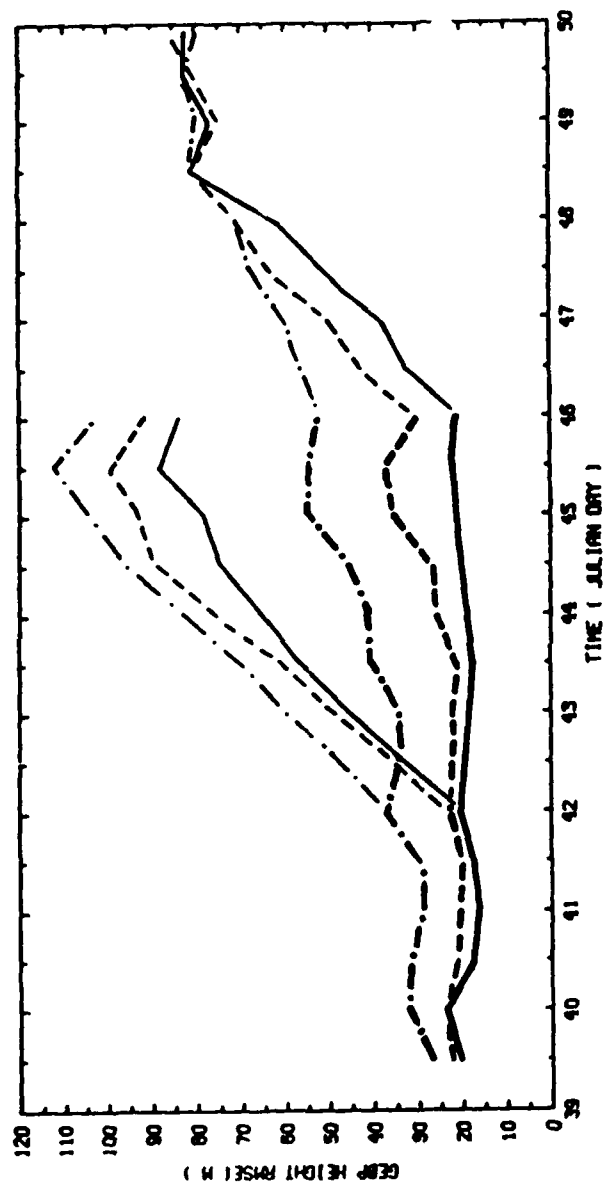
Fig. 5. Rms 1000 to 500 mb thickness differences between analyses (thick lines) or forecasts (thin) and radiosonde observations, for the February experiments. STATSAT: solid lines; NOCOR: dashed; NOCON: dash-dots.

-21-
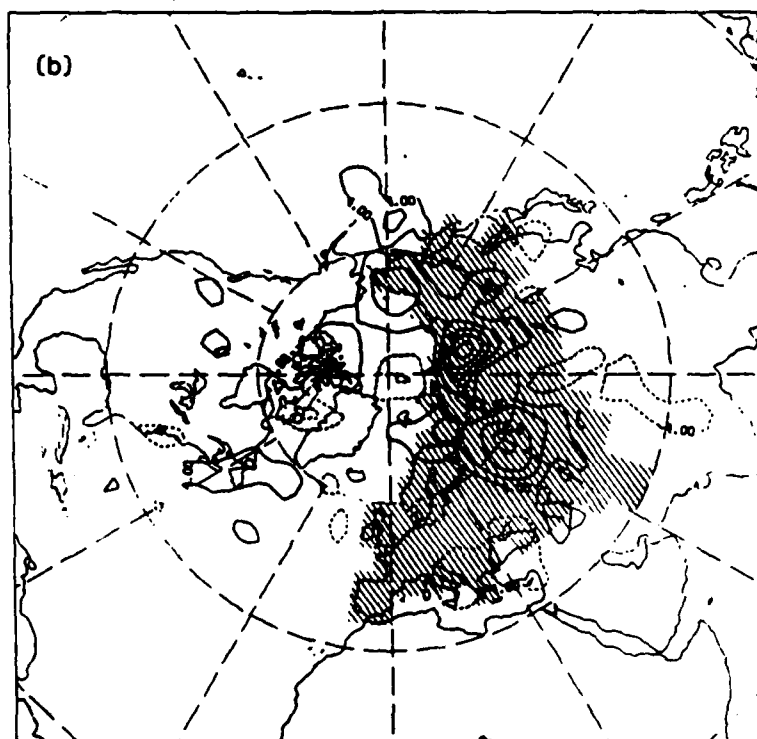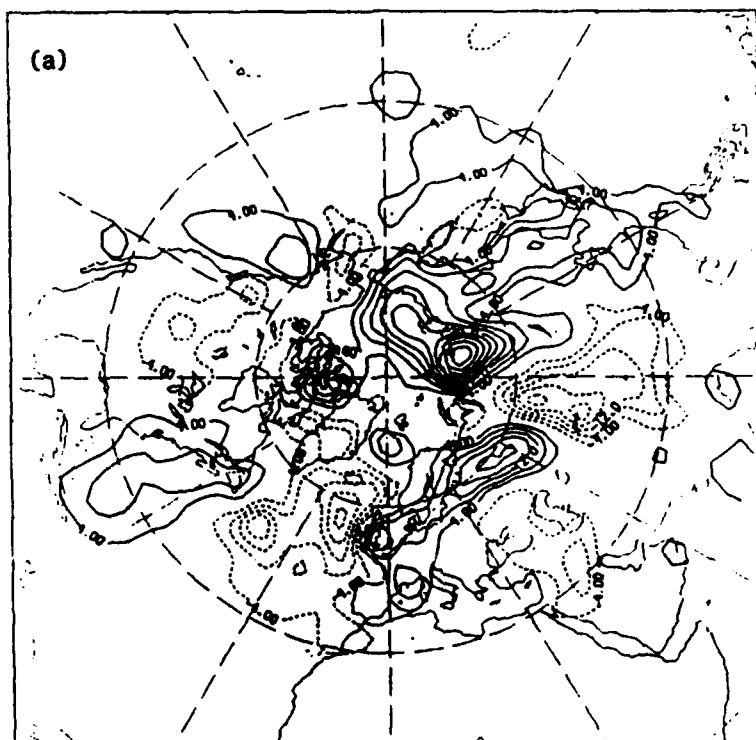
Fig 6. 500 mb height differences (dm) at 00 GMT, on 15 February 1979.
(a) NOCON - STATSAT analyses; (b) NOCOR - STATSAT analyses. The
shaded area shows where conventional data were denied in the NOCOR
analysis.

3.5 Physical temperature retrieval experiment: GLASAT

The effect of a different temperature retrieval method was tested by using the profiles provided by GLA, which are produced by a physical retrieval method (Susskind et al., 1984). All the other data were the same as in the baseline (STATSAT) experiment. The differences between the GLASAT analyses and the STATSAT analyses are generally fairly small. At the beginning of the assimilation experiment (not shown) the differences are clearly confined to the swath covered by the satellite during the 6 hour cycle. In the Northern Hemisphere winter, and in both seasons in the Southern Hemisphere, these differences start spreading after a few days. This is evidently due to the influence of the first guess, which advects the differences from the previous 6 hour period. This is particularly clear in June in the Antarctic Ocean, where different shapes and phases of the waves are apparent. On June 23 the maximum difference is 240 m at 500 mb (Fig. 7).

Some systematic behaviors should be pointed out. The differences are always smaller at 1000 mb than at 500 mb, although the patterns are fairly similar. In the Northern Hemisphere summer, the GLASAT height analysis tends to be lower than the STATSAT analysis. This tendency can also be seen at the beginning of the experiment in the Southern Hemisphere summer, but it is soon overwhelmed by the larger phase differences of the medium waves. In the winter hemisphere this systematic difference is not apparent, except in the tropics where, again, the GLASAT analysis is lower.

It is difficult to determine whether one analysis is better than the other. We have compared the forecasts made from the two analyses to the radiosonde data. This seems to us to be the most objective way to determine the quality of the analyses. Comparing the analyses themselves to the radiosondes would not be so good because the analyses are expected to differ most from each other far away from the radiosonde stations. Based on forecast-radiosonde differences there is not much to choose between the two systems. The rms differences are practically identical, both for the whole globe and for separate hemispheres.
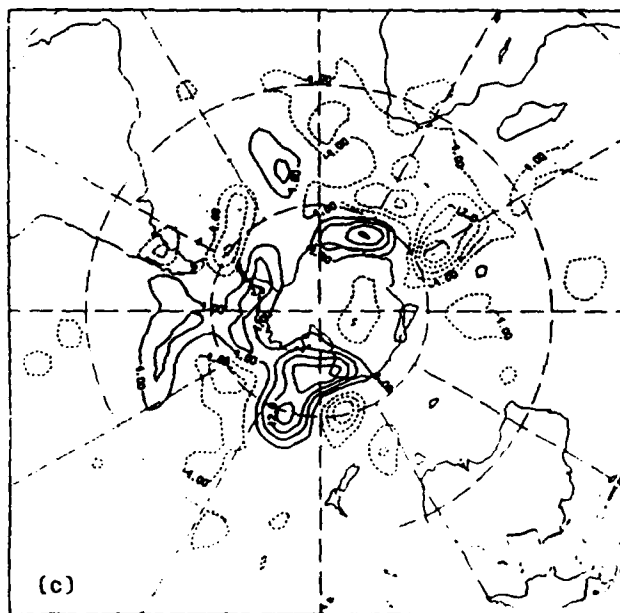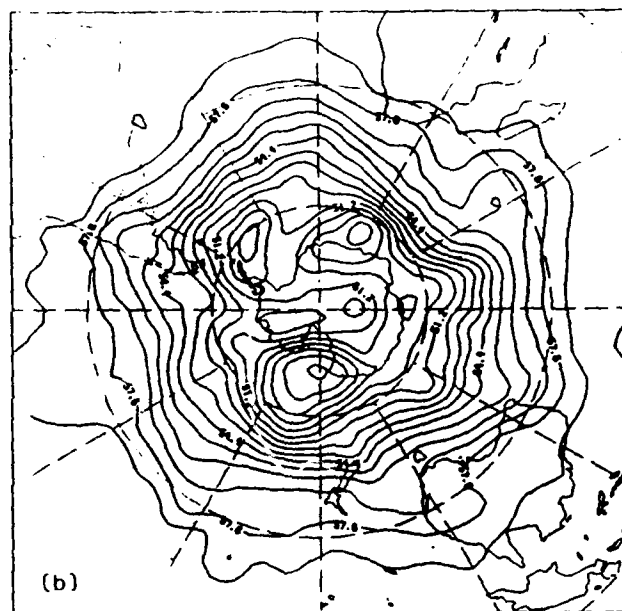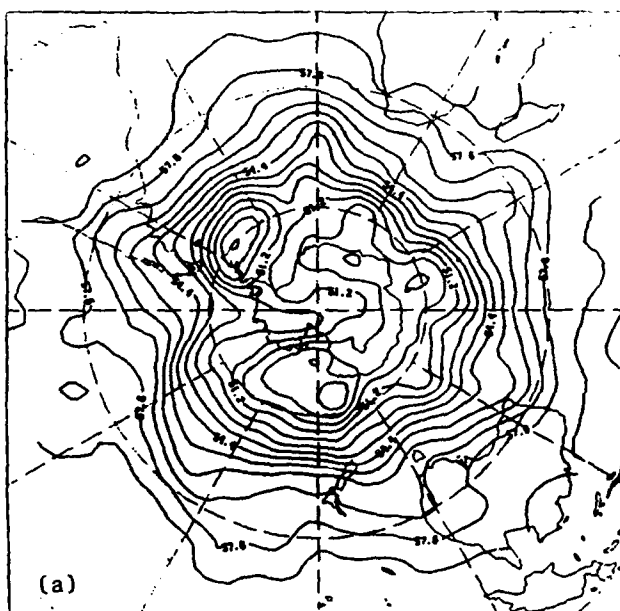
Fig. 7. 500 mb height (dm) on 23 June 1979, at 00GMT. (a) GLASAT
analysis; (b) STATSAT analysis; (c) difference between (a) and (b).

This lack of sensitivity to the retrieval method for HIRS is antici-
pated by the results of the Phillips et al. (1988) comparison. However,
our results are different from results obtained at GLA using higher resolu-
tion in the retrievals and the model (R. Atlas, 1988, personal communica-
tion). In addition GLA used their 6 hour model forecast as the first
guess. Possible advantages of such an interactive system are discussed by
Eyre (1989).

3.6 Physical moisture retrieval

In all the experiments performed so far only conventional obser-
vations of moisture are used. No satellite data is included. The re-
sulting relative humidity analysis is quite poor, with mean squared errors
approaching the climatological variance. An attempt was made to use satel-
lite data to supplement the conventional dataset. A detailed description
of this effort can be found in appendix A. Only a brief summary of the
main points is presented in this section.

The method developed at AER uses a physical retrieval algorithm, fol-
lowing Smith et al. (1986). A rapid transmittance computation tailored to
the HIRS/MSU channels is used for the forward problem, avoiding expensive
line by line computation. The cloud clearing method is based on Eyre
(1989). The temperature profiles needed in the algorithm are the same as
in the GLASAT experiment. Climatological moisture profiles are used as the
first guess in the iterative procedure.

Moisture retrievals were performed for the period 8-15 February
1979. A colocation with radiosonde data was perform to test the accuracy
of the retrievals. The results of this study were disappointing, with re-
trieval errors only slightly smaller than the climate variability. Since
the impact of these data in a data assimilation experiment is likely to be
insignificant, it was decided not to proceed with the planned observing
system experiment.

3.7 Moisture profiles from cloud data: NEPHSAT

Moisture and temperature retrievals, at least from infrared radi-
ances, are difficult in cloudy atmospheres. There is a possibility that

-25-

inference of humidity data from cloudiness information may·successfully supplement moisture retrievals from radiance data. In an experiment called NEPHSAT, we studied the impact of adding humidity profiles inferred from cloud data, using the 3DNEPH data base as the source of the cloudiness information. The cloud to relative humidity specification developed for this study is unique in that the relative humidity profiles are expressed in terms of EOFs and the coefficients of the EOFs are related to the cloud parameters. We note that this impact test differs from the one reported in Norquist (1988), in which bogus RH data was used to replace, rather than supplement, radiosonde measurements.

The 3DNEPH (now RTNEPH) is a high resolution cloud data base produced operationally by the US Air Force Global Weather Central (AFGWC). It is a global gridded data set with a resolution of 47.6 km on a polar stereographic grid. The data at each gridpoint consist of percent cloud cover for total sky cover and for 15 layers in the vertical, as well as several other parameters. In addition, a vertically compacted set of cloud cover values, which correspond to boundary layer clouds and cloud cover for layers surrounding the 6 mandatory levels between 1000 hPa and 300 hPa, was derived from the 15 layer values to reduce the amount of data. These data were used at full horizontal resolution and in the form of horizontal averages over 5x5 3DNEPH gridpoints.

The principle of the method is to derive a regression equation between the cloud data and moisture profiles. We used the empirical orthogonal functions (EOF) of relative humidity (RH) to determine the important features of the observed RH profiles. The EOF coefficients were related to colocated cloud data by means of multivariate linear regression equations. The data used in the regression study was restricted to the North American continent, resulting in a homogeneous sample of high quality radiosonde measurements. Only 00 UTC data was used. The multiple regression equations developed for EOF 1 and 2, based on horizontal averages of the vertically compacted cloud data, proved to be significant.

-26-
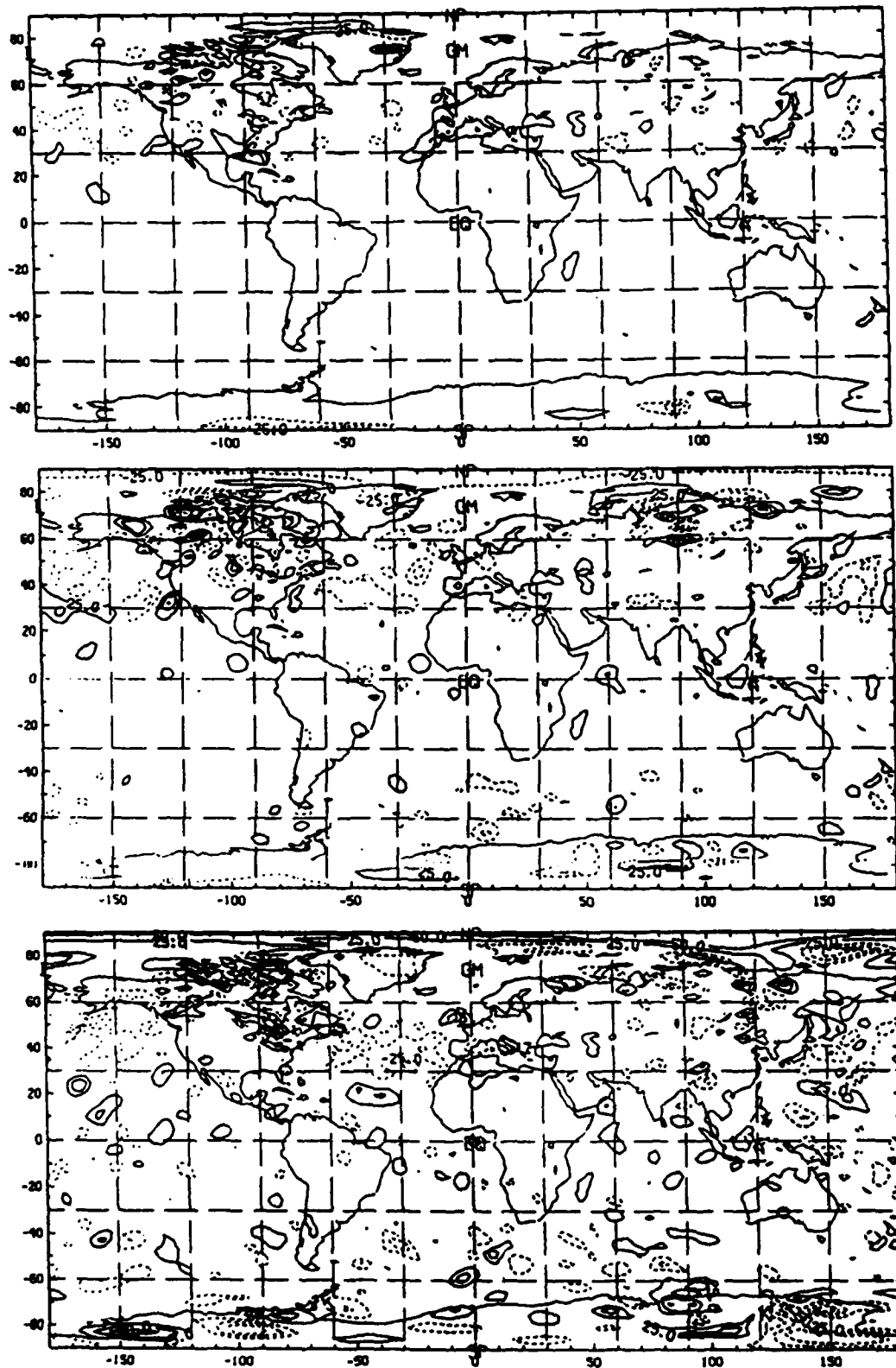
*Fig. 8. Difference between STATSAT and NEPHSAT 850 mb relative humidity
analyses for 8 February 1979 at 00 GMT (top), 9 February at 00 GMT
(middle) and 10 February at 12GMT (bottom).*

-27-

Stratification based on latitude or the use of the full resolution data did not improve the regression coefficients. We compared the results of these equations to existing methods that relate cloudiness to moisture locally, level by level. These methods are summarized in Norquist (1988). The errors of the regression equations are smaller for June than for February, and they are consistently smaller than those of climatology or the existing methods.

The OI analysis program had undergone some minor changes between the STATSAT and NEPHSAT experiments: changes were made to the quality control procedures of drop-windsonde data, and to the procedures to solve the normal equations of the analysis program (see Section 2.2, or Hoffman et al., 1988).

In the assimilation experiment, the 3DNEPH based bogus RH data were generated for all half-mesh points located between 30°N and 50°N   The impact of the 3DNEPH data is most clearly seen in Fig. 8, which shows differences of analyzed RH at 850 hPa between NEPHSAT and STATSAT for the February assimilation run. At the beginning (February 8 at 06Z) differences are essentially confined to the region where the 3DNEPH data were used, i.e. 30°-50°N; an exception to this are the high latitude regions of both hemispheres, where sizable differences occur, due most probably to differences in the OI program between the two experiments. The succeeding panels in Fig. 8 show that the differences within the 30°-50°N latitude band as well as outside it grow with time. After 2.5 days of assimilation the region where bogus RH data were used in NEPHSAT is no longer visibly different from the rest of the world in these difference maps. At that time the impact of the 3DNEPH data has become hidden by the noise level of the system. Nevertheless, plots of the 700 hPa 3DNEPH cloud cover (Fig. 9), which is the predictor with most influence on bogus RH data for the 850 hPa level, reveal some areas of little cloud cover, particularly over the middle and East Atlantic and North East Pacific, which correspond to areas where the 850 hPa NEPHSAT analyses are drier than the control.

The analyses and forecasts have also been compared to radiosonde observations. Fig. 10 shows the global rms error of relative humidity for NEPHSAT and STATSAT analyses and forecasts for February. The NEPHSAT
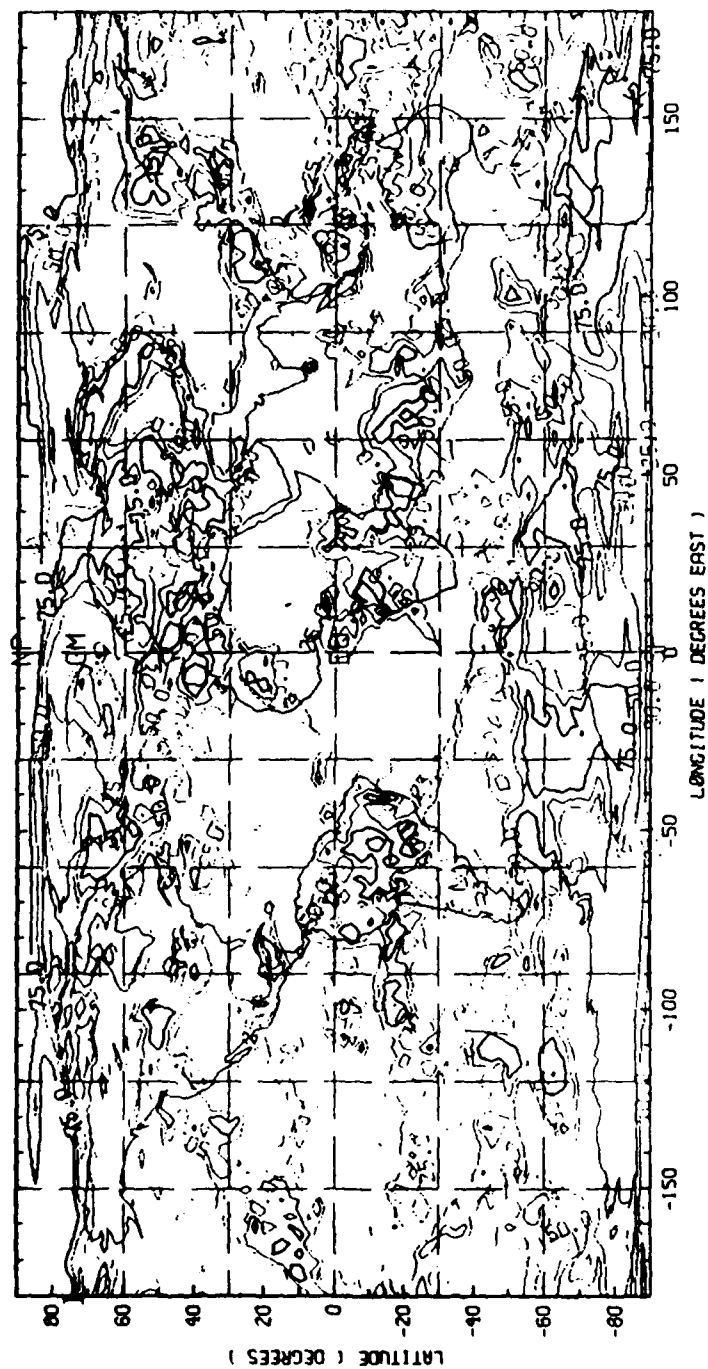
Fig. 9. 700 mb cloud cover from the 3DNEPH data, for 10 February 1979, 00GMT.
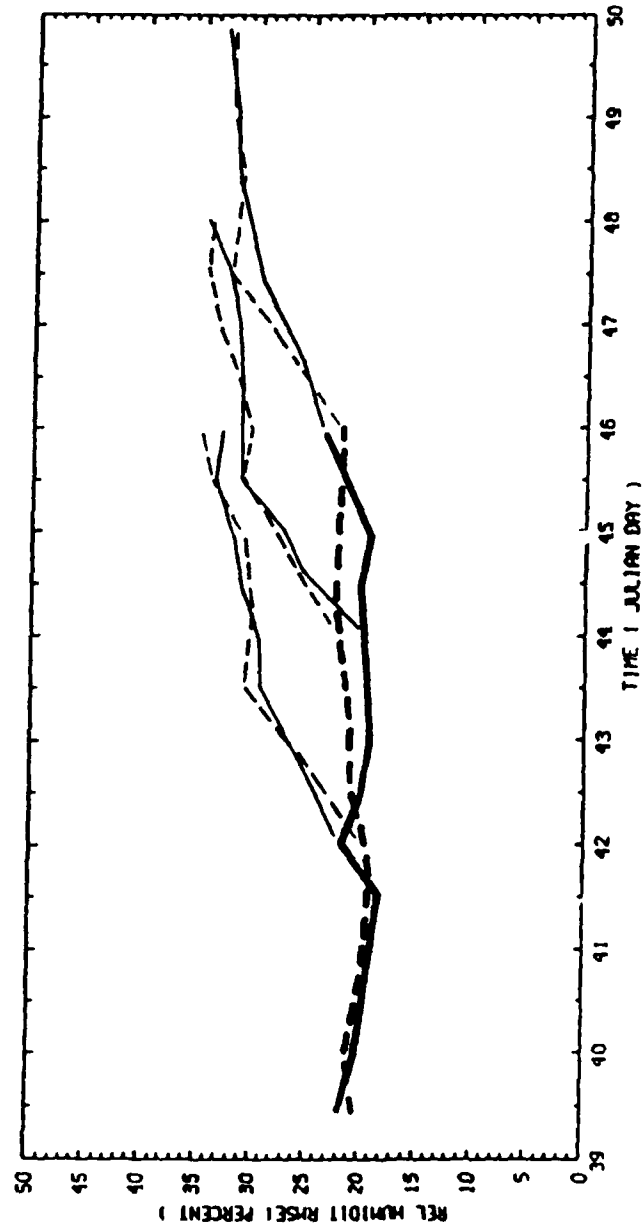
Fig. 10. *Global rms differences of 850 mb relative humidity between radiosonde observations and NEPHSAT (solid lines) or STATSAT (dashed) analyses (thick lines) or forecasts (thin) for the February period.*

errors are marginally smaller than those of STATSAT for most of the analyses. The forecasts do not indicate one to be superior to the other. The same general conclusions hold if one computes these statistics of just the Northern Hemisphere extratropics, or even just over North America where the beneficial impact of the 3DNEPH data would be expected to be largest. It should be noted that the quality of the bogus RH data is about the same as that of the first guess, resulting in only a small positive impact even in radiosonde-void regions. The generally inconclusive results of the 850 hPa radiosonde statistics hold for other levels, as well.

One of the obvious shortcomings of the present OSE is the limited geographical extent of the bogus RH data. In future studies, the RH profile approach could be extended to produce a global bogus RH data set by repeating the regression study performed here for different regions of the globe. Different EOFs, and different regression equations would then be used in different regions.

Other limitations of the OSE are related to the data assimilation system itself. Significant obstacles to an effective assimilation of moisture data include the relatively coarse resolution of the analysis and forecast, the use of an adiabatic NMI, and a very simple moist physics package in the GSM. There are several potential remedies to these shortcomings: using a diabatic NMI, in conjunction with a moisture initialization procedure as suggested in Donner (1988) would minimize the rejection of initial moisture data by the forecast model. Improvements to the physics package of the GSM are also necessary to limit the error growth during the assimilation cycle and the longer range forecasts produced from the analyses; the physics package currently being implemented and tested by AFGL is expected to improve this aspect of the GDAS.

Probably the most serious limitation to this technique is the relatively large observation errors of the bogus RH, which are larger than the globally averaged errors of the current RH analyses. It may be possible to reduce the errors of the bogus RH data with changes in the regression approach, such as the definition of the EOFs or the preprocessing of the cloud data. However, for significant reductions of the observation errors, it will be necessary to take account of the fact that there is no one-to-

one correspondence between relative humidity and cloud cover, and to include other atmospheric parameters (e.g., static stability, vertical motion) in the problem.

3.8 New model physics: GSMSAT

A set of new physics routines has been developed and incorporated in the AFGL GSM. These routines include a radiation scheme (Ou and Liou, 1988), a new planetary boundary layer scheme (Mahrt et al., 1987) which also includes a computation of the soil temperature and moisture, and a new version of the Kuo convection scheme (Soong et al., 1985). The purpose of our GSMSAT experiment was to test the influence of this new model in the data assimilation system, by using it with the same data as in the STATSAT experiment.

3.8.1 Preliminary tests and tuning

Before proceeding with the experiment, we performed some tests to check the new GSM itself. These consisted in some one-dimensional tests and a preliminary 3 day forecast.

The one-dimensional tests are described in Appendix B. We ran a few cases with arbitrary initial conditions, and some simulation of observed situations. These include day 33 of the Wangara data, period 5 of the O'Neill experiment, and a case of cold air advection over the Norwegian Sea.

The main conclusion of these experiments is that the new physics package performs in a reasonable fashion, but may need considerable tuning. Some of the deficiencies of the diurnal cycle simulations are due to the fact that the model imposes the initial values of the soil temperature and moisture on the basis of the atmospheric quantities. This constraint was removed for our data assimilation experiment. Another problem is the fact that the radiative fluxes are kept constant between radiative time steps. This makes the temperature rise in the morning lag behind by several hours, and damps the diurnal cycle. Some kind of time interpolation of the radiative fluxes appears necessary, at least for the computation of the surface energy budget. This could be done, for example, by computing the

atmospheric transmissivity and holding it constant between radiation steps, but using the correct solar angle to recompute the surface flux at each time step.

We also found that in case of very unstable boundary layer, the scheme tends to dry out the level above the PBL. This seems to be a numerical problem. A detailed reexamination of the algorithm that solves the vertical diffusion equation (an implicit, finite element scheme) should be done, including the treatment of the upper boundary condition.

A preliminary 3 day forecast was run with the new GSM to determine how to handle the initializing of the soil temperature and moisture (see Appendix C). As the code was written, these quantities would have been reset to the atmospheric values at each 6 hour cycle of the data assimilation. Since these soil variables seemed well behaved during the forecast, we decided to keep the first guess values as initial state for the next cycle.

During the course of this contract, we made a theoretical study of the model tuning problem. This study is presented in appendix D. As a problem in optimization, tuning of model parameters is difficult because it is a large, multidimensional, nonlinear problem. We have looked at the problem first in the context of a single column model. We proposed to reduce the dimensionality of the optimization problem by separating all the forecasts into different weather regimes, within which some forecast variables are sensitive to only a few parameters. This procedure would also increase the sensitivity of the objective function (a measure of the forecast error) to the parameters. It would also reduce the effect of errors in the initial conditions, which tend to corrupt the evaluation of the objective function sensitivity.

In a three-dimensional model the same procedure could be applied to the optimization of the parameters that have a global value, but would be prohibitive for the ground parameters that vary geographically.

3.8.2 The GSMSAT OSE

The evolution of surface quantities used by the new physics package during the assimilation experiment showed a generally stable and

reasonable behavior of the forecast model. Some possible problem areas were isolated: precipitation, in particular convective precipitation, is predicted by the model as excessive; the specification of a globally uniform sub-ground temperature leads to unrealistic soil temperature changes, particularly in regions covered by snow.

The impact of the enhanced forecast model on atmospheric analyses is rather small, with GSMSAT-STATSAT differences increasing slowly to values that are still well below typical STATSAT-NMC differences at the end of the assimilation. The quality of the analyses and forecasts, when measured as an rms difference from radiosondes, is not consistently improved (or degraded) in GSMSAT, although there are some indications that forecasts of 850 mb RH are slightly worse.

4. OBSERVING SYSTEM SIMULATION EXPERIMENTS (OSSE)

OSSEs are often used to estimate the impact of proposed advanced observing systems and as an aid in their design by examining various implementation scenarios. In the present study we examined the impact of the SSM/T-1,2 and the proposed Doppler wind lidar (DWL) instruments. In all five OSSEs were performed. These are listed along with the data sources used in each experiment in Table 4. Detailed descriptions of these experiments may be found in Hoffman et al. (1989) and Grassotti et al. (1989).

Table 4. OSSE data sources summary

OSSE	Ground-based	CDW	TOVS	SSM	DWL
STATSAT	X	X	X		
NOSAT	X				
SSMSAT	X	X		X	
SSM+TOVS	X	X	X	X	
WINDSAT	X	X	X		X

4.1 General OSSE strategy

There are four components common to any OSSE: (1) A four dimensional reference atmosphere, often called the nature run. This is

considered to be the "TRUTH". (2) A sampling procedure to obtain observations. (3) A data assimilation system, composed of a forecast model and analysis procedure. (4) A quantitative verification procedure. These components will be detailed in suceeding sections, except for the data assimilation system which is the same system which we have described in Section 2. Minor modifications to the assimilation system for the purpose of these experiments are detailed in Hoffman et al. (1989) and Grassotti et al. (1989).

Usually, the nature run is simply a long forecast made by an advanced NWP model or Global Circulation Model. The more sophisticated the nature model, the better. Remotely sensed data are influenced by many geophysical parameters, including sea surface temperature, atmospheric aerosol, clouds, etc, which should be included in the nature run to the extent possible. For example, SSM/T should provide less accurate retrieval of atmospheric boundary layer humidity when wind speeds are high since in this case surface emission dominates the atmospheric emission.

From the point of view of numerical weather prediction (NWP), the most important characteristics of any proposed remote sensing system are its geographical coverage, horizontal and vertical resolution and its error characteristics. In a simulation study these characteristics must be properly accounted for. The procedures for simulating data from the nature run should consider the following items: (1) representational errors; (2) sampling; (3) geophysical local bias; (4) random error, which might contain vertical and horizontal correlations; and (5) sensor filtering. When a sensor uses a statistical retrieval method, all its observations should be filtered by projecting onto the vertical basis functions which are used in the retrieval. This is also true for so-called physical retrieval methods. Note that (2)-(5) above can be achieved by simulating the sensor and its retrieval scheme (e.g. Atlas et al., 1985). This is costly and, in our study, we have, instead, performed selected sensor simulation/retrieval studies described below to define the error characteristics.

By representational errors we mean that it is not just accuracy of the measurement which is important, the measurement must be representative

as well. NWP is really concerned with the spatially and temporally smoothed behavior of the atmosphere. Variations on scales up to kilometers and minutes are generally considered to be averaged over and are parameterized within the model. Consequently, that part of the measured signal attributable to these scales is considered to be noise from the NWP point of view. This source of error can in some cases be predominant as, for example, in radiosonde observations. One implication of this is that as models improve in resolution, this source of error decreases. No existing global model has fine enough resolution to represent all scales of motion which exist in nature. In fact the smallest scales represented by models are usually severely damped for computational reasons. A method to unfilter the nature run was suggested by Hoffman (1988), but not used here.

Verification of OSSE results is easy because we have total knowledge of the "TRUTH". In these experiments we may legitimately use the word error instead of difference when we compare an experiment to the nature run. Interpretation of these results, however, is not so easy. For these reasons it is desirable to calibrate the OSSE results to OSE results. In the present case we conduct two OSSEs, NOSAT and STATSAT, for which we have previously conducted analogous OSEs. We use only a very simple calibration procedure in Section 4.5.3. Basically we assume OSSE impacts relative to STATSAT are proportional to corresponding OSE impacts in deriving our estimates of actual SSM/T-1,2 and WINDSAT impacts.

4.2 Nature run

It would be possible to use a series of real analyses for the reference atmosphere, but the results of such experiments would be even more difficult to interpret for the following reasons. In this situation the "TRUTH" is the actual atmosphere, not the reference atmosphere. Therefore, in data rich areas, the reference atmosphere would agree well with the "TRUTH" while in data voids it would not. Consequently, simulated observations in data rich areas would add correct information, but have little impact because of the concentration of other observations already available, while simulated observations in data poor areas would add erroneous information, which would be carried by the model during the data

-36-

assimilation cycle to other areas. If the results are then verified in data rich areas we might obtain a negative impact by adding a new observing system. Greater accuracy in the simulated observing system would not avoid adding erroneous data in data poor areas. For these reasons it is more advantageous to use a long range forecast for the nature run.

ECMWF generated the nature run used in this study. The nature run is simply a 20 day forecast from the FGGE IIIb analysis produced at ECMWF at 00 GMT 10 November 1979 (Bengtsson et al., 1982). The model used in the nature run forecast was a version of the 15 layer, 1.875 degree grid point model (Hollingsworth et al., 1980). This model included fairly complete physics (Tiedtke et al., 1979) with a diurnal cycle.

4.3 Simulated observations

NMC simulated the FGGE Level IIb and WINDSAT data for the period, in the NMC (Office Note 29) format from the ECMWF nature run (Dey et al., 1985). Almost all Level IIb data were simulated. Later GLA converted the NMC data to the standard FGGE format (WMO, 1986). All this work was completed by early 1984. We received copies of the nature run and FGGE format Level IIb data from GLA, courtesy of R. Atlas. We simulated the SSM/T-1 and SSM/T-2 data based on careful simulation studies of the operational retrieval methodology.

4.3.1 Conventional data

The simulated standard FGGE Level IIb data were created by replacing all the observed atmospheric variables in the real FGGE Level IIb data with values interpolated from the nature run, corrupted by adding a simulated observing error. Therefore if a particular radiosonde report is missing in the real data, it is missing in the simulated data, if it is present in the real data, it is present in the simulated data and has the same quality control marks and missing data flags as the real observation. The simulated observational error which is added to the value of the nature run at the observing location is composed of a random Gaussian error which is not correlated with anything else and in the case of the TOVS data a bias depending on the diagnosed cloud cover. The size of the

random error, or observing error standard deviation (OESD) is appropriate for the particular observation.

4.3.2. WINDSAT data

The basis of DWL is the measurement of the Doppler shift of a laser pulse backscattered by aerosols and other particles in suspension in the atmosphere. Two measurements of the same atmospheric volume from two different angles provide an unambiguous ~ id determination, with the reasonable assumption that the vertical velocity can be neglected.

The simulated WINDSAT data are created at all TIROS reporting locations in a manner similar to that described above for the other data types. At each TIROS location for which NESDIS performed a retrieval, a WINDSAT profile is produced. This profile extends from 10 mb down to the surface in relatively clear conditions or down to cloud top in cloudy conditions. Typically there are 2000 to 4000 WINDSAT profiles per six hour time period. We note that error levels assigned to the lidar winds are approximately half that of the RAOB winds. This characteristic combined with the full TIROS coverage and uncorrelated error structure should lead to greatly improved analyses and forecasts.

4.3.3. SSM/T data

In preparation for the OSSE using SSM data, we performed a review of the existing studies of moisture profile retrieval using microwave data, with special emphasis on the 183.31 Ghz water vapor resonance line. Our main findings are the following: Work performed at ERT, NASA, MIT, JPL and AER demonstrate the feasibility of using millimeter wave brightness temperature to retrieve water vapor profiles with a useful accuracy (about 20% in relative humidity). The main advantage of longer wavelengths, over infrared retrievals, is the low emissivity of the ocean surface, making possible the retrieval of low layer moisture. Another advantage is that clouds are not opaque to millimeter waves. However the effect of clouds on the transmission of these waves cannot be neglected and is a field of continuing research. In particular, the role of ice clouds as attenuators may have been underestimated. For measurements

taken over the oceans the main difficulty is estimating the effect of high winds on the surface emissivity. Over land, the variable, and in general higher surface emissivity makes millimeter wave retrieval of moisture less attractive.

In view of the results of the review, which showed relatively poor performance over land, we simulated SSM retrievals over oceans only. The best way to simulate an observing system such as the SSM is to use the nature run to generate the radiances (or brightness temperatures) that would be observed, add appropriate measurement errors and simulate the retrieval of temperature and moisture profiles. It would be very expensive to do this for every data point. Instead, we solved the forward problem, i.e. the brightness temperature computation, only for two subsets of the data. The first subset was used to derive regression equations between nature run temperature and moisture profiles and computed brightness temperatures. The coefficients of these equations form the so-called D matrix. Several D matrices were derived, depending on whether the ocean was ice covered or not, and whether the sky was clear or cloudy.

These D matrices were then used to simulate statistical retrieval on the second subset of data. This procedure enabled us to derive a set of retrieval error statistics. Again, these error statistics were stratified according to geophysical criteria, which included the latitude and strength of the wind in addition to cloudiness and the presence of ice. In the OSSE, then, we used these error statistics to modify the nature run profiles and create pseudo-retrievals. This work is described in detail by Grassotti et al. (1989).

4.4 Experimental design

For the experiment to be realistic, we must start the first assimilation cycle with a state reasonably far away from the nature run. Otherwise the baseline analyses would be too good, and any impact of a new observing system hard to detect. We started by running a four day forecast from the state given by the nature run for 00 GMT, 11 November 1979. The error growth was relatively slow but, after four days, the level of error was at least as large as that inferred for the GDAS

analyses by the OSEs described in Section 3. We then performed an additional three day spin-up assimilation run, using the STATSAT data configuration. The resulting fields, at 00 GMT on 18 November, where then used as the starting point for all assimilation experiments.

Each OSSE runs from 00 GMT 18 November through 00 GMT 25 November. For each OSSE 96 h forecasts are made from 00 GMT 21, 23 and 25 November. The OSSEs described make use of all the Level II data which were simulated by NMC as described in Section 4.3.1 except that surface observations are not used and satellite temperature soundings over land are not used. In NOSAT the satellite temperature soundings and CDW observations are excluded, in WINDSAT, the Doppler wind lidar observations described in Section 4.3.2 are added, in SSM + TOVS the SSM data described in Section 4.3.3 are added and in SSMSAT the SSM data replace the TOVS data. (Refer to Table 4.)

4.5 Results

There are some important caveats that apply to the results reported here. As is the case with all OSSEs and OSEs, the measured impacts apply to the particular data assimilation system used here. While the assimilation system is reasonably "state of the art", some aspects, in particular the anchoring of the satellite thicknesses and the limitations of data selection, may limit the extent to which our conclusions are generally valid. Since we used simulated data in our experiments, the realism of our OSSE results depends on how realistic the simulated observation errors were. We took particular care that they were of sufficient size and had the appropriate error correlation structure. Finally, the calibration of our OSSE results with OSE RAOB statistics suffers from the usual problems of the uneven distribution of radiosondes, particularly the bias toward land areas (where no satellite data were used), and the small sample sizes over the Southern Hemisphere.

4.5.1 Synoptic evaluation

The OSSE analyses were all found to be noisier than the corresponding nature data. The Northern Hemisphere WINDSAT analyses of 500 mb

height had errors smaller roughly by half than the corresponding STATSAT analyses. The SSMSAT and SSM+TOVS analyses in the Northern Hemisphere tended to be more like STATSAT. In the Southern Hemisphere, on the other hand, it is possible to differentiate between the different satellite based observing systems. Also in the Southern Hemisphere we see that NOSAT is quite poor.

In a number of cases synoptic features were better analyzed by SSMSAT than by SSM+TOVS. The somewhat surprising result that adding TOVS data led to a degradation of the analyses was caused by a combination of factors, some of which were related to the anchoring of satellite thick-nesses, others to data selection and quality control procedures.

Moisture analyses, either in terms of RH or cloud cover, are notice-ably improved by the addition of SSM data. It is quite clear that the moisture analyses created with the SSM retrievals are much improved over the STATSAT analyses. In addition, the results for SSMSAT and SSM+TOVS are very similar with only minor differences seen between the two experi-ments. It seems that the presence or absence of TOVS retrievals has only a small impact upon the humidity analysis and that nearly all the improve-ment stems from inclusion of the additional SSM data. Improvement is most dramatic over Southern Hemisphere ocean areas where conventional moisture data is practically nonexistent. Smaller, but equally clear improvements are seen in the Northern Hemisphere as well. Also noteworthy is the fact that the analyses are improved over some land areas even though SSM re-trievals were confined to the oceans.

In terms of forecast error, the Northern Hemisphere 500 mb height field is better forecast by WINDSAT relative to STATSAT by roughly 24 h. Interestingly, although the analysis error patterns of WINDSAT and STATSAT are quite different, the forecast error patterns tend to have the same shape. This suggests that model errors are significant in these experiments. The quality of the three sounder based forecasts are all similar; in some cases one is better, but in other cases the roles might be reversed.

4.5.2 Statistical evaluation

The conclusions of the previous section are summarized by the objective statistical measures we have examined. Fig. 11 shows the global 500 mb rms height error of the NOSAT, STATSAT, SSMSAT and WINDSAT analyses and forecasts. The NOSAT analysis errors increase from the 35 m typical for the STATSAT analysis to 50 m by day 4 of the assimilation, whereas the STATSAT analysis errors decrease by 1-2 m over the assimilation period. This is an indication of how well the spin-up process has performed. The forecast error growth is more rapid in STATSAT, but errors remain smaller than those of the NOSAT forecast for the length of the forecasts. The SSMSAT analysis errors are consistently smaller than those of STATSAT, but by only 2-3 m. SSM + TOVS (not shown) is quite similar to STATSAT. The WINDSAT data have a definite and dramatic impact on the analysis error; by 24 h the error has dropped to 20 m and continues to slowly decline thereafter. The WINDSAT forecast errors, since they start from such good initial conditions are the smallest of all the experiments. Results at other levels largely mirror those at 500 mb.

Considering the 850 mb relative humidity field (Fig. 12), we see that SSMSAT provides the best analyses yet the best forecasts are obtained from WINDSAT. This is more so in the extratropics than the tropics; presumably the relative humidity forecasts are determined largely by the large scale fields of temperature and winds in the extratropics and the WINDSAT analyses of these are superior. SSMSAT is always better than NOSAT which in turn is somewhat better than STATSAT. With regard to RH SSM + TOVS is very similar to SSMSAT. The particularly low growth rate of relative humidity errors for STATSAT is an indication that the errors have already saturated and that the STATSAT humidity analyses are nearly worthless.

We examined zonal cross sections of u and v wind components, temperature and relative humidity at individual synoptic times and averaged over the last five 0000 GMT analyses of the experiments. Generally good agreement with the GFDL (Lau, 1984) time and zonal averaged cross sections for November, 1979 was obtained. Considering the amount of high quality wind data available to WINDSAT, the small improvements to the zonally averaged wind fields are disappointing.
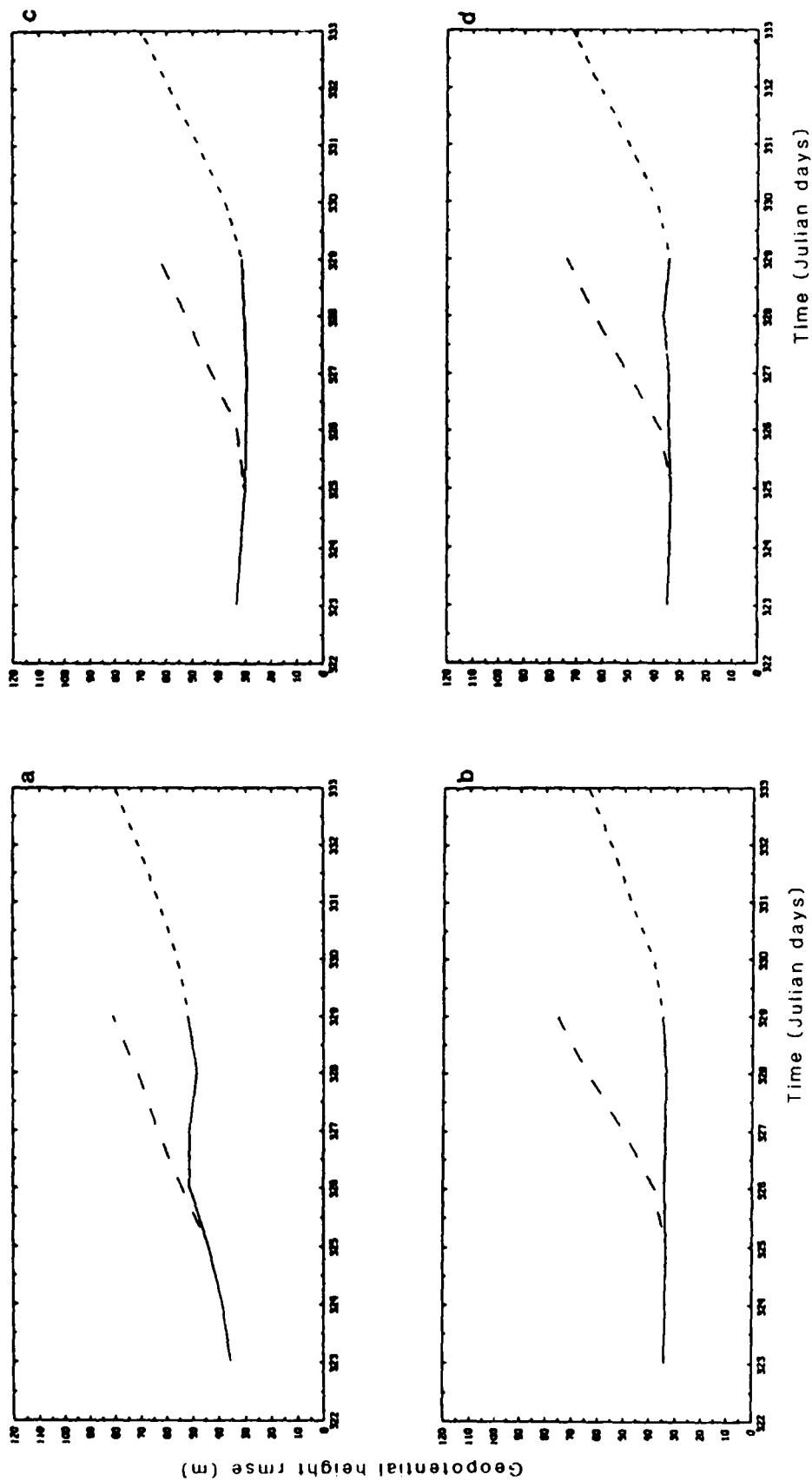
Fig. 11 Rms height errors at 500 mb. (a) NOSAT, (b) STATSAT, (c) SSM, (d) WINDSAT. Analysis errors are shown in solid curves, forecast errors in dashed curves. Julian day 322 corresponds to 00 GMT 13 November.
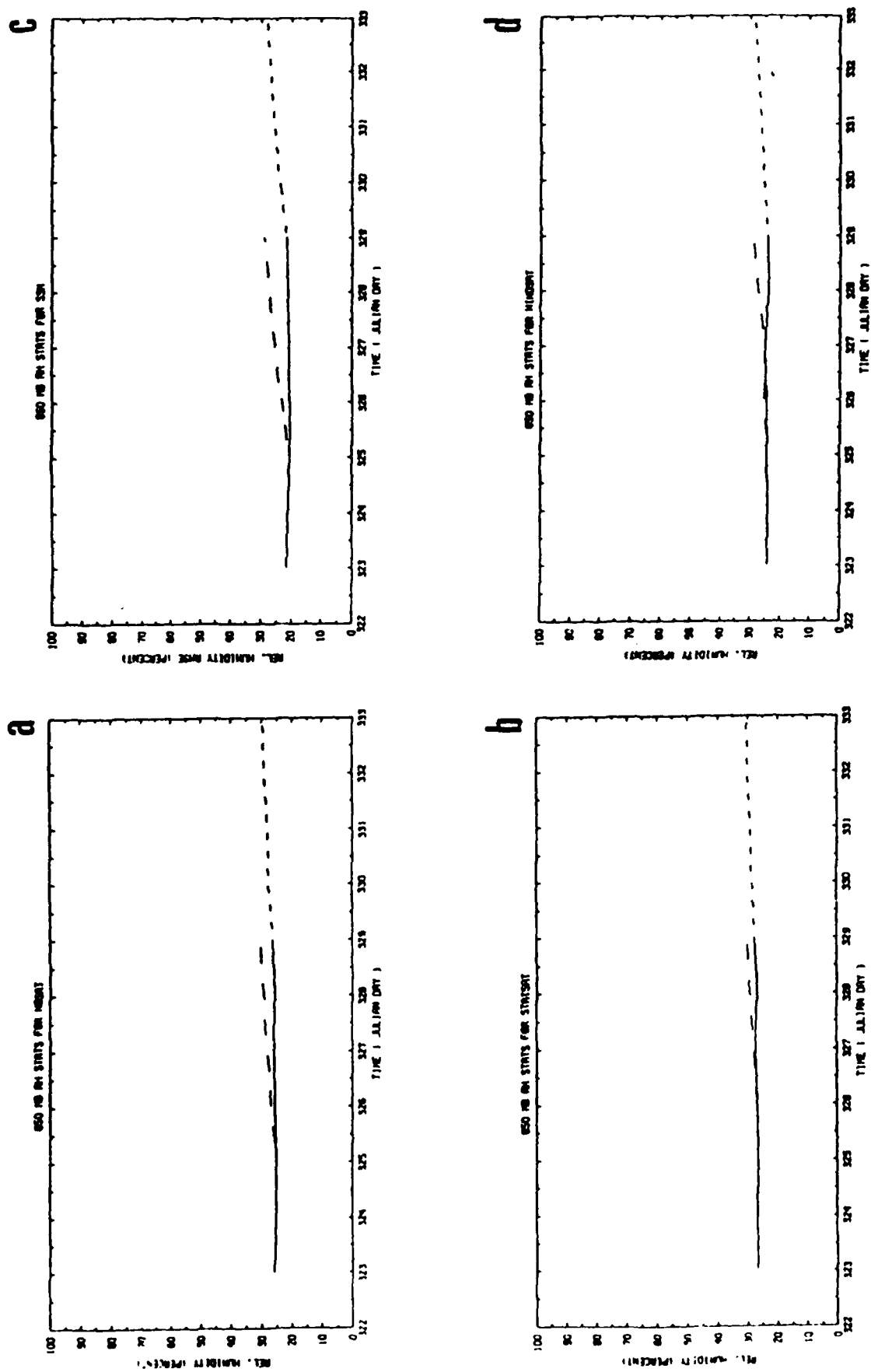
Fig. 12 Global rms analysis/forecast errors for 850 mb relative humidity. (a) NOSAT, (b) STATSAT, (c) SSM. (d) WINDSAT. Solid curves denote analysis, broken curves are forecasts.

Time mean zonally averaged RH errors are relatively large in all assimilations. Generally, the boundary layer is too cold and too dry and the tropical free atmosphere is too moist. We may contrast nature with the cross sections for STATSAT and SSMSAT (Fig. 13). In STATSAT the low level averaged relative humidity analysis below 850 mb is consistently too dry while at higher levels relative humidity is too high in the tropics and too dry near the poles. As a result the northern and Southern Hemisphere mid-level minima are greatly increased in magnitude in STATSAT. Additionally, the asymmetry seen in the nature run with respect to height is gone and both features now occur at 650 mb. In SSMSAT the averaged relative humidity analysis is improved almost everywhere. As in STATSAT, the averaged SSMSAT analysis does not retain the asymmetry in the moisture field which is seen in averaged nature data, although the magnitudes of the minima are better analyzed. In short, as determined from differences in averaged analyses, SSMSAT relative humidity analyses are closer to nature at most latitudes and at all vertical levels. However, in general, the polar regions and boundary layer are too dry and the midlatitude and tropical atmosphere above the PBL is too moist.

In general the biases during the forecast are small compared to the rms differences. However in many cases the biases grow very steadily with time indicating that the AFGL model is warming and drying relative to the ECMWF nature. For example Fig. 14 shows the evolution of bias for the 500 mb height in the OSSEs.

4.5.3 Calibration to OSE results

Forecasts made within OSSEs are often much better than any real forecast. There are two principle reasons for this behavior: First, the model used to produce the nature run is inevitably more similar than the true atmosphere to the forecast model. Second, the errors used in many simulation experiments are easy for the analysis system to handle.

For these reasons it is desirable to calibrate the OSSE results. In order to have a closer correspondence with the real world and to simplify our calibration procedure we have calculated rms difference between the forecasts and the simulated radiosondes for different regions and for several variables at each layer in the atmosphere. The variables examined
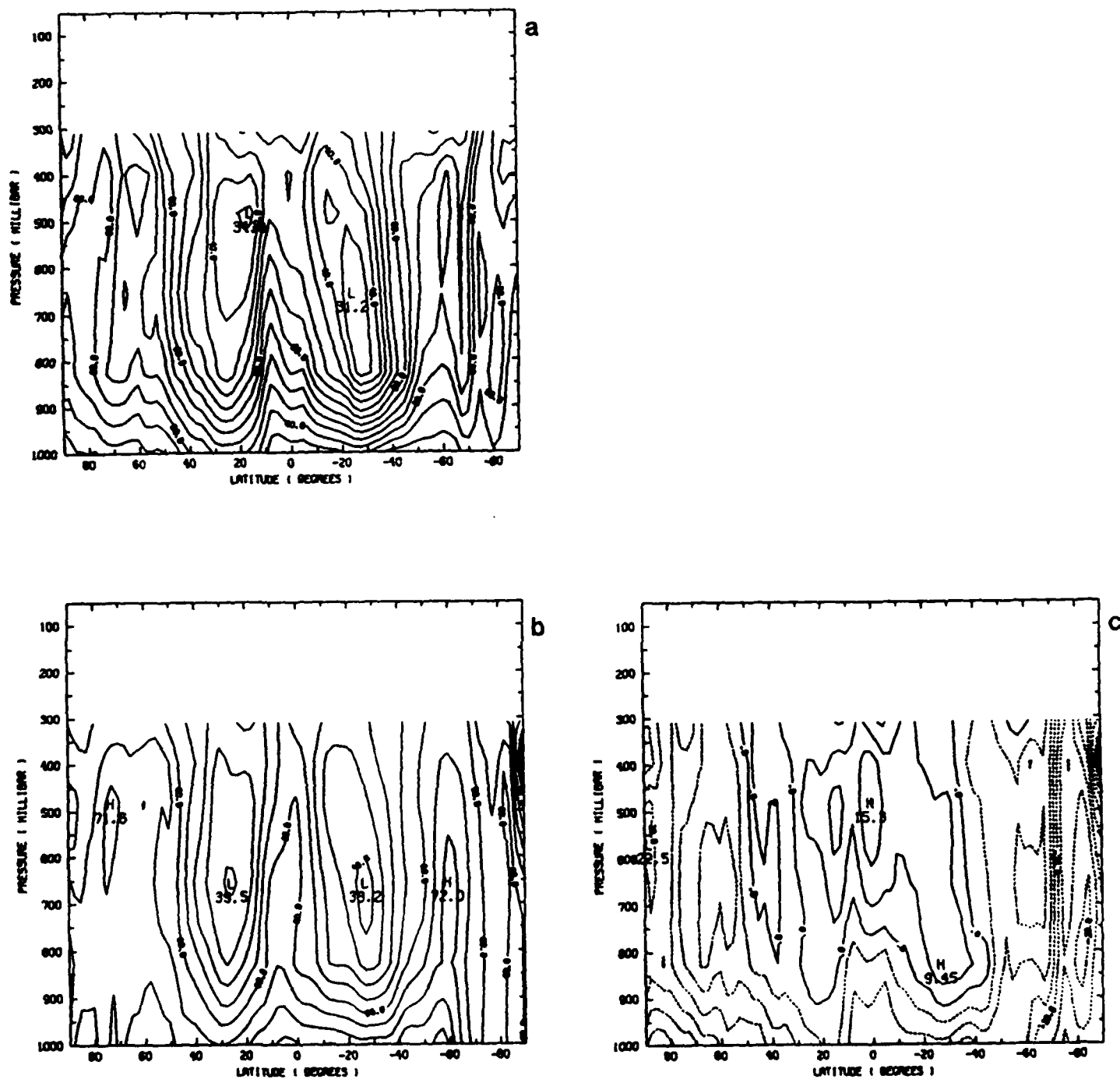
Fig. 13 Zonal time averaged relative humidity. (a) The nature run,
(b) STATSAT, (c) STATSAT - the nature run, (d) WINDSAT, (e) WINDSAT -
the nature run, (f) SSM, (g) SSM - the nature run. Contour interval
is 5 percent, negative values are dashed.

-46-
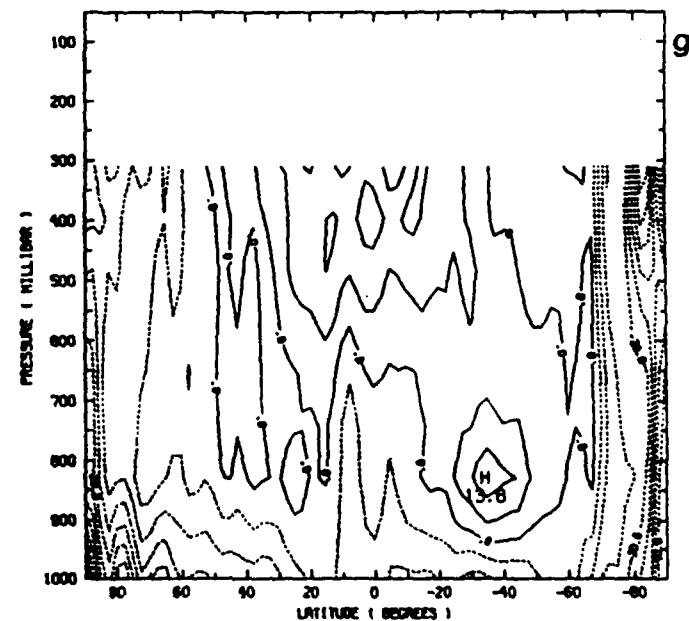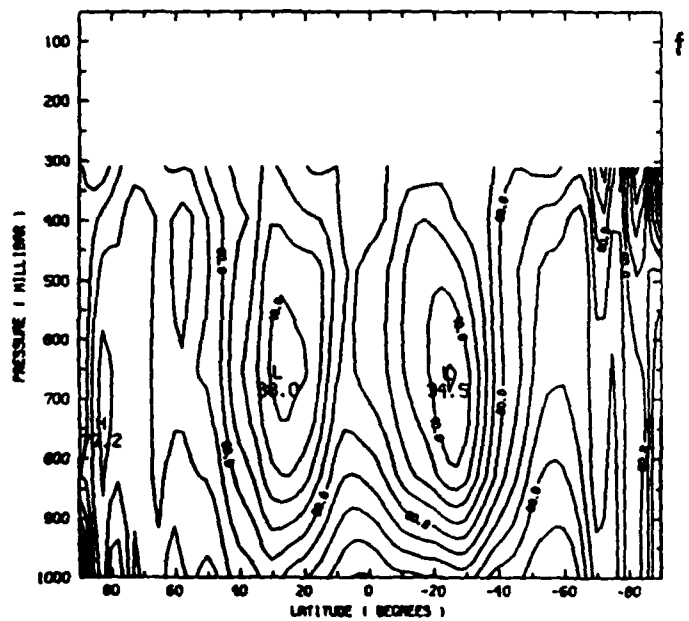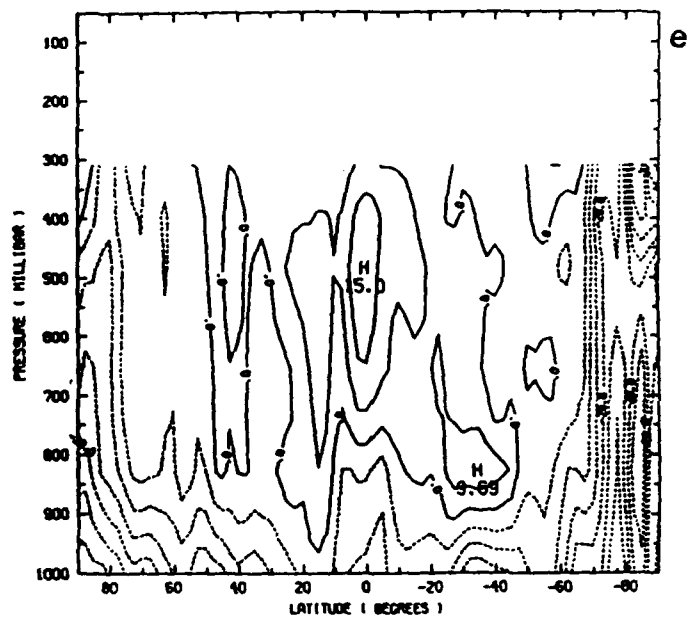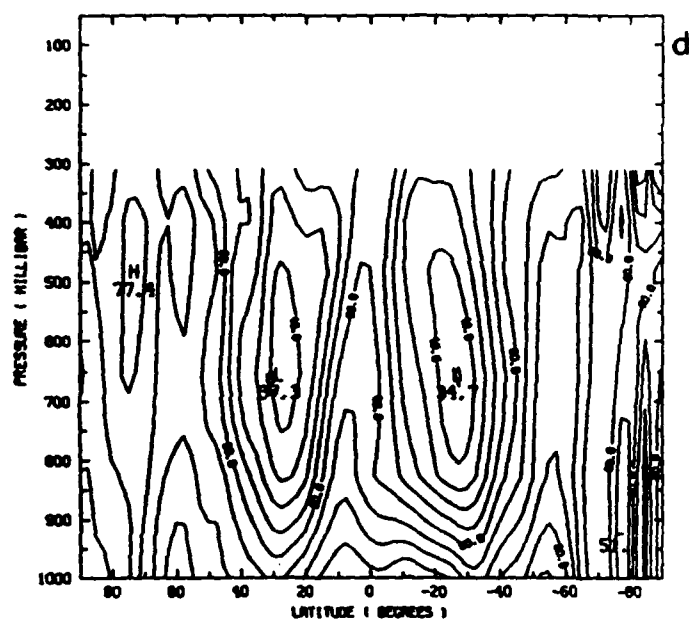
Fig. 13   (cont'd)
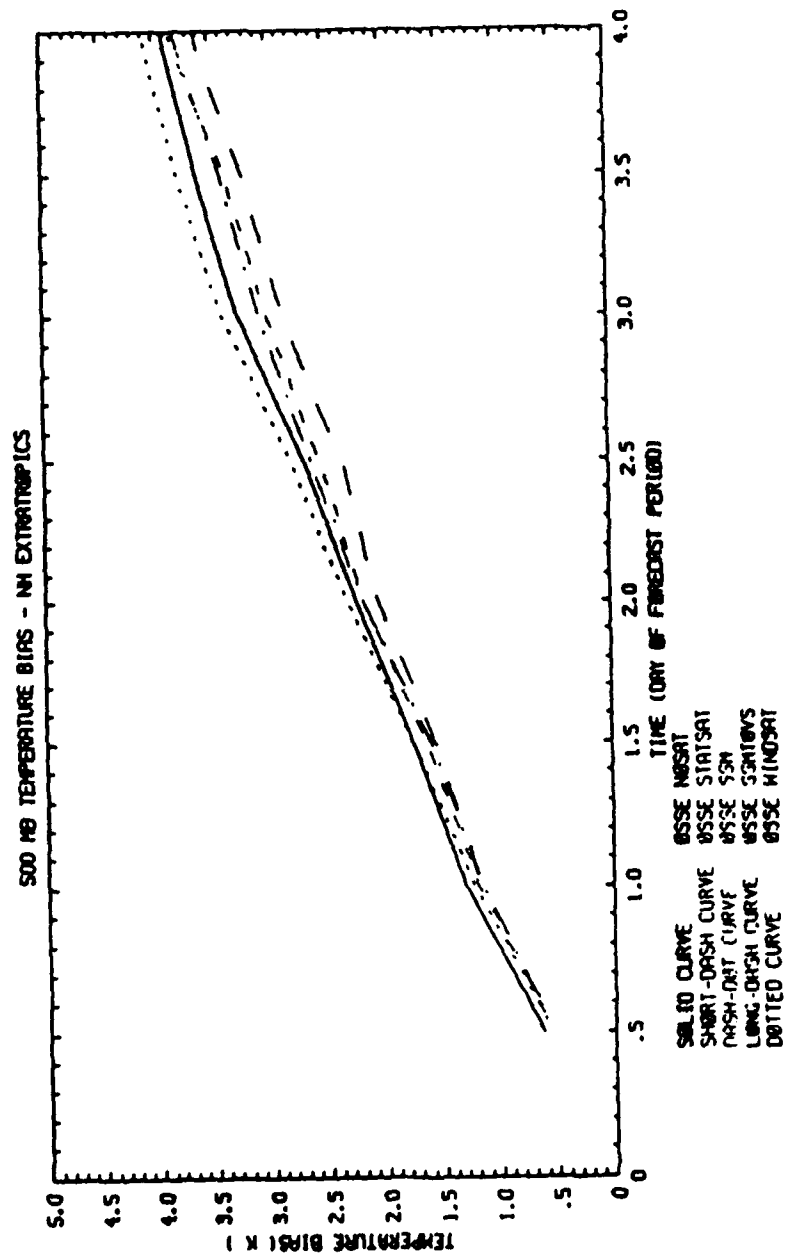
Fig. 14 Forecast 500 mb temperature bias for Northern Hemisphere extratropics.

include geopotential height, temperature, vector wind, relative humidity, and cloud cover. We then developed a procedure to calibrate these differences using the NOSAT - STATSAT impact observed in the OSEs as a yardstick. However, for the present experiments we find that the OSSE impacts are fairly similar to the OSE impacts and the calibration procedure does not greatly alter the conclusions one might draw from the OSSE results directly. We also calculated rms difference in cloud cover layer by layer. Invariably, the corresponding relative humidity and cloud cover plots look very similar.

Our principal calibration assumption is that the OSE impact of adding or removing an observing system is proportional to the corresponding OSSE impact. In our calibrations we always take STATSAT to be our standard. We use the NOSAT - STATSAT difference to determine the constant of proportionality. However, in the Northern Hemisphere, STATSAT and NOSAT OSSE results are often so nearly equivalent that impacts expected from advanced observing systems cannot be calibrated. We measure impact in terms of predictability time, i.e. we define impact to be the change in the useful length of the forecast. We then took advantage of the observation that our rms difference curves grow nearly linearly, at least during the forecast period from 12 to 48 hours, to fit these data with a series of straight lines having a common slope. In the Northern Hemisphere, these fits were very good. They are less reliable in the Southern Hemisphere and tropics, presumably because the number of radiosondes in these regions is small.

The impact of the SSM and DWL data was found to be generally small in the Northern Hemisphere, but quite substantial in the Southern Hemisphere. This result is in agreement with numerous previously conducted OSSEs and OSEs which measured the impact of satellite data. Consider the rms difference for 500 mb geopotential averaged over three forecasts (Fig. 15). WINDSAT is 36 hours better than STATSAT, which is in turn more than 36 hours better than the NOSAT forecasts. The three sounder based systems are roughly equivalent with SSM and SSM+TOVS better than STATSAT by 12 and 8 hours respectively. At 200 mb compared to STATSAT, WINDSAT provides 1, 2 and 2.75 day improvements in rms wind vector forecast skill in the Northern Hemisphere tropics and Southern Hemisphere,
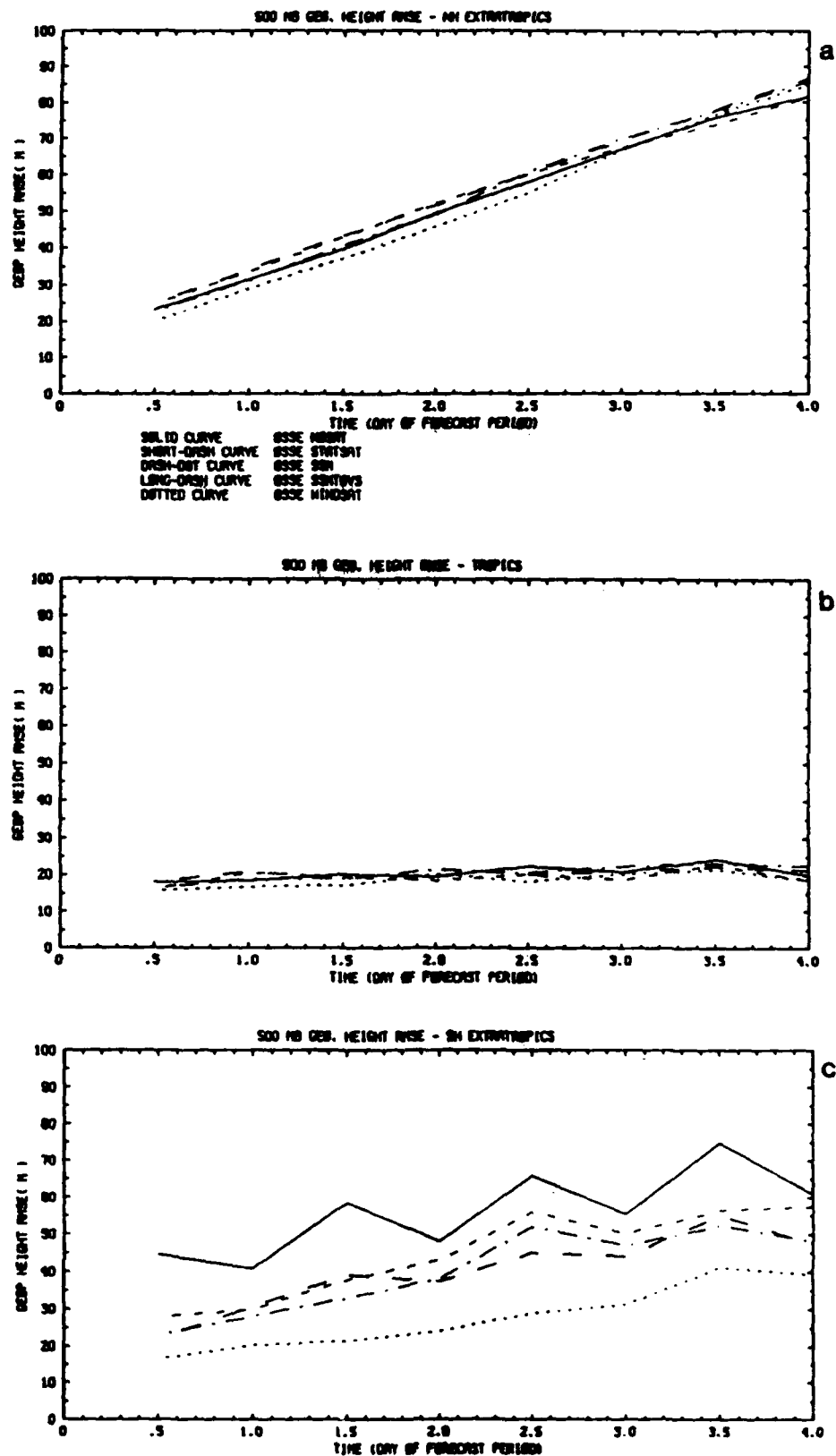
Fig. 15 Forecast rms error growth, 500 mb height. (a) Northern Hemisphere extratropics, (b) Tropics, (c) Southern Hemisphere extratropics.

-50-

respectively. In the Southern Hemisphere, the rms difference curves sometimes exhibit a sawtooth pattern due to sampling problems; there are usually about 60 RAOBs at 00 GMT and only about 40 at 12 GMT in the Southern Hemisphere. Most of the non-reporting RAOBs are in the Australian sector.

The RH forecast errors were not appreciably affected in the Northern Hemisphere extratropics, but a marked improvement could be seen in the tropics and the Southern Hemisphere extratropics (Fig. 16). Interestingly, though, WINDSAT moisture forecasts were superior to SSMSAT in the extratropics, even though only RAOB moisture data were used in WINDSAT, reflecting the dominant role of the mass and wind fields in forcing the moisture field in the extratropics. Overall the ranking is WINDSAT, SSM, NOSAT, SSM+TOVS and STATSAT. It appears that using TOVS degrades the moisture analysis.

5. **SUMMARY AND CONCLUDING REMARKS**

Satellite observations hold great promise for improving NWP (Isaacs et al., 1986a). During the current contract effort we have made substantial progress in several areas related to improving NWP, especially the analysis and forecast of moisture and cloudiness, through space borne observing systems. We have developed, implemented and evaluated techniques for retrieving geophysical parameters from satellite observations for use in NWP. We have tested the impact of these data and data from other sources in a series of OSEs. We have developed and applied a comprehensive data simulation methodology. We have conducted a series of OSSEs to evaluate the impact of new observing systems. In the course of this work, we have made some small but significant improvements to the AFGL GDAS. Some of the techniques we have developed are worthy of further refinement. In addition, our study has suggested future avenues of research and prompted us to speculate on how the moisture analyses and forecasts might be improved.

Our work in the area of techniques development has focused on retrieving relative humidity profiles from three distinct sources:
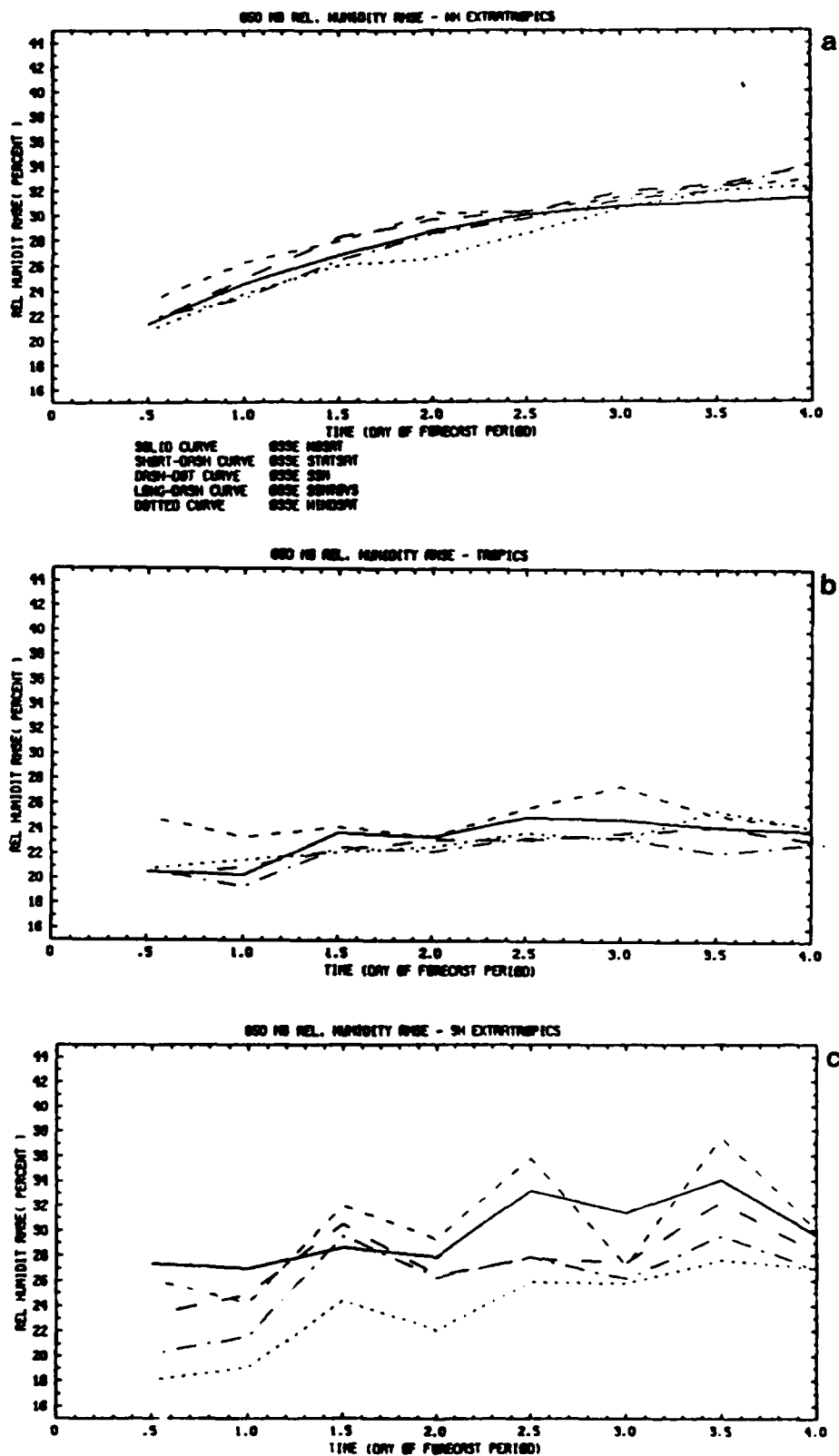
Fig. 16   Forecast rms error growth, 850 mb relative humidity.
(a) Northern Hemisphere extratropics, (b) Tropics, (c) Southern
Hemisphere extratropics.

-52-

- The TIROS Operational Vertical Sounder (TOVS) has several channels specifically chosen for their sensitivity to atmospheric water vapor. We developed a simultaneous retrieval methods for these channels. Evaluation of these retrievals showed these data to be poor quality.

- The operational Air Force three dimensional nephanalysis (3DNEPH) is based on a variety of sensors but principally on the OLS, the visible/infrared imager which is the primary DMSP sensor. The 3DNEPH cloud parameters were used to specify via a linear regression relationship complete relative humidity profiles, represented by EOF basis functions. The retrieved relative humidity profiles were seen to be superior to other data derived from cloud cover parameters. However in terms of NWP utility, these data are of marginal quality.

- The millimeter wave moisture profiler, SSM/T-1,2, should be launched soon. Simulation retrieval studies of this sensor were carried out using the statistical or D-matrix approach as well as the simultaneous approach.

In summary the real data sources, TOVS and 3DNEPH, provided only marginally useful moisture data, while in simulation, the SSM/T retrievals were significantly useful. We believe that the methods developed for TOVS and 3DNEPH are promising but must be extended to be truly useful and could be applied to other sensors. It must be recognized that TOVS and OLS will not provide high quality relative humidity profiles, no matter what retrieval and data analysis techniques are used. On the other hand SSM/T-2 is designed specifically for the retrieval of humidity profiles. Our experience, based so far only on simulation studies, is that SSM/T-1,2 provides useful humidity profiles even when using a simple D-matrix retrieval approach.

A series of real data OSEs were conducted using the AFGL GDAS to measure the impact of retrieved humidity profile data described above as well as the impact of other satellite data. These experiments are summarized

in Table 3. The STATSAT experiment demonstrates that the AFGL GDAS is operating satisfactorily. In particular, the agreement between the STATSAT and NMC analyzed 500 mb height fields is typically 40 (50) m in the Northern (Southern) Hemisphere.

The differences between NOSAT and STATSAT analyses are mainly found over the oceans, since only conventional data are used in NOSAT. However, the first guess advects errors from the data-sparse to data-rich regions. Thus, there are substantial analysis differences throughout the Southern Hemisphere and smaller differences along the western edges of the Northern Hemisphere land masses. We noted that when data are scarce (as in the NOSAT experiment) the resulting forecasts may be poor yet agree very well with the analyses of the same experiment.

The GLASAT experiment showed only small sensitivity of the AFGL GDAS system to different temperature retrievals, with no clear indication whether physical or statistical retrievals result in better forecasts. However, this particular result may be sensitive to the details of our experiments. Greater impact from satellite temperature (and humidity) observations may require not only a physically based retrieval method but also improved model and analysis resolution, improved physical parameterizations in the model and the incorporation of the retrieval method within the data assimilation system so that the best possible background field is used. For example, Dey et al. (1988) report generally favorable impacts due to physical retrievals in the Southern Hemisphere using higher resolution and more complete physics.

Finally, the NOCOR experiment, in demonstrating the speed at which the errors propagate across the Pacific in winter, highlights the global nature of numerical weather prediction. A novel aspect of these experiments is that we have denied both Type 1 (radiosondes) and Type 2 (surface) data in the NOCON and NOCOR experiments to simulate a situation (e.g. a breakdown in communications) that would force reliance solely on satellite data. In other studies, surface observations are usually retained in data denial experiments. In the present experiments, no use is made of surface observations to anchor the satellite thicknesses. Without conventional data in the Northern Hemisphere, the quality of the analyses

decays over the 7 day assimilations. Beyond 7 days this quality would probably continue to decay further. Such a data system is clearly inferior to the STATSAT (full FGGE) system. However for the first few days of data denial such a system may be adequate.

In the NEPHSAT OSE, the impact of the 3DNEPH is clearly visible at the beginning of the assimilation experiment, but it is lost in the noise of the system after 2-3 days. Moisture analyses have marginally smaller errors in NEPHSAT when compared to radiosonde observations, but forecast errors of RH show no sensitivity to the addition of 3DNEPH data. This lack of sensitivity is caused primarily by the relatively large observational errors of the 3DNEPH-based RH data.

The GSMSAT OSE was used to analyze the performance and the impact of the new physics package of the GSM. Quantities predicted by the physics package were found to behave reasonably, except that the convective precipitation rate seemed excessive. Impact on the analyses was rather small, with GSMSAT-STATSAT differences growing slowly to values that are still well below typical STATSAT-NMC differences after 7 days of assimilation. No clear improvement was visible in the forecast errors.

We have conducted a series of realistic observing system simulation experiments (OSSEs) to assess the impact of a Doppler Lidar Wind (DLW) sounder and the SSM/T-1,2 microwave and millimeter wave temperature and humidity profiles. These experiments are summarized in Table 4. The addition of DLW profiles in our WINDSAT experiment significantly improved the initial state specification, especially in the Southern Hemisphere extratropics relative to our control STATSAT experiment. The addition of the SSM data significantly improved the moisture analyses in the tropics and Southern Hemisphere extratropics.

In all the forecasts the AFGL model has a tendency to warm and dry out relative to the ECMWF nature model. We note that the version of the AFGL model which we used has no radiation parameterization. The results of our WINDSAT experiment are consistent with previous studies. However it should be noted that the error characteristics chosen for the simulated DLW seem optimistic.

WINDSAT improvements in forecasting ability were quite large in the Southern Hemisphere. These differences are expected to increase the length of the useful forecast by 36 hours in the height field at 500 mb and by 48 hours in the wind field at 200 mb. Details of the analyzed tropical wind field using WINDSAT were somewhat disappointing. Improvements to the assimilation procedures might enable the WINDSAT data to have greater impacts in the tropics. In the Southern Hemisphere, SSM data improved the 500 mb height forecasts by 8-12 hours.

Moisture analyses were substantially improved in the SSM OSSEs. Typical rms errors were decreased by 1/5 from 27% to 22%. Cloud cover estimates derived from the relative humidity fields are too high. Either the model is too moist or the relative humidity to cloud cover algorithm needs to be tuned. The forecasts of relative humidity were also significantly improved. The comparisons of rms difference of cloudiness yield the same results as comparisons of rms difference of relative humidity. However we note that improved wind data also improved the analyzed and forecast moisture and cloudiness fields. Relative humidity forecasts are best in WINDSAT although SSM had better relative humidity analyses. This is to be expected since the relative humidity field adjusts to the large scale mass-motion fields, which are better analyzed and forecast in WINDSAT.

We developed a calibration procedure to translate our simulation results into realistic estimates of forecast impact. The calibration indicates that the improvements seen in the OSSEs in the Southern Hemisphere and tropics are realistic, but in the Northern Hemisphere extratropics, the fact that satellite data has little impact as seen in our NOSAT versus STATSAT comparisons implies that any novel observing system will have limited impact.

While we are proud of our accomplishments under the current contract, much remains to be done. It is clear that a good humidity analysis/ forecast system requires improvements in the model, the retrieval/analysis techniques and the sensors. Progress on all three fronts is called for. With regard to the model, higher spatial resolution and better parameter- izations of physical processes are clearly desireable. Also improved

-56-

numerical schemes for positive quantities may have a significant impact. Current retrieval/analysis techniques could be improved by using more accurate forward problems and making use of a Bayesian approach to formulate the retrieval problem, thus optimally incorporating all prior data. Finally, new sensors with improved error characteristics will surely improve the initial specification of the humidity field.

There are a number of specific areas investigated under the current contract which could be enhanced.

- The simultaneous retrieval of moisture and temperature should be investigated in order to improve the accuracy of current statistical methods. These physical concepts can be extended to the incorporation of image data based cloud property information using the uniform retrieval approach (Isaacs, 1987).

- The specification of relative humidity profiles based on cloud parameters should be extended to include the notion that cloud formation depends on other parameters besides relative humidity, including for example, vertical velocity and atmospheric lapse rate.

- Model tuning and optimization could be undertaken on a more rational basis following the approach described in Appendix D.

- In the realm of simulation studies there is considerable opportunity to improve and refine the experiments reported here. Such efforts would allow the quantification of the relative impact of proposed advanced temperature sounders and DWLs. In addition cost benefit analyses of observational accuracies could be supported by such studies. In future studies it will be important to carefully simulate the geographical coverage and error characteristics of proposed instruments. In particular, natural phenomena which give rise to correlated observational errors shoud be included to the extent possible. Future experiments might use a more recently generated nature run from ECMWF, based on a T106 truncation model, in conjunction with a method to enhance the small scale features in the nature run (Hoffman, 1988).

APPENDIX A   **Physical retrievals of water vapor for global data**
assimilation

1. GENERAL APPROACH

1.1 Retrieval Method

The purpose of this effort was to implement an enhanced physic-
ally based method for the retrieval of water vapor from TIROS/HIRS infra-
red measurements. There are a variety of approaches one might take in
solving the generalized inverse problem, that is, solving for the vertical
profile of temperature and/or constituent concentrations from a set of
frequency dependent radiance observations. Given the nature of the prob-
lem we have adopted Smith's so called simultaneous physical method (Smith
et al., 1986) and tailored it for our purposes.

Basically, this approach employs a perturbation form of the radiative
transfer equation which relates changes in atmospheric parameters to cor-
responding changes in sensor channel radiances. These changes are defined
with respect to a selected a priori first guess or climatological mean.
By inverting this relationship in a least squares sense, residuals in
channel radiances (defined as the difference between observed radiances
and those evaluated assuming the first guess atmospheric profiles) can be
used to evaluate the most likely atmospheric profile. In its most general
form, this approach uses all channels and retrieves such parameters as the
temperature and water vapor profiles, surface temperature, and surface
emissivity simultaneously. In practice, however, we retrieve only water
vapor and have pared down the number of channels used to only those which
are sensitive to variations in tropospheric temperature and water vapor
(see section 2 for further details.)

The simultaneous retrieval method has been extensively documented
elsewhere (Smith et al., 1986; Isaacs, 1988) and so we highlight only its
most relevant features. The method relies on a linearization of the radi-
ative transfer equation which is written in perturbation form. When dis-
cretized this may be written as

$$\delta R_i - \sum_j \frac{\partial R_i}{\partial q_j} \delta q_j ;$$

where δR_i is the measured radiance residual in channel i,

$\frac{\partial R_i}{\partial q_j}$ is the derivative of the ith channel radiance with respect to water vapor amount at the jth vertical level,

δq_j is the water vapor perturbation of the jth level.

The perturbation amounts are computed with respect to some a priori value (or current estimate) and the normal vertical integral which appears in the radiative transfer equation is folded into the derivative term.

Assuming the number of radiometric channels is greater than the number of retrievables, the solution of q (now the vector of water vapor residuals) may be determined from R (now the vector of channel radiance residuals) using a least squares method. That is,

$$\delta q - (A^T A + \gamma I)^{-1} \cdot (A^T \delta R)$$

where A is the matrix of radiance derivatives such that $A_{ij} - \frac{\partial R_i}{\partial q_j}$,

    I is the identity matrix,

    γ is a scalar smoothing parameter.

The term γI acts to stabilize the matrix inversion against numerical and physical noise in the system of equations.

Since each update of the moisture profile allows a new calculation of the channel radiances the method is iterative. That is, each moisture iteration yields a new set of radiance residuals which in turn produce a new estimate of the update to the moisture profile, and so on. One may apply various criteria in deciding when to terminate the iterations.

Table A.1 illustrates application of the physical water vapor retrieval to simulated TOVs infrared water vapor channels (8, 10, 11, 12) assuming the availability of equivalent clear column radiances. Tabulated are root mean square (rms) error statistics of layer abundances (g/cm^2)

for a set of 100 midlatitude soundings initiating the calculation with a water vapor first guess profile based on: (a) climatology, or (b) a statistical retrieval. For each case the climatology variance of the first guess is indicated and rms results are presented depending on the source of required accompanying temperature retrieval, i.e. an exact knowledge of the temperature, climatology, or a statistical retrieval. Note that with a temperature retrieval at best as accurate as that available from a statistical retrieval, the physical water vapor retrieval can significantly improve over climatology with a simple climatological first guess (see Table A.1.a). As the quality of the water vapor first guess improves, the ability of the physical retrieval to further improve the result diminishes. For example, using statistical retrievals of both temperature and water vapor or first guesses (see Table A.1.b), there is only a small improvement over climatology in the physical retrieval result.

Table A.1

(a) Climatology First Guess for Water

Pressure Level (mb)	Temperature Profiles			Climatology First Guess
	Exact	Climatology	Stat Fg	
1 - 300	0.00395	0.00472	0.00406	0.00264
1 - 500	0.03415	0.03840	0.03334	0.04051
1 - 700	0.10198	0.18260	0.10686	0.18817
1 - 1000	0.35384	0.67542	0.37005	0.55847

(b) Statistical First Guess for Water

Pressure Level (mb)	Temperature Profiles			Climatology First Guess
	Exact	Climatology	Stat Fg	
1 - 300	0.00193	0.00819	0.00183	0.00069
1 - 500	0.01198	0.01475	0.01324	0.01388
1 - 700	0.05298	0.16163	0.06268	0.05806
1 - 1000	0.28749	0.62183	0.31472	0.32740

1.2 Forward Problem

All physically based infrared retrieval methods require calculation of the expected upwelling radiation at predefined frequencies based upon an estimate of the atmospheric temperature and constituent profile, and the nature of the underlying surface. Computational expediency is required for such methods because the forward problem must be solved repeatedly. This then precludes algorithms which rely on highly accurate line by line calculations of the transmission functions. Rather we have used the so called rapid transmittance algorithm developed at NASA Goddard which allows extremely fast and relatively accurate calculation of the forward problem (McMillin et al., 1979; Susskind et al., 1982). This algorithm has been tailored to the HIRS/MSU channel set aboard the TIROS series of satellites.

1.3 Cloud Clearing

Given the high occurrence of cloudiness globally, it is quite likely that a large number of the HIRS fields of view (FOVs) were at least partially filled with clouds. Some fraction of the channel measurements made in these footprints would be contaminated, making them unusable by the retrieval scheme if used directly. This contamination must be corrected before applying the simultaneous retrieval algorithm.

To address this problem we have employed the cloud clearing method utilized by Eyre (1989) which is a variation of the CO_2 absorption technique (Menzel et al., 1983). The objective here is to find that combination of cloud top pressure and cloud fraction which minimizes the difference between the measured radiance and that computed using our forward model. The essential approach is to select a small group of channels sensitive to the presence of cloud. The forward model we use contains a set of predefined cloud top pressure levels which may be determined by the user. If a first guess temperature, water vapor and ozone profile is specified we may compute the expected HIRS radiances for a cloud top which occurs at each of the pressure levels (assuming 100 percent cloudiness.) A least squares approach yields the following relationships:

$$N_c = \frac{\sum\limits_{i} (\check{R}_i - R_i^c)(R_i^o(p_c) - R_i^c)}{\sum\limits_{i} (R_i^o(p_c) - R_i^c)^2} \qquad (1.1)$$

where N_c — cloud fraction (0-1),

$\check{R}_i$ — measured radiance in channel i,

R_i^c — computed clear radiance in channel i,

$R_i^o(p_c)$ — computed overcast radiance in channel i for cloud top

at

pressure p_c.

Defining δ_i as the difference between the measured and computed (cloud contaminated) radiance in channel i,

$$\delta_i = \check{R}_i - R_i$$

where

$$R_i = (1-N_c) R_i^c + N_c R_i^o(p_c) \qquad (1.2)$$

we obtain

$$\sum\limits_{i} \delta_i^2 = \sum\limits_{i}(\check{R}_i - R_i^c)^2 - N_c^2 \sum\limits_{i}(R_i^o(p_c) - R_i^c)^2 \qquad (1.3)$$

Thus the best guess cloud fraction, N_c, is that determined by finding the value of p_c which, via equations (1.1) and (1.3), minimizes $\sum\limits_{i} \delta_i^2$. The cloud-cleared radiances may then be computed from equation (1.2).

2.   APPLICATION/ALGORITHM TUNING

2.1   Data

The data we have used for our retrievals consist of two parts. First, HIRS radiance measurements obtained from the FGGE level I data set

for February 1979 were chosen. Second, since the retrieval scheme requires an estimate of the vertical temperature profile we used the physically based temperature retrievals created at the Goddard Laboratory for Atmospheric Science (GLAS) which were based on the same set of HIRS/MSU radiances. The creation of both data sets has been described by Susskind et al. (1984).

As discussed by Susskind et al. (1984) the HIRS FOVs are organized into 250 km by 250 km boxes each of which are further subdivided into 125 km by 125 km quadrants. Each quadrant contains two sets of radiances corresponding to the average of the warmest and coldest FOVs (based on the 11 μm window radiance) within the quadrant. In addition to the radiance data other ancillary FOV information is included such as latitude, longitude, time, satellite zenith angle, land/ocean flags, etc.

The GLAS temperature retrievals are defined on 66 pressure levels which correspond to the levels used by our forward model and also include relevant information such as retrieval time and location.

The two data sets were in slightly different formats (one sorted by time, the other by latitude) so to avoid a time consuming search procedure during the retrieval stage the data sets were preprocessed to create one file which contained the colocated GLA temperature retrievals and HIRS radiances. Finally, since retrievals were to be performed over ocean areas only, all land data were discarded.

2.2 **Retrieval Application and Tuning**

The following processing sequence was repeated for every water vapor retrieval: First the appropriate quadrant was selected from the set of HIRS radiances by simply choosing the warmest of the four (as defined by HIRS channel 8, the window channel.) Presumably this FOV would be the most cloud free and/or contain the lowest cloud tops. Second, a first guess water vapor profile was chosen based on FOV latitude. This guess was one of five latitude dependent profiles determined from the February 1979 climatology of the NMC FGGE III-a relative humidity analyses. The five profiles correspond to the following latitude zones: north polar,

northern midlatitude, tropical, southern midlatitude, and south polar.
Third, the radiances were cloud-cleared using the algorithm described
earlier with radiances from HIRS channels 7 and 8. Since the cloud
clearing algorithm computes an effective cloud fraction and cloud top
pressure a quality control criterion was applied at this stage such that
any profile with an effective cloud fraction greater than or equal to 90%
was skipped. Assuming this step is completed the iterative process begins
and terminates when a solution is reached.

A certain amount of tuning was done during the development of the re-
trieval procedure. This was primarily driven by limitations on our com-
puter resources which, in turn, were exacerbated by the large number of
profiles which needed to be processed.

We ultimately chose HIRS channels 5 through 12 (see Table A.2) mainly
because we only needed to retrieve water vapor. These channels are most
sensitive to variations in tropospheric temperatures and water vapor. The
forward model was modified so that the radiative transfer calculations
were performed for only the reduced channel set. It was initially hoped
that some of the nonlinearity in the water vapor retrieval problem could
be addressed by updating the radiance derivatives (the A matrix elements)
at each iteration. However, this proved computationally prohibitive
since, in a multiparameter retrieval, each update of the A matrix elements
requires running the forward model once for each retrievable. Instead the
derivatives were calculated once at the initial iteration, with respect to
the first guess profile. Lastly, we addressed the question: when to stop
iterating? Normally, iterations may proceed until a suitable convergence
criterion is satisfied. For example, one may stop if channel radiance re-
siduals fall below some predefined level, or if the rate of convergence
based on updates to the solution vector becomes small enough. In practice
these methods required too many iterations so that a strict limit was re-
quired. Since preliminary experiments using simulated satellite data in-
dicated that a minimum of 5 iterations were usually required to obtain a
reasonable solution, we adopted this as a limit during our retrievals.

Even with all these streamlining efforts the method required approx-
imately 7.5 cpu seconds per retrieval on the AFGL Cyber 860. For a file

containing 6 hours of satellite data which contains approximately 1300 retrievals this amounted to nearly 3 hours of computer time.

3. RESULTS

Once the specific algorithm was settled upon, retrievals were conducted for the period 8-15 February 1979. The scheme actually retrieved a water vapor mixing ratio profile on 6 mandatory levels which were converted to relative humidities upon output. An internal check for supersaturation would reset values back to 100%.

As a means of assessing the relative accuracy of the water vapor retrievals, output relative humidity profiles were compared with radiosondes from the FGGE data base for 0 GMT and 12 GMT time periods. Colocation was performed using a time window of 3 hours and a search radius of 300 km. In Table A.3 root mean square (rms) errors of the retrievals are shown for selected latitude zones (corresponding to the first guess bins) and for the globe as a whole. For comparative purposes the climate variability of relative humidity determined from NMC analyses is also shown. These statistics show that for select latitudes at certain vertical levels the retrieval errors are somewhat smaller than the climate variability. However globally, the rms errors are comparable to climatology, implying that the retrievals improve little upon a climatological guess. It was largely on the basis of these statistics that we decided not to conduct the assimilation experiment since the impact of adding such retrieval data would be insignificant.

There are several likely reasons for the relatively poor quality of the retrievals. As we stated earlier computational limitations required modifications of the algorithm which could adversely affect the retrievals. Limiting the number of iterations to 5 might not allow sufficient convergence, especially if the initial guess profile was far from the true profile. Additionally, calculation of the radiance derivatives at only the first iteration may slow the rate of convergence. Other factors could be equally important. Our first guess was relatively crude, leading us to believe that a guess based upon either a model forecast or a

statistical retrieval would represent a clear improvement. The algorithm used to perform cloud clearing of the radiance data may have also degraded results in manner dependent upon cloud amount and type. Certainly retrievals using cloud-cleared radiances, rather than clear column measurements have been shown to be less accurate (Susskind et al., 1984). An analysis which stratifies errors according computed cloud fraction would help answer this.

As much as anything these results highlight a primary drawback associated with physical retrievals: the large computational demands associated with such schemes make implementation within an operational environment difficult. At the same time statistically based methods, while inherently limited, are seen to yield reasonable results (LeMarshall, 1988) and place a far lighter burden upon computer resources. Increases in available computer power can only make physically based schemes more attractive.

Table A.2 Characteristics of HIRS channels used for water vapor retrievals

Channel	Central Wavelength (μm)	Central Wavenumber (cm^{-1})	Peak of dr/dlnp (mb)	Peak of bd$r$/dlnp (mb)
5	14.00	716.10	475	575
6	13.70	732.40	725	875
7	13.40	748.30	Surface	Surface
8	11.10	897.70	Window, sensitive to water vapor	
9	9.90	1027.90	Window, sensitive to O_3	
10	8.30	1217.10	Lower tropospheric water vapor	
11	7.30	1363.70	Middle tropospheric water vapor	
12	6.70	1484.40	Upper tropospheric water vapor	

Table A.3   Relative Humidity Retrieval Errors Compared with Climate Variability (%)

	Retrieval rms error Latitude Bin				
P(mb)	NP	NM	T	SM	Global
300	32.1	34.6	20.7	32.1	29.8
400	30.9	27.6	23.7	31.1	28.8
500	30.1	28.9	29.5	36.4	32.7
700	38.1	28.6	23.5	34.0	31.5
850	34.3	23.6	29.1	31.2	28.7
1000	40.0	22.5	16.9	27.2	25.6

	Climate Standard Deviation Latitude Bin				
P(mb)	NP	NM	T	SM	Global
300	37.0	34.2	30.7	33.7	34.0
400	33.1	29.1	25.7	28.4	29.2
500	28.4	27.4	26.8	27.6	28.1
700	28.6	28.3	25.7	25.9	27.5
850	25.9	26.3	21.0	21.8	23.8
1000	31.8	28.4	22.6	21.9	25.8

NP: Latitude $\geq$ 50°
NM: 50° > Latitude $\geq$ 25°
T :   25° > Latitude > -25°
SM: -25° $\geq$ Latitude $\geq$ -50°

APPENDIX B One-dimensional tests

To be able to study the behavior of the physical parameterization under controlled conditions, and the interactions between its various elements we have developed a one-dimensional version of the GSM. We decided to use as much as possible of the GSM code, so that our tests would simulate the actual GSM as closely as possible. Since the physics routines are written for single grid points, the task did not present too much difficulty. However, the structure of the code imposed some constraints which are described here.

Throughout the code, the computation is done within a loop over the two hemispheres. Instead of changing each and every loop, we decided to have at least one point in each hemisphere. Furthermore, within the dynamics (and also some physics routines) there is a succession of transformations between spectral coefficients and grid points. The inputs are spectral coefficients, but the physics works in grid point space. The spectral coefficients are complex numbers and the grid point values are real numbers, but both share the same storage space. We found that the easiest way to deal with this problem was to use 2 identical *points* in each hemisphere, corresponding to 2 spectral coefficients, one being $(x/\sqrt{2},0)$ and the other one $(0,0)$, where x is the grid point value.

Most of the dynamics can be eliminated, although we go through the motions of the transformations from spectral to grid point space, and vice versa. The time stepping becomes a simple leap-frog, but we keep the time filtering. The horizontal diffusion is eliminated.

The geostrophic wind is computed at the first time step by

$$u_g - u - \frac{1}{f} \left(\frac{\partial v}{\partial t}\right)_{v.d.}$$

$$v_g - v + \frac{1}{f} \left(\frac{\partial u}{\partial t}\right)_{v.d.}$$

where the terms ( )$_{v.d.}$ are the tendencies due to the vertical dif-
fusion. This minimizes the excitation of inertial waves by an initial im-
balance. In the current version, the geostrophic wind is kept constant
during the integration. We can also impose a vertical velocity.

In total, less than 100 lines of code needed to be changed. Almost
no changes were made to the physics routines, except in the radiation code
where we had to replace some of the dimension specifications by edit sym-
bols, and some of the numerical constants by symbolic constants.

We ran several experiments with the one-dimensional version of the
model. Some used arbitrary initial data. In others, we tried to simulate
observed data. Here we present the results of three of these experi-
ments: Wangara, day 33 (16 Aug. 67) (Clarke et al., 1971), O'Neill, fifth
period (24 Aug. 53) (Lettau and Davidson, 1957) and a case of cold air ad-
vection over ocean (Økland, 1976). A description of the results follows.

1. WANGARA

These data were obtained 16 February 1967 at Hay, New S. Wales,
Australia (34° 31' S, 144° 31' E). We start the experiment with the data
for midnight, local time, and run for 30 hours. The only external forcing
is the geostrophic wind, which is assumed to be constant during the ex-
periment. Considering the coarse vertical resolution of the model, the
results are reasonably good, but some deficiencies are also apparent, as
can be seen on Fig. B.1, which shows the observed and computed surface
temperature.

The initial nighttime cooling of the surface is well reproduced. Un-
fortunately, the model ignores the initial surface temperature and re-
places it, instead, with the lowest atmospheric level temperature. This
is done in subroutine BLFUX, which also imposes the ground moisture arbit-
rarily. Hence the surface temperature, throughout the night, is about 4K
too high. After sunrise, the simulated surface temperature increase is
much slower than observed. This appears to be due mainly to the fact that
the solar radiation is kept constant for 2.5 hours after each radiation
step. In this particular experiment there happens to be a radiation time
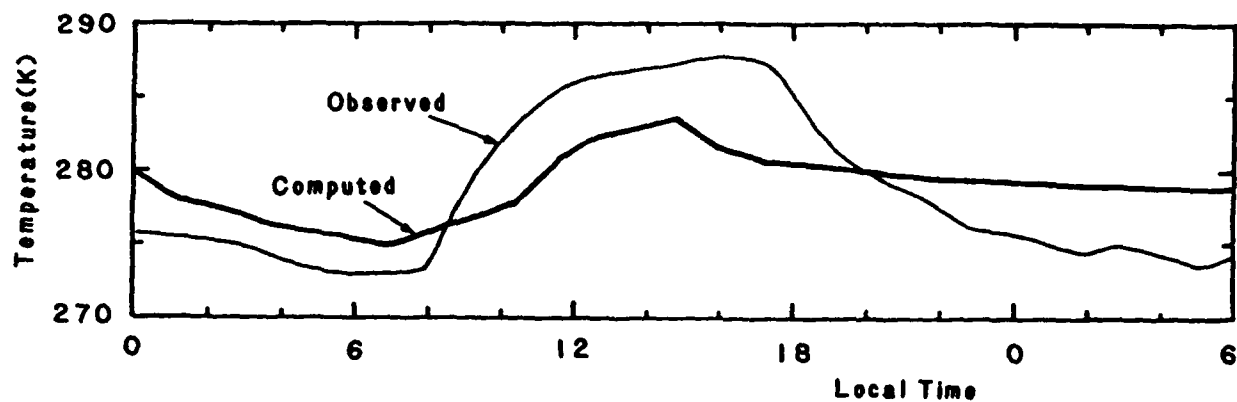
Fig. B.1. *Wangara data, day 33, surface temperature, observed (thin line) and computed (thick line).*
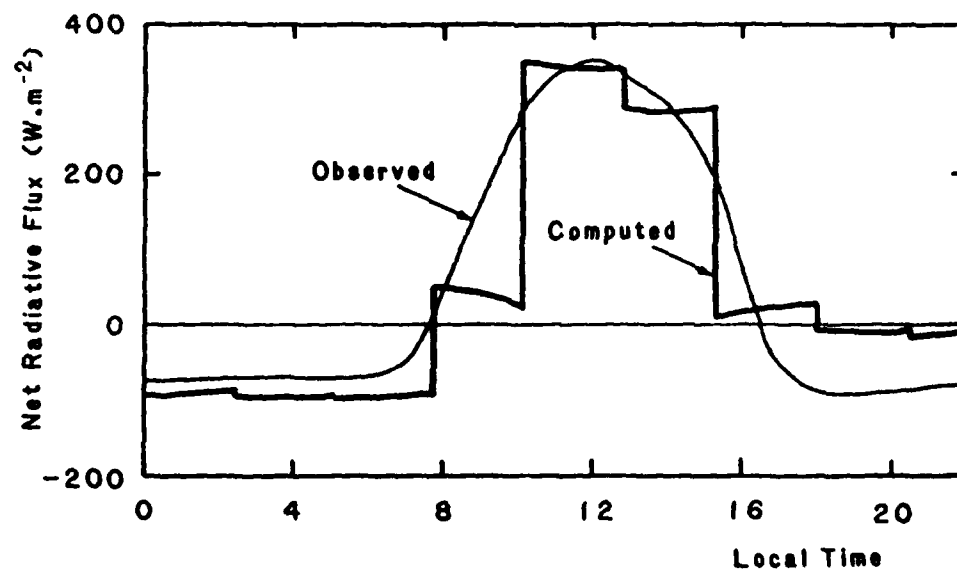

Fig. B.2. *Wangara data, day 33, surface net radiative flux, observed (thin line) and computed (thick line).*

-70-

step shortly after sunrise. The net radiation (Fig. B.2) then remains very small until about 10 am, local time, producing a large deficit in the energy reaching the ground. It would be very useful to devise some time interpolation method that would account for the changing solar zenith angle.

Because of this deficiency, the peak surface temperature is about 4K too low, making the amplitude of the daily cycle only about half of what it should be. During the next night the cooling is much too small, with the net radiation at the ground nearly zero instead of the observed -40 W m^{-2}.

2. O'NEILL

These data were observed 24 August 1959 at O'Neill, Nebraska (42° 28' N, 98° 38' W). This is also a diurnal cycle experiment, but in summer and at higher latitude than Wangara. This time we start from 12:00 local time, and integrate for 30 hours. The results, using the observed profiles as initial conditions, are not good at all. The nighttime cooling, both at the surface (Fig. B.3) and in the atmosphere, is much too weak, and the second day cycle amplitude is almost nonexistent. Similarly, the surface relative humidity (not shown) stays around .35 through the night while it should reach .70 at sunrise. These problems may be due to a cloud layer that forms at the top of a very deep boundary layer in the model. In reality the atmosphere remained cloudless. In the 1D model there is no mechanism for the moisture to escape from the PBL. We tried to reduce the initial moisture profile artificially. In this case the amplitude of the temperature diurnal cycle becomes much better (Fig. B.3). In both cases the latent heat flux is positive (upwards) during the night while, according to all the estimates (by Suomi, Lettau, or Halstead as reported by Lettau and Davidson,1957), it should be downwards. Apparently the model does not allow the formation of dew. In Fig. B.4 only the original computation is shown. The dry case is similar, with a slightly larger flux on the second day.

The surface stress (Fig. B.5) also has an odd behavior, too large during the night (.35 to .4 N.m^{-2} instead of .15 N.m^{-2}), and decreasing

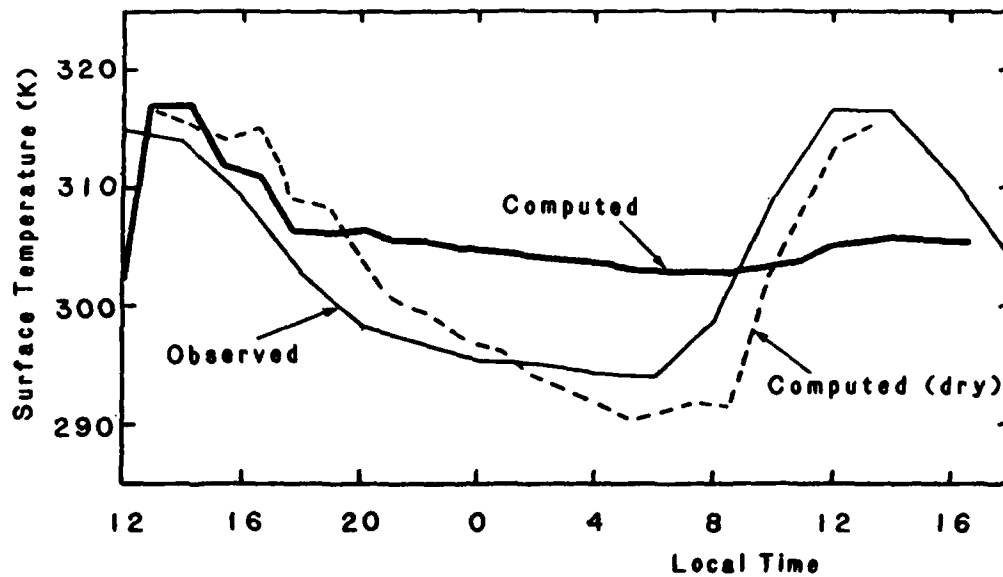Fig. B.3. O'Neill data, period 5, surface temperature, observed (thin line) and computed (thick).
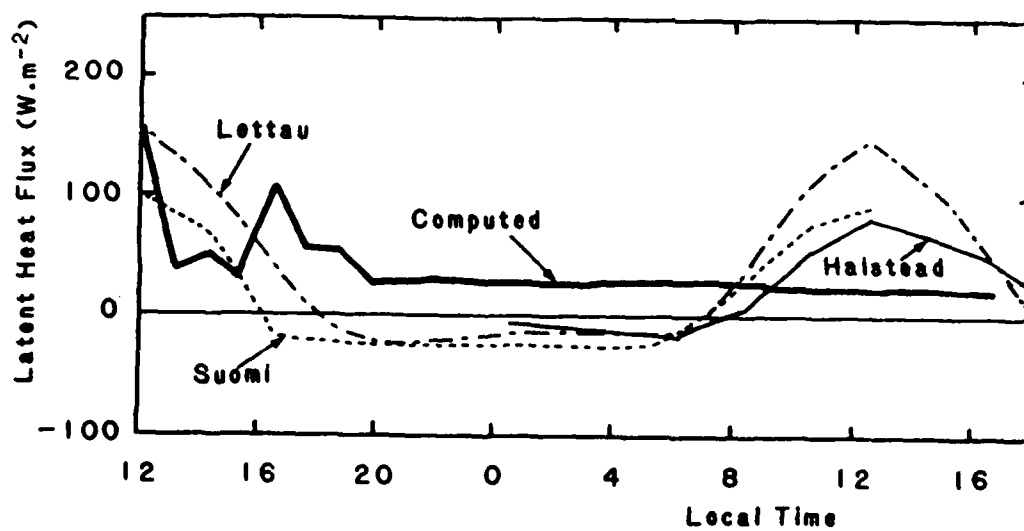
Fig. B.4. O'Neill data, period 5, surface latent heat flux, estimated by Lettau (dot-dash), Suomi (dash) and Halstead (thin solid), and computed (thick).
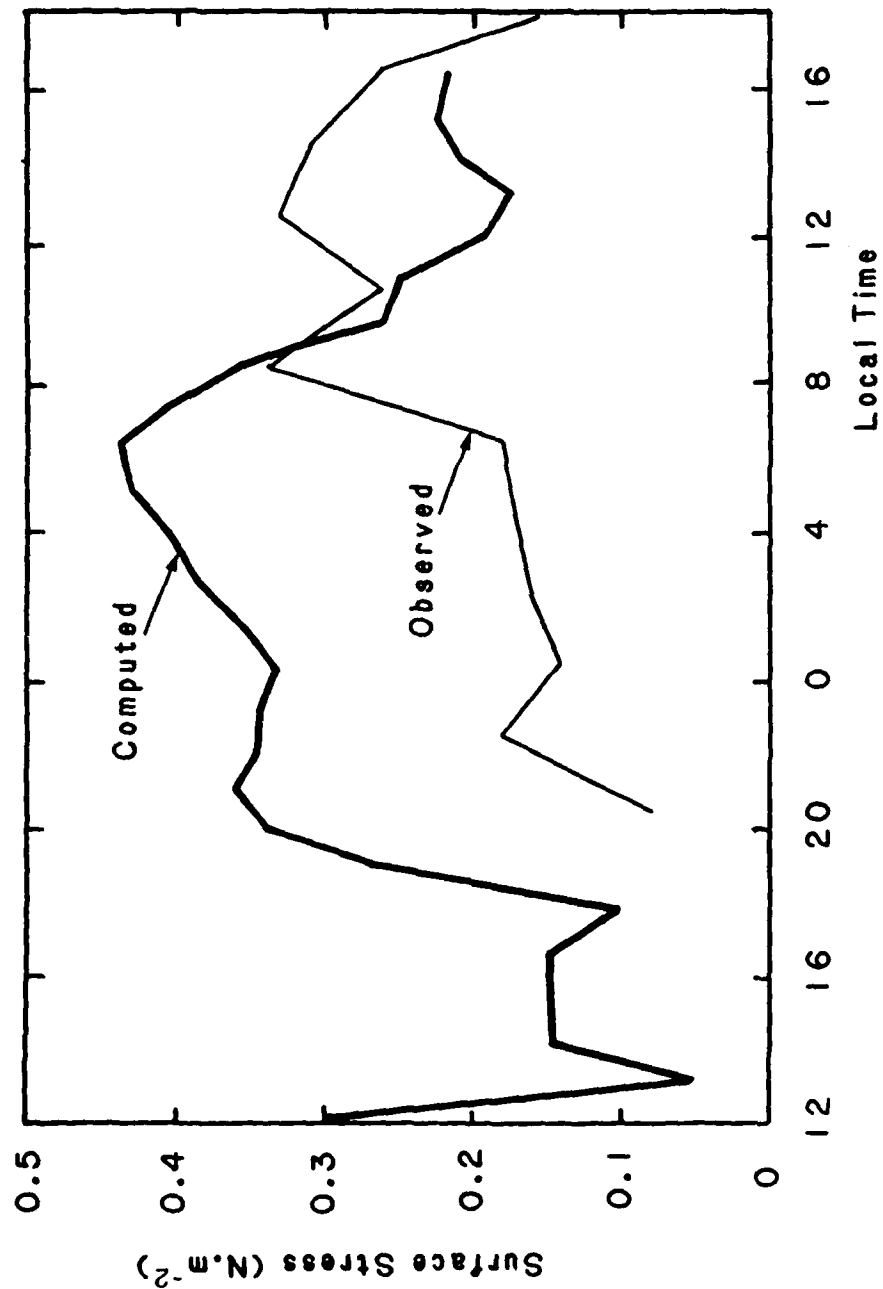
Fig. B.5. O'Neill data, period 5, surface stress, measured by drag plate (thin line) and computed (thick). Measurements are not available before 20:00 local time.

during the morning of the second day. This is the reverse of what was ob-
served and of what should be expected, i.e. that the stress should in-
crease when the PBL becomes unstable during the day.

3.   COLD AIR ADVECTION

In the case documented by Økland (1976), very cold, stable air
crossed the ice edge near Bear Island, south of Spitzbergen, and travelled
south over the Norwegian Sea. Since there was little wind shear, a column
of air can be followed as it crossed the Norwegian Sea. Soundings at
Björnöya (Bear Island), 74° 30' N, 19° 00' E and Bodö (Norwegian coast,
67° 18' N, 14° 26' E) provide profiles for the initial condition and final
verification respectively. For the sake of simplification, we assumed a
constant sea surface temperature of 278 K, instead of trying to simulate
the SST gradient. We also assumed a constant latitude of 70° N.

Results of the experiment are shown on Fig. B.6. The heating and
moistening of the boundary layer during the 18 hours of the experiment are
simulated quite well. Considering the approximation made and the fairly
coarse vertical resolution of the model, the temperature error of 1.5 K in
the PBL and the slightly too high cloud layer are quite acceptable.

A disturbing feature of this experiment, however, is the behavior of
the moisture above the PBL. At the level immediately above the top of the
PBL cloud layer, the humid'-y mixing ratio quickly becomes zero. To try
to understand this problem, we isolated the part of the code that computes
the vertical diffusion in the PBL (in subroutine PBL). We imposed a high
moisture and high diffusion coefficient in the 4 bottom layers, and a low
moisture and zero diffusion coefficient above. After only one time step,
the fifth layer moisture became negative and an oscillation with a $2\Delta z$
wave length appeared above.

-74-
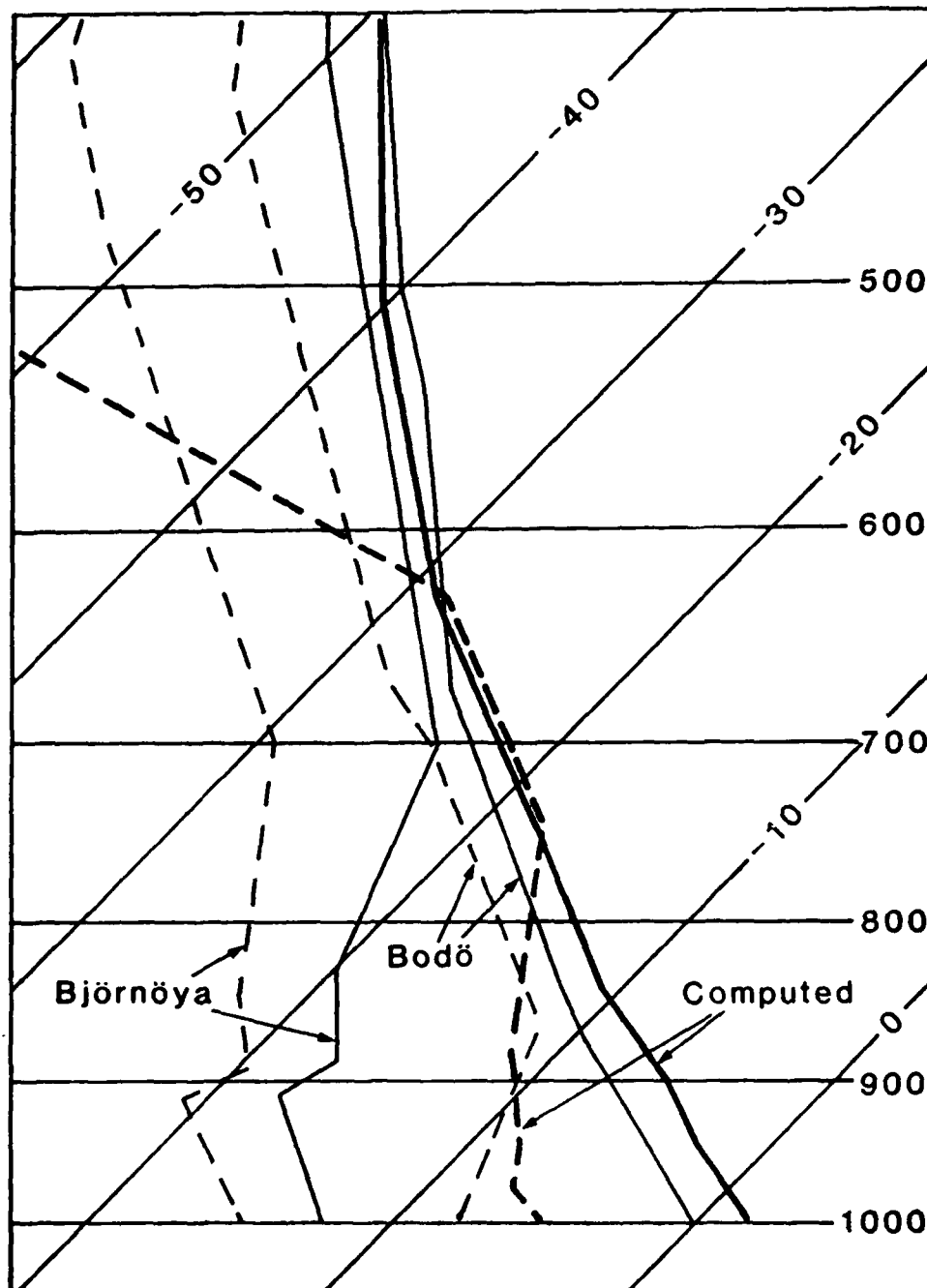
Fig. B.6. Skew T/log p diagram of soundings at Björnöya (21/11/75 at
00GMT) and Bodö (21/11/75 at 12 GMT), and computed profiles
corresponding to the Bodö sounding. Solid lines: temperature;
dashed: dew-point temperature.

APPENDIX C   GSMSAT EXPERIMENT

The impact of the new physics package is the subject of the GSMSAT OSE. In the following, the preparation of codes and data sets for a preliminary 3-day test forecast are described in sections 1 and 2. Implications from this forecast are discussed in section 3, which describes the design of the GSMSAT OSE. The results of the OSE are described in Section 4.

1.   PREPARATION OF CODES AND INITIAL DATA SETS

After consultation with the contract technical monitor, it was decided to use the version of the GSM that has undergone extensive testing by AFGL, rather than the revised version of the new physics package, because the latter was not well enough tested out and understood to be used in the OSE. Because of machine limitations on the Cyber, 12 rather than 18 or more layers were used; the vertical structure is the so-called "12A" structure, with sigma level interfaces at sigma = .1, .2, .325, .45, .575, .7, .8, .875, .925, .965, .99, and 1.0. The GSM code had to be changed slightly for use on the Cyber, to overcome the restrictive limits on extended memory on the Cyber.

The initial data sets for the OSEs were created by interpolating the uninitialized analyses from the spinup run, valid at February 8, 00Z , and June 17, 00Z, respectively, to the 12A sigma structure. Program SITOSI was written for that purpose; it interpolates the spectral coefficients directly from one sigma coordinate system to the other. Note that surface pressure is unaffected by this transformation. To assess the effects of this interpolation, both the original and interpolated analyses were interpolated to mandatory pressure levels and differences were calculated. These differences were found to be quite small: standard deviations of the differences in u and v were less than 1 m/s below 300 mb, less than 4 m/s below 70 mb, and less than 7 m/s at 70 and 50 mb; height differences were on the order of 10 m or less below 200 mb, increasing to 25 m at 150 mb and 140 m at 50 mb; RH differences were on the order of 12 % or less. The resulting spectral coefficient files were then initialized with the NMI appropriate for the 12A sigma structure. The NMI procedure uses

eigenvalues/vectors created by a run of the GECVAR program; the nonlinear tendencies computed by the GSM include boundary layer physics, but not radiative tendencies. Initialization increments were quite small: rms differences of temperature were 0.5 K or less in the lowest 5 layers, and 0.8 K or less at the upper layers, those of rotational wind velocities were 1 m/s or less, those of the divergent wind velocities were 1.7 m/s or less. The analysis procedure was also changed to reflect the 12A vertical structure.

2. THREE-DAY TEST FORECAST

A three-day preliminary test forecast was performed, starting from the initial data set at February 8, 00Z. The primary purpose was to determine the evolution of the surface fields, and whether these surface fields should be allowed to evolve over the course of the one-week assimilation runs. The surface fields were initialized with monthly mean fields. They are: surface temperature, surface roughness, surface albedo, snow depth (expressed as water-equivalent depth), soil temperature and volumetric water content at the two layers of the soil model (the first layer is 5cm thick, the second 95 cm), the surface moisture (defined as an effective surface moisture; over water it is equal to the saturation specific humidity at the surface), and the canopy water content (also expressed as a water depth). Initially, the soil temperature in the two layers is equal, and the surface moisture and the canopy water content is zero everywhere. Of these variables, only the surface roughness over land and surface albedo are left unchanged by the model. The temperature of the shallow soil layer and the moisture content of the two soil layers are reset by the model on the basis the other atmospheric variables at the first time step.

History output containing the surface data was saved every 6 hours during the forecast; the surface data were interpolated to regular latitudes that were very close to the corresponding Gaussian ones, and the resulting data sets were then plotted and analyzed.

The evolution of the surface fields was examined at 6-hourly intervals over the first 30 hours of the forecast, and compared to the change

over the last two days of the forecast (i.e. the difference between 00Z on February 11 and 00Z on February 9). A striking feature of the initial fields are the extreme values used for the snow depth: the snow depth increases from zero to values on the order of 100 cm water equivalent (i.e. 10 m of snow!) in a narrow strip, essentially eliminating the possibility of any significant shifts in the snow line. In fact these values are 10 times too large due to an error in scaling the original snow depth data. In any case, we observed computed 6-hour changes of the snow depth as large as 1 cm of water equivalent in some locations, and decreases over the last two days of the forecast were as large as 12.3 cm water equivalent along the southern edge of the snow sheet.

In general, the changes in the other surface fields are noisy, i.e. they are quite localized; some general observations are possible, however. Changes in surface temperature are confined to the land surfaces, where they reach maximum amplitudes of 40-50 K over 6 hours; the sign of the changes follows the diurnal cycle. The global standard deviation of the 6-hr changes is approximately 8 K. The changes over the first 6 hours of the forecast are slightly larger than, but generally of the same order of magnitude as the changes during the corresponding 6-hour period of the second day of the forecast, thus indicating that there do not seem to be any wild adjustments taking place from the initial values of the surface temperature. The changes over the last two days of the forecast, which are not due to diurnal effects since the same time of day is compared, also reach magnitudes of 40-50 K over snow-covered surfaces, and 25 K elsewhere. The mean change is much smaller than the standard deviation, indicating no systematic trend in the surface temperature. The surface moisture shows extremely noisy 6-hour changes (magnitudes up to 36 g/kg, standard deviation approximately 2 g/kg); the two-day change looks similar to the 6-hour changes, except that there is a large maximum over Australia (+126 g/kg). At the end of the forecast (2/11, 00Z), evaporation in that region is near potential evaporation with a surface temperature of approximately 53°C. While such a high surface temperature might be achieved in a desert, it seems highly unrealistic that it could happen with a saturated soil.

Changes in the shallow soil temperature are diurnally driven, with magnitudes up to 30 K over 6 hours (standard deviation is approximately 4 K); large changes are confined to the snow-free land surfaces. Changes over the last two days of the forecast reach maxima of 18 K (standard deviation=2.4 K), with no significant mean trend. The deep soil temperature has a slower response and shows changes of much smaller amplitude: 6-hourly changes of less than 1 K (standard deviation approximately 0.16 K); the trends are largest in areas with snow cover, where consistent warming occurs. The two-day trend in those areas reaches magnitudes of 4.3 K, with the snow-free areas showing mostly cooling (up to 3.8 K); the globally averaged change is +0.5 K, the global standard deviation is 1.0 K. Of course, the deep soil temperature change in snow-covered areas does not affect the atmosphere at all. The changes in the deep soil temperature are dominated by warming or cooling from below (the bottom temperature is held fixed at +10°C everywhere); noisy patterns in this warming were caused by a noisy initial soil moisture field, which strongly affects the soil thermal diffusivity and which was initialized from atmospheric RH.

The shallow soil moisture, which is expressed as the fraction of volume filled by water, shows 6-hour changes of up to 20%, with a standard deviation of apprimately 2%. No clear patterns are discernible from the plots. The two day changes have magnitudes of up to 30%, with a standard deviation of 3%; the mean change over the 2 days is only 0.6%. As with the soil temperature, the deep soil moisture changes are smaller: 6-hour minima/maxima of -1.7%/+3.7%, with a standard deviation of about 0.4%; the changes seem to decrease during the first forecast day. The two-day changes reach magnitudes of 13% with a standard deviation of 1.1%; the global mean change is -0.07%. Finally, the canopy moisture (which is not allowed to exceed 2mm), shows non-zero changes only over snow-free areas. The standard deviation of 6-hour changes decreases from 0.4mm to 0.3mm over the first forecast day; areas of positive 6-hour changes follow the diurnal cycle, apparently corresponding to dew deposition in the early morning hours. The two day changes, with a standard deviation of 0.7mm, show small scale patterns of increases and decreases; the global mean change is +0.03mm, i.e. essentially zero.

On the basis of this experiment it was decided to use the ground variables history output from the previous 6-hour forecast as the boundary data set for the data assimilation cycle. In addition, the GSM was further modified to not initialize the soil moisture from atmospheric data, but rather use the climatologic values from the boundary data set at the first time step. For this purpose, the monthly mean boundary data sets were updated to contain nonzero soil moisture fields; at the same time, the monthly mean equivalent snow depth data were replaced with more reasonable values. Both of the new surface fields were acquired recently from NMC. The soil moisture fields are far less noisy than those that were derived from the atmospheric RH. The new snow cover fields have significantly larger values over some areas (e.g. over Greenland, Antarctica, and the Himalayas), and a more gradual increase from zero to nonzero values. The only exception to this rule is the shallow layer temperature which remains set to the lowest atmospheric layer temperature at the start of each forecast. This is to avoid any possible large discrepancy between ground and atmospheric temperatures, resulting from the analysis of the atmospheric temperature. Such a discrepancy could produce unrealistic radiation fluxes, which would be held approximately constant until the next radiation step, three hours later. For a similar reason, during the very first cycle the radiation routines do not use the monthly mean surface temperature as the skin temperature during the first 3 hours of the forecasts, but use a temperature extrapolated from the lowest two sigma levels instead.

On the advice of Don Norquist, who found only small differences between the adiabatic and diabatic NMI for this version of the GSM, it was decided to use the adiabatic NMI for the OSE.

3. ANALYSIS METHODS

The ASAP mass and moisture OI programs with the updates described in Hoffman et al. (1988) were used, along with an updated version of the pre-analysis program which was modified to ensure first guess RH values between 0 and 1. The only other changes to the analysis codes involved the different vertical structure of the data sets. The initial analysis error

fields were those from the spinup OSE, interpolated to the new sigma structure.

Unfortunately, one change which should have been made to the analysis procedure was overlooked. As a consequence, the final interpolation of analysis increments from the layer-layer temperatures to the layer temperatures made use of the old vertical structure. When this was discovered, time did not permit rerunning the assimilation experiments. We did estimate the effect of this error by rerunning one cycle. The lowest layer, which is very thin in the 12A vertical structure, exhibited large impacts. Differences in the other layers were quite small. In temperature, the lowest layer changes were of $O(1.5 K)$, changes in other layers were $O(0.2 K)$ and the effect on heights was of magnitude 2.5 m throughout the atmosphere. The temperature changes are also reflected in specific humidity, although relative humidity is unchanged. Changes in the lowest layer were $O(1 g/kg)$ which is about 15% of the variability of specific humidity in that layer. At higher levels the changes were small.

There was one other difficulty involved with converting the analyzed height increments to temperature increments. Because the lowest layer is so thin, in cases when there were significant differences between the two lowest height analyses, the hydrostatic computation of the layer-layer temperatures could yield very large values. Layer-layer temperature corrections with magnitudes in excess of 30 degrees were reset to zero.

4. OSE RESULTS

4.1 Evolution of the surface quantities

The evolution of the surface quantities during the assimilation period and during the day 3 and day 7 forecast was carefully analyzed for the February time period, mainly to assess the performance of the new physics package of the GSM. As described earlier, the surface quantities are allowed to evolve during the assimilation period, with successive 6-hour forecasts being initialized with the surface quantities predicted by the previous forecast. The effect of the analysis and initialization steps on the surface fields are through the changes made to the lowest level atmospheric variables.

An examination of maps of surface fields over one diurnal cycle (not shown) shows the expected changes in surface temperature, effective surface humidity (a measure of evaporation), and shallow soil temperature, which all follow the changes in the sun's radiation impinging on the surface. The largest changes in soil moisture occur underneath the snow cover, caused by melting and subsequent percolation and diffusion of water. These changes in soil moisture result in greater soil thermal diffusivities, which in turn result in large changes in the deep soil temperature. None of these changes are relevant to the atmosphere, however, at least as long as the snow cover does not melt away completely. Since the initial snow cover has a rather steep gradient at its edges, however, the snow line does not change appreciably throughout the assimilation or forecast periods.

Fig. C.1 shows globally averaged values of the surface temperature, shallow and deep soil temperature, and 6 h accumulated precipitation throughout the assimilation period. Fig. C.2 shows the corresponding plots for the day 3 and day 7 forecasts. A marked diurnal cycle is present in the global values of surface and shallow soil temperature, a reflection of the uneven distribution of land masses: maxima (minima) occur at 12 GMT (00 GMT), which corresponds to early to late afternoon (midnight to early morning) local time over the large land masses of Western Europe, Western Africa, and South America. The shallow soil temperature can be seen to lag the surface temperature by approximately 3 hours. The diurnal cycle of surface temperature has a smaller amplitude in the forecasts than in the assimilation, mostly because the minima are not as pronounced.

Neither the assimilation nor the forecasts exhibit any noticeable trend in surface temperature. The shallow soil temperature, however, shows a definite warming trend during the assimilation, which is even stronger than the warming trend of the underlying deep soil layer; in contrast, the shallow soil temperature during the forecasts shows no systematic warming, even though the deep soil layer warms faster than during the assimilation. The warming trend of the deep soil layer is an artifact of the assumption of a globally uniform temperature of 283 K underneath the

-82-
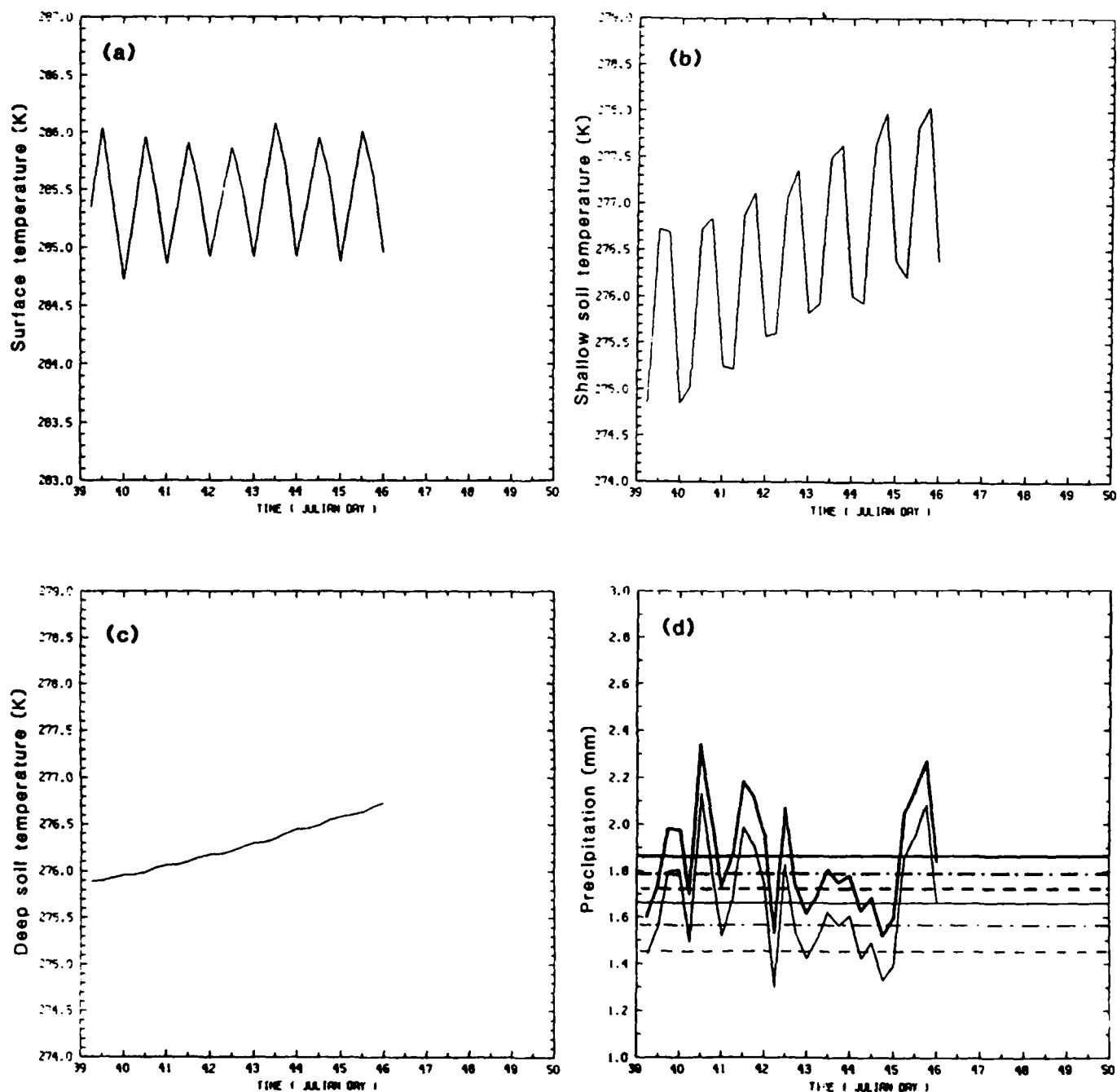
Fig. C.1: Evolution of globally averaged surface and soil quantities
during the assimilation period in February. Values are plotted every
6 hours. Surface temperature (a), shallow (b) and deep (c) soil tem-
perature, and 6-hourly precipitation (d). In (d), the solid thick
(thin) line is the total (convective) precipitation; the straight
horizontal lines are the corresponding average values over the as-
similation period. Also shown are the average values from the day 3
(day 7) forecast, as dashed (dash-dotted) lines.

-83-
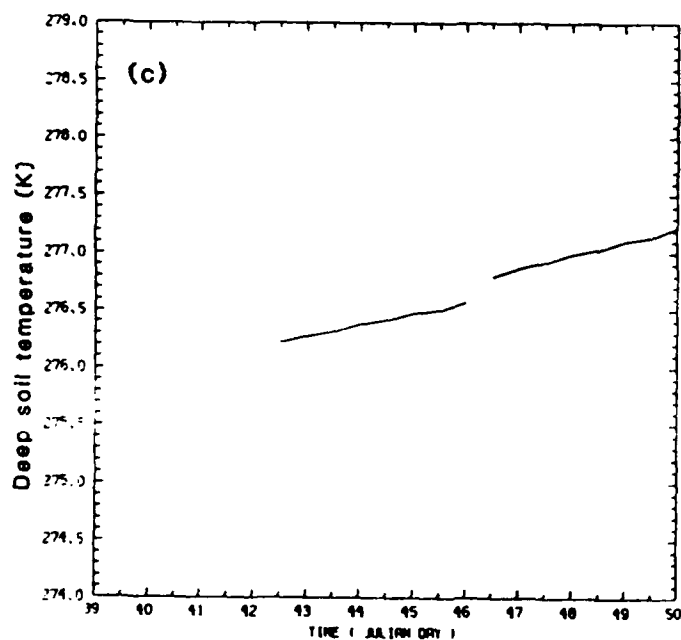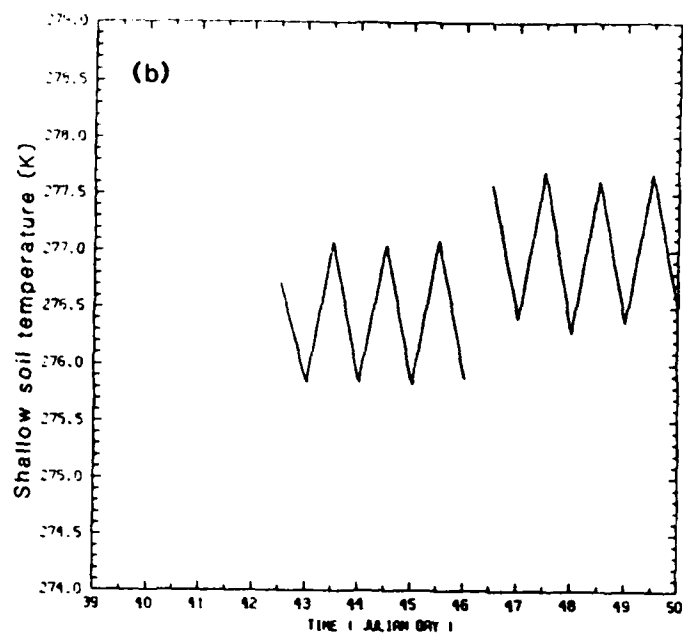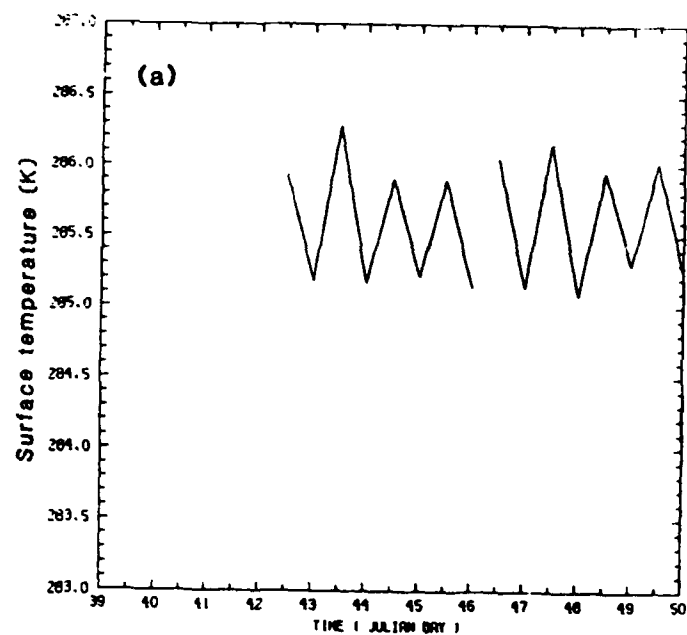
Fig. C.2: As Fig. C.1a-c, except for the day 3 and day 7 forecasts.
Values are plotted every 12 hours.

deep soil layer; since a large part of this change takes place underneath snow, it is not a serious problem. The differences in the shallow soil temperature evolution are caused by the fact that during the assimilation, it is reinitialized every 6 hours (in the absence of snow cover) to the lowest level atmospheric temperature, which was modified by the analysis. The 6-hourly precipitation during the assimilation is quite variable, but shows no trends or definite periodicities. The absolute values are quite large (the climatological value for February given by Jaeger, 1976, is 2.8 mm/day), with the largest fraction being due to convective precipitation. The average 6-hourly precipitation during the assimilation is larger than the corresponding value for the day 3 and day 7 forecasts, indicating a possible overprediction of precipitation during the spin-up period of any forecast; however, the precipitation at the end of the 6-hour forecast started from day 3 is less than the average values for the 4-day forecast from day 3, or even the first 12 hours of that forecast, indicating the reverse spin-up problem. Thus, any spin-up variations of globally averaged precipitation are well within the day-to-day variability.

The evolution of the effective surface humidity (not shown) closely parallels that of the surface temperature. The changes in the globally averaged soil moisture values are strongly influenced by the areas covered by snow and are not shown here. Finally, the canopy water content, which is initialized with a globally uniform zero value, increases steadily throughout the first 1 1/2 days of assimilation to a value of approximately 0.65 mm, and does not depart much from that for the rest of the assimilation or the forecasts.

In summary, the model produces reasonable diurnal variations of surface values, and no large scale trends. The soil variables show a great sensitivity to snow melt, which does not affect the assimilation or the short term forecasts, but which might be problematic in forecasts long enough to have significant shifts in the snow line. The globally averaged precipitation, which is mostly convective, is too large (compared to climatology) by a factor of approximately 2.5. Maps of precipitation all show a distribution similar to that shown in Fig. C.3, which is the pre-
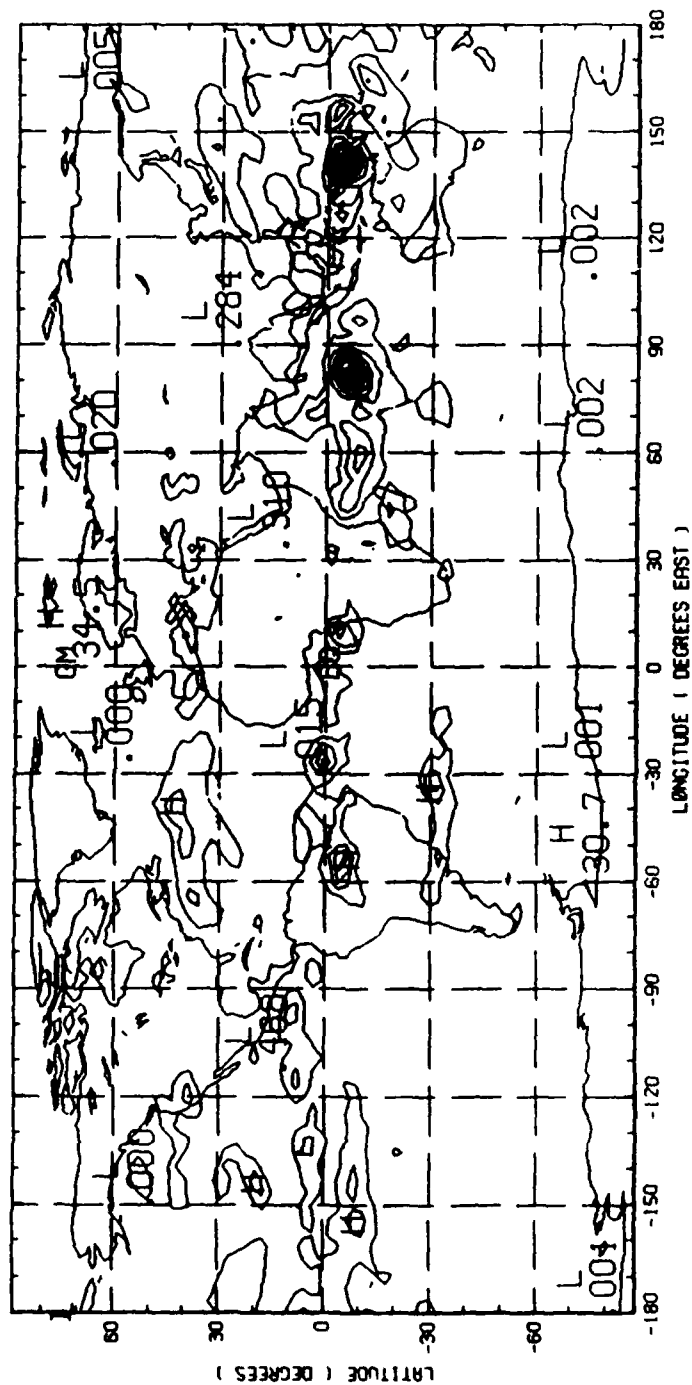
Fig. C.3: Map of total precipitation at the of the 4-day day 3 forecast. The contour interval is 50 mm. the maximum value is 610 mm.

cipitation at the end of the day 3 forecast. The significant precipitation is concentrated along the ITCZ, but the picture is dominated by very localized maxima. The maximum value shown in the figure is 619 mm over the 4 days of the forecast.

4.2 Atmospheric variables

We made a comparison of the GSMSAT and STATSAT analyses of 500 and 1000 mb heights, 850 mb relative humidity and 200 mb winds. These two experiments used exactly the same data, including statistically retrieved temperature profiles from TOVS. The small differences in the analysis systems, described in Section 2.2, were shown to have very small effects on the analyses. Thus the main difference between the two experiments is the use of the new physics parameterization in the GSMSAT experiment. A subjective evaluation of the fields indicate that the differences are very slow to establish themselves in the analyses. Only toward the end of the assimilation periods can one identify differences that could have synoptic significance.

Examples of GSMSAT analyses and GSMSAT-STATSAT differences are shown in Fig. C.4, for 00 GMT 15 February. The largest differences in the Northern Hemisphere are associated with the trough-ridge-trough structure over the North Pacific, which has a larger amplitude in GSMSAT, and the cut-off low off the east coast of North America, which has a more pronounced trough axis in GSMSAT. In the Southern Hemisphere, the largest differences are due to more pronounced ridges in GSMSAT, one to the east of South America, one south of Madagascar. Figs. C.4e and C.4f show that these features are present in the GSMSAT-NMC difference maps, as well; there are, however, numerous instances where GSMSAT and STATSAT agree with each other, but differ from NMC. Presumably, the former are dominated by model differences, the latter by analysis system differences, although such a separation is not strictly possible in as strongly a nonlinear a system as the atmosphere.

This subjective evaluation is confirmed by the rms differences between the GSMSAT and STATSAT analyses. Fig. C.5 shows the rms differences between the 500 mb height analyses, separately by season. Values for the

Fig. C.4: GSMSAT 500 mb height analyses (a,c) and GSMSAT-STATSAT differences (b,d) for the Northern (a,b) and Southern (c,d) Hemisphere. GSMSAT-NMC difference maps are shown in panels e and f for the Northern and Southern Hemisphere, respectively. The contour interval in the analyses is 80 m, in the difference maps it is 40 m.
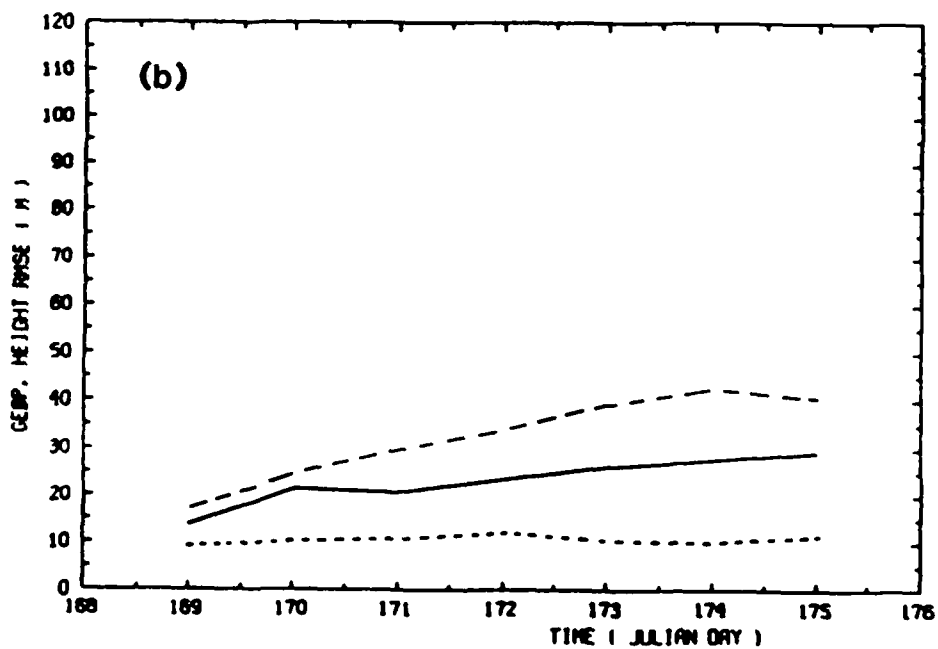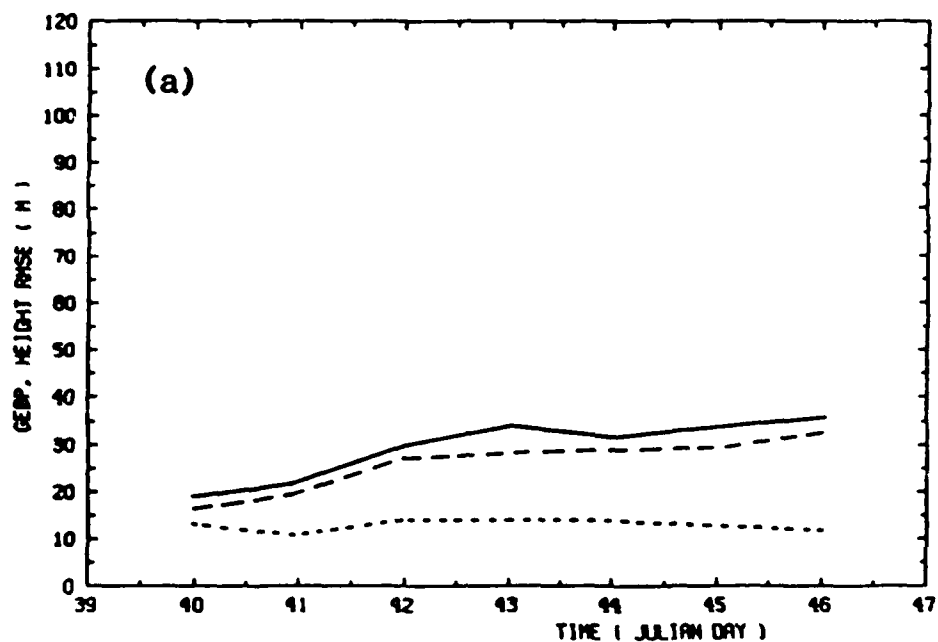
Fig. C.5.   Rms differences between GSMSAT and STATSAT for February (a)
and June (b) for the Northern Hemisphere (solid), Southern Hemisphere
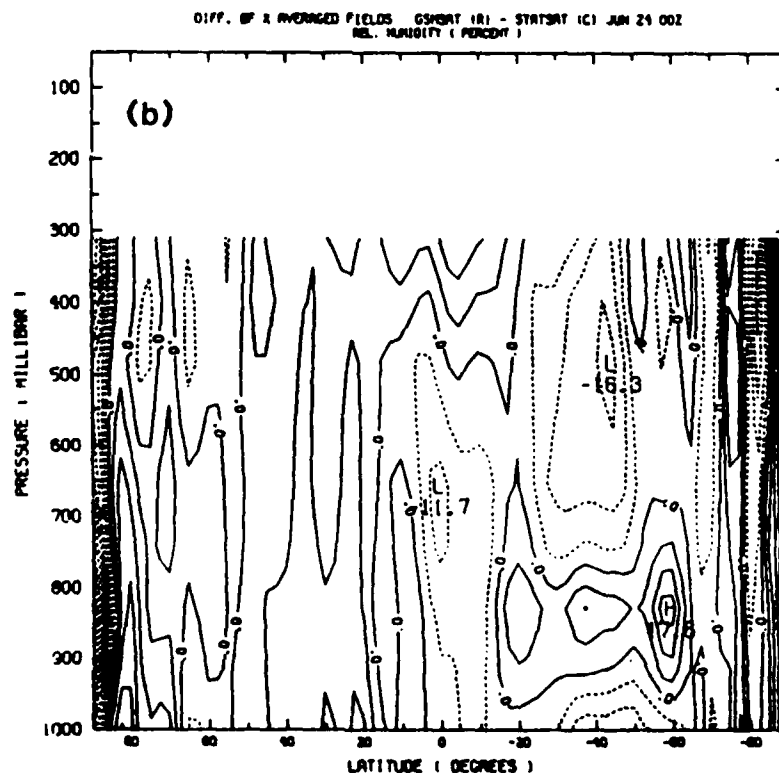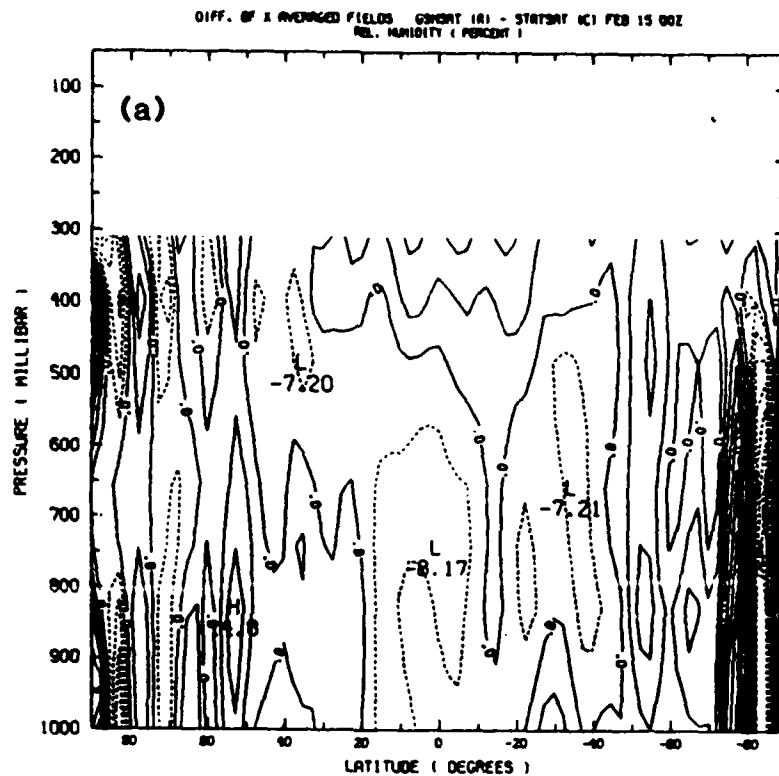(long dashes) and tropics (short dashes).

Fig. C.6. Meridional cross sections of zonal mean RH differences between GSMSAT and STATSAT for 00 GMT 15 February (a) and 00 GMT 24 June (b). The contour interval is 5%.

tropics and both hemispheres' extra-tropics are shown. Except in the tropics, where differences are relatively small and constant, the rms differences keep growing during the assimilation periods. The differences are larger in the winter hemisphere. Differences between the two OSE analyses at the end of the assimilation period are still smaller than typical STATSAT-NMC differences.

Difference maps of the 850 mb relative humidity for February (not shown) indicate that GSMSAT tends to be drier at that level than NMC, particularly over the oceans, more so than STATSAT. Examination of zonal mean meridional cross-sections (Fig. C.6) reveals negative GSMSAT-STATSAT differences below 500 mb for almost the entire 20° N to 40° S latitude band in February; positive differences are most pronounced at 850 mb, near 50° N. In June, GSMSAT is again more humid at 850 mb in the winter hemisphere (60° S); large negative differences dominate near the equator (700 mb) and at 50° S (500 mb).

If we compare the analyzed fields to radiosonde observations, (Fig. C.7) it is hard to declare one analysis better than the other. This is consistent with the difference maps (Fig. C.4), which shows generally small differences over areas covered by radiosondes. The height forecast errors are inconclusive, as well, showing STATSAT more skillful than GSMSAT for some forecasts (day 3, 7 in June), less skillful for some (day 3 in February), and roughly the same for others (day 7 in February). One aspect of our procedure for computing RAOB statistics is that analyzed heights were anchored to RAOB heights, so that "height" statistics actually refer to thickness statistics. Forecasts of 850 mb RH are of roughly the same quality or less skillful in GSMSAT in February and June. To the extent that the new GSM physics is more sophisticated than that used in the STATSAT experiment, it is disappointing to see that the forecasts of relative humidity are, if anything, worse in the GSMSAT experiment. However, it should be stressed that the new physical parameterizations have only recently been integrated with each other and the model, our tuning has been relatively superficial and the effect of the 12A vertical structure on the analysis procedures may be detrimental. In view of all this and in view of the importance of the model first guess in data sparse regions, we conclude that the current experiments are in fact encouraging.
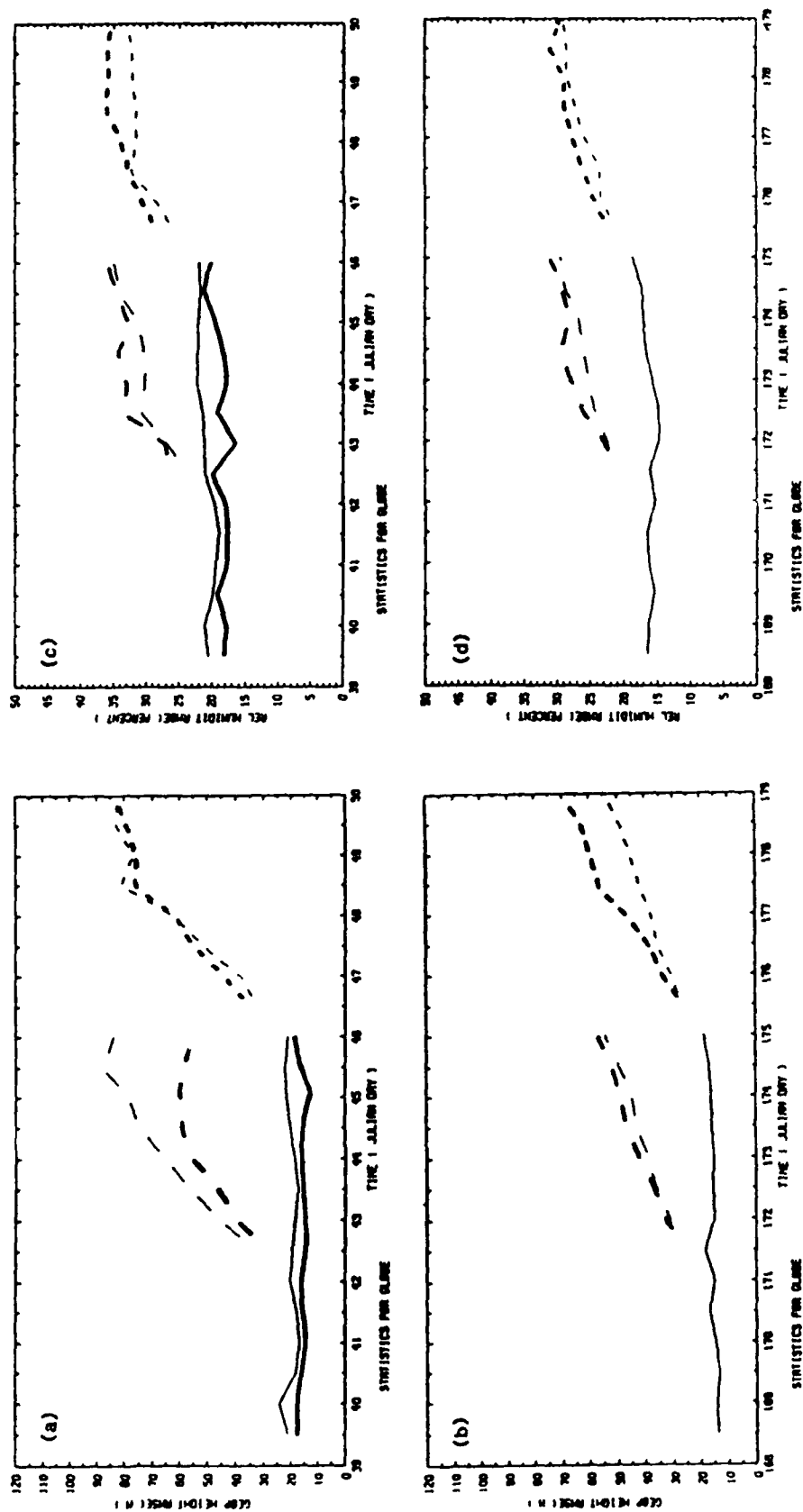
-91-

Fig. C.7. Rms differences from RAOBs for analyses (solid lines) and forecasts (dashed lines). Geopotential height at 500 mb for February (a) and June (b), 850 mb RH for February (c) and June (d). STATSAT is shown in light, GSMSAT in heavy lines. For June, GSMSAT analysis statistics are not shown.

APPENDIX D   PARAMETER OPTIMIZATION

The problem of model parameter optimization or, as it is commonly called, model tuning is to find the set of model parameters such that, over a large ensemble of forecasts, some measure of the forecast error is minimized. The parameters are physical constants that are not well known, such as the heat capacity of the ground or the entrainment rate of cumulus clouds, quantities that appear because of a closure problem, such as eddy diffusivities, or because of a mathematical simplification, such as broad band emissivities. The model state is defined by the prognostic variables (temperature, wind, moisture, etc) or any quantity that can be derived from them, and the forecast error could be, for example a combination of root mean square errors for several variables.

This is a problem in optimal control: the control variables (the set of model parameters) affect the output variables (the forecast state) and must be adjusted so that a cost or objective function (the forecast error) is minimized. There is however, a difference with the usual optimal control problems. Normally, there is a direct feed-back between the objective function and the control variable, either instantaneous, or with a small delay. In the weather forecast problem the feed-back is not direct and cannot be instantaneous. Once parameters are set at the beginning of a forecast, they are not changed during that forecast. Only after many forecasts with one set of parameters can one compute meaningful error statistics that can be used as a feed-back to modify the parameters.

There are a number of major difficulties associated with model tuning. First of all, it is a multidimensional problem since the forecast models have many parameters that need adjusting. In addition, the relationships between the model state and the parameters are highly non-linear; in fact they cannot be written analytically since they are the result of the integration of a set of nonlinear equations. Finally, imperfect parameters are not the only sources of errors in the forecasts: the initial conditions for the integration are also not perfect, and there may be atmospheric processes which the model is incapable of simulating. In the following we examine these difficulties and suggest ways of approaching them.

Let us call $v - v_1, \ldots, v_n$ the vector of all the variables that define a forecast state, and $\hat{v}$ the corresponding observed atmospheric state. Since we will be considering ensembles of forecasts, we will use superscripts to designate individual forecasts. So, $v_i{}^j$ is the ith variable of the jth forecast. Let us define a measure of the forecast error:

$$S - \sum_{i-1}^{n} \sum_{j-1}^{J} w_i{}^j \ (v_i{}^j - \hat{v}_i{}^j)^2 \Big/ \sum_{i-1}^{n} \sum_{j-1}^{J} w_i{}^j \tag{1}$$

where J is the total number of forecasts and $w_i{}^j$ is a weighting factor needed to scale the different variables, but which will also be used later to select forecasts.

Let us now call $\alpha - \alpha_1, \ldots, \alpha_m$ the vector of all the adjustable parameters of the forecast model. Formally, we can write that the forecast is a function of these parameters:

$$v^j - v^j(\alpha), \tag{2}$$

even though we cannot write this relationship explicitly since the forecast equations do not have an analytical solution. Nevertheless we see that, through v, S is a function of α.

Our purpose is to find the value of α which will minimize S. There are many methods to search for the minimum of a function, but implementing any method blindly would be very expensive. Most methods require the estimation of the gradient of the function S with respect to the parameters. In this case, since the relationship between S and α is not analytical, we have to estimate the gradient by finite difference, repeating each forecast m times, varying slightly each parameter in turn. So, at every step of the optimization, we have to do $(m+1) \times J$ integrations. This is quite large and, for a three-dimensional model, would be prohibitive. In addition, the sensitivity of a function such as S, which is an average over many forecasts, to the model parameters is generally rather small.

That is because this sensitivity depends on the meteorological situation. For one situation in which the influence of a particular parameter α_k is large there may be many in which this influence is very small. In other words, the minimum of the function S, in the parameter space, is rather flat, the convergence toward the minimum will be slow, and its result uncertain. We need to find ways to decrease the number of computations needed, and to sharpen the minimum of the objective function.

We note, first of all, that it would be very advantageous to do this optimization in one dimension. Indeed, because of the stratification of the atmosphere, the physical processes mainly involve exchanges between different layers or between the atmosphere and the ground. Hence it seems reasonable to assume that parameters optimized for a one-column model would also be applicable to a three-dimensional model. That is true at least for the atmospheric parameters. The ground parameters, which vary geographically, would obviously need a global optimization, but techniques developed in one dimension may be useful in three dimensions as well.

Using (1) for the definition of S, its gradient with respect to α can be written:

$$\nabla_\alpha S = \sum_{i=1}^{n} \sum_{j=1}^{J} w_i^j \, (v_i^j - \hat{v}_i^j) \, \frac{\partial v_i^j}{\partial \alpha_k} \tag{3}$$

Instead of computing $\nabla_\alpha S$ by finite difference as suggested above, we can use these sensitivity runs to compute the derivatives $\partial v_i^j/\partial \alpha_k$. What can we expect for the derivatives? Some variables v_i may be quite insensitive to the parameter α_k. For these variables, $\partial v_i^j/\partial \alpha_k$ would be small or zero for all j forecasts. Some other variables, which are sensitive to α_k, will have nonzero $\partial v_i^j/\partial \alpha_k$. However, in general, it will depend on the meteorological situation and may have quite different values depending on the forecast case. If we could plot $\partial v_i^j/\partial \alpha_k$ vs. v^j for a particular α_k (this would be an (n+1)-dimensional plot), we would find a lot of scatter. If the parameterization methods are well chosen, we should be able to find, for each α_k, a subset of the v space, defining what we call

a weather regime, in which $\partial v_i{}^j / \partial \alpha_k$ for at least one of the v_i's is large but with a small scatter. This is the essence of our proposed method.

Let us call δv such a subspace. Large values of $\partial v_i{}^j / \partial \alpha_k$ mean that the variable v_i, in that weather regime, is sensitive to the parameter α_k, and a small scatter in those values indicates that this sensitivity is fairly systematic. If such a δv can be found, then we can select only the forecasts for which v falls within δv to compute the gradient (3) needed for the minimization.

The optimization method can be summarized as follows in the one-dimensional problem. The first task is to identify all the adjustable parameters in the forecast model and their possible range. Then, two types of variables need to be chosen: the variables that will be used to verify the forecast, which we will call output variables, and the ones that will be used to define weather regimes, which will be called state variables. Output and state variables do not necessarily need to be the same. Both types, however, need to be routinely observed, or at least such that they can be derived from routine observations, and they must also either be model variables or easily derived from them. It is imperative to choose the variables that define a weather regime very carefully, in such a way that relationships between them and the parameters are as clear as possible. Just defining an atmospheric state by the values of temperature, moisture, etc., at a number of levels may not be the best way. Some gradients, like a temperature inversion, or the temporal change of a quantity, such as the amplitude of the daily cycle, may be more appropriate variables to use. Also, variables that are independent from each other, such as empirical orthogonal functions, should be considered.

Once parameters and variables have been defined, a large number of sensitivity experiments must be undertaken. Using archived data as initial conditions, a series of forecasts are made, changing slightly each parameter in turn, and the derivatives of every variable with respect to every parameter are computed.

Each derivative $\partial v_i{}^j / \partial \alpha_k$ is then examined to find which output variables (v_i) are most sensitive to the parameters (α_k). Since different

variables (say, temperature and relative humidity) cannot be directly compared, the derivatives need to be scaled and reduced to a common dimension An appropriate scaling factor could be the root mean square change of the variable during the forecasts. If several variables are found to be sensitive to the same parameter, a correlation analysis should be performed, and only variables that are independent from each other should be retained. This examination of the sensitivity of the variables to parameters leads to the definition of the weather regimes which are then used for the optimization.

Before the changes in the parameters are implemented, forecast tests should be made to verify that the changes actually reduce the mean forecast error. In these tests, all forecasts should be used (or a random sampling of them), and not just those in the selected weather regimes. If an improvement has been achieved, then one can proceed with the next step of the minimization procedure. If no such improvement occurs, this indicates a problem with the parameterization scheme itself.

The use of different weather regimes to optimize different sub-sets of parameters achieves the goals that were set out. It separates a large multidimensional problem into several simpler problems. Only at the first step of the optimization do we need to compute the sensitivities of all the variables for all the forecasts. Once the weather regimes are chosen, only the forecasts that fall within these regimes need to be run for the successive steps, thus reducing the computing cost considerably. Finally, by defining the weather regimes as those which are sensitive to the parameters, the sharpness of the minimum of the objective function should be much improved.

Let us now go back to the three-dimensional problem. There are two kinds of parameters: some are global constants, and others depend on the geographical location. The latter kind are the parameters that describe the physical characteristics of the ground. Since a typical model has thousands of grid points, there is little hope of optimizing these thousands of parameters independently from each other by a systematic method as described above. The only possibility is to relate these parameters to some known quantities such as the orography or the vegetation type and

cover. The constants in these relationships then become global adjustable parameters.

Can we use the method developed here to optimize global constants in a three-dimensional model? As far as the definition of the weather regimes is concerned, one can consider a global model as the juxtaposition of a large number of one-column models. Hence sensitivity tests on a few global forecasts should be sufficient to define all the necessary weather regimes. By making the weights $w_i{}^j$ in (1) non-zero only for the grid points that fall within a defined weather regime, we build a measure of the forecast error which is very sensitive to a particular (or a few) parameter(s). Thus, instead of using a mean error defined over many forecasts, with a poor sensitivity, we define it over many selected points of a few forecasts and increase its sensitivity. We still need to do $(n+1) \times J$ forecasts to compute the gradients of the objective functions, but now the number of forecast cases, J, is very small. One or two global forecasts may be enough.

In a way, the optimization method developed here is similar to the traditional tuning in which, by trial and error, one tries to adjust some parameters until one simulates correctly a case which is believed, a priori, to be sensitive to these parameters. The main difference is that the proposed method performs the adjustment systematically, and on the basis of routinely available data.

ACRONYMS

3DNEPH	three-dimensional nephanalysis
AER	Atmospheric and Environmental Research, Inc.
AFGL	Air Force Geophysics Laboratory
AFGWC	Air Force Global Weather Center
AMTS	advanced moisture and temperature sounder
ASAP	AFGL Statistical Analysis Program
CDW	cloud drift wind
DMSP	Defense Meteorological Satellite Program
DWL	Doppler wind lidar
ECMWF	European Centre for Medium-Range Weather Forecasts
EOF	empirical orthogonal function
ERT	Environmental Research and Technology, Inc.
FGGE	First GARP Global Experiment
FOV	Field of view
GDAS	Global data assimilation system
GFDL	Geophysical Fluid Dynamics Laboratory
GLA	GFSC Laboratory for Atmospheres (NASA)
GLAS	GSFC Laboratory for Atmospheric Sciences
GLASAT	OSE using GLA physical retrievals
GSM	global spectral model
GSMSAT	OSE using new physics parameterizations
GWE	Global Weather Experiment
HIRS	high-resolution infrared sounder
ITCZ	Inter-tropical convergence zone
JPL	Jet Propulsion Laboratory
MIT	Massachusetts Instititute of Technology
MSU	microwave sounder unit
NASA	National Aeronautics and Space Administration
NEPHSAT	OSE using cloud derived RH profiles
NESDIS	National Environmental Satellite and Data Information Services (NOAA) (Washington, DC)
NMC	National Meteorological Center
NMI	normal mode initialization
NOCON	OSE using no conventional data
NOCOR	OSE using no conventional data in a region
NOSAT	OSE or OSSE using NO SATellite data

NWP	numerical weather prediction
OESD	observing error standard deviation
OI	optimal interpolation
OLS	Operational Linescan System
OSE	observing system experiment
OSSE	observing system simulation experiment
PBL	planetary boundary layer
RAOB	radiosonde observation
RH	relative humidity
RTNEPH	real time nephanalysis
SASC	Systems and Applied Sciences Corporation
SOP	Special Observing Period
SSM	special sensor microwave
SSM+TOVS	OSSE using SSM and TOVS data
SSM/T-1,2	special sensor microwave/temperature
SSMSAT	OSSE using SSM data
SST	sea surface temperature
STATSAT	OSE or OSSE using conventional satellite data
TOVS	TIROS operational vertical sounder
VAS	VISSR atmospheric sounder
WINDSAT	OSSE using simulated DWL data
WMO	World Meteorological Organization

REFERENCES

Atlas, R., E. Kalnay, W.E. Baker, J. Susskind, D. Reuter, and M. Halem, 1985: Simulation studies of the impact of future observing systems on weather prediction. In _Preprint volume. Seventh Conference on Numerical Weather Prediction_. American Meteorological Society, Boston, MA, pp. 145-151.

Aune, R. M., L. W. Uccellini, R. A. Petersen, and J. J. Tuccillo, 1987: A VAS-numerical model impact study using the Gal-Chen variational approach. _Mon. Weather Rev._, 115, 1009-1035.

Ballish, B., 1980: Initialization, theory and application to the NMC spectral model. Ph.D. Thesis, Dept. of Meteorology, University of Maryland, College Park, MD, 151 pp.

Barwell, B. R. and A. C. Lorenc, 1985: A study of the impact of aircraft wind observations on a large-scale analysis and numerical weather prediction system. _Q. J. R. Meteorol. Soc._, 111, 103-129.

Bengtsson, L., M. Kanamitsu, P. Kåallberg and S. Uppala, 1982: FGGE 4-dimensional data assimilation at ECMWF. _Bull. Amer. Met. Soc._, 63, 29-43.

Bergman, K. H., 1979: Multivariate analysis of temperature and winds using optimum interpolation. _Mon. Weather Rev._, 107, 1423-1444.

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_, AFGL, Hanscom AFB, MA, 65 pp. [NTIS AD A129283]

Brenner, S., C.-H. Yang, and K. Mitchell, 1984: The AFGL global spectral model: Expanded resolution baseline version. _AFGL-TR-84-0308_, AFGL, Hanscom AFB, MA 80 pp. [NTIS ADA 160370]

Clarke, R.H., A.J. Dyer, R.R. Brook, D.G. Reid and A.J. Troup, 1971: The Wangara Experiment: Boundary layer data. Technical Paper No. 19 [Division of Meteorological Physics], CSIRO, Australia.

Dey, C.H., and L.L. Morone, 1985: Evolution of the National Meteorological Center Global Data Assimilation System: January 1982-December 1983. Mon. Weather Rev., 113, 304-318.

Dey, C., C.P. Arnold, and W. Bostelman, 1985: Design of a windsat observing system simulation experiment. In Global Wind Measurements, W. E. Baker and R. J. Curran (Eds.), A. Deepak Publishing, Hampton, VA, pp. 73-79.

Dey, C., R. Petersen, B. Ballish, P. Caplan and G. White, 1988: A comparison of model performance using statistical and physical NESDIS TOVS retrievals. In Preprints, Eighth Conference of Numerical Weather Prediction, February 22-26, 1988, Baltimore, Maryland, American Meteorological Society, Boston, pp. 252-259.

Donner, J.L., 1988: An initialization for cumulus convection in numerical weather prediction models. Mon. Wea. Rev., 116, 377-385.

Eyre, J.R., 1989: Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation: Theory and simulation for TOVS. Submitted to Q.J.R. Meteorol. Soc.

Gilchrist, A., 1982: JSC study conference on observing systems experiments, Exeter, 19-22 April 1982. GARP/WCRP Numerical Experimentation Program, Rep. 4, 83 pp.

Grassotti, C., R. Isaacs, R.N. Hoffman, J.-F. Louis, T. Nehrkorn and M. Mickelson, 1989: A study of the impact of simulated 183 GHZ water vapor retrievals on numerical weather prediction. GL-TR-89-0093, GL, Hanscom AFB, MA, 173 pp.

Halbertstam, I., C. Johnson, D.C. Norquist, and S.-L. Tung, 1984: Two methods of global data assimilation. AFGL-TR-84-0260, AFGL, Hanscom AFB, MA., 91 pp. [NTIS ADA155981].

Halem, M., E. Kalnay, W.E. Baker, and R. Atlas, 1982: An assessment of the FGGE satellite observing system during SOP-1. Bull. Am. Meteorol. Soc., 63, 407-426.

Hoffman, R.N., 1986: Documentation of the AFGL statistical analysis program (ASAP) for the global multivariate analysis of heights and winds. AFGL-TR-87-0155, AFGL, Hanscom AFB, MA, 92 pp. [NTIS ADA182600]

Hoffman, R., 1988: Sampling data for OSSEs. In Preprints, Eighth Conference on Numerical Weather Prediction, February 22-26, 1988, Baltimore, MD, American Meteorological Society, Boston, MA, pp. 831-833.

Hoffman, R.N., M. Mickelson, and T. Nehrkorn, 1988: Enhancements to the AFGL Statistical Analysis Program (ASAP) for the global multivariate analysis of heights and winds. AFGL-TR-88-0279, AFGL, Hanscom AFB, MA, 110 pp. [NTIS ADA202912]

Hoffman, R.N., J.-F. Louis, T. Nehrkorn, M. Mickelson and R. Isaacs, 1989: A study of the impact of simulated satellite lidar wind observations. GL-TR-89-0094, GL, Hanscom AFB, MA, 77 pp.

Hollingsworth, A., K. Arpe, M. Tiedtke, M. Capaldo and H. Savijärvi, 1980: The performance of a medium range forecast model in winter - impact of physical parameterization. Mon. Wea. Rev., 108, 1736-1773.

Isaacs, R.G., 1987: Review of 183 GHz moisture profile retrieval studies. AFGL-TR-87-0127, Air Force Geophysics Laboratory, Hanscom AFB, MA, 45 pp. [NTIS ADA182417]

Isaacs, R.G., 1988: A unified retrieval methodology for the DMSP meteorological sensors. In Remote Sensing Retrieval Methods. Deepak Publishing.

Isaacs, R. G., R.N. Hoffman, and L.D. Kaplan, 1986a: Remote sensing of geophysical parameters for numerical weather prediction, Rev. Geophys., 24, 701-743.

Isaacs, R.G., Y.-Q. Jin, G. DeBlonde, R.D. Worsham, and L.D. Kaplan, 1986b: Remote sensing of hydrological variables from the DMSP microwave mission sensors. Paper presented at the Beijing International Radiation Symposium, 26-30 August, Beijing, China.

Isaacs, R.G., R.N. Hoffman, J.-F. Louis, and T. Nehrkorn, 1988: A study of the impact of 183 GHz water vapor retrievals. In Preprints, Eighth Conference on Numerical Weather Prediction, February 22-26, 1988, Baltimore, MD, American Meteorological Society, Boston, MA, pp. 285-289.

Jaeger, L., 1976: Monatskarten des Niederschlags für die ganze Erde. Ber. Dtsch. Wetterdienstes, 18, (139), 38 pp.

Knowlton, L., Norquist, D. and R. Hoffman, 1989: Dynamic normal mode initialization using the AFGL global spectral model. GL, Hanscom AFB, MA, in press.

Lambert, S. J., 1988: A comparison of operational global analyses from the European Centre for Medium Range Weather Forecasts (ECMWF) and the National Meteorological Center (NMC), Tellus, 40A, 272-284.

Lau, N.-C., 1984: Circulation statistics based on FGGE Level III-b analyses. NOAA Data Report ERL GFDL-5, NOAA, Washington DC, 427 pp.

LeMarshall, J.F., 1988: An intercomparison of temperature and moisture fields derived from TIROS operational vertical sounder data by different retrieval techniques. Part I: Basic Statistics. _J. Appl. Meteorol._, _27_, 1282-1293.

Lettau, H.H., and B. Davidson, 1957: _Exploring the Atmosphere's First Mile_. Vol. 2. New York, Pargamon Press, 578 pp.

Louis, J.-F., R.N. Hoffman, T. Nehrkorn, and D. Norquist, 1989: Observing systems experiments using the AFGL four-dimensional data assimilation system. _Mon. Weather Rev._, submitted.

Louis, J.-F., R.N. Hoffman, T. Nehrkorn, T. Baldwin and M. Mickelson, 1987: Baseline observing system experiments using the AFGL global data assimilation system. _AFGL-TR-87-0327_, AFGL, Hanscom AFB, MA, 64 pp. [NTIS ADA188861]

Louis, J.-F., R.N. Hoffman, T. Nehrkorn, T. Baldwin, and M. Mickelson, 1988: Four-dimensional data assimilation experiments using the AFGL GDAS. In _Preprints, Eighth Conference on Numerical Weather Prediction, February 22-26, 1988, Baltimore, MD_, American Meteorological Society, Boston, MA, pp. 260-267.

Mahrt, L., H.-L. Pan, P. Ruscher, and C.-T. Chu, 1987: Boundary layer parameterization for a global spectral model. _AFGL-TR-87-0246_, AFGL, Hanscom AFB, MA, 188 pp. [NTIS ADA199440]

McMillin, L.M., H.E. Fleming, and M.L. Hill, 1979: Atmospheric transmittance of an absorbing gas 3: A computationally fast and accurate transmittance model for absorbing gases with variable mixing ratios. _Appl. Opt._, _18_, 1600-1606.

McPherson, R.D., 1984: Cloud-drift and estimates during FGGE. _N.M.C. Office Note 288_, 13 pp.

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. Weather Rev., 107, 1445-1461.

Menzel, W.P., W.L. Smith, and T.R. Stewart, 1983: Improved cloud motion wind vector and altitude assignment for VAS. J. Clim. Appl. Meteorol., 22, 377-384.

Mostek, A., A. Siebers, R. Petersen, D. Keyser, C. Hayden, W. P. Menzel and A. Schreiner, 1988: VAS model impact studies at NMC. In Preprints. Eighth Conference of Numerical Weather Prediction. February 22-26, 1988, Baltimore, Maryland, American Meteorological Society, Boston, pp. 277-284.

Nehrkorn, T., and R. Hoffman, 1988a: Inference of relative humidity profiles from 3DNEPH cloud data. In Preprints. Eighth Conference on Numerical Weather Prediction. February 22-26, 1988, Baltimore, MD, American Meteorological Society, Boston, MA, pp. 290-292.

Nehrkorn, T., and R. Hoffman, 1988b: Inferring relative humidity profiles from 3DNEPH cloud cover data: Technique development and data impact study, AFGL-TR-88-0327, AFGL, Hanscom AFB, MA, 47 pp. [NTIS ADA205825]

Norquist, D.C., 1982: Adaptation and application of NMC nonlinear normal mode initialization scheme. In Objective Analysis and Prediction Techniques, A.M. Gerlach (ed.), AFGL-TR-82-0392, AFGL, Hanscom AFB, MA, pp. 21-32 [NTIS ADA131465].

Norquist, D.C., 1986: Alternative forms of moisture information in 4-D data assimilation. AFGL-TR-86-0194, AFGL, Hanscom AFB, MA, 144 pp. [NTIS ADA179792]

Norquist, D.C., 1988: Alternate forms of humidity information in global data assimilation. Mon. Wea. Rev., 116, 452-471.

Økland, H., 1976: An example of air-mass transformation in the Arctic and connected disturbances of the wind field. University of Stockholm, Dept. of Meteorology, Report No. DM-20.

Ou, S.-C., and K.-N. Liou, 1988: Development of radiation and cloud parameterization programs for AFGL Global Models. AFGL-TR-88-0018, AFGL, Hanscom AFB, MA, 88 pp. [NTIS ADA193369]

Paegle, J., W. E. Baker and J. N. Paegle, 1986: The analysis sensitivity to tropical winds from the Global Weather Experiment. Mon. Weather Rev., 114, 991-1007.

Phillips, N., J. Susskind and L. McMillin, 1988: Results of a joint NOAA/NASA sounder simulation study. J. Atmos. Ocean. Tech., 5, 44-56.

Ramamurthy, M. K. and F. H. Carr, 1987: Four-dimensional data assimilation in the Monsoon region. Part I: Experiments with wind data. Mon. Weather Rev., 115, 1678-1706.

Reuter, D., J. Susskind and A. Pursch, 1988: First-guess dependence of a physically based set of temperature-humidity retrievals from HIRS2/MSU data. J. Atmos. Oceanic Tech., 5, 70-83.

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

Smith, W.L., H.M. Woolf, and A.J. Schreiner, 1986: Simultaneous retrieval of surface and atmospheric parameters, a physical and analytically direct approach. Advanced in Remote Sensing Retrieval Methods. A. Deepak Publishing, 221-230.

Soong, S.-T., Y. Ogura and W.-S. Kau, 1985: A study of cumulus parameterization in a global circulation model. AFGL-TR-85-0160, AFGL, Hanscom AFB, MA, 113 pp. [NTIS ADA170137]

Sugi, M., 1986: Dynamic normal mode initialization. _J. Meteorol. Soc._
Japan, _64_, 623-636.

Susskind, J. and D. Reuter, 1985: Retrieval of sea surface temperatures
from HIRS2/MSU. _J. Geophys. Res._, _90_ (C6), 11602-11608.

Susskind, J., J. Rosenfield, D. Reuter, and M.T. Chahine, 1982: _The GLAS_
Physical Inversion Method for Analysis of HIRS2/MSU Sounding Data.
NASA-TM-84936, 101 pp.

Susskind, J., J. Rosenfield, D. Reuter, and M.T. Chahine, 1984: Remote
sensing of weather and climate parameters from HIRS2/MSU on
TIROS-N. _J. Geophys. Res._, _89_, 4677-4697.

Tiedtke, M., J.-F. Geleyn, A. Hollingsworth and J.-F. Louis, 1979: ECMWF
model parameterisation of subgrid scale processes. _ECMWF Tech. Rep._
_No. 10_. Reading, Berkshire, U.K.

Trenberth, K. E. and J. G. Olson, 1988: An evaluation and intercomparison
of global analyses from the National Meteorological Center and the
European Centre for Medium Range Weather Forecasts. _Bull. Amer._
Meteorol. Soc., _69_, 1047-1057.

Tung, S.L., 1983: Generalization and application of Machenhauer
initialization scheme. In _Objective Analysis and Prediction_
Techniques, A.M. Gerlach (ed.), _AFGL-TR-83-0333_, AFGL, Hanscom AFB,
MA. [NTIS ADA142441]

WMO, 1986: Formats for International Exchange of FGGE Level II Data Sets,
International Council of Scientific Unions/World Meteorological
Organization, The World Climate Research Programme, _WMO/TD-No. 100_,
WMO, Geneva, 86 pp.