

AD-A275 674



2

PL-TR-93-2203

A Global Forecast Model Comparison Study - Year 1

Thomas Nehr Korn
Mark Mickelson
Lawrence W. Knowlton
Marina Živković

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

September 15, 1993
Scientific Report No. 1

DTIC QUALITY INSPECTED 8,


DTIC
ELECTE
FEB 15 1994
S E D

Approved for public release; distribution unlimited




PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010


94 2 14 062

94-04966


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


DOUGLAS C. HAHN
Contract Manager


DONALD A. CHISHOLM
Chief, Atmospheric Prediction Branch
Atmospheric Sciences Division


DONALD A. CHISHOLM
Acting Director, Atmospheric Sciences Laboratory

This document has been reviewed by the ESC 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 PL/TSI, Hanscom AFB, MA 01731-3010. This will assist us in maintaining a current mailing list.

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 15 September 1993	3. REPORT TYPE AND DATES COVERED Scientific No. 1	
4. TITLE AND SUBTITLE A Global Forecast Model Comparison Study - Year 1			5. FUNDING NUMBERS F19628-92-C-0092 PE: 62101F PR 4026 TA 01 WU LD	
6. AUTHOR(S) Thomas Nehrkorn Lawrence W. Knowlton Mark Mickelson Marina Zivkovic				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Atmospheric and Environmental Research, Inc. 840 Memorial Drive Cambridge, MA 02139			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Phillips Laboratory 29 Randolph Road Hanscom AFB, MA 01731-3010 Contract Monitor: Douglas Hahn/GPAP			10. SPONSORING/MONITORING AGENCY REPORT NUMBER PL-TR-93-2203	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This report describes the results of the first 15 months of a two-year effort aimed at evaluating the advanced physics global spectral model (APGSM) developed by the Phillips Laboratory, in comparison with the currently operational GSM at Global Weather Central (GWC). Preliminary results from the verification of forecasts from the GWC GSM and APGSM, using the GWC analyses as a reference, are presented. The compaction and analysis of RTNEPH cloud data in preparation for generating and verifying cloud forecasts from the APGSM are described. The development of a cloud forecast scheme based on an application of the cloud curve algorithm to synoptically defined weather regimes is described.				
14. SUBJECT TERMS Numerical weather prediction RTNEPH data Global spectral model Cloud forecasting			15. NUMBER OF PAGES 46	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT SAR	

Table of Contents

1.	Introduction	1
2.	Forecast model comparison	2
2.1.	The GWC GSM	2
2.2.	The APGSM	3
2.3.	Preliminary Results	5
3.	RTNEPH data handling	21
3.1.	RTNEPH data compaction	21
3.2.	Vertical stacking	27
3.3.	Coordinate transformations	32
4.	Cloud Forecast Scheme Development	33
5.	References	38

Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

List of Figures

- Figure 1: Vertical distribution of the 15 analysis levels (HIRAS) and σ -layers and interfaces for the GWC GSM (12-layers) and APGSM (18 layers). 4**
- Figure 2: Mean forecast error (bias) of 1000 hPa geopotential height for January 1-20, for the GWC GSM (solid line) and the APGSM (dashed line). Errors are shown for the global region, Northern and Southern Hemisphere extratropics, tropics, North America and Europe. 7**
- Figure 3: As Figure 2, except for the 1000 hPa geopotential height root mean square forecast error (RMSE). 8**
- Figure 4: As Figure 2, except for the 500 hPa geopotential height bias. 9**
- Figure 5: As Figure 2, except for the 500 hPa geopotential height RMSE. 10**
- Figure 6: As Figure 2, except for the 1000 hPa horizontal vector wind RMSE. 11**
- Figure 7: As Figure 2, except for the 850 hPa horizontal vector wind RMSE. 12**
- Figure 8: As Figure 2, except for the 500 hPa horizontal vector wind RMSE. 13**
- Figure 9: As Figure 2, except for the 300 hPa horizontal vector wind RMSE. 14**
- Figure 10: As Figure 2, except for the 1000 hPa relative humidity bias. 15**
- Figure 11: As Figure 2, except for the 1000 hPa relative humidity RMSE. 16**
- Figure 12: As Figure 2, except for the 850 hPa relative humidity bias. 17**
- Figure 13: As Figure 2, except for the 850 hPa relative humidity RMSE. 18**
- Figure 14: As Figure 2, except for the 500 hPa relative humidity bias. 19**
- Figure 15: As Figure 2, except for the 500 hPa relative humidity RMSE. 20**
- Figure 16: Frequency distribution of the sum of weights for horizontal compaction of cloud cover. 23**
- Figure 17: Frequency distribution of the total cloud cover. 23**

Figure 18: Frequency distribution of layer 1 (1000 hPa) cloud cover.	24
Figure 19: Frequency distribution of the layer 2 (850 hPa) cloud cover.	24
Figure 20: Frequency distribution of the layer 3 (700 hPa) cloud cover.	25
Figure 21: Frequency distribution of the layer 4 (500 hPa) cloud cover.	25
Figure 22: Frequency distribution of the layer 5 (400 hPa) cloud cover.	26
Figure 23: Frequency distribution of the layer 6 (300 hPa and above) cloud cover.	26
Figure 24: The overlap parameter r as a function of the layer separation. Shown are values used by MH during step 1 (circles), step 2 (+), and step 3 (x), as well as from the formula for $r_{min} = 0.2$ (top curve) and $r_{min} = 0.13$ (bottom curve).	29
Figure 25: Frequency distribution of temperature on the model σ -layers for the preprocessed HIRAS data.	34
Figure 26: As Figure 25, but for RH.	34
Figure 27: As Figure 25, but for zonal velocity u .	35
Figure 28: As Figure 25, but for meridional velocity v .	35
Figure 29: Percent of variance explained by the first 10 EOFs..	36
Figure 30: Structure of the first 8 EOFs (see text for explanation).	36

List of Tables

Table 1: Statistics of total cloud cover for the month of July, 1989. Shown are observed half-mesh values (cloudy grid points only), computed values (from layer amounts, using different values of r_{min}), and the difference between computed and observed. Shown are the mean (%), standard deviation (%), and root mean square (%); the last column contains the correlation coefficient between computed and observed total cloudiness.	31
---	----

1. Introduction

This report describes the results of the first 15 months of a 24-month study designed to demonstrate that the advanced weather prediction model developed at the Geophysics Directorate of the Phillips Laboratory (PL) is capable of fulfilling the forecasting needs of the US Air Force. Air Force operations are influenced in a significant way by the weather, and there is a need for accurate forecasts of not only the conventional forecast variables (pressure, temperature, moisture, and winds), but also other weather related variables such as aerosols, haze and clouds. Forecasts of these non-conventional variables must be based on sophisticated models that simulate the physical processes involved.

The Air Force Global Weather Central (GWC) currently uses a global spectral model (GSM) obtained from the National Meteorological Center (NMC) as far back as 1984, with a very simple set of physical parameterizations. Operational cloud forecasts are based on trajectory models with highly parameterized physics, and utilize only the wind fields of the forecast model output. The Phillips Laboratory has developed a replacement for this GWC GSM with advanced physics parameterizations (APGSM), and demonstrated its potential usefulness in preliminary tests. Simultaneously, cloud forecast schemes based on the forecast model output have been developed and tested as a potential replacement or complement to the present trajectory models.

In the present study, forecasts of conventional variables from both models (GWC GSM and APGSM) are compared in side-by-side tests, using data from all four seasons of 1989. The experimental design and preliminary results from this evaluation are described in the following section. To evaluate the usefulness of the APGSM for cloud forecasts, cloud forecast schemes are being developed, and verified against RTNeph data. Aspects of the RTNeph data handling and analysis are discussed in section 3, and the cloud forecast scheme development is described in section 4.

2. Forecast model comparison

As part of this task, 1 month of twice daily GSM forecasts out to four days are produced, using the GWC GSM and the APGSM. The operational analyses from GWC (HIRAS) for those times are used as the initial state, and for verification. To assess how much the results vary with season, three additional months of forecasts are produced, but only one forecast every three days.

2.1. The GWC GSM

The GSM currently operational at GWC is based on the model developed at NMC (Sela, 1980). The hydrodynamics were completely redesigned (Brenner *et al.*, 1982). The physics package was last updated in 1984, and it consists of a very simple set of physical parameterizations. It does not simulate the transfer of radiation through the atmosphere, and includes only a drag-law type boundary layer parameterization. The adjustment physics consists of large-scale precipitation when model cells become saturated, dry adiabatic adjustment to avoid instability, and a version of the Kuo (1965) convection scheme, which is disabled at most points through the use of high threshold values and various criteria that must be met before moist convection is allowed to take place. This physics package is described more fully in Yang *et al.* (1989).

Model initial states are derived from gridded analyses by preprocessing (vertical interpolation to the model's σ -surfaces, and spectral transformation from gridpoint to spectral space), and a subsequent nonlinear normal mode initialization (NMI). The NMI is a standard adiabatic Machenhauer (1977) scheme obtained from NMC (Ballish, 1980).

In the present study, parameters of the GSM are used that closely correspond to current operational practice at GWC. In particular, the horizontal resolution is set at rhomboidal truncation 40, and 12 layers are used in the vertical (moisture is carried at only 7 layers). The sigma layer interfaces for the 12 layers are at 1.0, .925, .800, .650, .500, .375, .300, .250, .200, .150, .100, .050, 0. A time step of 12 minutes is used, and the coefficients for the time stepping are 0.04 (for the time filter coefficient) and 0.5 (for the time stepping coefficient, implying a semi-implicit step). The coefficient for the horizontal ∇^4 diffusion is $6 \times 10^{15} \text{ m}^4 \text{ s}^{-1}$. The terrain data set is a mean terrain height derived from a coarse input

gridded field ($2.5^\circ \times 2.5^\circ$), and the boundary data sets of sea surface and temperature and drag coefficients use monthly mean values from NMC datasets. The NMI uses two iterations of the Machenhauer scheme, initializing only the four gravest modes with periods less than 48 hours.

The HIRAS analyses used for initialization and verification consist of geopotential height, zonal and meridional wind on 15 mandatory pressure levels (1000, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa), and relative humidity on the lowest 6 mandatory levels. Figure 1 shows the distribution of the analysis levels in the vertical, alongside those of the GWC GSM. It can be seen that while the resolution of the GSM is slightly higher than that of the analysis in the lower troposphere, it is much coarser in the stratosphere (the top 4 analysis levels are all contained in a single, the topmost, σ -layer).

2.2. The APGSM

The APGSM hydrodynamics code is based on the code described by Brenner et al. (1982), but it was recoded by Nehr Korn et al. (1992) to allow more general horizontal truncation, and to make use of vectorization and multiprocessing. The physics package of the APGSM was developed by several research groups and tested and integrated by PL personnel. The version of the GSM used in this study is described in Norquist et al. (1992). It contains a planetary boundary layer (PBL) parameterization (Mahrt *et al.*, 1984), which includes two soil layers and makes use of geographic databases of surface roughness, soil type, albedo, and other surface fields. A gravity wave drag parameterization (Vernekar *et al.*, 1991) is included, as is a radiative transfer package developed by Liou et al. (1984), Ou et al. (1988), and Schattel (1992). The dry adiabatic adjustment and the large scale precipitation parameterization remain essentially unchanged from the GWC GSM, but moist convection is parameterized with the European Centre for Medium-Range Weather Forecasts (ECMWF) mass flux scheme (Tiedtke, 1989).

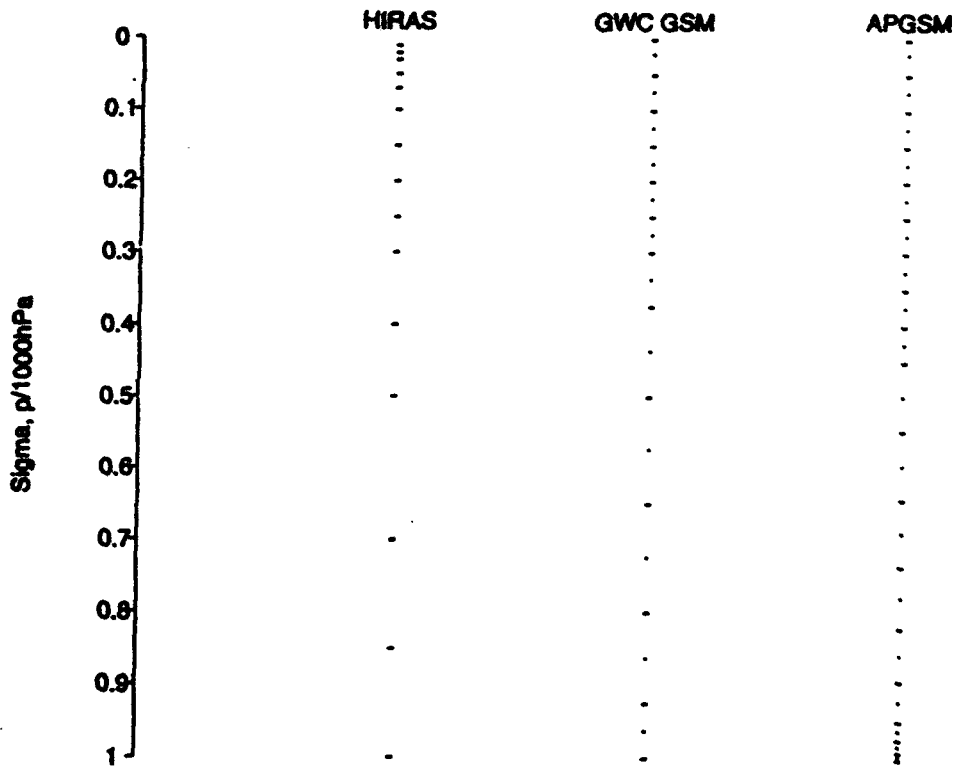


Figure 1: Vertical distribution of the 15 analysis levels (HIRAS) and σ -layers and interfaces for the GWC GSM (12-layers) and APGSM (18 layers).

The parameters of the GSM used in this study closely correspond to those used previously by Nehr Korn et al. (1993) and Norquist et al. (1992). Specifically, the horizontal resolution is set at rhomboidal truncation 40, and 18 layers are used in the vertical (moisture is carried at all 18 layers). The sigma layer interfaces for the 18 layers are at 1.0, .990, .973, .948, .893, .820, .735, .642, .546, .450, .400, .350, .300, .250, .200, .150, .100, .050, 0. The vertical distribution is also shown in Figure 1. It can be seen that the additional layers are mainly near the ground (to accommodate the planetary boundary layer parameterization), in addition to a slightly higher resolution in the middle troposphere. Above $\sigma=.3$, the vertical structure is identical to that of the GWC GSM. A time step of 15 minutes is used, and the coefficients for the time stepping are 0.04 (for the time filter coefficient) and 0.5 (for the time stepping coefficient, implying a semi-implicit step). The coefficient for the horizontal ∇^4 diffusion is $2.55 \times 10^{15} \text{ m}^4 \text{ s}^{-1}$. The terrain data set is a silhouette terrain derived from the Navy 10' dataset; boundary data sets for the physical parameterizations use monthly mean values compiled from a number of sources.

The NMI uses two iterations of the Machenhauer scheme, initializing only the four gravest modes with periods less than 48 hours. For the forecasts of the first half of January, the preprocessor was used unchanged from the version used in the GWC GSM, and the NMI used diabatic tendencies including all physical processes. Because of problems encountered in the initialization (see section 2.3), the preprocessor was modified for later forecasts (the remainder of January, and April, July, and October), and adjustment processes (dry adiabatic adjustment and large-scale precipitation) were excluded from the diabatic NMI.

Throughout its evolution, the PL GSM has undergone testing and comparison with the GWC GSM. Results from tests of a number of preliminary versions can be found in Yang et al. (1989). The most recent version of the physics package was tested by Norquist et al. (1992). They found that for a series of six January and July forecasts (out of 10 days), the APGSM performed better than the GWC GSM. The present study extends these tests in two ways: comparisons with the GWC GSM are performed over a much larger number of forecasts, and the quality of cloud forecasts from the APGSM is studied and compared with operational GWC cloud forecasts.

2.3. Preliminary Results

The APGSM forecasts encountered a problem during the initialization after half the forecasts for January had been completed. Further analysis revealed that for the 00Z 15 January time period, a large amount of stratiform precipitation is predicted by the GSM over Greenland when started from the preprocessed, uninitialized state (13 mm /time step, corresponding to 52 mm/hr). During the diabatic NMI, this problem is exacerbated because the latent heat release is reflected in strong temperature tendencies. The final profiles of temperature and height strongly differ from either the preprocessed or the original HIRAS fields.

Part of the problem lies with the preprocessor: RH at σ -layers above 300 mb (the top-most analysis level with RH data) is extrapolated, resulting in a large numbers of points either at saturation or with zero RH (this is also reflected in the histograms presented in section 4). This problem was rectified by replacing the extrapolation with a constant value. This change reduced the amplitude of the latent heat release, but the NMI still resulted in unreasonable adjustments. The diabatic NMI was then changed to exclude tendencies from the adjustment

physics. Diagnostics of the modified NMI increments indicate reasonable values over Greenland; maps of NMI increments at 1000 mb and 500 mb indicate smaller increments for the new system everywhere, but particularly over Greenland and other high altitude locations, with less energy at small scales. Differences from the original HIRAS analyses are not much affected outside the area over Greenland.

All forecasts have been completed for the GWC GSM, but only part of the forecasts for the APGSM have been completed at this time. Figures 2 - 15 show results for both models for slightly over half the forecasts in January (Jan. 1 - Jan 20). Shown are forecast error statistics (mean error, or bias, and root mean square error, or RMSE) for a number of geographic regions:

global: entire globe

Northern Hemisphere extratropics: all points between 20°N and 80°N

Southern Hemisphere extratropics: all points between 20°S and 80°S

tropics: all points between 20°S and 20°N

North America: all points between 25°N and 60°N, and 120°W and 70°W

Europe: all points between 35°N and 70°N, and 10°W and 40°E.

Comparison of the 1000 hPa and 500 hPa bias of geopotential height (Figures 2 and 4) show an increasing warm bias (thicknesses too large) for the GWC GSM, which is the result of a lack of radiative cooling (this was also noted in (Louis *et al.*, 1989); this bias is absent in the APGSM. Height RMSEs (Figures 3 and 5) are generally smaller at the later lead times for the APGSM than the GWC GSM, but tend to be slightly larger at the beginning of the forecast. RMSEs of the horizontal wind (Figures 6 - 9) are uniformly larger for the APGSM than the GWC GSM. Results for the RH field (Figures 10 - 15) are mixed, with better results for the APGSM at lower levels (1000 hPa and 850 hPa) and later lead times, and the opposite at the beginning of the forecast, and at 500 hPa.

These preliminary results are being investigated more thoroughly at the moment. A comparison of results for January 1-20 with those for January 15-20 (not shown here) indicated that the change in initialization procedure of the APGSM had no appreciable effect on the error statistics.

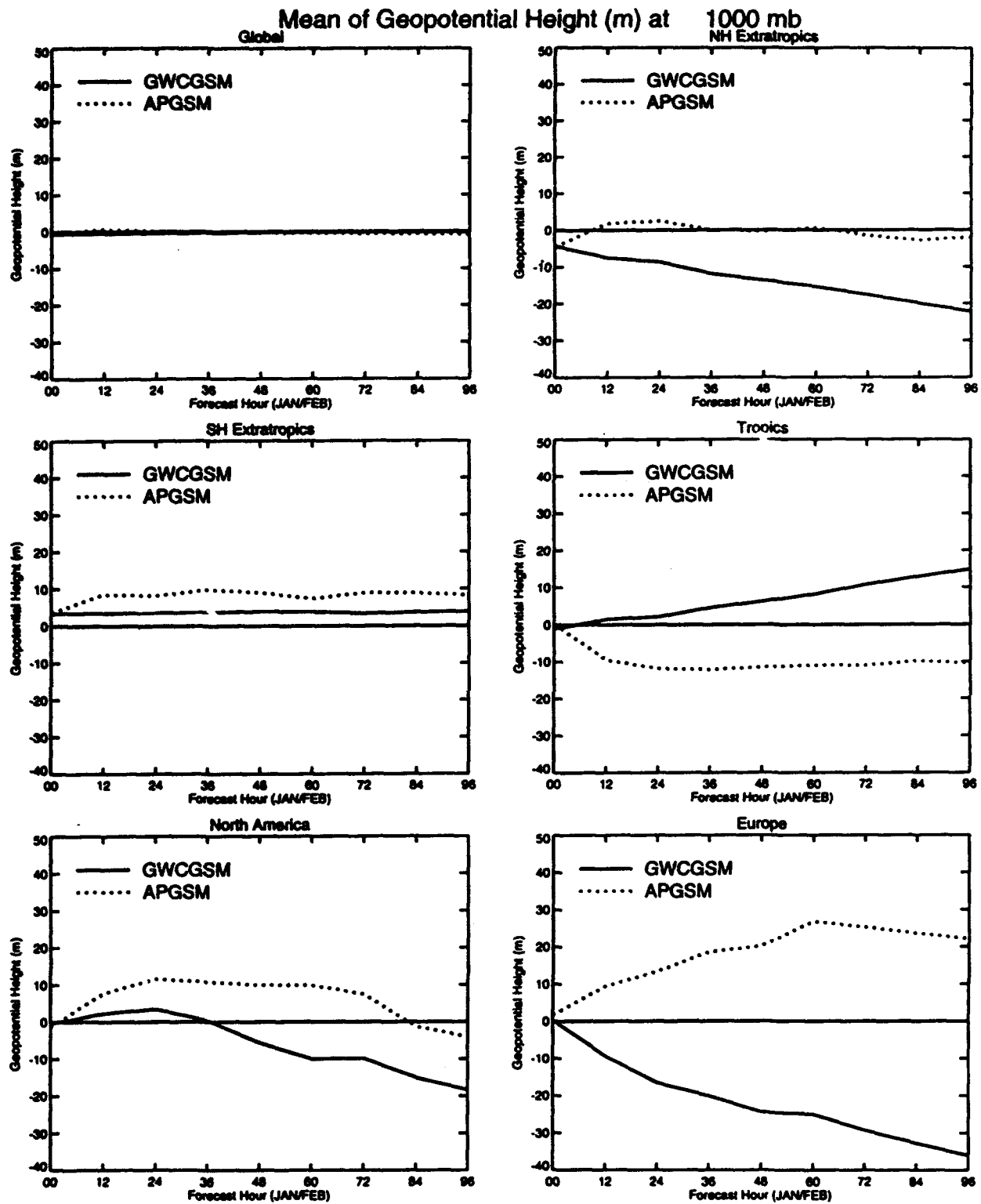


Figure 2: Mean forecast error (bias) of 1000 hPa geopotential height for January 1-20, for the GWC GSM (solid line) and the APGSM (dashed line). Errors are shown for the global region, Northern and Southern Hemisphere extratropics, tropics, North America and Europe.

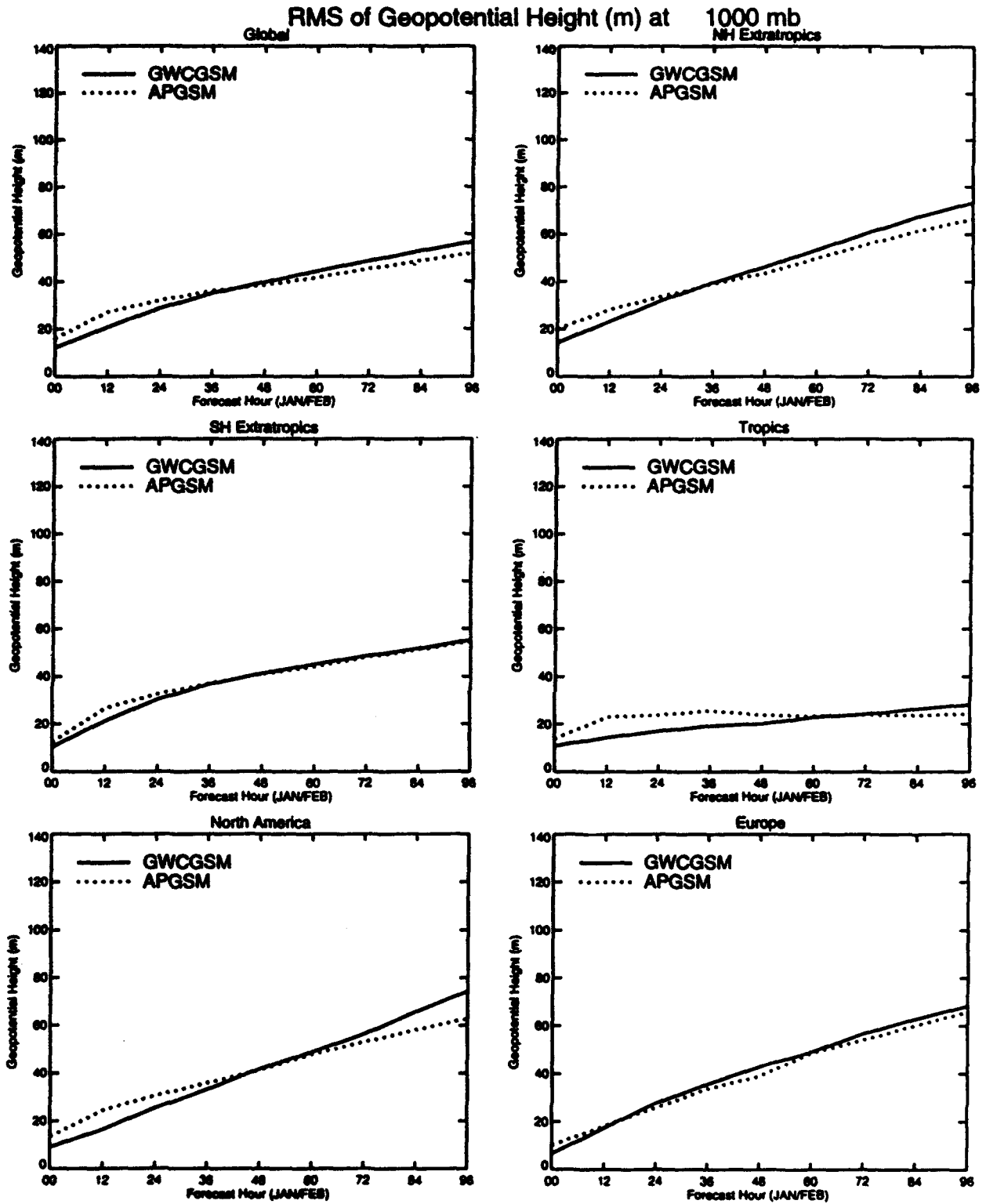


Figure 3: As Figure 2, except for the 1000 hPa geopotential height root mean square forecast error (RMSE).

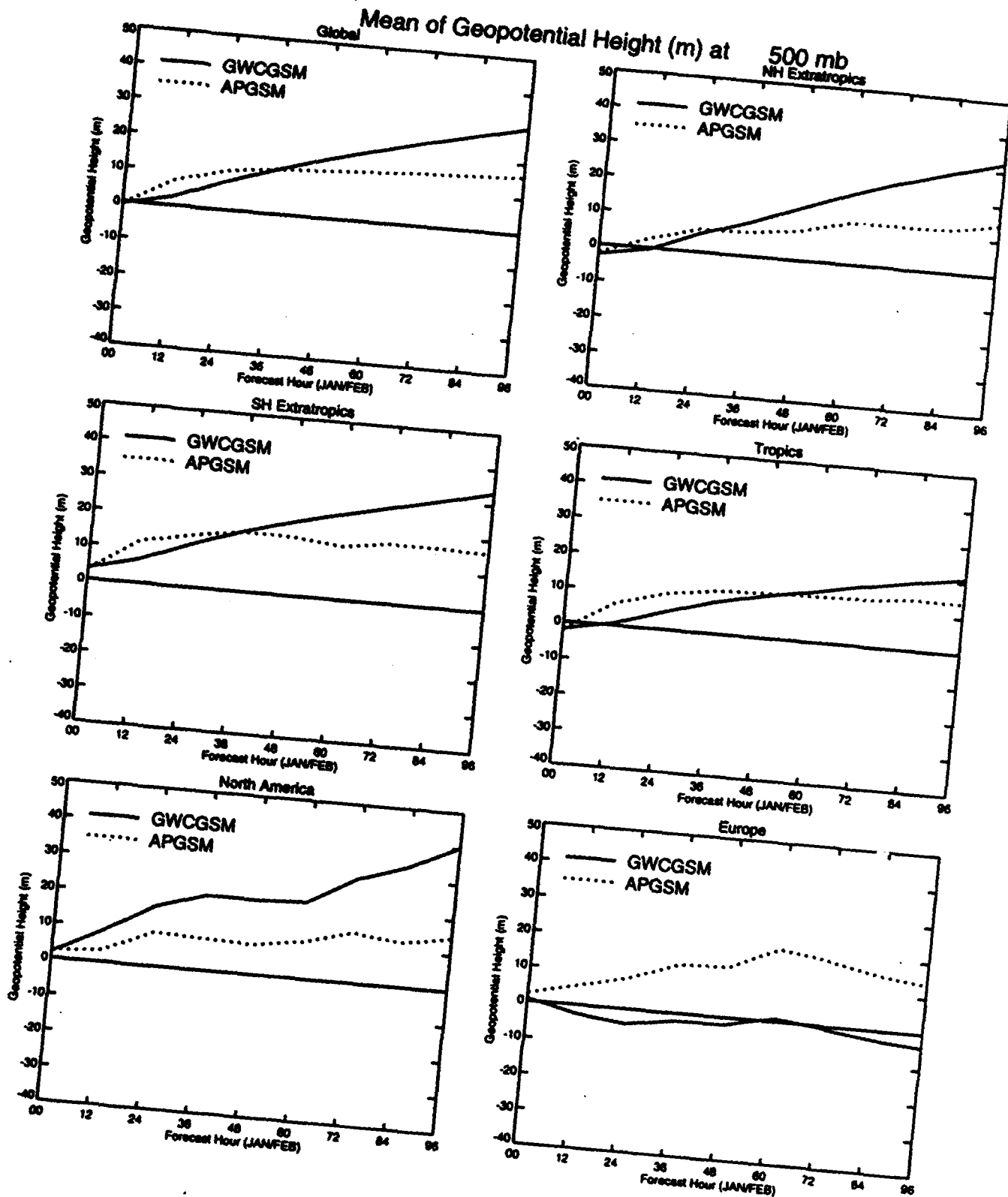


Figure 4: As Figure 2, except for the 500 hPa geopotential height bias.

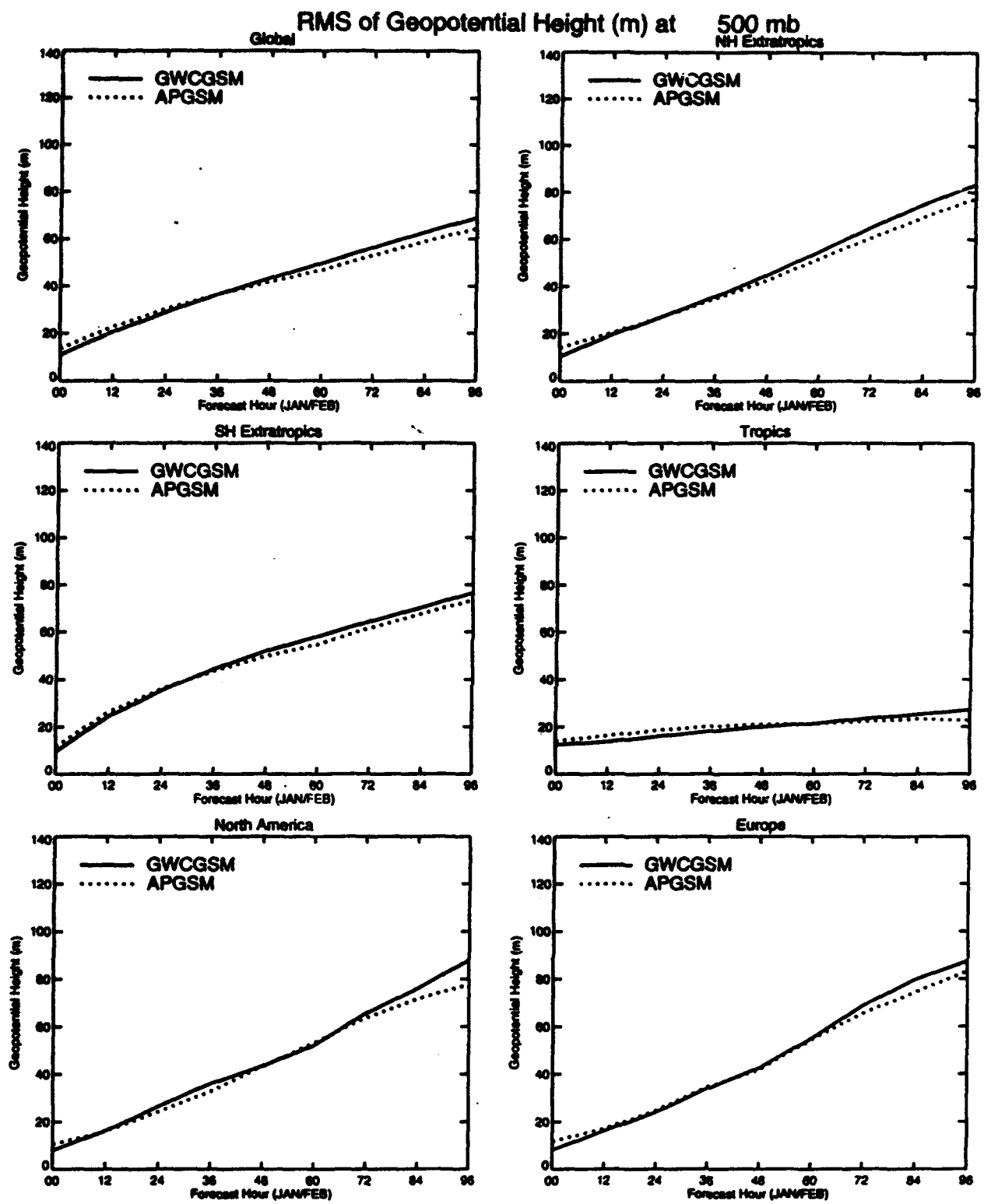


Figure 5: As Figure 2, except for the 500 hPa geopotential height RMSE.

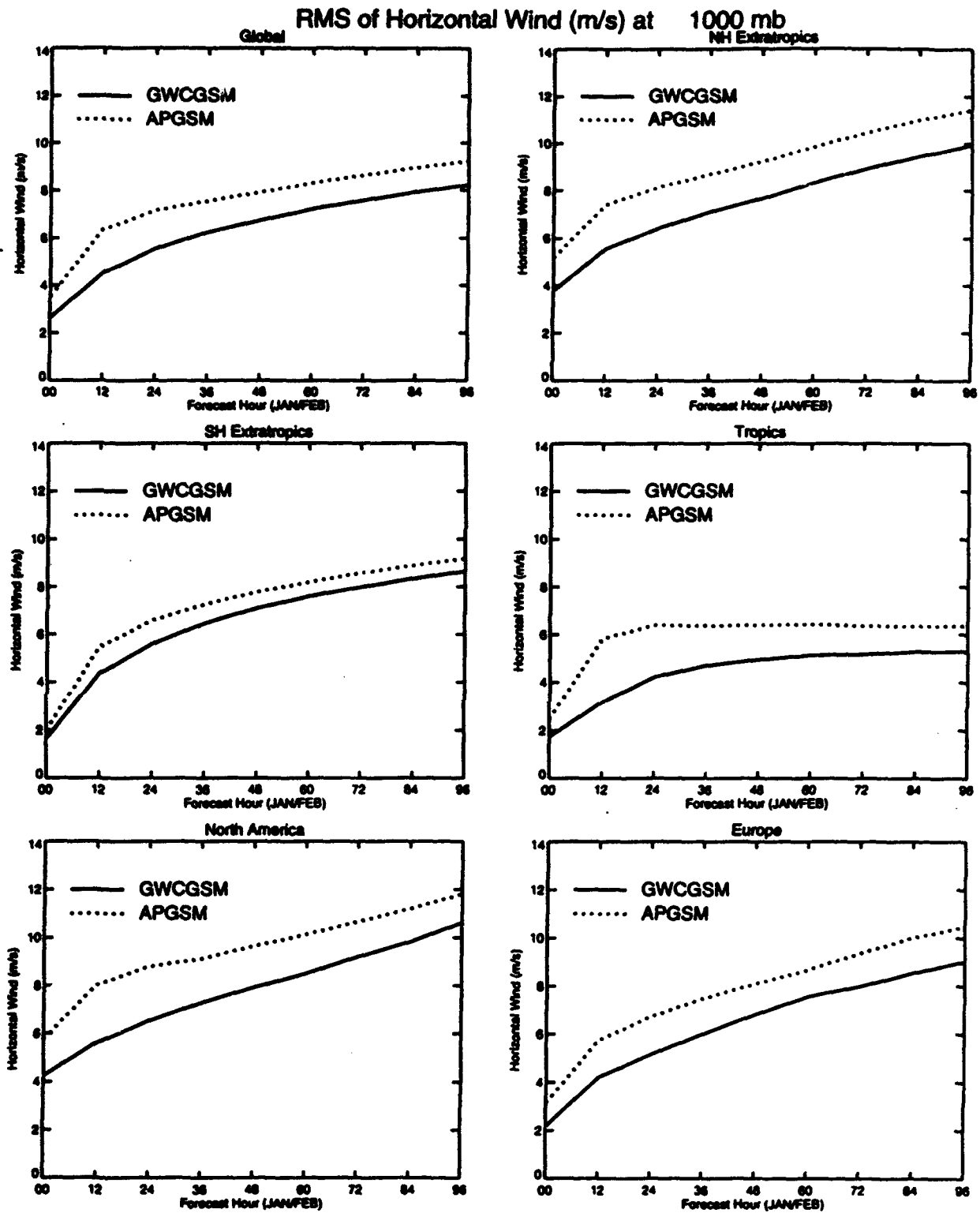


Figure 6: As Figure 2, except for the 1000 hPa horizontal vector wind RMSE.

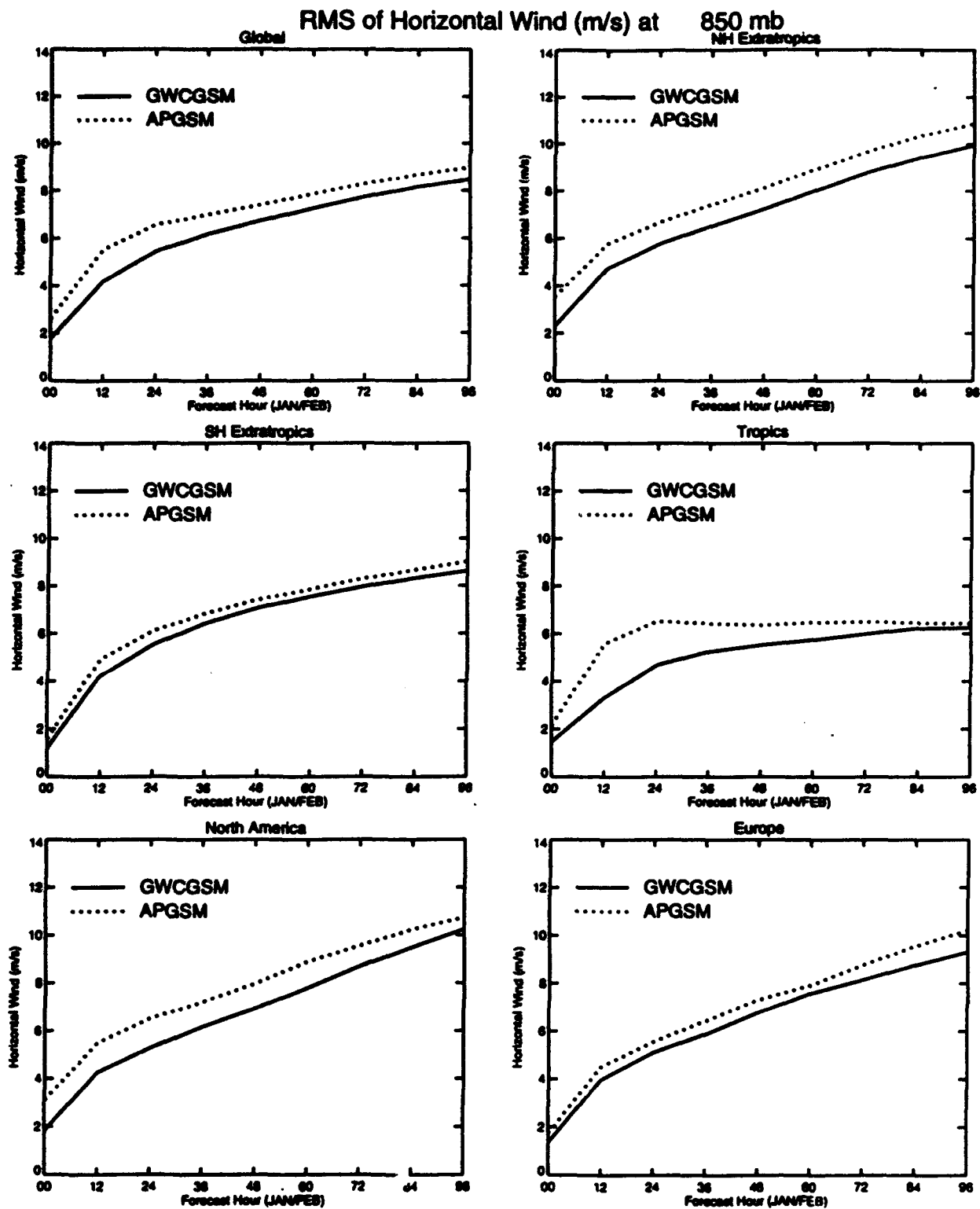


Figure 7: As Figure 2, except for the 850 hPa horizontal vector wind RMSE.

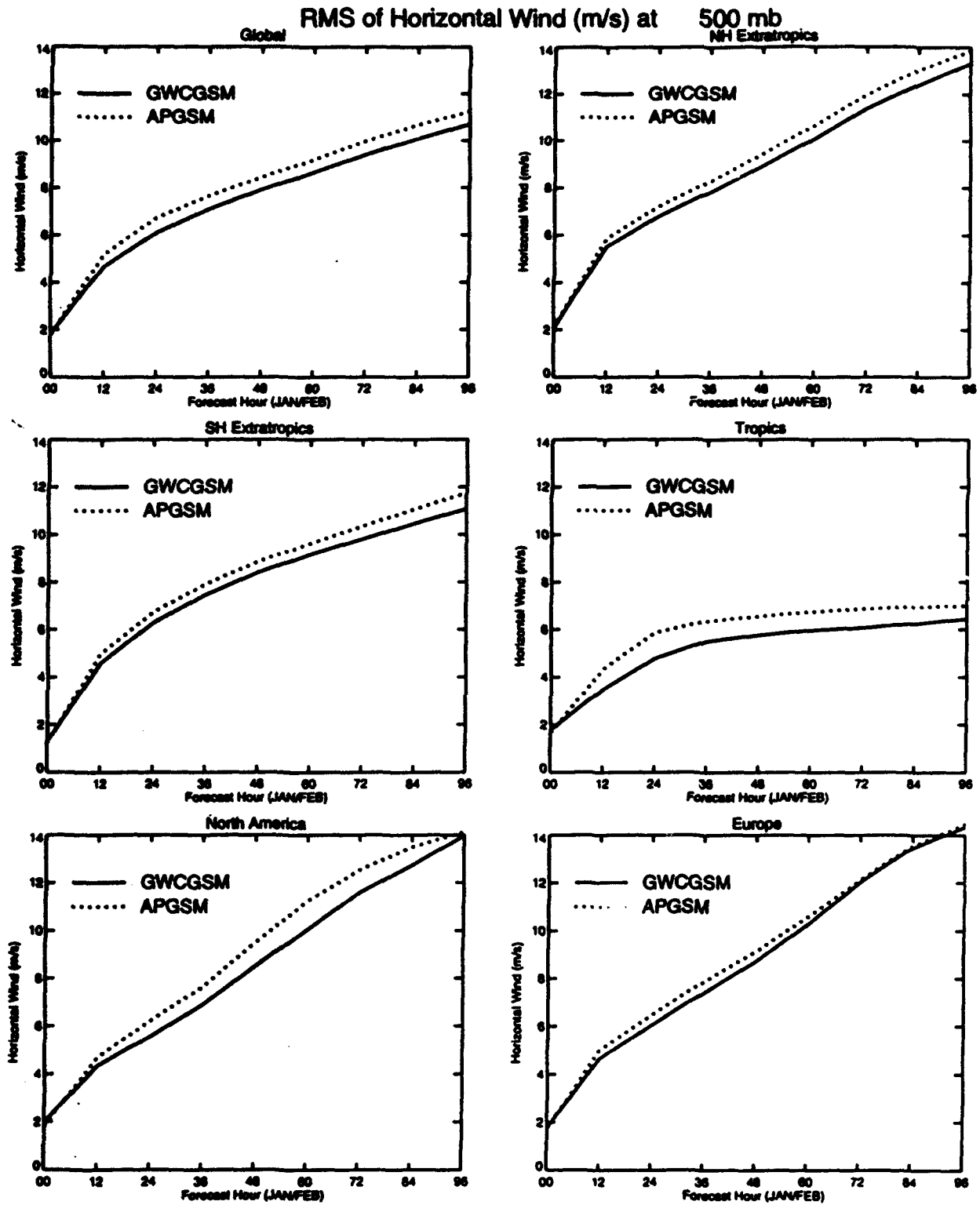


Figure 8: As Figure 2, except for the 500 hPa horizontal vector wind RMSE.

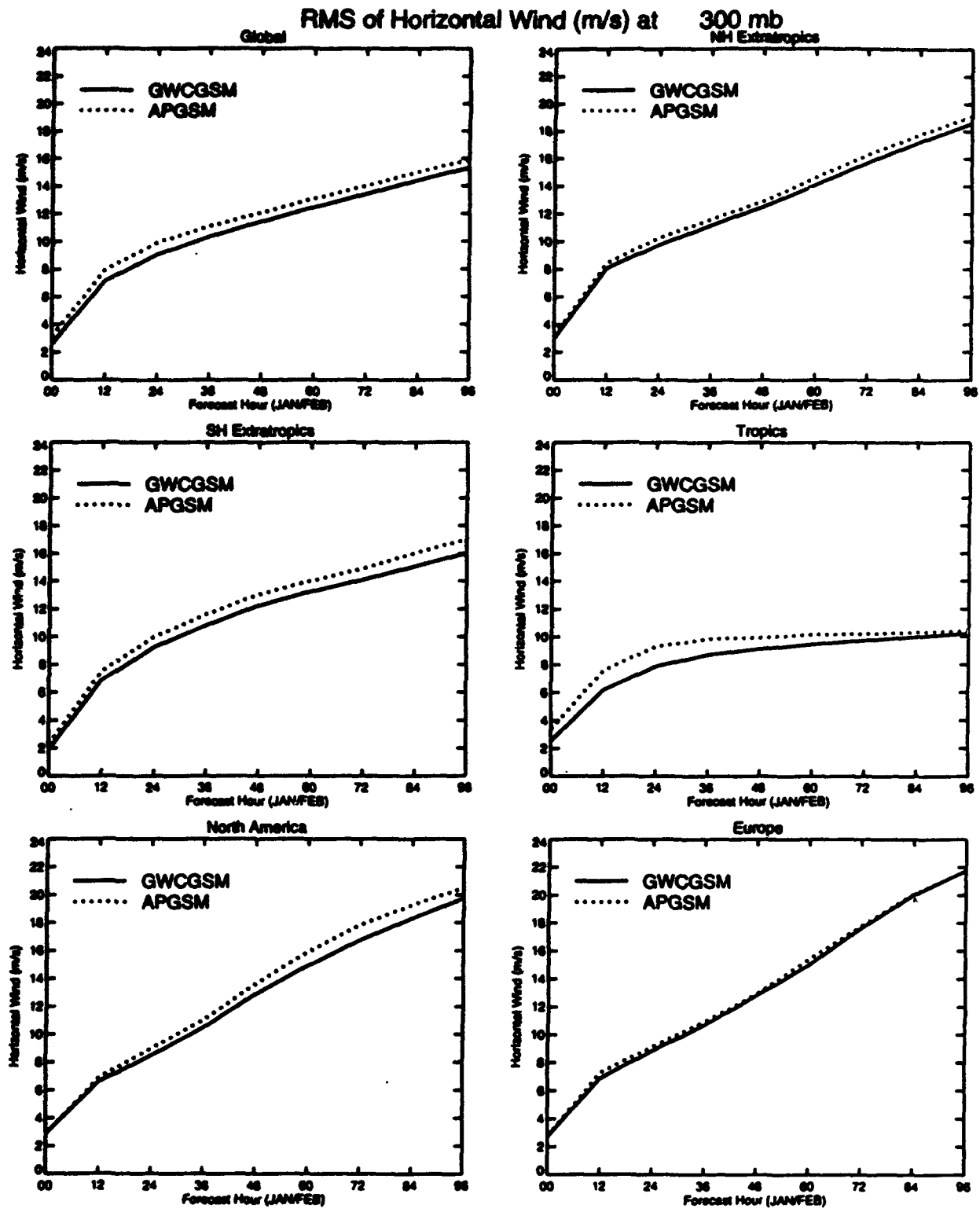


Figure 9: As Figure 2, except for the 300 hPa horizontal vector wind RMSE.

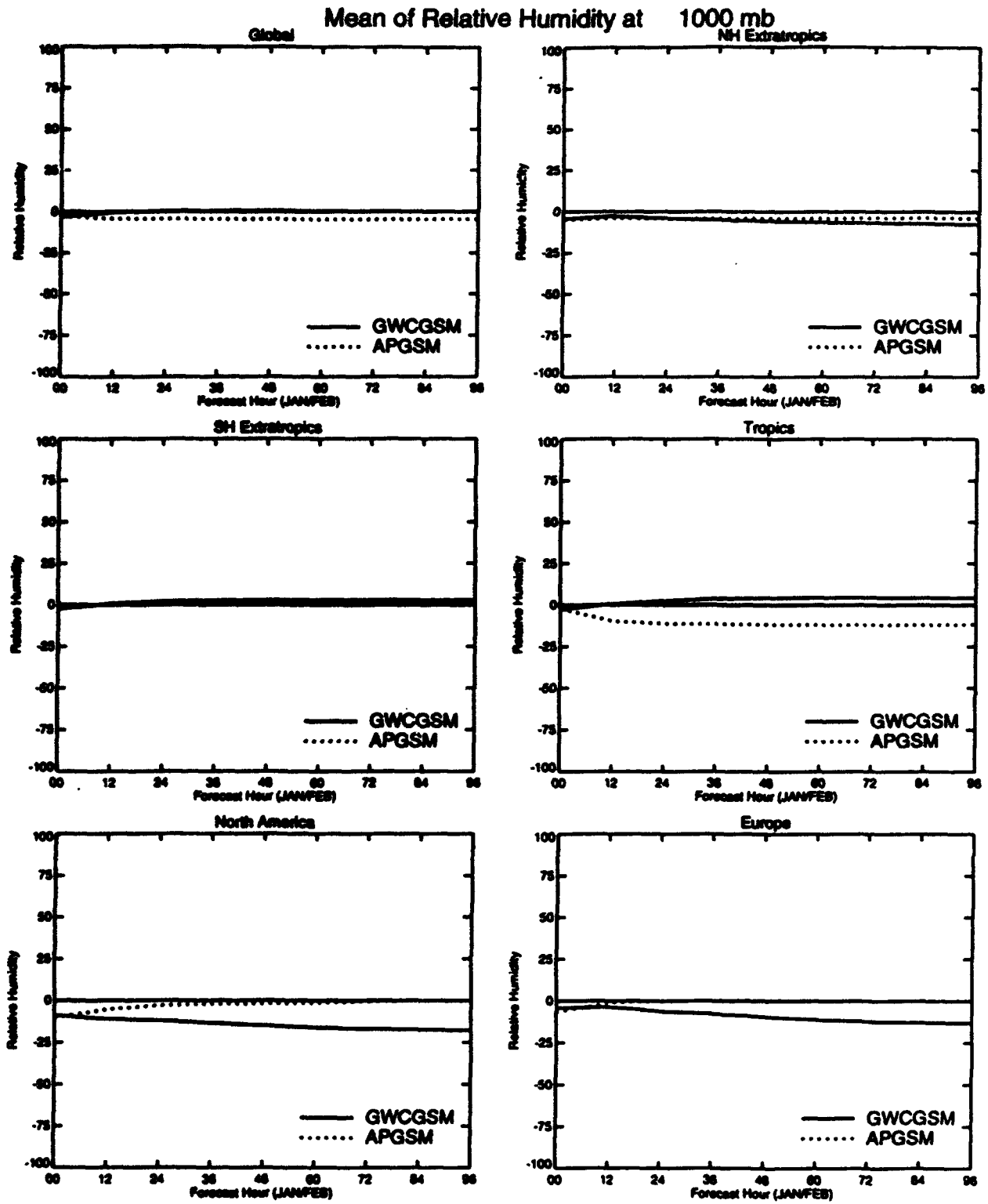


Figure 10: As Figure 2, except for the 1000 hPa relative humidity bias.

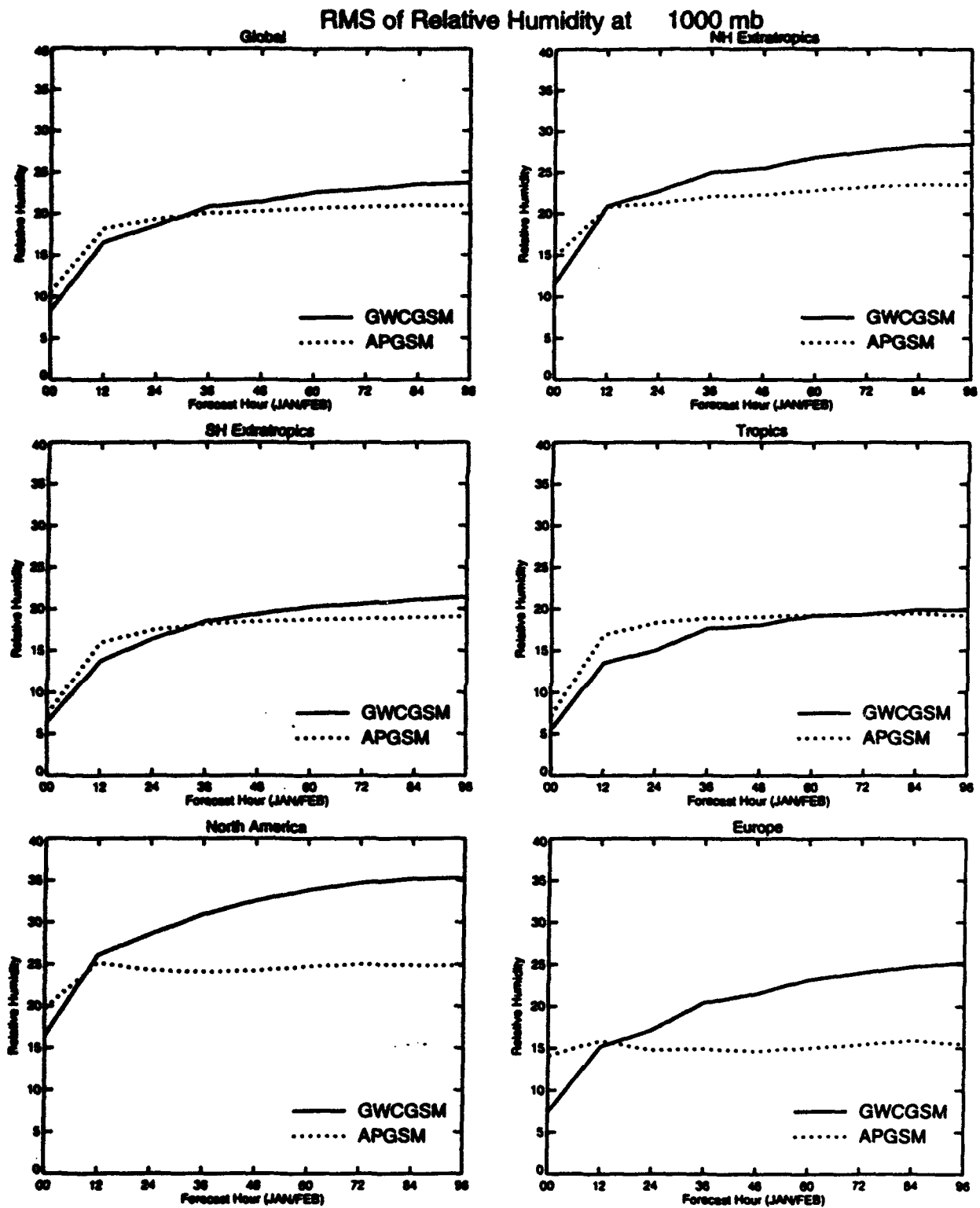


Figure 11: As Figure 2, except for the 1000 hPa relative humidity RMSE.

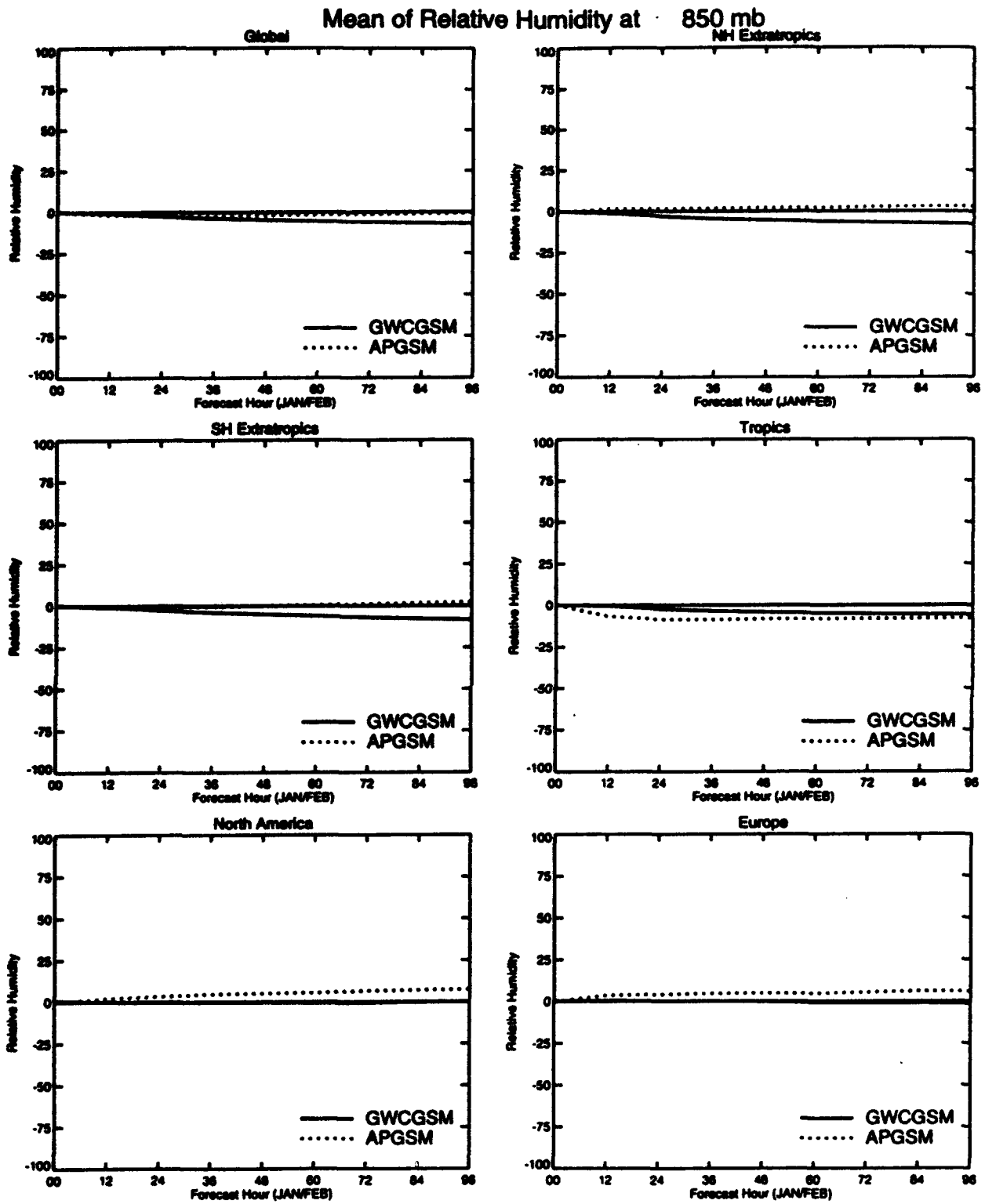


Figure 12: As Figure 2, except for the 850 hPa relative humidity bias.

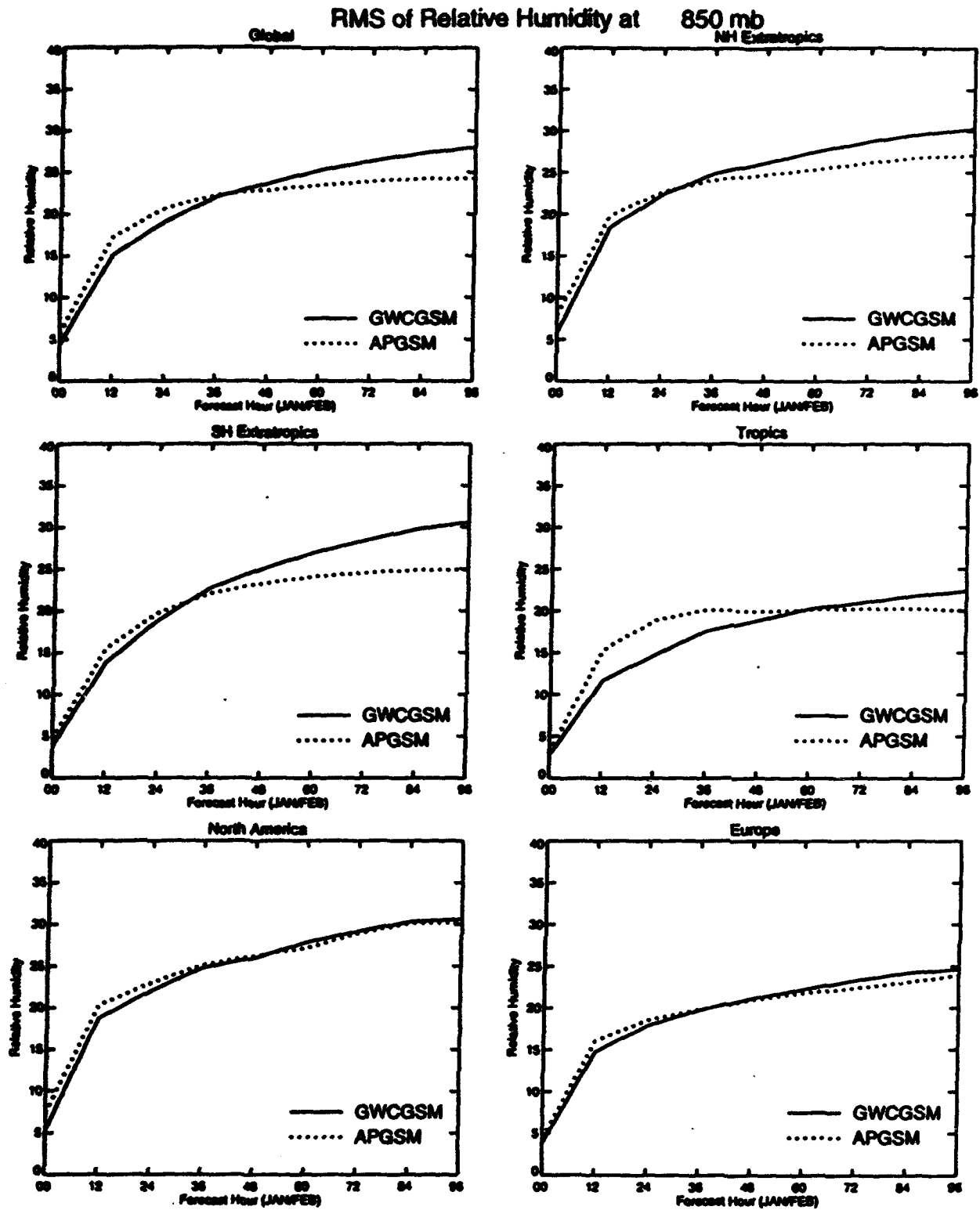


Figure 13: As Figure 2, except for the 850 hPa relative humidity RMSE.

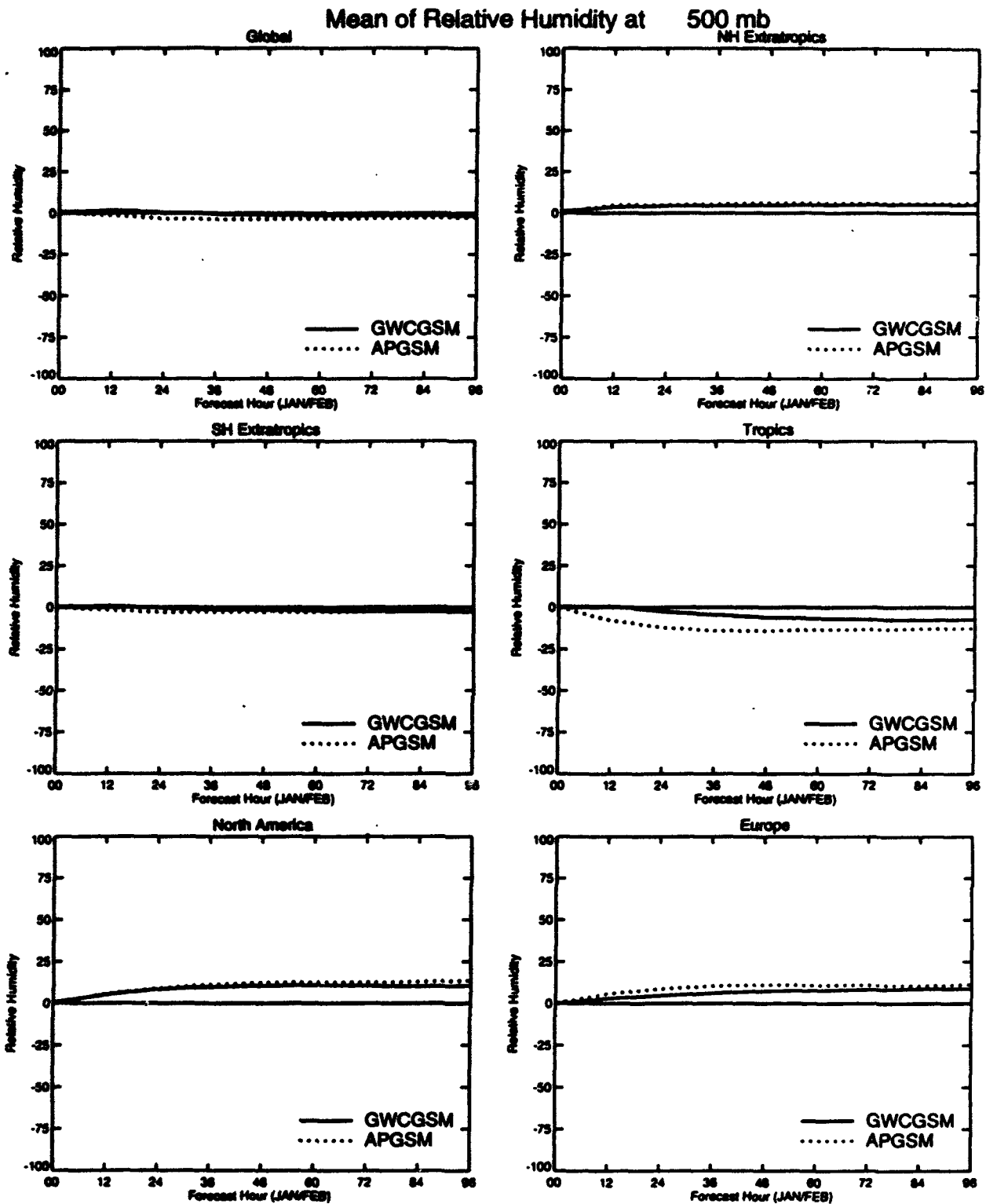


Figure 14: As Figure 2, except for the 500 hPa relative humidity bias.

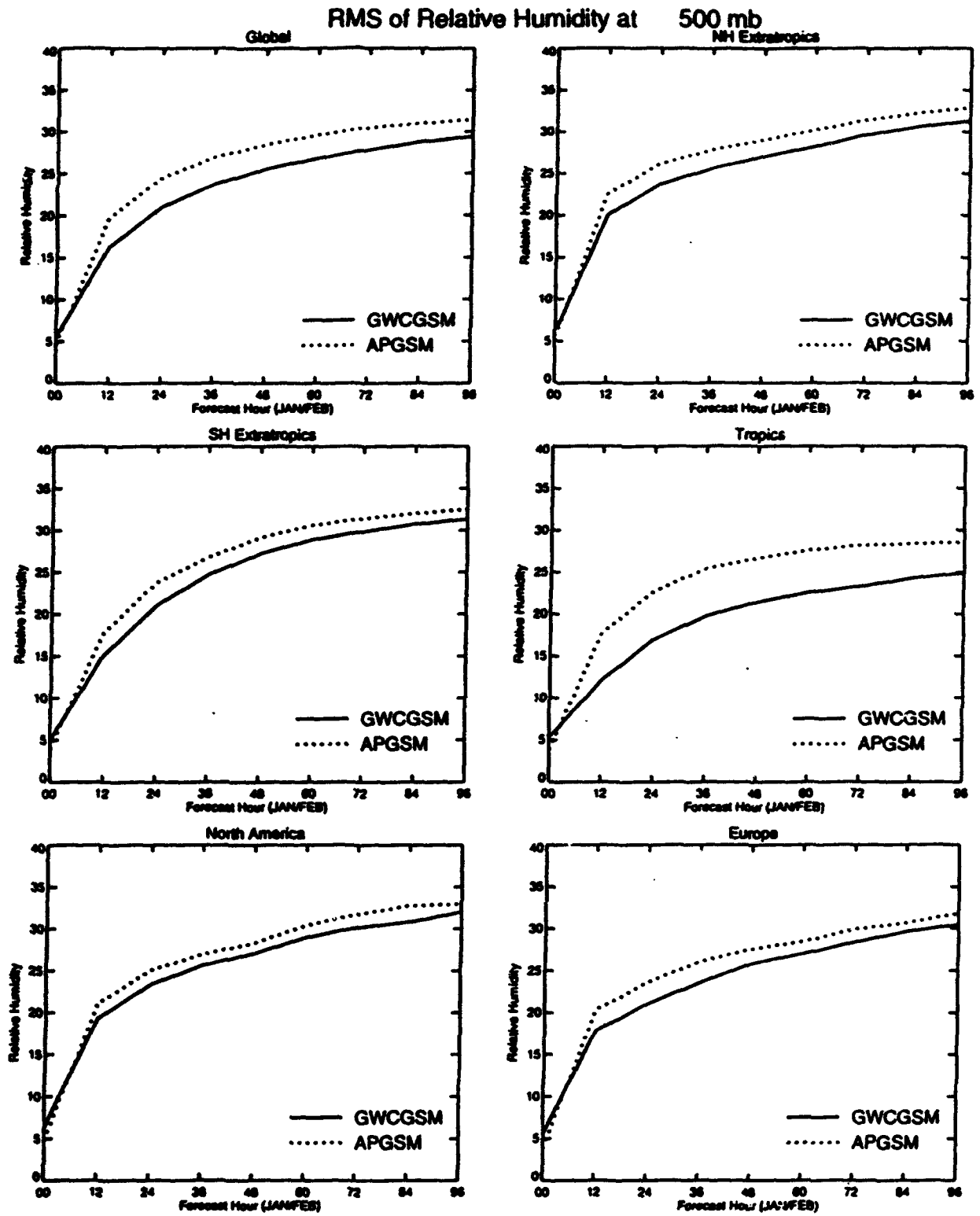


Figure 15: As Figure 2, except for the 500 hPa relative humidity RMSE.

3. RTNEPH data handling

Cloud data are needed for the development and testing of the cloud forecast schemes. For this purpose, Northern Hemisphere RTNEPH (Hamill *et al.*, 1992) data (for the four months for which forecasts are produced) are used. The raw data consist of total cloudiness, layer cloud cover of up to four layers, and various data source, timeliness, and cloud type information on the so-called 1/8-mesh grid (this is a regular grid on a polar stereographic projection, with a grid spacing of 47.625 km at 60° N). Before these data are used in our study, they are compacted to the 1/2-mesh (grid spacing of 190.5 km at 60° N), so that they represent scales resolved by the forecast model. Data at this resolution are also used by GWC in the verification of the operational cloud forecast models. In the following, three different aspects of the RTNEPH data handling are discussed: data compaction, consistency between layer and total cloud amounts, and coordinate transformations between the RTNEPH grid and the latitude/longitude grid systems used by the analysis and forecast model.

3.1. RTNEPH data compaction

The 1/8-mesh RTNEPH are unpacked and compacted to the 1/2-mesh all in one step. The final output from this compaction are working sums for averaged cloud cover (total and 6 layers). The working sums consist of the sum of weights and the sum of weight times cloud cover.

The methodology is as follows: The up to 4 floating RTNEPH layers are assigned to 6 fixed MSL layers. The MSL layer tops are chosen to correspond approximately to the 6 mandatory pressure levels between 1000 hPa and 300 hPa, except that the top layer includes all high clouds. Layer boundaries are at 0, 1.07, 1.98, 4.27, 6.71, 7.92, and 25 km. Working sums for horizontal averages for the total cloud amount, and the 6 MSL layer amounts, are then formed by using a 25-point weighted average with a 1-2-2-2-1 weighting applied in both the *i* and *j* directions. (The weights are between 1 and 4, and the maximum for the sum of weights is 64.) Layer clouds identified as thin are set to zero cloud cover. If more than one floating layer contributes to a fixed MSL layer at one 1/8-mesh point, maximum overlap is assumed.

Only points with valid cloud cover data satisfying the timeliness criterion (data used for cloud analysis are within +/- 2 hours of the valid time) are used.

Figure 16 is an example of the resulting distribution of the sum of weights (for 00 UTC 15 January 1989). Most 1/2-mesh gridpoints (59%) have no data in any of the surrounding 1/8-mesh gridpoints, either because they are off-world points or because no timely data existed, and 25% have data in all 25 surrounding 1/8-mesh gridpoints; only 16% have sum of weights between 0 and the maximum possible number (64). This implies that most points lie either wholly within or outside regions with data coverage, and not much is to be gained (in terms of sample size) by using points with less than complete averages.

Figure 17 shows the distribution of total cloud cover for the same time period. The distribution is bimodal, with 12% of the points clear, 28% overcast (cloud cover > 90%), and a nearly uniform distribution for intermediate values. The mean total cloud cover is 52%. For layer cloud amounts (Figures 18 - 23), more points are totally clear, and fewer are totally cloudy. Layer cloud amounts are largest for layer 3 (700 hPa), and smallest for layer 6 (high clouds); average values for layers 1 - 6 are : 28%, 25%, 30%, 20%, 11%, 10%.

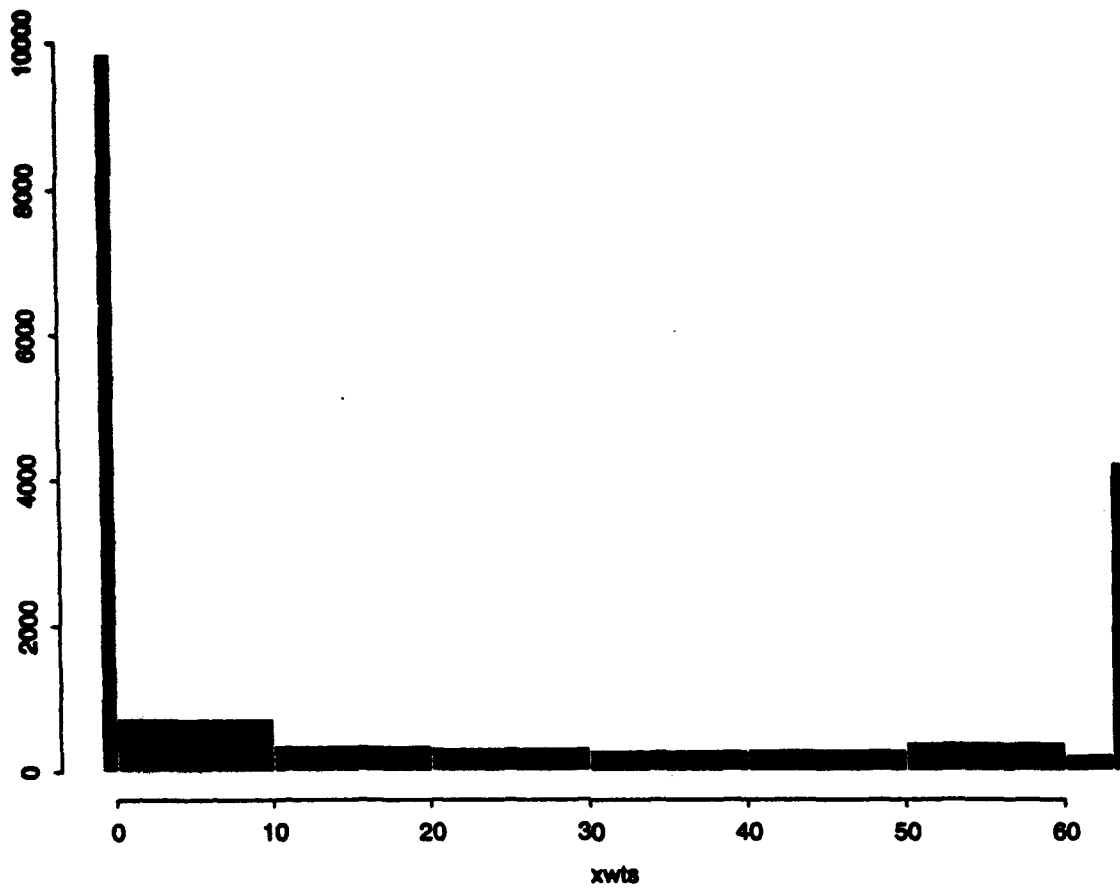


Figure 16: Frequency distribution of the sum of weights for horizontal compaction of cloud cover.

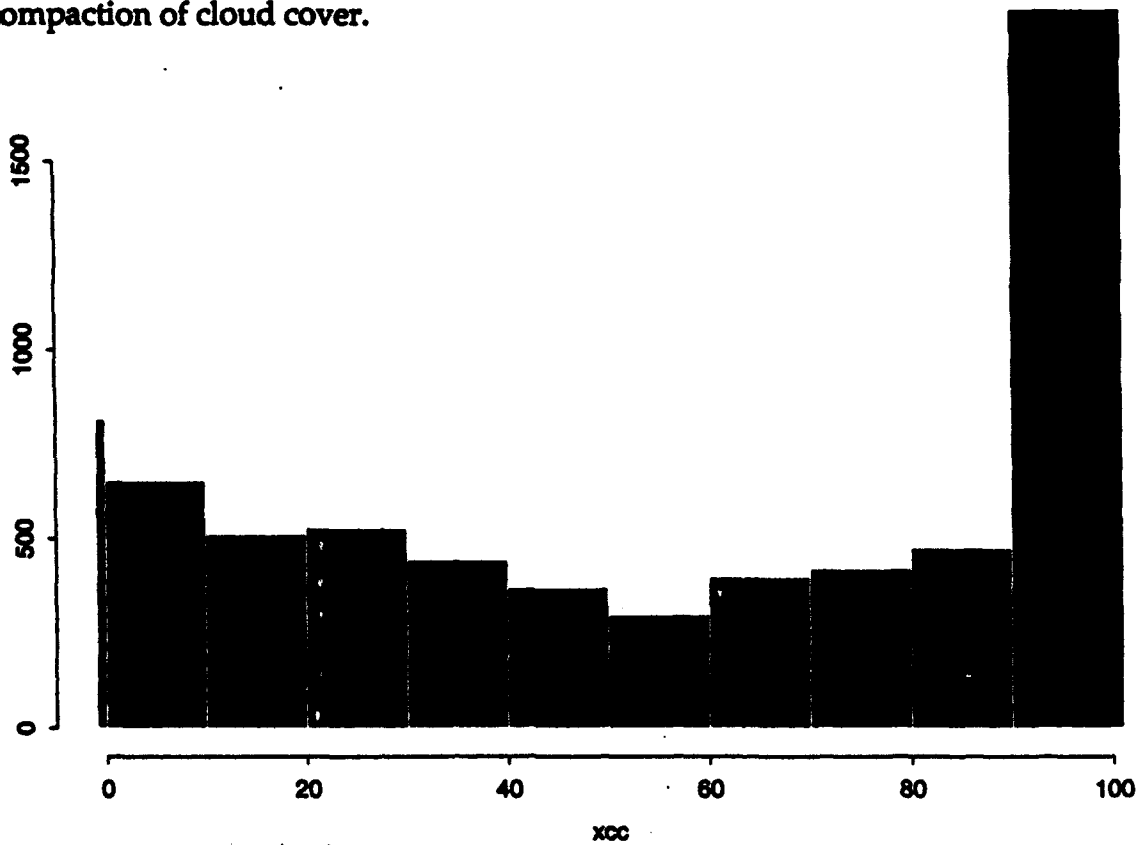


Figure 17: Frequency distribution of the total cloud cover.

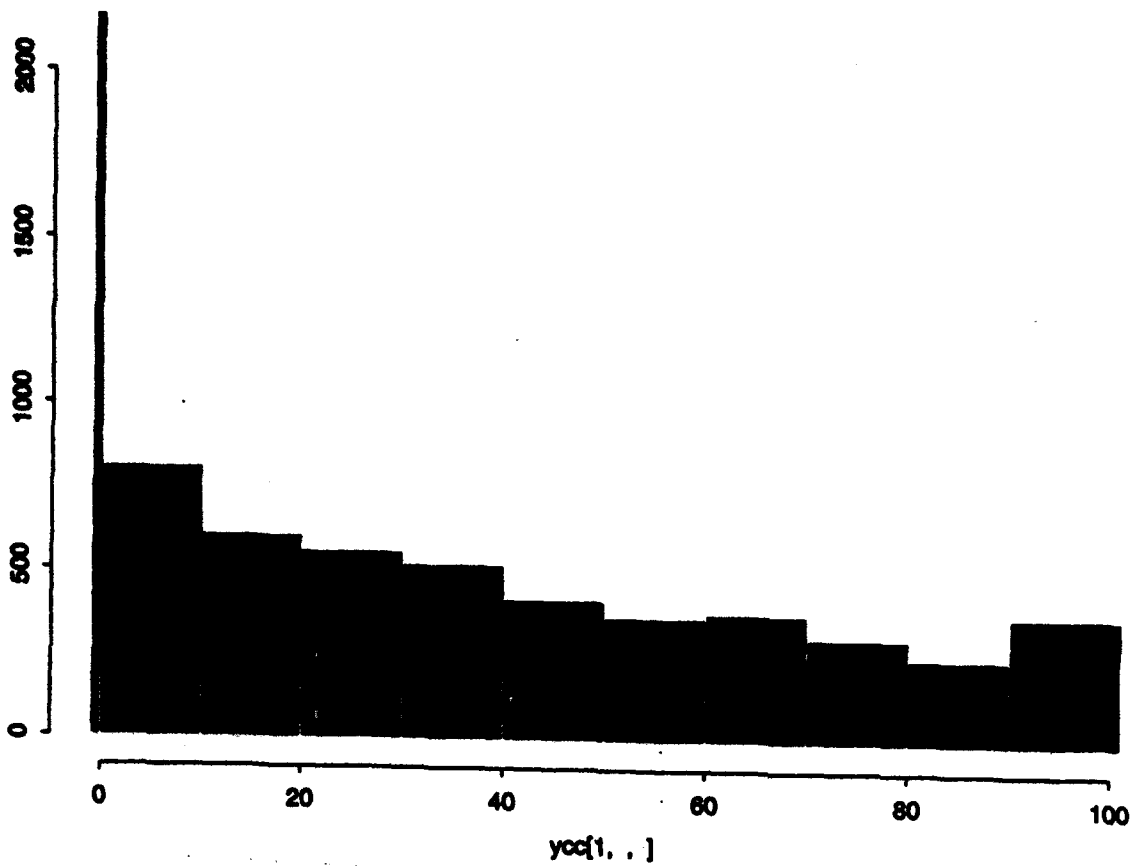


Figure 18: Frequency distribution of layer 1 (1000 hPa) cloud cover.

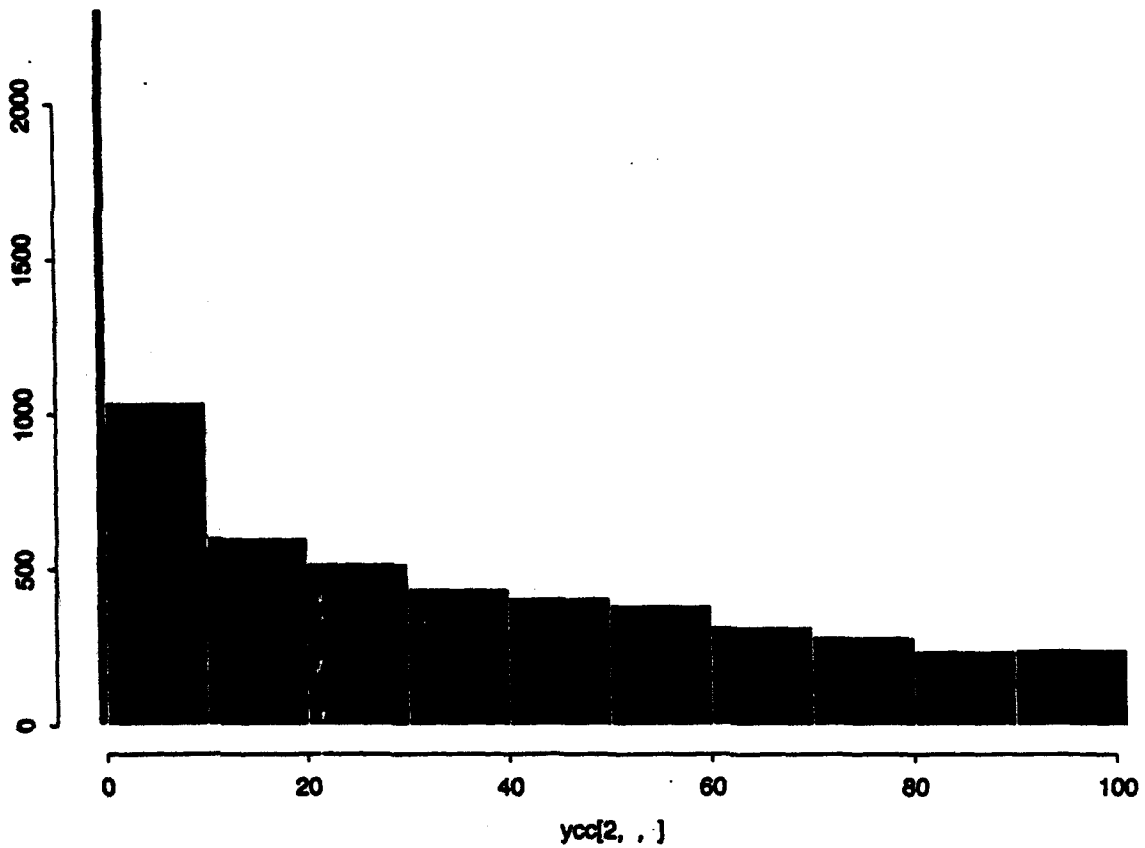


Figure 19: Frequency distribution of the layer 2 (850 hPa) cloud cover.

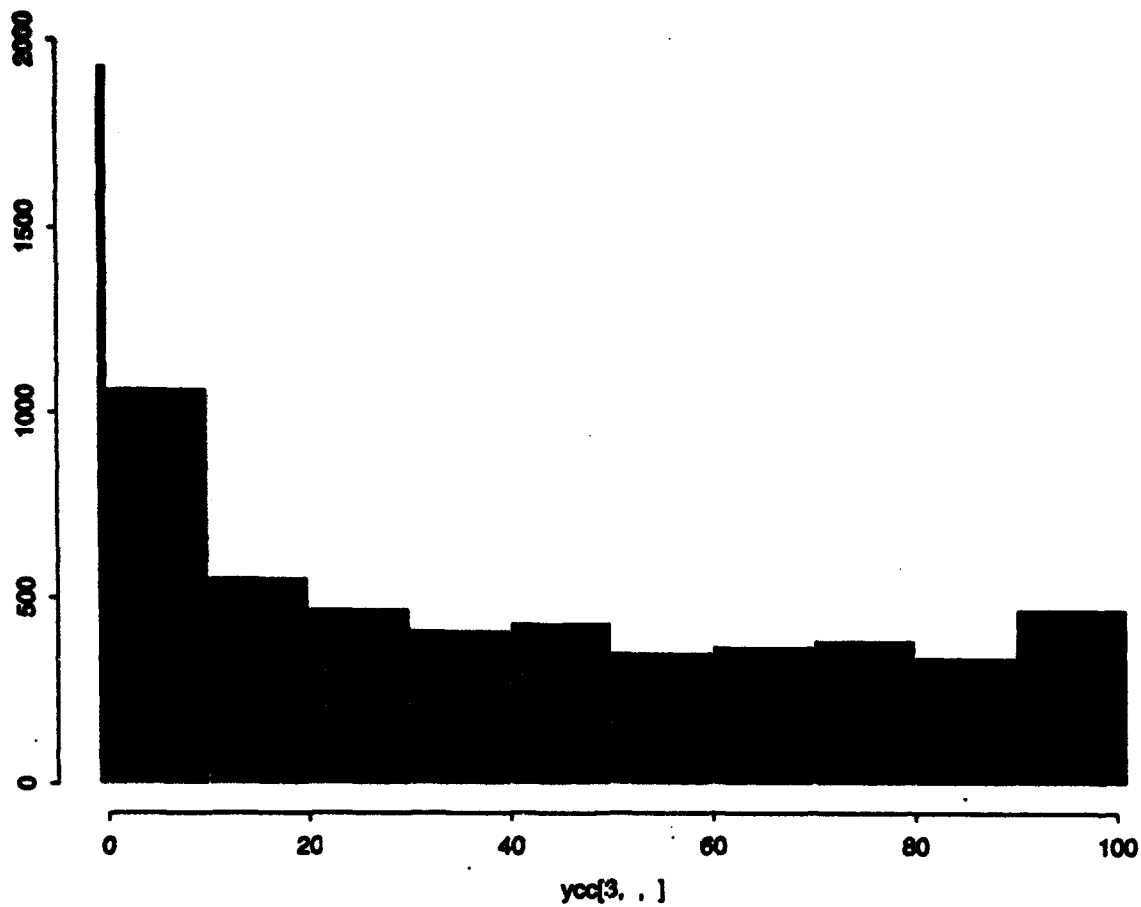


Figure 20: Frequency distribution of the layer 3 (700 hPa) cloud cover.

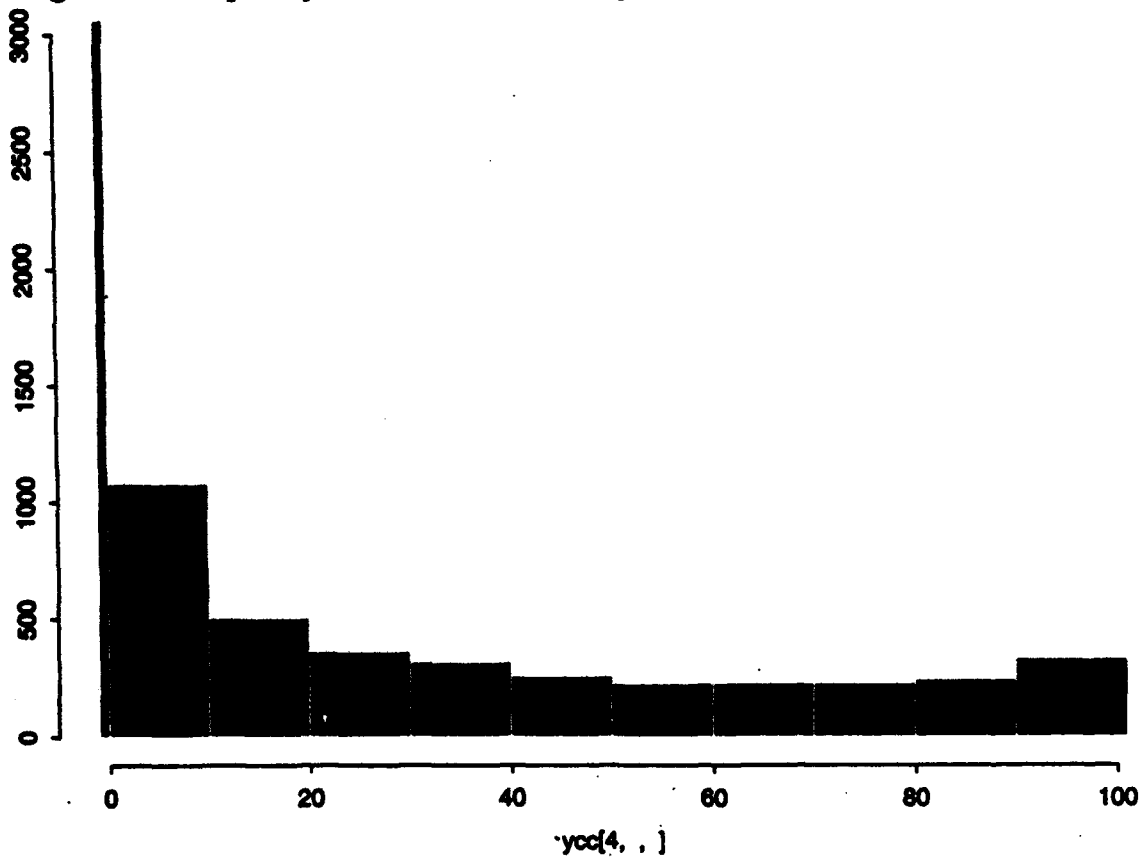


Figure 21: Frequency distribution of the layer 4 (500 hPa) cloud cover.

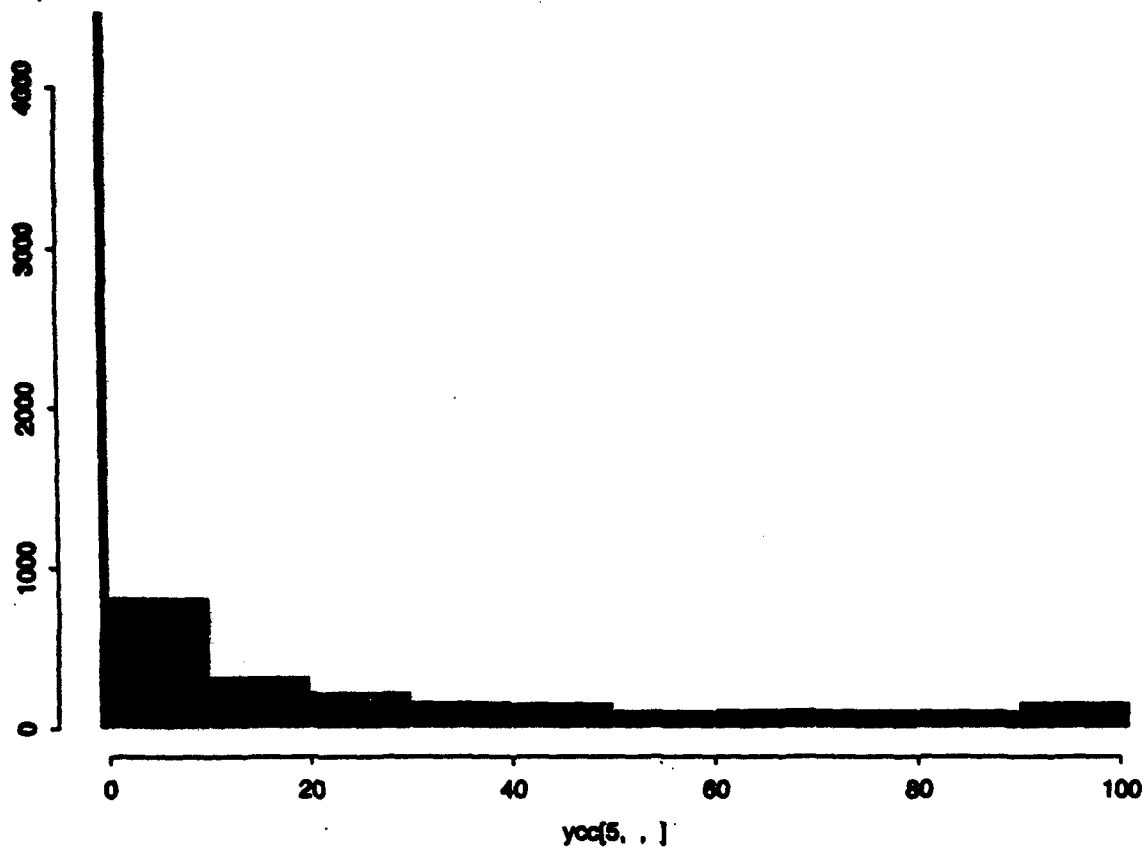


Figure 22: Frequency distribution of the layer 5 (400 hPa) cloud cover.

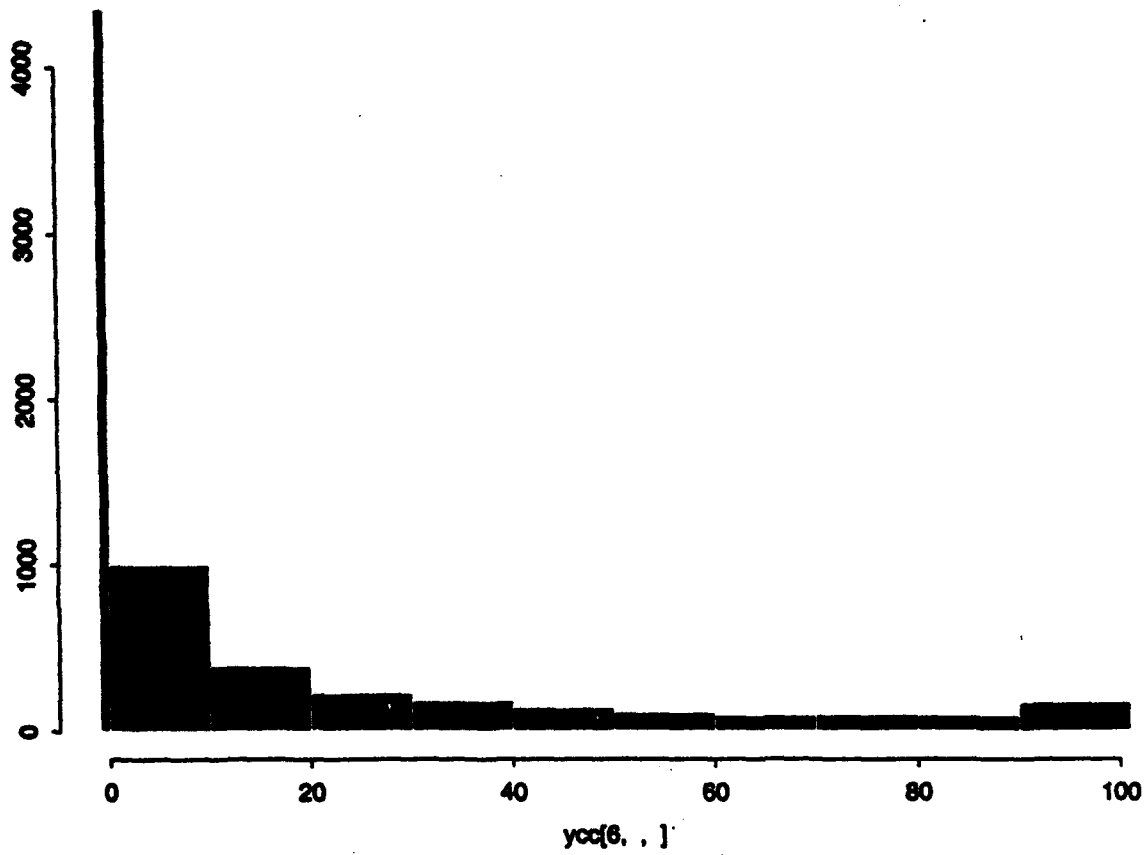


Figure 23: Frequency distribution of the layer 6 (300 hPa and above) cloud cover.

3.2. Vertical stacking

Although total and layer cloud amounts are treated independently from one another during data compaction, they are, of course, related. The exact relation depends on the amount of overlap between the layer clouds. Since at least one of our cloud forecast schemes is designed to predict layer cloud cover, knowledge of the amount of overlap in the compacted data set is important for the accurate prediction of total cloud cover. We therefore investigated what kind of overlap assumption results in the best agreement between the observed and computed (from layer cloud amounts) total cloud cover in the 1/2-mesh RTNEPH data.

The two limiting cases are random overlap (location of layer clouds completely uncorrelated, resulting in largest total cloud amount), and maximum overlap (location of layer clouds perfectly correlated, resulting in smallest total cloud amount). The formulas relating total cloud cover (c_{tot}) to the cloud cover of N cloud layers (c_{lay}) for these two cases are given by (all cloud amounts normalized to the interval [0,1]):

$$\text{Random overlap: } c_{tot} = 1 - \prod_{k=1}^N (1 - c_{lay}(k)) \quad (1)$$

$$\text{Maximum overlap: } c_{tot} = \max(c_{lay}(k)) , k = 1, N \quad (2)$$

If only two layers are combined, these two formulas can be combined into (see p. A1-A4 of (Mitchell and Hahn, 1989), hereafter referred to as MH):

$$c_{A,B} = c_A + (1 - c_A)c_B(1 - r), \quad (3)$$

where $c_{A,B}$ represents the total cloud cover of layers A and B, and where it is assumed that $c_A \geq c_B$. The parameter r ($0 \geq r \geq 1$) determines the amount of overlap: $r=0$ ($r=1$) correspond to random (maximum) overlap (note that MH use the quantity $R=1-r$ in their formula). Intermediate values of r correspond to partial correlations of layer clouds (note, however, that the quantity r is not equal to the correlation coefficient: since cloud amounts are confined to the

interval [0,1], the combination of the amounts is highly nonlinear and cannot be simply related to the correlation coefficient).

When combining more than 2 layers, the above 2-layer equation must be applied successively until the total cloud amount has been computed. In the case of the compacted 1/2-mesh RTNEPH data, the 6 layers are compacted in 3 steps (as in MH):

Step 1: combine amounts c_1 and c_2 into $c_{1,2}$, c_5 and c_6 into $c_{5,6}$

Step 2: combine amounts c_3 and $c_{1,2}$ into $c_{1,3}$, c_4 and $c_{5,6}$ into $c_{4,6}$

Step 3: combine amounts $c_{1,3}$ and $c_{4,6}$ into c_{tot}

The notation $c_{i,j}$ denotes the amount derived from compacting layers i through j ; the layer indices correspond roughly to the mandatory pressure levels (1-1000 hPa; 2-850 hPa; 3-700 hPa; 4-500 hPa; 5-400 hPa; 6-300 hPa).

MH used different values of r for the different layer combinations. They are plotted in Figure 24, which shows the value r as a function of the distance between the layers, expressed as the absolute value of the natural logarithm of the ratio of the two layer pressures ($\ln(P_A/P_B)$). Also shown are curves of r from the formula

$$r = \frac{1}{1 + \alpha (\ln(P_A/P_B))^2} \quad (4)$$

for two values of α . The parameter α can be expressed in terms of the minimum value of r , r_{min} , which is the value of r at the maximum layer separation (layers 1 and 6, $\ln(1000/300) = 1.204$). The upper curve ($r_{min} = 0.2$) was fitted to the value used by MH for the maximum layer separation. Note that random overlap corresponds to a value of $r_{min} = 0$, and maximum overlap to $r_{min} = 1$.

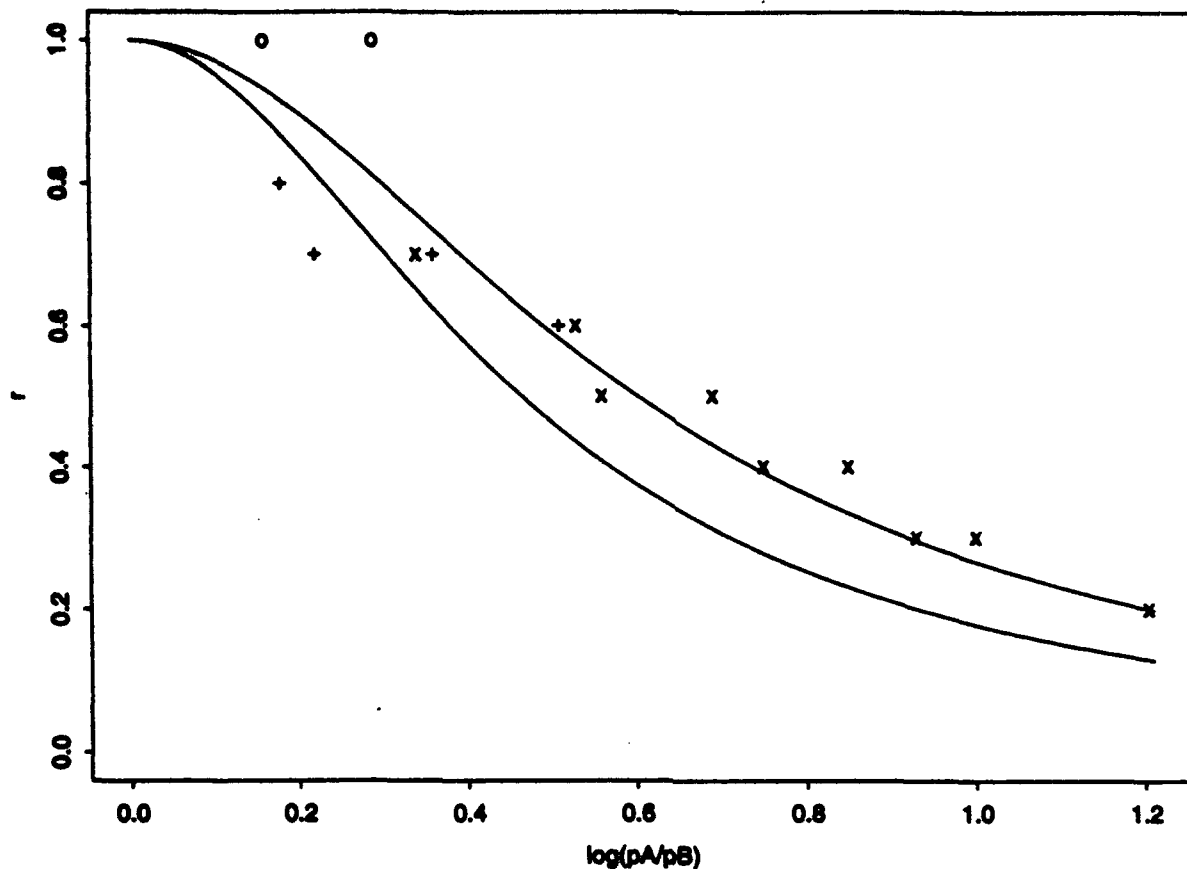


Figure 24: The overlap parameter r as a function of the layer separation. Shown are values used by MH during step 1 (circles), step 2 (+), and step 3 (x), as well as from the formula for $r_{min} = 0.2$ (top curve) and $r_{min} = 0.13$ (bottom curve).

We tested different values of r_{min} with the compacted 1/2-mesh data for one time period (00 UTC 15 January 1989). For the purpose of computing the appropriate value of r in steps 2 and 3 and of the stacking algorithm, a weighted (by cloud amount) average value of the layer pressure logarithm is computed for the compacted cloud amounts. For each value of r_{min} , the total cloud amount was computed from the observed layer cloud amounts through the stacking algorithm described above, and compared with the total amount obtained by horizontal compaction of the reported total cloud amounts. Best agreement between the computed and observed total cloud amounts (using only points with nonzero cloud cover) was found for values of r_{min} between 0.12 and 0.14. These tests were repeated for various cutoff values of the sum of weights in the 1/2-mesh working sums (using values of 1 - the minimum possible, 16, 32, 48, and 64 - the maximum possible). The agreement between computed and observed total cloud amount increases with the value of the cutoff (minimum rms errors decrease from 6.42% to 5.35%, maximum

correlation coefficients increase from .984 to .987); however, optimum values of r_{min} are not affected by the value of the cutoff.

The above computations were repeated, for cutoff=64 (i.e., using only half-mesh points without any missing 1/8-mesh data), using all time periods in July. Results from this computation (see Table 1) confirm the conclusion for the January case, with an optimum value of r_{min} of between 0.12 and 0.14. Random overlap ($r_{min} = 0$) leads to a substantial positive bias, and maximum overlap ($r_{min} = 1$) to a substantial negative bias. Between those extreme values, the bias is monotonically decreasing, being closest to zero for $r_{min} = .135$ (for both the January case and the July run). However, the standard deviation of the error is slightly smaller for nearby values of r_{min} , resulting in minimum RMSE values for $r_{min} = .130$ in the January case, and $r_{min} = .125$ for the July run. The value $r_{min} = .13$ is chosen on the basis of these results for vertical stacking of layer cloud amounts in our cloud forecast schemes. The corresponding curve of R is plotted as the lower curve in Figure 24; we note that this corresponds to assuming slightly less overlap between layers than MH.

Table 1: Statistics of total cloud cover for the month of July, 1989. Shown are observed half-mesh values (cloudy grid points only), computed values (from layer amounts, using different values of r_{min}), and the difference between computed and observed. Shown are the mean (%), standard deviation (%), and root mean square (%); the last column contains the correlation coefficient between computed and observed total cloudiness.

r_{min}	Field	Mean	Std Dev	RMS	Corr
	Observed	59.41	35.63	69.28	
0.000	Comp	69.93	34.33	77.90	
0.000	Comp - Ob	10.51	10.97	15.20	0.951
0.010	Comp	67.61	34.41	75.86	
0.010	Comp - Ob	8.19	9.19	12.31	0.966
0.050	Comp	63.16	34.23	71.84	
0.050	Comp - Ob	3.74	6.37	7.39	0.984
0.100	Comp	60.56	33.92	69.41	
0.100	Comp - Ob	1.14	5.40	5.52	0.989
0.120	Comp	59.86	33.81	68.75	
0.120	Comp - Ob	0.44	5.28	5.29	0.989
0.125	Comp	59.70	33.79	68.60	
0.125	Comp - Ob	0.28	5.25	5.26	0.989
0.130	Comp	59.55	33.77	68.46	
0.130	Comp - Ob	0.13	5.27	5.27	0.989
0.135	Comp	59.40	33.75	68.32	
0.135	Comp - Ob	0.00	5.27	5.27	0.989
0.140	Comp	59.27	33.72	68.19	
0.140	Comp - Ob	-0.14	5.26	5.26	0.989
0.160	Comp	58.76	33.65	67.71	
0.160	Comp - Ob	-0.65	5.27	5.31	0.990
0.180	Comp	58.32	33.58	67.30	
0.180	Comp - Ob	-1.09	5.33	5.44	0.989

Table 1 (continued)

r_{min}	Field	Mean	Std Dev	RMS	Corr
0.200	Comp	57.93	33.51	66.93	
0.200	Comp - Ob	-1.47	5.40	5.60	0.989
0.300	Comp	56.52	33.28	65.59	
0.300	Comp - Ob	-2.89	5.87	6.54	0.987
0.400	Comp	55.61	33.13	64.73	
0.400	Comp - Ob	-3.79	6.36	7.40	0.985
0.500	Comp	54.97	33.02	64.13	
0.500	Comp - Ob	-4.43	6.73	8.06	0.983
0.600	Comp	54.49	32.95	63.68	
0.600	Comp - Ob	-4.91	7.06	8.60	0.981
0.700	Comp	54.12	32.90	63.33	
0.700	Comp - Ob	-5.29	7.34	9.05	0.980
0.800	Comp	53.82	32.85	63.05	
0.800	Comp - Ob	-5.59	7.55	9.40	0.978
0.900	Comp	53.57	32.82	62.83	
0.900	Comp - Ob	-5.83	7.74	9.69	0.977
1.000	Comp	53.37	32.79	62.64	
1.000	Comp - Ob	-6.04	7.90	9.95	0.976

3.3. Coordinate transformations

Use of the RTNEPH data in conjunction with the forecast model requires coordinate transformations between the 1/2-mesh grid, which is a regularly spaced grid on a polar stereographic projection, and the grids used by the forecast model and analysis, which are regularly spaced in longitude, at either regularly spaced (analysis) or Gaussian latitudes (GSM transform grid). Routines exist to convert longitude/latitude coordinates to coordinates in the polar stereographic projection; however, for the purpose of converting grid-box average quantities, the locations of all lat-lon grid boxes overlapping a given 1/2-

mesh box are needed. Lookup tables have been generated that contain, for each RTNEPH point, the locations of all GSM transform grid (128 longitudes by 102 Gaussian latitudes) boxes overlapping the RTNEPH grid box, and the fraction of area covered by the RTNEPH grid box. RTNEPH cloud cover can thus be properly transformed to the GSM transform grid by forming weighted averages.

4. Cloud Forecast Scheme Development

At least one cloud forecasting scheme will be developed for the APGSM and implemented. Candidate techniques are the CCA technique and the Slingo-type cloud parameterization used in the APGSM. In addition, the scheme developed by PL will be implemented as it becomes available.

We have begun the cloud forecast scheme development for a modified CCA scheme, in which the curves are derived for samples stratified by synoptic regime rather than by geography and season. We are presently deriving the synoptic weather regimes from HIRAS analysis data. For this purpose, the HIRAS analyses are preprocessed, i.e. vertically interpolated to the GSM σ -layers, and spectrally truncated to R40; this was done to allow easier application of the synoptic weather regimes (which are defined in terms of vertical profiles of atmospheric quantities) to the GSM output. The preprocessed analyses are sampled at locations spaced approximately 1000 km apart in the Northern Hemisphere, for four time periods, spaced one week apart, in each of the four months (January, April, July, and October). To reduce the number of degrees of freedom, the vertical profiles of all atmospheric quantities (temperature, wind speed, relative humidity), along with the values of surface pressure and precipitable water, are represented by empirical orthogonal functions (EOFs), and weather regimes are then defined in terms of clusters of EOF coefficient values.

Some preliminary results of this analysis are shown in Figures 25 - 30. Figures 25 - 28 show frequency distributions of the atmospheric variables at the model σ -layers (the numbering convention adopted here is such that 1 corresponds to the lowest layer, 18 to the highest). This is based on a sample of 4672 points (292 points each for 16 time periods). Temperature at the lowest layers is skewed toward the high end of the spectrum, and the situation is reversed in the upper layers, with more nearly symmetrical distributions at

Probability Distribution of Temperature at 18 Sigma Levels

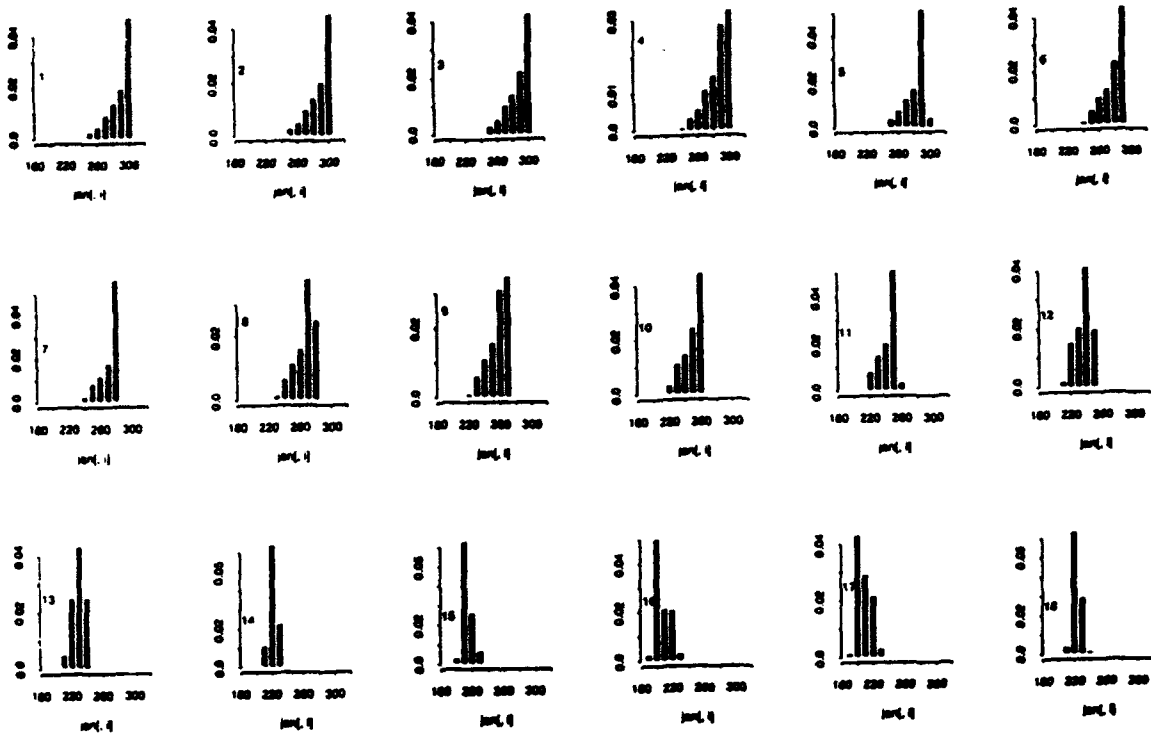


Figure 25: Frequency distribution of temperature on the model s-layers for the preprocessed HIRAS data.

Probability Distribution of Relative Humidity at 18 Sigma Levels

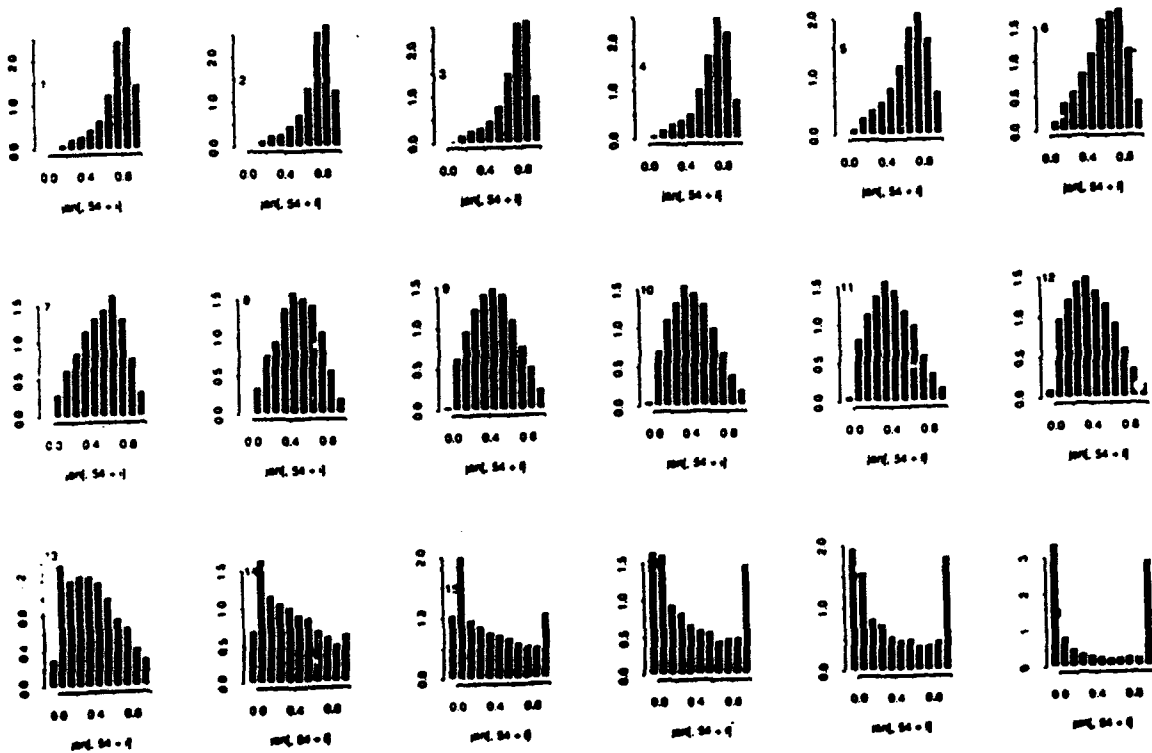


Figure 26: As Figure 25, but for RH.

Probability Distribution of U-velocity at 18 Sigma Levels

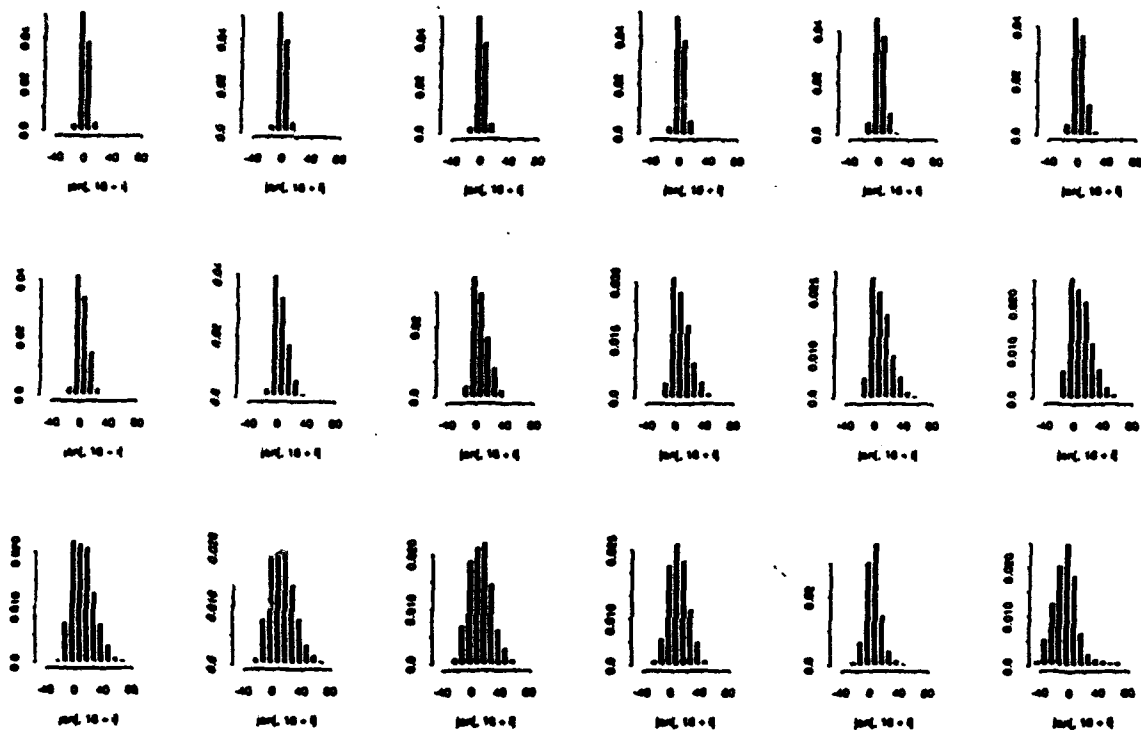


Figure 27: As Figure 25, but for zonal velocity u .

Probability Distribution of V-velocity at 18 Sigma Levels

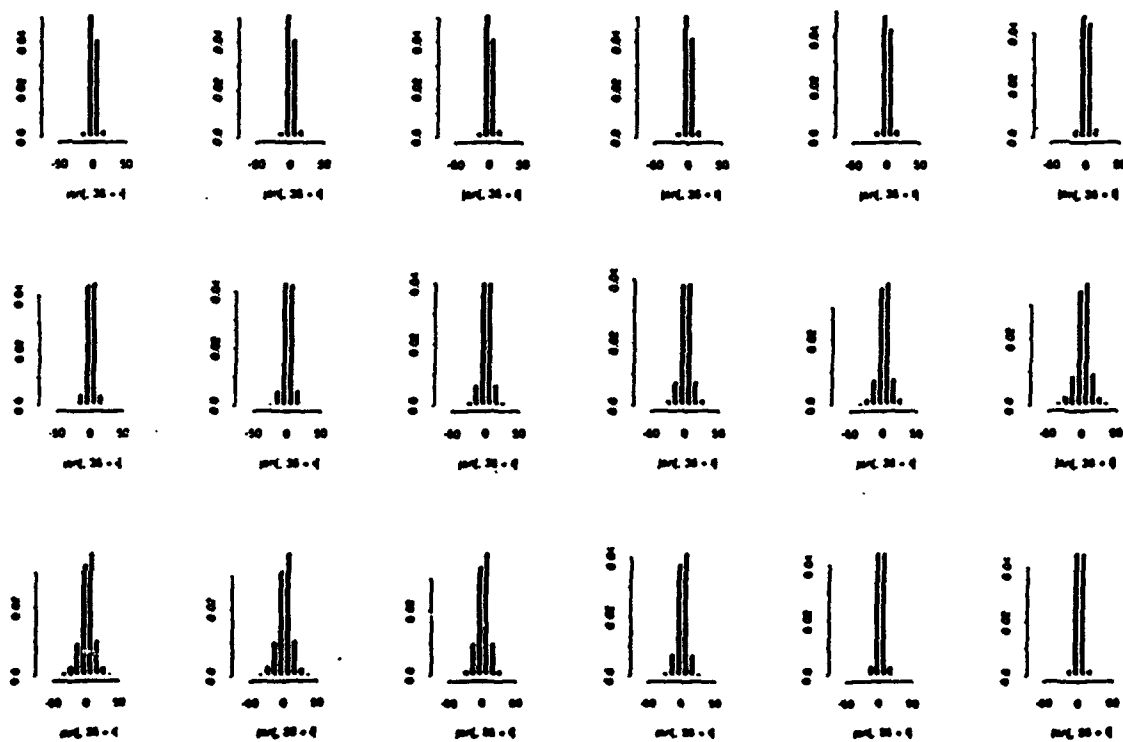


Figure 28: As Figure 25, but for meridional velocity v .

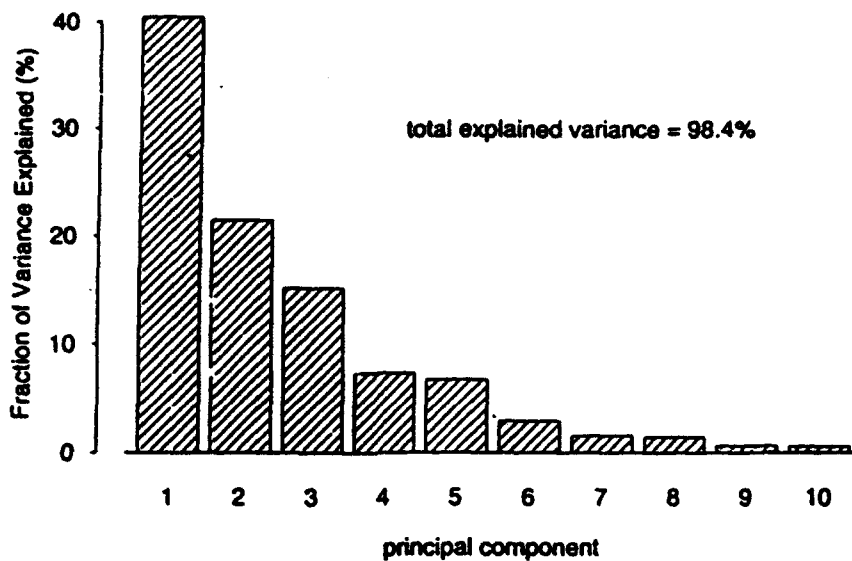


Figure 29: Percent of variance explained by the first 10 EOFs..

First Eight Principal Components Resolve 97% of Variance

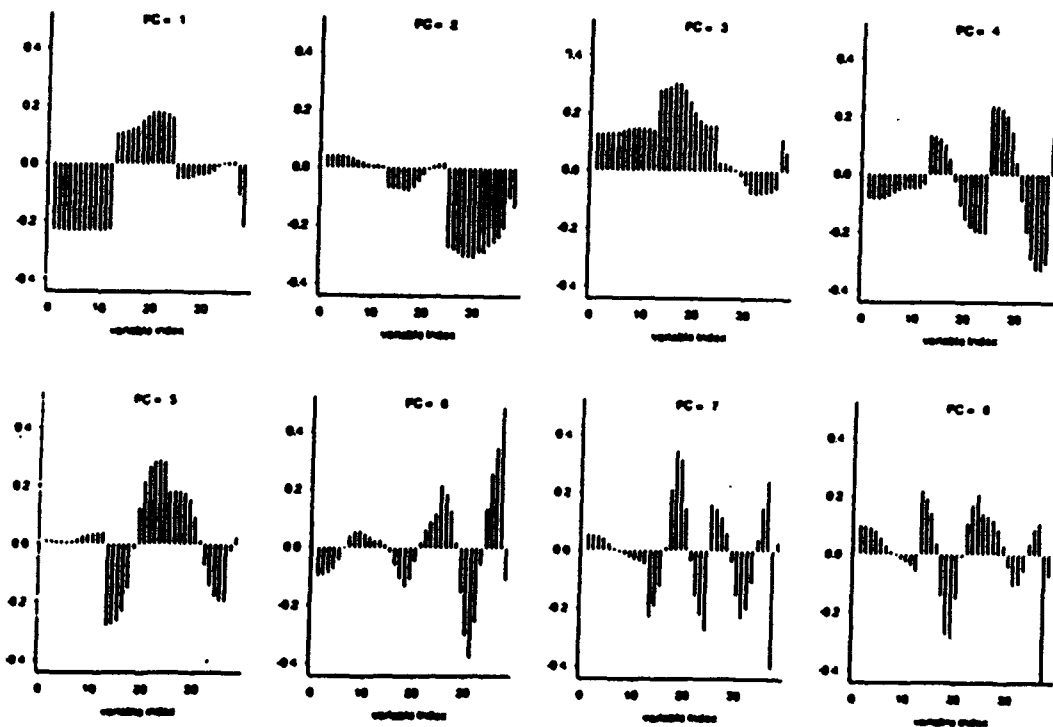


Figure 30: Structure of the first 8 EOFs (see text for explanation).

intermediate altitudes. The distribution of relative humidity reflects the commonly observed decrease of relative humidity with height. Another interesting feature of the RH distributions is the progressively bimodal distribution at layers 14-18. This feature is an artifact of the preprocessor, which extrapolates RH above 300 mb (layers 13-18). This problem was discussed previously in section 2.3. Distributions of u reflect the predominance of westerlies in the troposphere, those of v the generally smaller values and overall near-zero mean.

For the EOF analysis, the data volume was reduced by considering only the lowest 12 σ -layers (roughly corresponding to data at 325 hPa and below), and by combining the u - and v -velocity into total wind speed. This restricts attention to the vertical domain of greatest relevance for cloud cover, and it avoids use of extrapolated values of RH. Thus each data point is represented by 12 values of temperature, wind speed, and RH, and the value of surface pressure and precipitable water, for a total of 38 values. The data was normalized for each level by subtracting the mean, and dividing by the standard deviation (mean and standard deviation were computed from the entire sample for each variable and layer). Figure 29 shows the fraction of variance explained by the first 10 principal components of the covariance matrix of the normalized variables. The total variance explained by the first 10 EOFs is 98.4%, but it is evident that even the first 3 EOFs explain over 3/4 of the variance. Figure 30 shows the structure of the first 8 EOFs. In these plots, the first 12 values correspond to (normalized) temperature, the next 12 to wind speed, the next 12 to RH, and the last 2 to surface pressure and precipitable water. The first EOF is dominated by the variance of temperature and wind speed, and only a weak RH signal. EOF 2 is dominated by the variance of RH. There is little or no vertical structure of each variable in the first two EOFs. EOFs 3 and 4 have only moderate to weak signals in temperature, and exhibit some vertical structure in wind speed and RH. Higher order EOFs contain vertical variations in temperature, as well as wind speed and RH.

We are currently investigating the definition of weather regimes in terms of clusters of EOF coefficient values. One of the possible areas of investigation is whether modifications of the definition of the EOFs (using different variables or

combinations of variables, or using different weights for different variables) will result in weather regimes that show a greater correspondence to cloud cover.

5. References

Ballish, B., 1980: Initialization, theory and application to the NMC spectral model. Ph.D. Thesis, Department of Meteorology, University of Maryland

Brenner, S., C. H. Yang and S. Y. K. Yee, 1982: The AFGL spectral model of the moist global atmosphere. AFGL-TR-82-0393. Air Force Geophysics Laboratory, Hanscom AFB, MA. 65 pp. [NTIS ADA129283]

Hamill, T. M., R. P. d'Entremont and J. T. Bunting, 1992: A description of the Air Force Real-Time Nephanalysis model. *Wea. Forecasting.*, 7, 288-306.

Kuo, H. L., 1965: On formation and intensification of tropical cyclones through latent heat release by cumulus convection. *J. Atmos. Sci.*, 22, 40-63.

Liou, K.-N., S.-C. Ou, S. Kinne and G. Koenig, 1984: Radiation parameterization program for use in general circulation models. AFGL-TR-84-0217. Air Force Geophysics Laboratory, Hanscom AFB, MA. [NTIS ADA148015]

Louis, J.-F., R. N. Hoffman, T. Nehr Korn and D. Norquist, 1989: Observing systems experiments using the AFGL four-dimensional data assimilation system. *Mon. Wea. Rev.*, 117, 2186-2203.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialization. *Beitr. Phys. Atmos.*, 50, 253-271.

Mahrt, L., H.-L. Pan, J. Paumier and I. Troen, 1984: A boundary layer parameterization for a general circulation model. AFGL-TR-84-0063. Air Force Geophysics Laboratory, Hanscom AFB, MA. [NTIS ADA144224]

- Mitchell, K. E. and D. C. Hahn, 1989: Development of a Cloud Forecast Scheme for the GL Baseline Global Spectral Model. GL-TR-89-0343. Geophysics Laboratory, Hanscom AFB, MA. 147 pp. [NTIS ADA231595]
- Nehrkorn, T., R. Hoffman, J.-F. Louis and M. Zivkovic, 1992: An enhanced global spectral model. PL-TR-92-2011. Geophysics Directorate, Phillips Laboratory, Hanscom AFB, MA 01731. 52 pp. [NTIS ADA251242]
- Nehrkorn, T., R. N. Hoffman, J.-F. Louis, R. G. Isaacs and J.-L. Moncet, 1993: Analysis and forecast improvements from simulated satellite water vapor profiles and rainfall using a global data assimilation system. *Mon. Wea. Rev.*, **121**, 2727-2739.
- Norquist, D. C., C.-H. Yang, S. Chang and D. C. Hahn, 1992: Phillips laboratory global spectral numerical weather prediction model. PL-TR-92-2225. Phillips Laboratory, Hanscom AFB, MA. pp. 154 pp. [NTIS ADA 267293]
- Ou, S.-C. and K.-N. Liou, 1988: Development of radiation and cloud parameterization programs for AFGL Global Models. AFGL-TR-88-0018. Air Force Geophysics Laboratory, Hanscom AFB, MA. 88 pp. [NTIS ADA193369]
- Schattel, J., 1992: Refinement and testing of the radiative transfer parameterization in the PL global spectral model. PL-TR-92-2169. Phillips Laboratory, Hanscom AFB, MA. [NTIS ADA256840]
- Sela, J. G., 1980: Spectral modeling at the National Meteorological Center. *Mon. Wea. Rev.*, **108**, 1279-1292.
- Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.*, **117**, 1779-1800.
- Vernekar, A. D., J. Zhou and B. Kirtman, 1991: A comparison of systematic errors in AFGL and COLA forecast models. PL-TR-91-2164. Phillips Laboratory, Hanscom AFB, MA. [NTIS ADA244458]

Yang, C.-H., K. Mitchell, D. Norquist and S. Yee, 1989: Diagnostics for and evaluations of new physical parameterization schemes for global NWP models. GL-TR-89-0158. Geophysics Laboratory, Hanscom AFB, MA. [NTIS ADA228033]