One-Month Forecast Experiments—without Anomaly Boundary Forcings

K. MIYAKODA, J. SIRUTIS AND J. PLOSHAY

Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, NJ 08542

(Manuscript received 17 August 1985, in final form 19 May 1986)

ABSTRACT

A series of one-month forecasts were carried out for eight January cases, using a particular prediction model and prescribing climatological sea-surface temperature as the boundary condition. Each forecast is a stochastic prediction that consists of three individual integrations. These forecasts start with observed initial conditions derived from datasets of three meteorological centers. The forecast skill was assessed with respect to time means of variables based on the ensemble average of three forecasts. The time or space filter is essential to suppress unpredictable components of atmospheric variabilities and thereby to make an attempt at extending the limit of predictability. The circulation patterns of the three individual integrations tend to be similar to each other on the one-month time scale, implying that forecasts for the 10 day (or 20 day) means are not fully stochastic. The overall results indicate that the 10-day mean height prognoses resemble observations very well in the first ten days, and then start to lose similarity to real states, and yet there is some recognizable skill in the last ten days of the month. The main interests in this study are the feasibility of one-month forecasts, the adequacy of initial conditions produced by a particular data assimilation, and the growth of stochastic uncertainty. An outstanding problem turns out to be a considerable degree of systematic error included in the prediction model, which is now known to be "climate drift." Forecast errors are largely due to the model's systematic bias. Thus, forecast skill scores are substantially raised if the final prognoses are adjusted for the model's known climatic drift.

1. Introduction

A preliminary study of one-month forecasts (Miyakoda et al., 1983) has revealed that a pronounced blocking event of January 1977 was successfully reproduced by some general circulation models (GCM) but not by others, suggesting that the models with relatively high spatial resolution as well as refined subgrid-scale parameterizations are essential for simulating slowly varying planetary waves (Bengtsson, 1981). The reason for the need of such a high quality (realistic) GCM for monthly forecasts is that forecasts start from real initial data and that the initial conditions are dominant factors in the dynamics in a subsequent month or so. However, models are basically different from reality, and therefore prediction models should handle this discrepancy properly without damaging crucial components contained in the initial state. In other words, a model's imperfection is unavoidable, but the gap between the model and nature should not be excessively large.

Based on this consideration, a particular GCM was selected and frozen in 1980. Since then, this model has been applied to 30-day integrations for eight January cases. The main objective is to investigate the feasibility of one-month forecasts based on a dynamical approach.

In order to extend the forecast range beyond the limit of deterministic predictability, i.e., about two weeks (Lorenz, 1982), two factors, at least, have to be taken into account, that is, the *time average* and the *ensemble average* of stochastic forecasts. The time averaging process is essential to filter out the possibly unpredictable high-frequency components of atmospheric circulation and to retain slowly varying components (Smagorinsky, 1969; Gilchrist, 1977). As discussed by Blackmon et al. (1977) and Shukla and Mo (1983), the low-pass variances are particularly dominant in the blocking prone areas. In fact, the simulation of blocks is one of the major tasks of long-range forecasts (Gilchrist, 1982; Shukla and Gutzler, 1983). The stochastic forecast process is necessary to determine the *probabilistic mean* and *its scatter* (Epstein, 1969), though the stochastic process alone may not extend the predictability range. In this study, three different initial conditions are used for these processes. An aspect of interest is to see to what extent the three forecasts (realizations) are separated in the monthly time scale. Another objective of this study is to investigate the adequacy of initial conditions for the one-month forecasts. The assessment is made by comparing three forecasts based on the analyses produced by different four-dimensional data assimilations at three centers.

2. Design of experiments

a. Prediction system

The GCM is approximately the same as the one used in the pilot study (Miyakoda et al., 1983), i.e., the N48L9-E. The finite difference method is used with

the N48 horizontal resolution and nine vertical levels, where the N48 denotes 48 equally spaced gridpoints between a pole and the equator (1.875° meridional distance). The E-type subgrid-scale physics (Miyakoda and Sirutis, 1977, 1984) consists of the turbulent closure formulation with hierarchy level of 2.5 of Mellor–Yamada (1974), the Monin–Obukhov similarity scheme of the surface boundary layer flux, moist convective adjustment as the cumulus parameterization, nonlinear viscosity for the lateral diffusion, and heat conduction in the soil. Perhaps it may be useful to note that the important processes missing in this model are diurnal variation and cloud–radiation interaction. It is also noted that the following minor differences exist between this model and the one in the pilot study. The roughness over mountains is increased based on the formulation of C. T. Gordon (personal communication, 1980); the lower boundary condition for soil temperature at the 5 meter depth is specified by the climatological annual mean surface temperatures as a function of spatial coordinate, whereas the previous model employed a constant, i.e., 280 K everywhere; and the snow forecasts are based on the method described in Miyakoda and Strickler (1981).

Boundary forcing over the ocean is specified by the climatological norm of sea surface temperature that changes with season. In the initial condition, soil moisture and snow/ice cover are the climatological norms of January taken from the GCM climate study (Manabe et al., 1974).

TABLE 1. Cases and initial data.

number	initial time	Level III data A	B	C
1	00GMT 1 JAN 1977	GFDL–1		
	"		NMC	
	00GMT 2 JAN 1977			GFDL–1
2	12GMT 31 DEC 1977	GFDL–1		
	00GMT 1 JAN 1978		NMC	
	12GMT 1 JAN 1978			GFDL–1
3	00GMT 1 JAN 1979	GFDL–1		
	"		NMC	
	"			ECMWF
4	00GMT 16 JAN 1979	GFDL–1		
	"		NMC	
	"			ECMWF
5	00GMT 1 JAN 1980	GFDL–2		
	"		NMC	
	"			ECMWF
6	12GMT 1 JAN 1981	GFDL–2		
	00GMT 1 JAN 1981		NMC	
	"			ECMWF
7	00GMT 1 JAN 1982	GFDL–2		
	"		NMC	
	"			ECMWF
8	00GMT 1 JAN 1983	GFDL–2		
	"		NMC	
	"			ECMWF

b. *Cases and initial data*

Eight January cases are taken from 1977 to 1983 (Table 1). Seven cases have approximately 1 January as the initial time, except the year 1979. Each case includes three initial conditions, which were derived from the analyses produced by four-dimensional data assimilation schemes of GFDL (Geophysical Fluid Dynamics Laboratory), NMC (National Meteorological Center), and ECMWF (European Centre for Medium-Range Weather Forecasts). In Table 1, these initial conditions are listed by the A, B and C in the column of Level III dataset. Note that Level III denotes the so-called gridded analysis, in contrast to the original observations, denoted by Level II or I.

The initial condition A is based on the GFDL Level III data. For the first four cases (1977–79), the data assimilation scheme is the FGGE system (Ploshay et al., 1983; Stern et al., 1984), which is denoted by the GFDL-1, while for the last four cases (1980–83), the scheme is the post-FGGE system (Puri and Stern, 1984), denoted by GFDL-2. In the GFDL four-dimensional data assimilation, a continuous data insertion method has been used consistently. The post-FGGE scheme has been improved by applying a linear normal-mode initialization to the injected data increments and using a wider range of data collection for the determination of insertion data. Note that the GFDL-2 incorporates the observed SST (sea surface temperature), prepared by R. Reynolds, Climate Analysis Center, Washington, DC, for the lower boundary condition.

The initial condition B is based on the NMC Level III data, and the initial condition C is based mostly on the ECMWF Level III data, though the first two cases are substituted by the GFDL-1 (see Table 1).

The NMC and the ECMWF datasets are the operational products, except for the 1979 of ECMWF. The global (operational) Level III data have been available only since 1977 from NMC and since 1980 from ECMWF. These constraints led us to determine the selection of sample cases in this study. The data analysis schemes of these two centers use the intermittent data assimilation method and have been continually improved in these years. In particular, during the period before and after the FGGE, the schemes have undergone extensive modification (see Daley et al., 1985; Hollingsworth et al., 1985).

The case of 16 January 1979 is the one with which the ECMWF had a best 8-day forecast in 1980 (Bengtsson, 1981). For this reason, and because this FGGE case has been recommended for a careful and intensive study by Baumhefner and Bettge (1981) under the auspices of the Working Group on Numerical Experimentation, World Meteorological Organization, this case was added to this series of experiments.

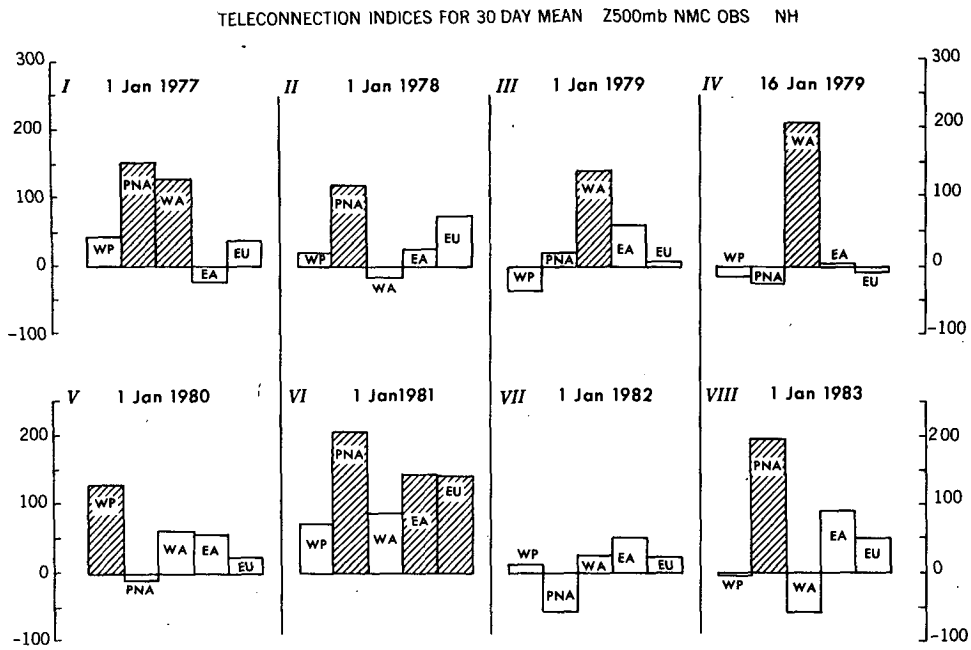

FIG. 1. Teleconnection indices for the eight January cases.

c. Teleconnection indices

In order to depict the characteristic feature of a general circulation pattern, Wallace and Gutzler (1980) proposed a teleconnection index based on the past 30-year data (see also Esbensen, 1984). These indices appear convenient and suitable for identifying and classifying teleconnection patterns. A set of five categories of indices characterizes a height field. The index is calculated on the basis of a specific formula for the respective category by the use of height values (see Miyakoda and Sirutis 1985).

Figure 1 is intended to give an overview of the eight January cases, using the index for the monthly mean 500 mb geopotential height fields. To explain the diagram, let us take as an example the first case, i.e., January 1977. The geopotential height field is characterized by a set of five indices, represented by vertical bars in the figure, labeled by WP (West Pacific), PNA (Pacific–North America), WA (West Atlantic), EA (East Atlantic) and EU (Eurasia). The values of the bar larger than 100 and less than −100 are hatched differently (see Fig. 2).

The observation of Fig. 1 together with the "weather and circulation" review of Wagner (1977; 1978; 1979; 1980; 1981; 1982) and Quiroz (1983) may suggest the following.

- Positive indices are dominant, particularly PNA and WA, indicating that these eight January cases may belong to a particular climate regime, which is biased relative to a long term (more than 30 years) climatological norm. In the period of 1977–81, winter temperatures at southwest cities of the United States, for example, were consistently lower than those in the previous 5 years, 1971–76 (Namias, 1982).
- Investigating the 89-year data, Karl et al. (1984) found that an uncharacteristic spell of abnormal winters occurred from 1975–76 through 1982–83 as defined by areally averaged temperature over the United States. Six of eight winters during this period were either abnormally warm or cold, relatively moderate winters being 1979–80 and 1981–82.
- The eight cases include the historic eastern U.S. cold event in 1977 and the most pronounced El Niño phenomenon of 1983. In many aspects, January 1983 was an extraordinary month (Quiroz, 1983).
- In all cases, blocking actions are included. The case of 1982 is the only month close to normal climatology in terms of the teleconnection index. Even in this case, there is a report about blocking activities in the Pacific and Atlantic.
- According to meteorological records, the midlatitude jets were stronger than normal in most of these cases with the highest in 1983 and 1977 and the lowest in 1979 and 1982. The positions of the jet axis were displaced southward. There were vigorous cyclone activities along the storm tracks.

Examples in Fig. 2 illustrate the variability of teleconnection indices within a month, which were calculated for the running 10-day mean of 500 mb geopotential height. These diagrams (not all shown here) suggested that the cases of steady evolution are January 1977, '78 and '81, while the cases of abrupt evolution are 1 January 1979, '80 and 16 January 1979. The

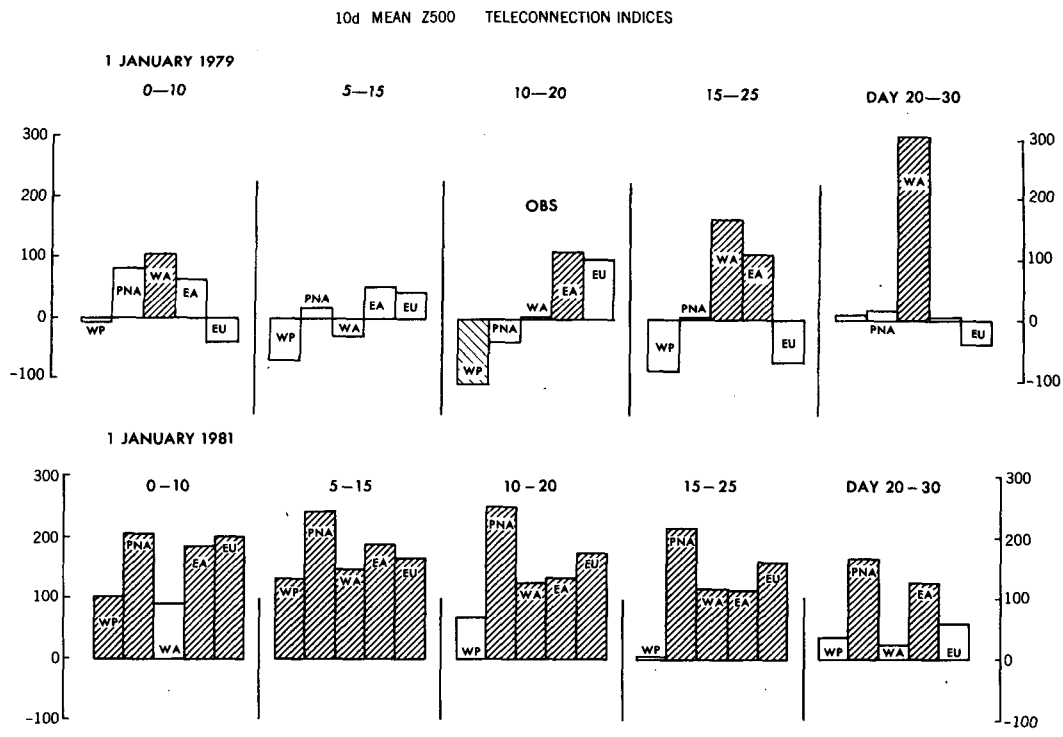

FIG. 2. Variation of teleconnection indices for 10-day running mean of 500 mb height within a month in two January cases.

January 1982 and 1983 cases do not fit into either of these two categories. It should be remarked that the abrupt evolution does not mean rapid fluctuation of small-scale features, but it corresponds to the sudden change of blocking activities from Pacific to Atlantic, for example. The case of January 1977 shows a record of remarkably persistent circulation patterns.

3. Prognostic maps

In order to present a comparison of the forecast and the observed maps, first a time-averaging is applied to the variables, and then ensemble averages are taken over three realizations. The observations for verification are provided by the NMC Level III data.

In stochastic forecasts, there are two fundamental items of prognoses, i.e., the ensemble (probabilistic) mean (centroid) and its standard deviation (scatter) (Epstein, 1969; Preisensdorfer and Barnett, 1983). Using the notation of arithmetic average over three realizations, $\langle\ \rangle$, and an arbitrary variable, x, the former is written by

$$\langle x \rangle \tag{3.1}$$

and the latter is written by

$$\sigma_f(x) = [\langle (x - \langle x \rangle)^2 \rangle]^{1/2} \tag{3.2}$$

where $\langle (x - \langle x \rangle)^2 \rangle = \frac{1}{M-1} \sum^{M} (x - \langle x \rangle)^2$, M being the number of samples, i.e., three in this case.

Before proceeding further, let us define other notations. The arithmetic mean over the eight cases is denoted by $E(\ )$. Using the time-mean geopotential height, $Z(t)$, in which t represents the middle of the averaging period, the *height error* is expressed by

$$\Delta Z = Z_p(t) - Z_o(t) \tag{3.3}$$

and the *height anomaly* is expressed by

$$\delta Z = Z(t) - Z_n \tag{3.4}$$

where the subscripts n, p and o denote the climatological norm, the prediction and the observation, respectively.

The predicted pattern, represented by the ensemble average of three realizations of the 20-day mean for Day 10–30, for example, is denoted by $\langle Z_p(20\ \text{day}) \rangle$. In general, forecasts in the ensemble mean height, $\langle Z_p \rangle$, bear some resemblance to the corresponding observations, Z_o. Another salient feature is that individual forecasts for Day 10–30 based on the A, B and C initial conditions are similar to each other on the one-month time scale. It would be worthy to note that similarity among these realizations is more striking than that of each forecast to the observation, as in Spar et al. (1978).

a. Eight case averages

The arithmetic averages of geopotential height and sea level pressure over the eight January cases are compared between forecasts and observed. Then the dif-

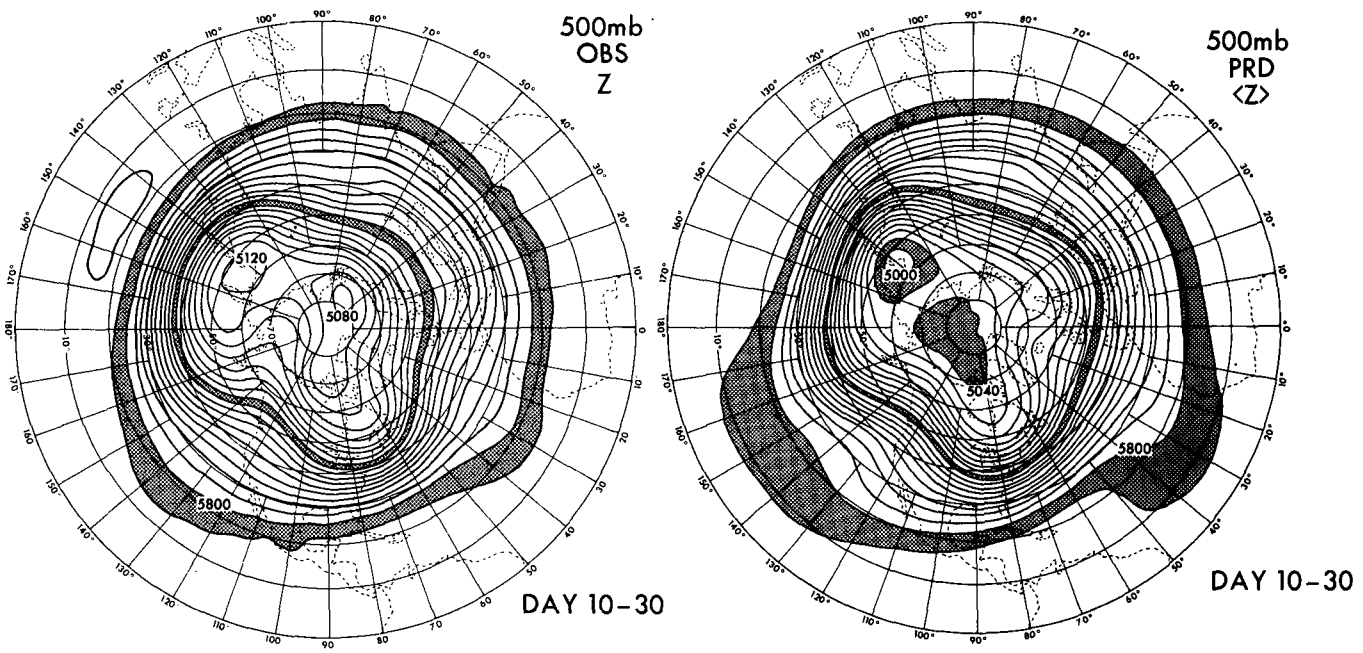

FIG. 3. Eight case averages of the 20-day mean maps of 500 mb geopotential height for the observations (left), and the forecasts (right). Contour interval is 40 meters. The shaded contour belts are 5000–5040, 5400–5440 and 5800–5840 meters.

ferences between the forecasts and the observations are calculated. Another aspect of interest is the rms forecast error of geopotential height. The displays are limited to only examples of the 20-day mean centered at Day 20 over the Northern Hemisphere.

500 mb height. Figure 3 is the comparison of the 20-day mean of geopotential height for the observation, $E(Z_o)$, and the forecast, $E\langle Z_p\rangle$. Certain contour belts of the height fields are stippled for making a pattern inspection easier.

- Compared with the observation, the forecast height field is more zonally symmetric, which is often the case in numerical predictions (see Miyakoda et al., 1972; Arpe and Klinker, 1986).
- The observation in Fig. 3 is more nonzonal than in the climatological norm, reflecting the fact that blocking activities were more frequent in these years than in the normal years.

500 mb height errors. Figure 4 displays two types of forecast errors. One is the arithmetic mean of height error, defined by

$$E(\langle Z_p\rangle - Z_o) = E(\Delta\langle Z\rangle), \tag{3.5}$$

which is simply the difference of the height at the right from that at the left in Fig. 3. Another is the rms error of forecasts, i.e.,

$$[E(\Delta\langle z\rangle)^2]^{1/2}, \tag{3.6}$$

both for the 20-day mean. The latter may represent errors for individual cases, which are the major interest in forecasts.

- In the arithmetic mean error, the negative area dominates in the Northern Hemisphere, particularly in the polar region, implying that the forecasts have a colder tendency compared to reality.
- There are three major geographical locations of error, i.e., the negative regions near the Aleutians and Novaya Zemlya, and the positive region south of the Aleutians.
- There is a mathematical relation that the panel at the right, $[E(\Delta\langle z\rangle)^2]^{1/2}$, is larger than or equal to the panel at the left, $E\Delta\langle z\rangle$. In the experiment described in this paper, the mean systematic error is as large as the rms. This is quite different from daily forecasting where the systematic error is about 20% of the rms (see Hollingsworth et al., 1980).

Sea level pressure, P_{SL}, and errors. A presentation for P_{SL} similar to that of 500 mb height in Fig. 3 is given in Fig. 5. A presentation similar to Fig. 4 is given in Fig. 6.

- The arithmetic mean error pattern for P_{SL} has some similarity to that for Z_{500}.
- Associated with the orography such as Himalaya, the Rocky mountains and Greenland, the positive errors appear to be generated in the immediate lee side and the negative errors follow further downstream.

b. Standard deviations

The standard deviations (stochastic scatter) are presented below for two variables, i.e., the 500 mb geopotential height Z_{500} and the 850 mb temperature T_{850}. The latter variable may be of practical interest.

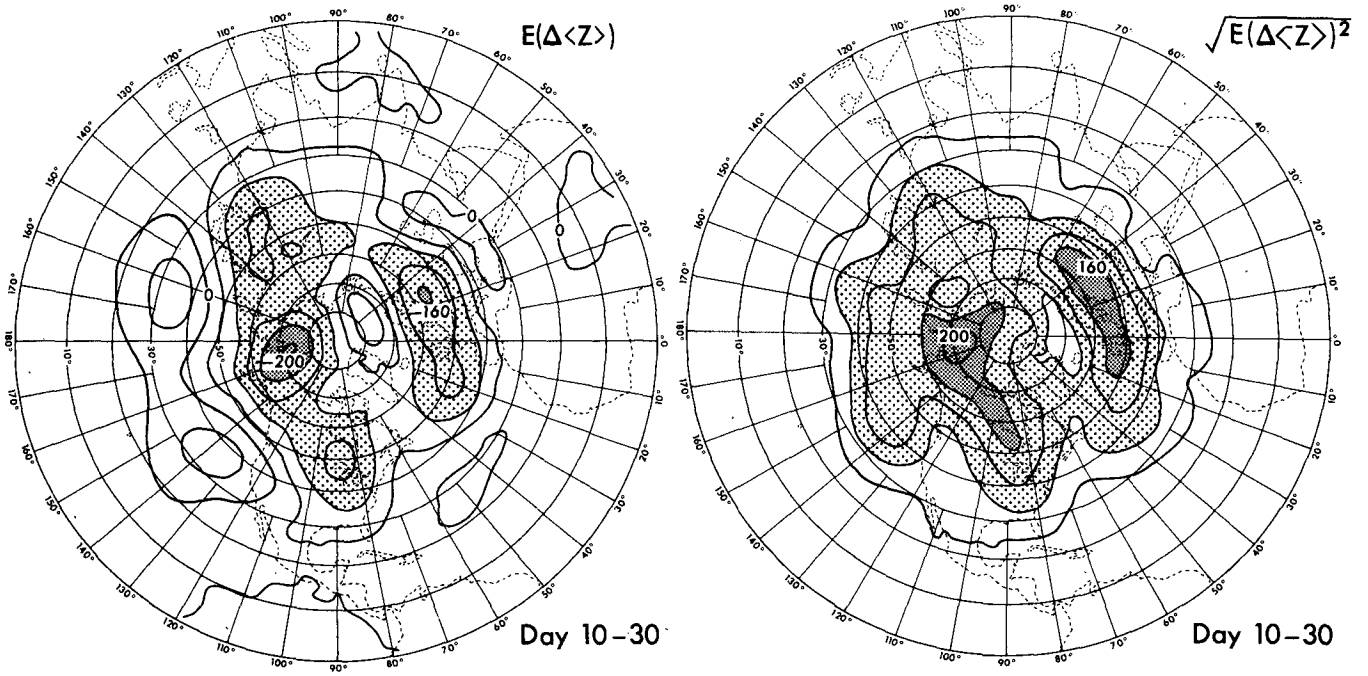

FIG. 4. Eight-case average of 500 mb height errors for Day 10–30. Arithmetic mean error (left), and rms error (right). Contour interval is 40 meters. The areas of lower than 80 meters (in left panel) and of higher than 80 meters (in right panel) are stippled.

500 mb heights. Figure 7 shows the geographical distribution of the scatter of 20-day mean (Day 10–30) geopotential height. The arithmetic average of the scatter over eight January cases is calculated by

$$\sigma_f(z) = [E\{\langle (z - \langle z \rangle)^2 \rangle\}]^{1/2}. \quad (3.7)$$

The large values of $\sigma_f(Z_{500})$ in Fig. 7a are located predominantly at high latitudes, so far as this variable is concerned.

850 mb temperature. The temperature at 850 mb is adopted as representative of thermal distribution in the lower part of the atmosphere. The standard devia-

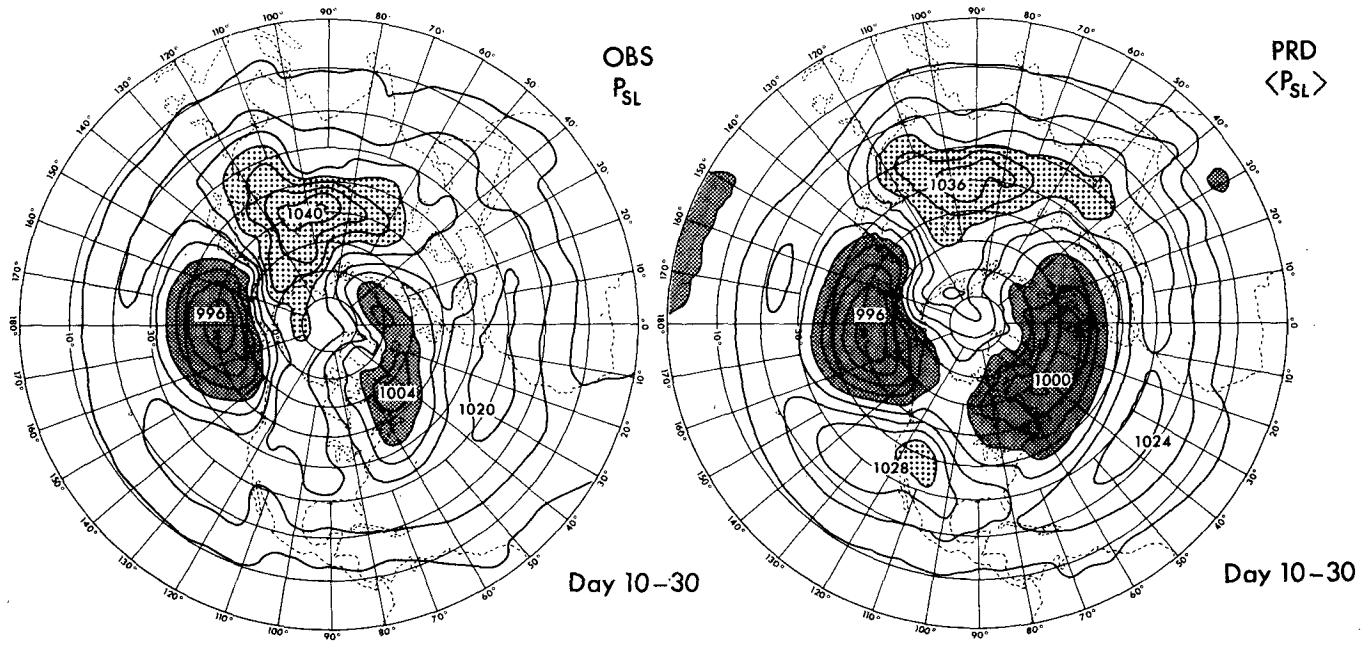

FIG. 5. As in Fig. 3 but for sea level pressure. Contour interval is 4 mb.

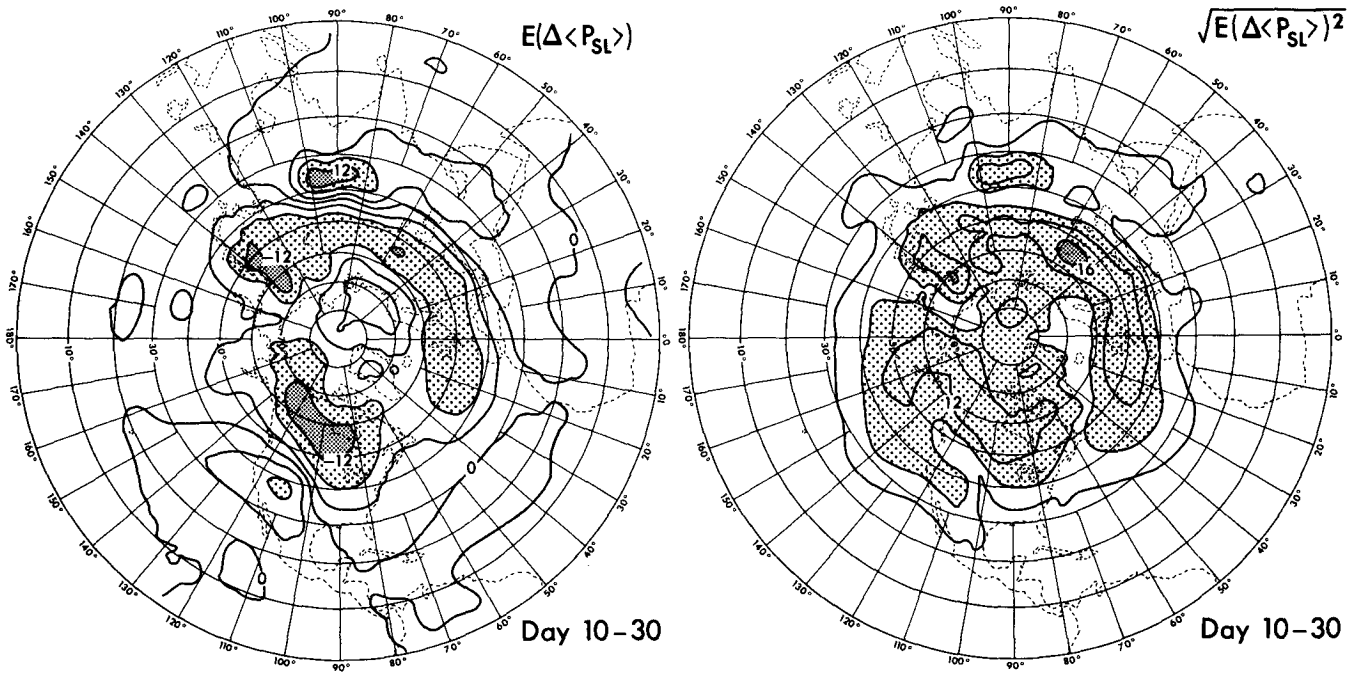

FIG. 6. As in Fig. 4 but for sea level pressure. Contour interval is 4 mb.

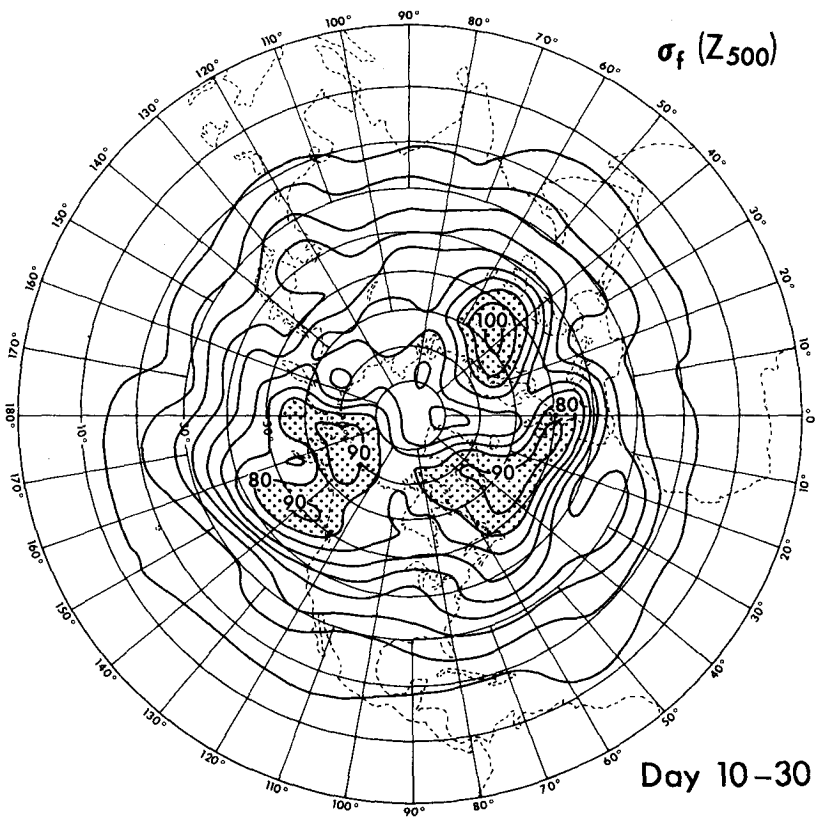

FIG. 7a. Eight-case average of stochastic scatter of geopotential height, $\sigma_f(Z_{500})$. Contour interval is 10 meters. The areas of $\sigma_f > 80$ meters are stippled.

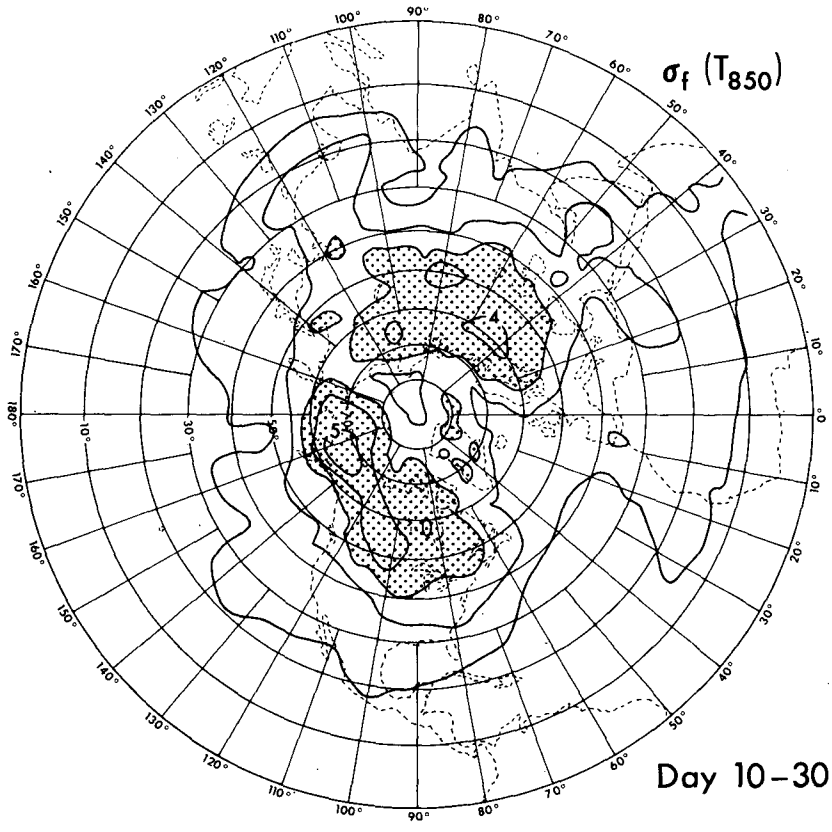

FIG. 7b. Eight-case averages of stochastic scatter of 850 mb temperature, $\sigma_f(T_{850})$. Contour interval is 1°C. The areas of $\sigma_f > 3$°C are stippled.

tion of the eight-case average, $\sigma_f(T_{850})$, is shown in Fig. 7b. The large values are predominantly located over land.

Comments. A salient feature is that the pattern of $\sigma_f(Z_{500})$ is similar to the distribution of low frequency variance in the upper tropospheric geopotential height shown by Blackmon et al. (1977) (see also Miyakoda and Sirutis, 1985). In other words, the areas of large stochastic scatter are located in the same regions as blocking ridges. It should be noted, however, that the scatter for individual cases emerges in a more disorganized way (not shown here), and so do the blocks for individual cases.

The agreement of $\sigma_f(Z_{500})$ with the blocking region is not surprising in view of the similar character of these variables. However, a delicate discrepancy is noticed in the precise position of the Atlantic Ocean and Ural Mountains maxima. This fact might be related to the climatic regime of this particular sample; it is not due to the climatological bias of the model, because the observational counterpart is similar to the forecast.

4. Verification scores

a. *Formulas of the scores*

The performance of one-month forecasts is assessed by two skill scores, i.e., the correlation coefficients of the geopotential height anomalies (deviation from the climatological norms) between the forecasts and the observation, and the root-mean-square (rms) error of the geopotential heights. The verification is the NMC Level III data.

This scoring system is exactly the same as that used in the deterministic medium-range forecasts (Miyakoda et al., 1972; Arpe et al., 1976). The only difference is that the time-mean is taken for both the forecast and the observed data. The time-mean is applied in an attempt to obtain positive skills by filtering out the unpredictable components (Shukla, 1981; Miyakoda and Chao, 1982; Mansfield and Palmer, 1983). As discussed in Miyakoda and Sirutis (1985), the idea is to suppress baroclinic eddies whose dominant peak in the power spectrum is 2–8 days. For this purpose, the 10-day mean is used, and as additional information, the 20-day and the 30-day means are taken as well.

For the verification calculation, a spatial average is used, and it is denoted by the bar operator as

$$\bar{X} = \sum_i \frac{1}{m^2} X_i \Big/ \sum_i \frac{1}{m^2} \tag{4.1}$$

where X_i is an arbitrary variable at the grid point i, and m is the map scale factor. Except for mountain areas

in the lower troposphere, the summation is made over all Northern Hemispheric gridpoints north of 25°N.

Using this operator, the *correlation coefficient* for the anomaly of an individual forecast is given by

$$\text{correlation anomaly} = \frac{\overline{\delta Z_p(t) \cdot \delta Z_o(t)}}{[\overline{(\delta Z_p)^2} \cdot \overline{(\delta Z_o)^2}]^{1/2}} \quad (4.2)$$

where the *climatological norm* for geopotential height, obtained by Oort and Rasmusson (1971) and Oort (1983), was used. As an auxiliary score, the correlation for 10-day mean *persistence* is used. It is a lag correlation, defined by the formula (4.2) but $\delta Z_p(t)$ is replaced by the observation $\delta Z_o(-5$ day), where $Z_o(-5$ day) denotes the average of Z_o for ten days from -10 to 0 days. In the same manner, the 20- and 30-day persistences are based on the 20- and 30-day means of the observed height from -20 to 0 days, and from -30 to 0 days, respectively.

The rms error is given by

$$\text{rms} = [\overline{(\Delta Z)^2}]^{1/2}. \quad (4.3)$$

For comparison, two auxiliary measures are also used. One is the *persistence,* defined by the formula (4.3) but $Z_p(t)$ is replaced by $Z_o(-5$ day). The other is the *climate departure,* defined by the formula (4.3) but with $Z_p(t)$ replaced by Z_n. These two measures are used as references of "no-skill forecasts."

b. *Scores for individual cases*

Figures 8 and 9 show the correlation coefficients of the anomalies, and rms error, respectively, for the 500 mb geopotential height in eight January cases. Thick lines with small circles represent the scores of the ensemble averages of three realizations, i.e.,

$$\text{correl. anom.} = \frac{\overline{\delta\langle Z_p\rangle \cdot \delta Z_o}}{[\overline{(\delta\langle Z_p\rangle)^2} \cdot \overline{(\delta Z_o)^2}]^{1/2}} \quad (4.4)$$

$$\text{rms} = [\overline{(\Delta\langle Z\rangle)^2}]^{1/2} \quad (4.5)$$

for the 10-day mean geopotential heights.

In Fig. 8, the vertical bars indicate twice the standard deviations of the correlation coefficients for three individual forecasts. These are calculated by

$$2\times\left[\frac{1}{M-1}\sum_{i=1}^{M}(\text{corr}_i - \langle\text{corr}\rangle)^2\right]^{1/2} \quad (4.6)$$

where corr is the correlation coefficient for an individual 10-day mean forecast, Eq. (4.2), and $M = 3$ in this case.

In Fig. 9, the degree of stochastic scatters is indicated by the error variances of three forecasts, e_3. According to Hayashi (1986), the error variance of an M-mean Monte Carlo forecast, e_M, is related to the variance among randomly perturbed forecasts, p_1, as

$$e_M = (1 + M^{-1}) \cdot p_1 \quad (4.7)$$

where p_1 is derived from the stochastic scatter as

$$\text{scatter} = \left[\frac{1}{M-1}\sum_{i}^{M}(Z_i - \langle Z\rangle)^2\right]^{1/2} \quad (4.8)$$

$$= [\langle (Z - \langle Z\rangle)^2\rangle]^{1/2}$$

$$= {p_1}^{1/2}. \quad (4.9)$$

There is another statistical quantity, as suggested by a reviewer, i.e., the differences between two pairs of forecasts among the N-perturbed forecasts, i.e.,

$$(\text{diff})^2 = \frac{2}{N(N-1)}\sum_{j=1}^{N}\sum_{i=j}^{N}(Z_j - Z_i)^2. \quad (4.10)$$

It is shown (Hayashi, personal communication) that

$$(\text{diff})^2 \equiv 2 \times (\text{scatter})^2, \quad (4.11)$$

and, for any M or N, therefore, we use here the scatter, Eq. (4.8), and the error variance e_M, Eq. (4.7). Note that the formula (4.7) is reduced to that originally derived by Leith (1974) (see Hayashi, 1986, appendix B). In Fig. 9, the spatially averaged scatters for the 10-day mean forecasts are shown by the rms of spatially averaged error variance, ${\bar{e}_3}^{1/2}$.

Examination of these diagrams may lead to the following conclusions.

- The lengths of the vertical bars in the diagram of correlation coefficients (Fig. 8) are not particularly large, indicating the individual forecasts of the 10-day mean are not widely scattered for the monthly time scale in terms of correlation coefficients.
- The error variance, ${\bar{e}_3}^{1/2}$, (Fig. 9) indicates not only the spread of individual stochastic forecasts, but also the theoretical forecast error due to the uncertainty in the initial condition (for sufficient statistical average); this theoretical error agrees with the real error only if initial conditions are the only error sources and if the initial errors are random. Obviously, however, this is not the case in reality. The rms errors $[\overline{(\Delta\langle Z\rangle)^2}]^{1/2}$ in Fig. 9 are almost always (except for several occasions) larger than the error variance ${\bar{e}_3}^{1/2}$. This may suggest that there is a large portion of other sources of errors.
- In three cases (1978, 1979 and 1980), stochastic scatters are not small but considerably large. In the case of January 1979, for example, as pointed out by Baumhefner and Bettge (1981) (see also Fig. 2), a rapid and vigorous evolution occurred on the Alaskan as well as European blocks during 8–20 January. In this particular situation, the model is extremely sensitive to initial states, so that three initial conditions create different courses of evolution in integrations.
- In the first 10 days, forecasts are better than persistence in both skill scores, and then they become worse in the second and the third 10 days, though not rapidly. So far as individual cases are concerned, the

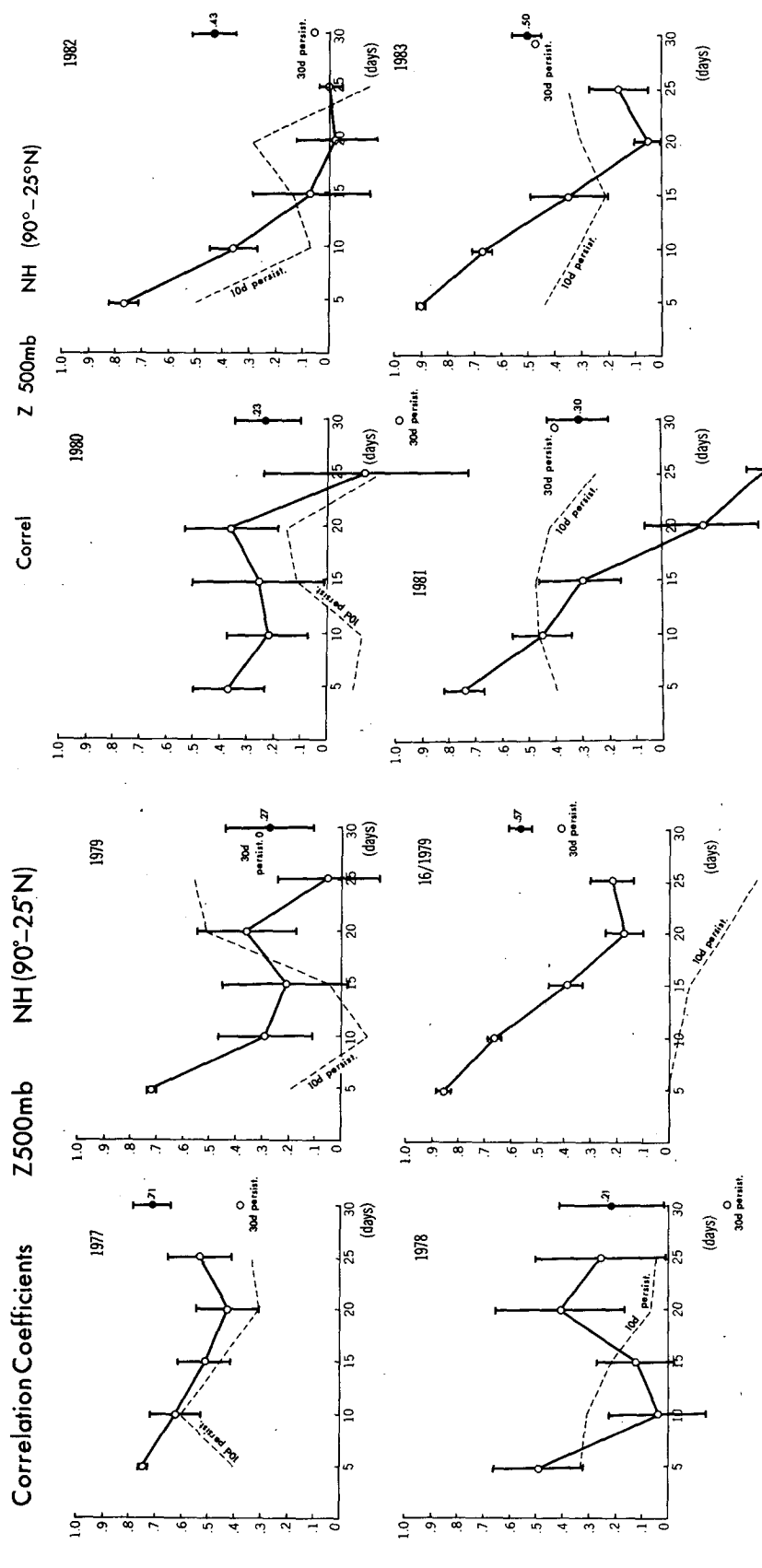

FIG. 8a. Correlation coefficients for the 500 mb geopotential height in the four January cases. Solid lines are for the 10-day mean height, and the vertical bars indicate the range of standard deviation of correlation coefficient for three realizations. Dotted lines are for the 10-day persistence. Coefficients for the 30-day mean height forecast and persistence are plotted at the right-hand side of each diagram.

FIG. 8b. As in Fig. 8a but in the latter four January cases.

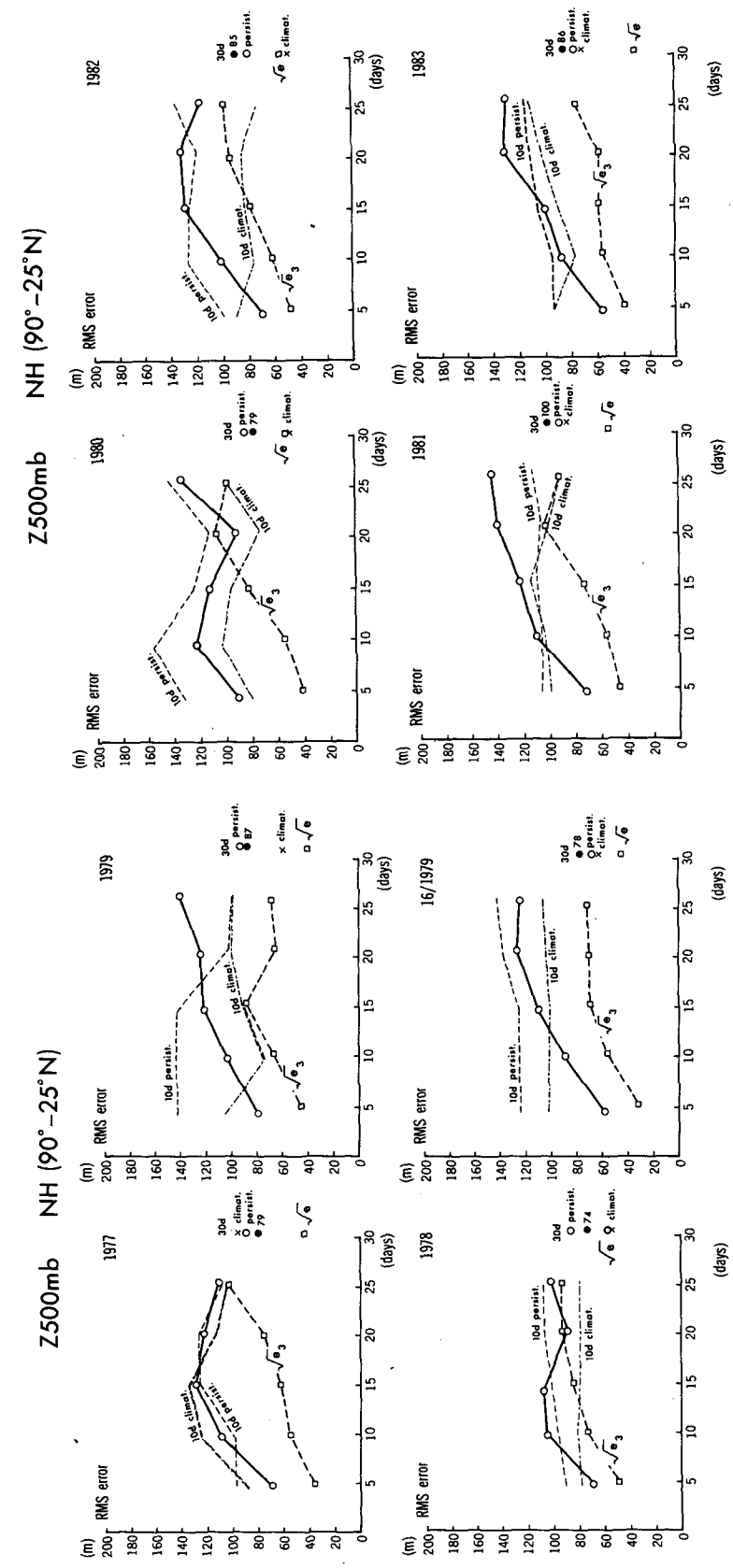

FIG. 9a. rms error of 500 mb geopotential height in the former four January cases. Thick solid lines connected with small circles are for the 3-ensemble 10-day mean height. Dotted lines are for the 10-day persistence, and dotted–dashed lines are for the climate departure of 10-day mean height. Thick dashed lines connected with small squares are the error variances of three ensemble 10-day mean forecasts, $\bar{e}_3^{1/2}$, which are derived from the stochastic scatter.

FIG. 9b. As in Fig. 9a but in the latter four January cases.

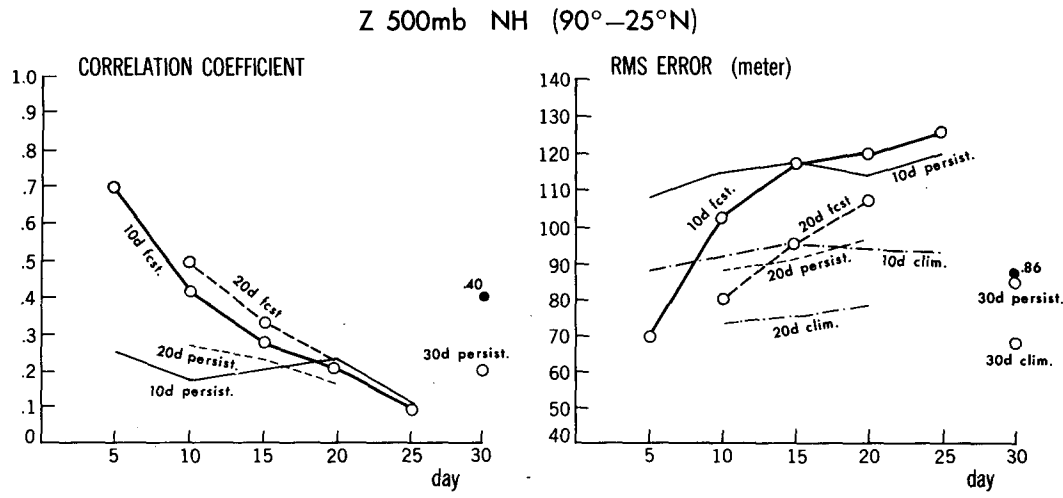

FIG. 10. Ensemble mean skill scores over eight January cases for 500 mb geopotential height. Thick lines connecting small circles are for forecasts, and thin lines are for persistences. Solid and dashed lines are for the 10-day and the 20-day mean height, respectively. The scores for the 30 day mean height are plotted at the right-hand side of each diagram. In the rms diagram, thin dashed–dotted lines are for the climate departure.

predictive skill appears to be independent of the score of persistence forecast.

• The eight cases include the two El Niño months of January 1977 and 1983. Yet the performance of the model is not bad, in spite of the fact that the climatological SST is used instead of the observed SST (see Miyakoda et al., 1983).

• The scores for ensemble averages of three realizations are better in the rms error than the arithmetic average of those for individual forecasts (not shown here). It is known that the greater the number of realizations employed, the more the rms values are reduced, though within a limit (Leith, 1973; Hayashi, 1986).

• Interestingly, correlation coefficients for ensemble averages of three realizations are also slightly better compared with the arithmetic average of three individual coefficients (not shown here), implying that random errors included in individual forecasts are canceled in the averaging process.

c. *Averaged scores*

Figures 10 and 11 are the arithmetic averages of individual scores over the eight cases, i.e., E(corr. coeff.) and E(rms) for the 500 and the 1000 mb levels, respectively.

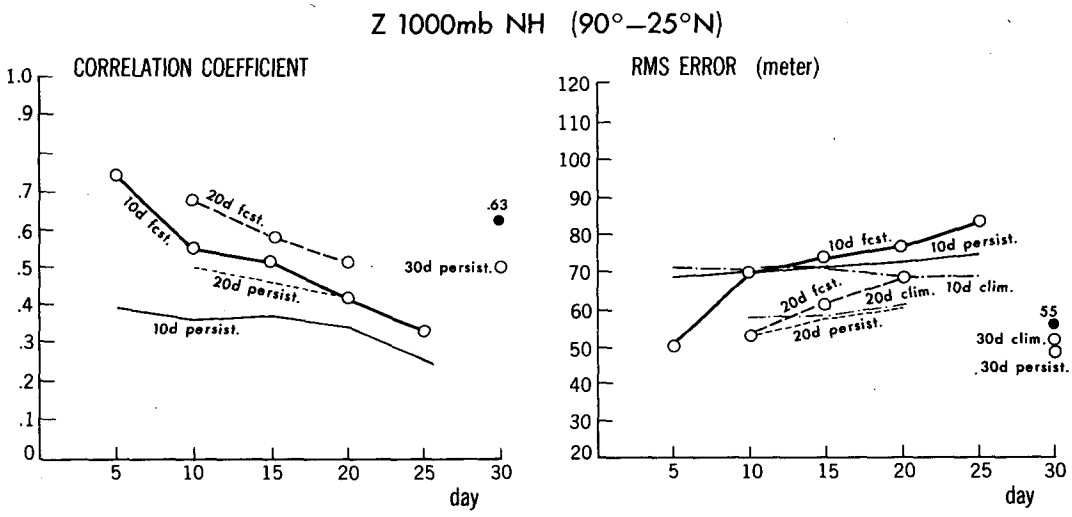

FIG. 11. As in Fig. 10 but for 1000 mb geopotential height.

1) CORRELATION COEFFICIENTS

• All scores except one case are above the persistence level for the entire month. Only the score for the 10-day mean of 500 mb height reaches the persistence at Day 20. The skill is marginal, i.e., 0.2 in the correlation coefficient, at the end of month.

• The correlation curves of the 20-day mean height are consistently but slightly higher than those of 10-day mean height, indicating that the longer time-mean extends the limit of predictability further.

• The correlation coefficients for 1000 mb height appear to be too good even at the end of month. Note, however, that the persistence is very high. A test was carried out, calculating the correlation coefficients based on the hypothetical climatology of zero, but the situation was not much different (in fact, even better in correlation scores). It is speculated that the good scores are due to the unusual climate regime for these eight cases.

2) RMS

• For the 500 mb height, the 10-day mean score is better than the 10-day mean persistence only until 14 days, and it is better than the 10-day mean climate only until about 7 days.

• Theoretically, the rms error should be lower than the persistence, if there is no systematic bias in the GCM (Hayashi, 1986). Figures 10 and 11 indicate, therefore, that the model must have an appreciable degree of systematic error.

d. Vertical distribution

Figure 12 shows the time evolution of correlation coefficients at various vertical levels. The coefficients are for the 10-day mean of geopotential height anomalies and are obtained by averaging the individual forecast scores over the Northern Hemisphere north of 25°N with respect to eight January cases. The chart indicates that the scores at the lowest level and in the 300–400 mb layer are relatively high, whereas the scores at the 500–700 mb layer and in the stratosphere (calculations were made up to the 50 mb level) are lowest.

e. Conclusions of sections 3 and 4

Summarizing sections 3 and 4, it may be stated that (i) overall a marginal skill is found for one-month forecasts in the correlation coefficient of the height fields; (ii) the predictive skill in terms of rms is only recognized for less than 10–15 days; (iii) initial conditions are crucial elements for monthly forecasts, suggesting that internal dynamics is dominantly important rather than external forcings (Shukla, 1981); (iv) the stochastic scatter is not large compared with the interannual variability (see section 7c), but not negligible either; (v) the 10-day mean filter appears to be effective to extend the predictability limit beyond two weeks; (vi) the 20-day mean score is slightly better than the 10-day mean in correlation coefficients; and (vii) the GFDL four-dimensional data assimilation scheme is equivalent to other analyses for the purpose of one-month forecasts (see also Lau, 1984, for the comparable qualities in other features).

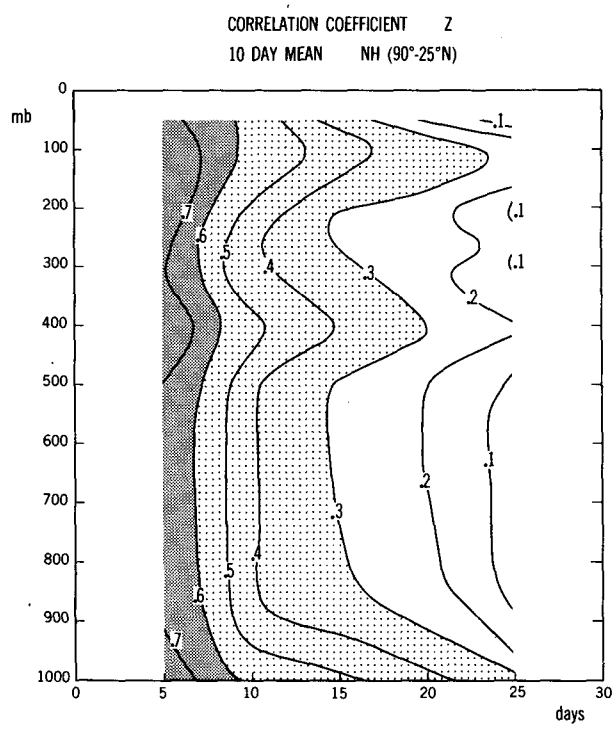

FIG. 12. Time evolution of the forecast skill with height in terms of correlation coefficients on the 10-day mean geopotential height. The ordinate is scaled in equal pressure interval (mb).

5. Diagnostic of errors

a. Hemispheric or zonal mean error

Figures 13 and 14 display the comparison of temperature and zonal component of wind between the forecasts and the observations. The variables are averaged monthly and zonally for three realizations and eight January cases.

The results show that the predicted temperature is overall lower than the observation, as is usually the case for almost all GCMs around the world. The cooling is so extensive and intense that a devastating impact on forecasts is anticipated. Associated with the temperature bias, the subtropical jet is shifted poleward in the winter hemisphere, whereas the Southern Hemispheric jet is spuriously displaced equatorward. A question is whether this error can be reduced by a proper treatment of subgrid-scale processes. According to Miyakoda and Sirutis (1984), which studied the effects of parameterizations for the cumulus convection,

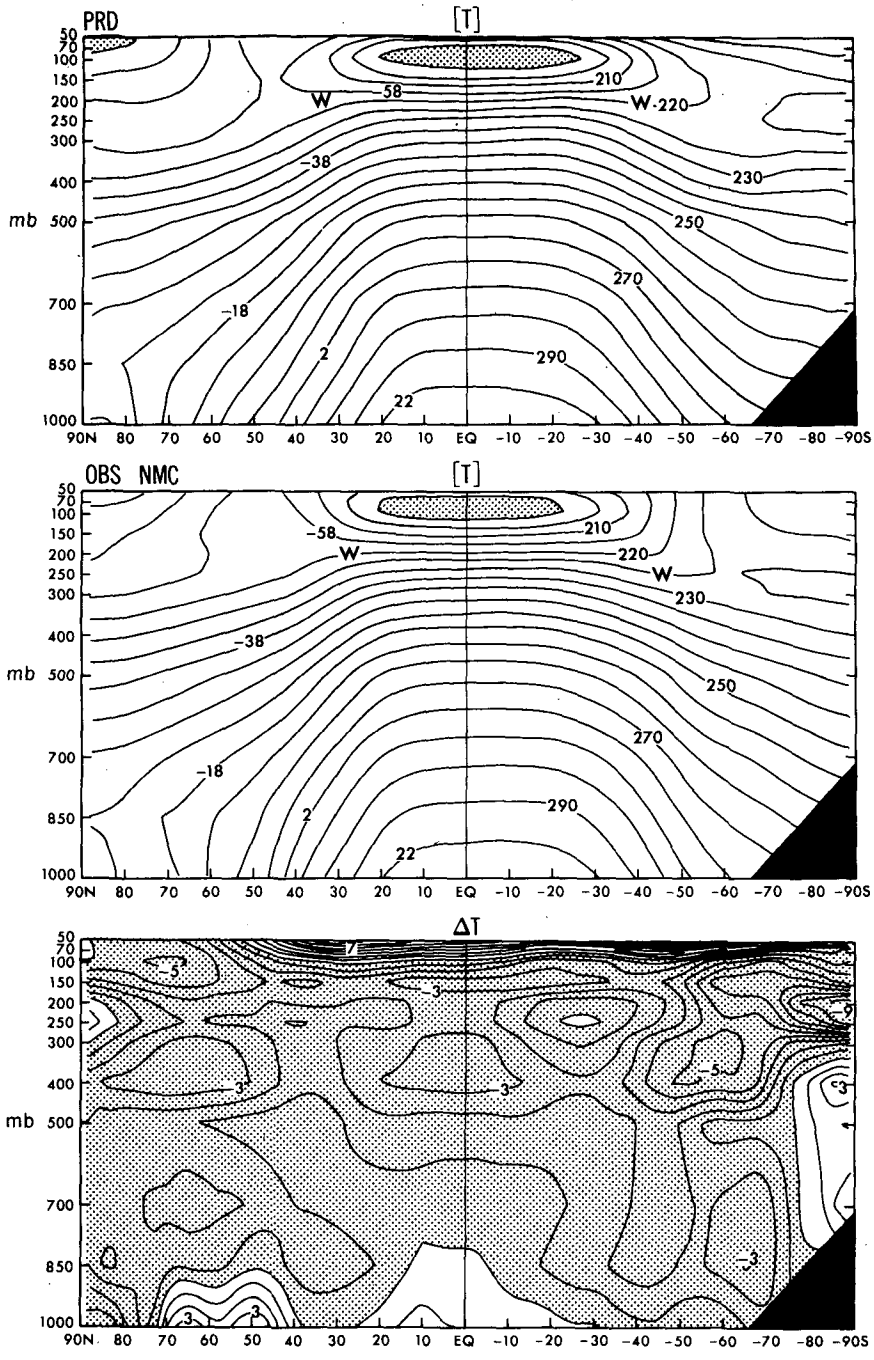

FIG. 13. The monthly and zonal mean of temperatures for prediction (upper), observation (middle), and difference, i.e., forecast minus observation (lower) with contour intervals of 5° (upper and middle) and 1° (lower). The abscissa is the latitude, and the ordinate is the pressure in mb, both at equal intervals. The negative regions are shaded. The positions of subtropic westerly jets are indicated by W.

the orographic representation, the boundary layer fluxes and the above-boundary-layer turbulent processes, these processes alone cannot eliminate this systematic error, and other considerations are necessary.

Figure 15 is the time evolution of temperature error, which is represented by the rms error of the 10-day mean temperature, i.e., $[E(\Delta\langle T\rangle)^2]^{1/2}$. It may be appropriate and interesting to compare this chart with the corresponding figure for the time evolution of instantaneous (day-to-day) temperature error in two-

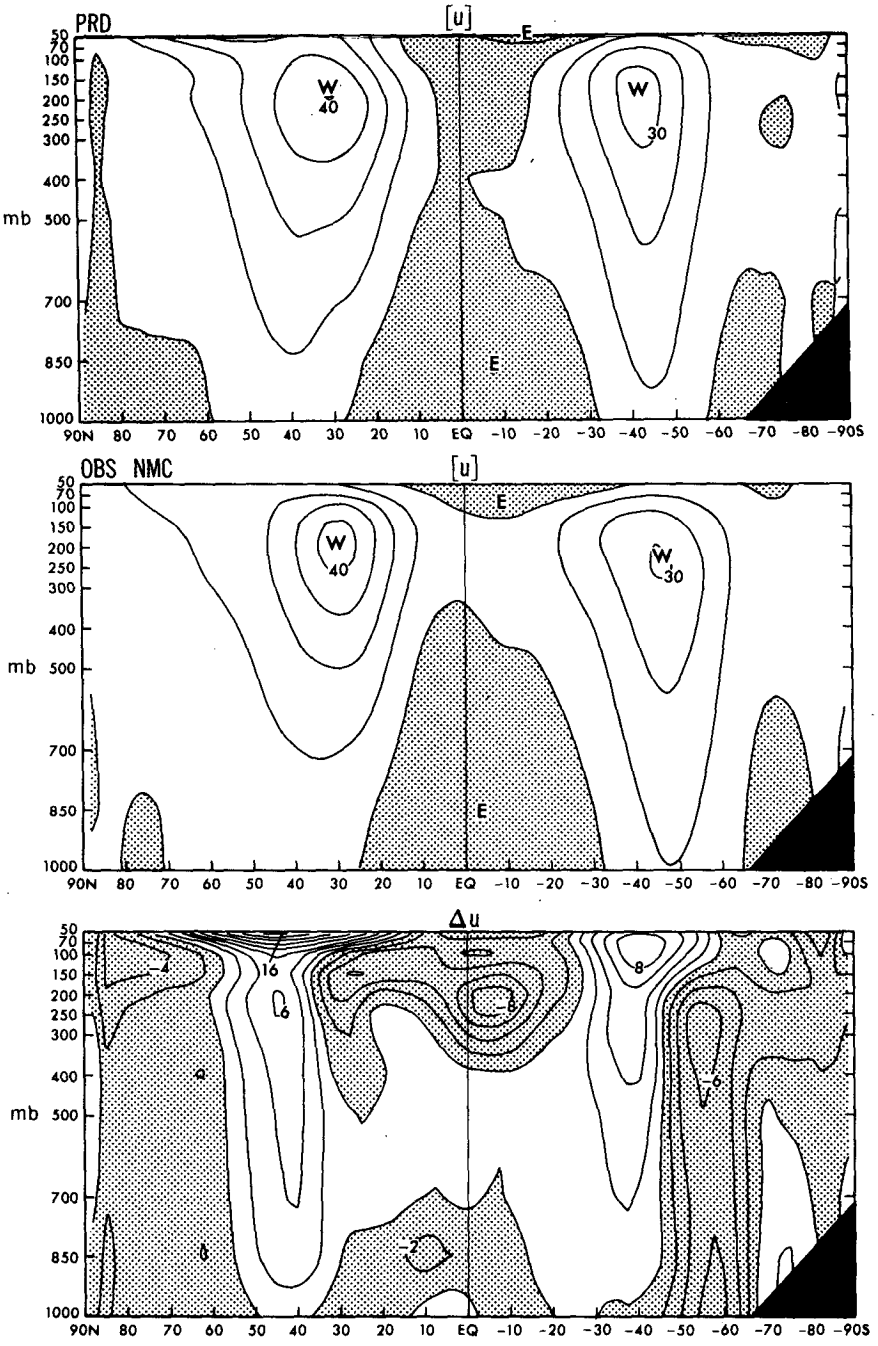

FIG. 14. The monthly and zonal mean of zonal wind for prediction (upper), observation (middle), and difference, i.e., forecast minus observation (lower) with contour interval of 10 m s^{-1} (upper and middle) and 2 m s^{-1} (lower). See Fig. 13 for further explanation.

week forecasts (e.g., Miyakoda et al., 1972; Druyan et al., 1975). The distributions in both charts resemble each other except the minimum at about 500 mb (1 ~ 5th level from the top). Other differences are that the magnitudes in Fig. 15 are overall lower by about 1°–2°C, and the time scales for the abscissa are, of course, stretched by a factor of 2.

Figure 16 is similar to Fig. 15 but for the normalized rms error for 10-day mean temperature, i.e., $[E(\Delta\langle T\rangle)^2]^{1/2}$/persistence. The major errors seem to start in the 300–400 mb layer and the stratosphere, though minor error starts near the surface and extends upward. In order to depict the detailed feature of development of error in the middle of atmosphere, Fig.

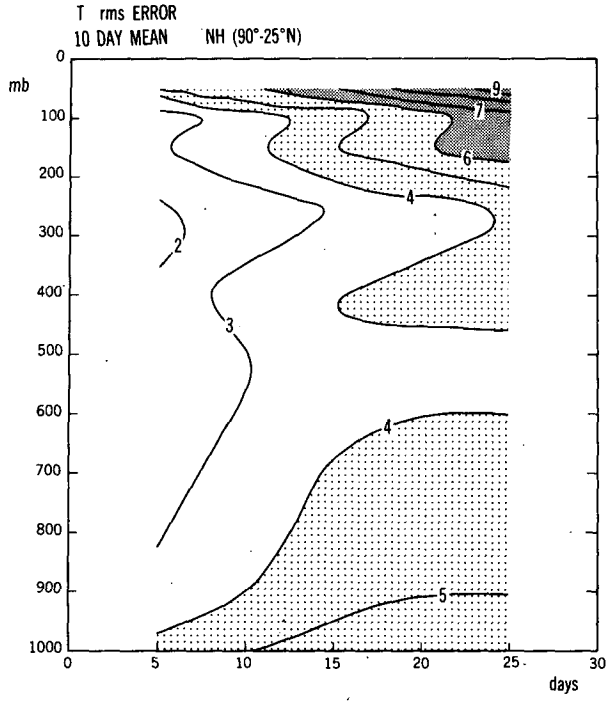

FIG. 15. Time evolution of the rms error of 10-day mean temperature (°C) as a function of height. The ordinate is scaled in equal pressure intervals (mb).

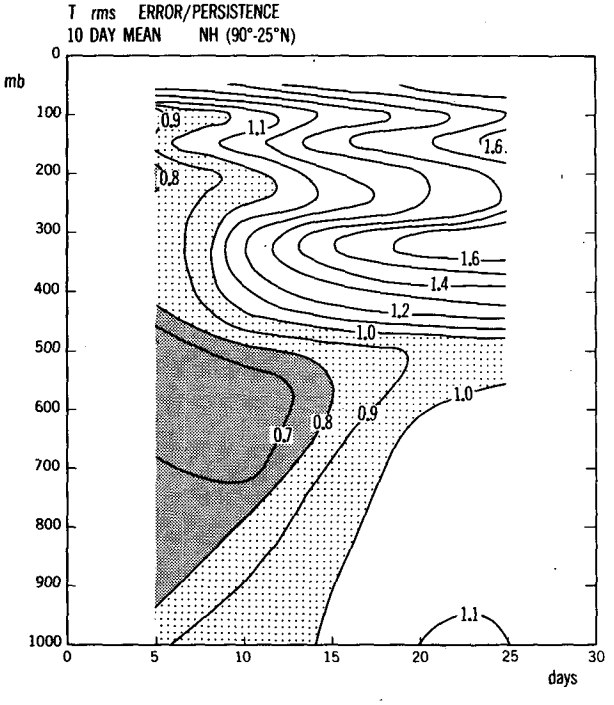

FIG. 16. Time evolution of the normalized rms 10-day mean temperature error. See Fig. 12 for further explanation.

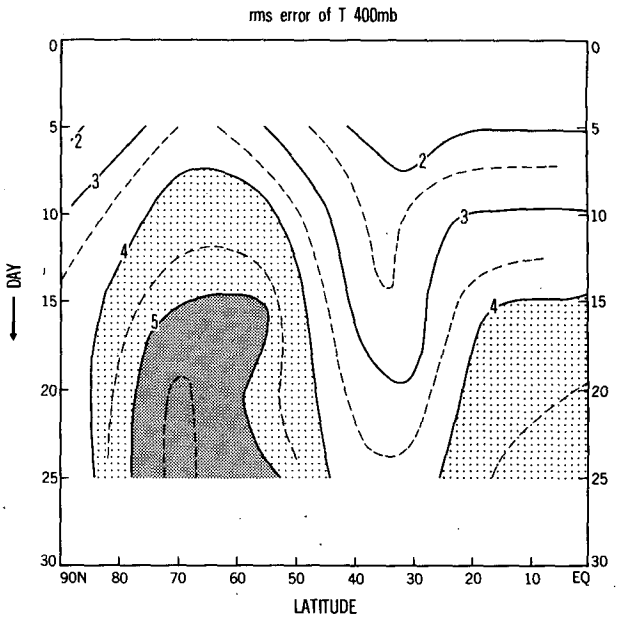

FIG. 17. Time evolution of the rms error of 10-day mean temperature for the 400 mb level as a function of latitude. The units are °C.

17 is presented, which shows the temperature error at the 400 mb level. The figure reveals that the error develops particularly at the latitudinal belt of 50°–80°N and the equator.

The tropical systematic errors have recently been discussed by Heckley (1984) and Kanamitsu (1985) with respect to the ECMWF forecasts. Kanamitsu found that this error has a gravest mode of zonal wavenumber 1, geographically locked over the Atlantic (in his particular case), which is symmetrical around the equator for temperature and antisymmetrical for streamfunction. He concluded that the cumulus parameterization has a considerable degree of connection with this tropical error. Considering these findings, it may be stated in our case that the equatorial errors may be associated with (i) the inadequate lower boundary conditions in the tropics, (ii) inaccurate verification data in this region, and (iii) the incomplete cumulus parameterization.

b. Geographical distribution of error

For the purpose of having a better understanding of error growth, the hemispheric distribution of temperature error patterns at various levels is shown in Figs. 18 and 19.

- The errors distributed in the lower atmosphere are predominantly over continents. This fact may be related to the finding of Barnett (1978), who showed that the variability of temperature is larger over continents than over sea by a factor of 2–6.

• It should be pointed out, however, that the temperature over land is excessively high, as is seen in ΔT at 1000 mb in Fig. 18b. This may be due to the particular bias of this model and partially due to the scheme of vertical extrapolation.

• The errors in the upper troposphere consist of two regions, i.e., storm tracks over oceans and the equatorial belt.

• The middle- and high-latitude errors could be associated with northward shift in jet and could be generated by the baroclinic and wind shear instabilities, associated with the jet streams, and they could develop no matter how the source of initial error is supplied.

Apart from the spatial configuration of error there is another important aspect. Let us temporarily go back to Figs. 4 and 6, in which errors of Z_{500} and P_{SL} are presented. An intriguing fact is that the rms error patterns (right in these figures) are highly correlated with the arithmetic mean error patterns (left in these figures), despite the fact that the former has only positive areas and the latter has both positive and negative areas.

In order to investigate this point further, the temperature rms error patterns are compared with the temperature's arithmetic mean errors in Figs. 18 and 19. Once again it is pronounced that major rms error patterns roughly correspond to those of arithmetic mean errors.

Different from the zonal mean errors in Figs. 13 and 14, the longitudinal distributions of error appeared to depend on the subgrid-scale physics employed in the GCM (Miyakoda and Sirutis, 1984). In the paper of Miyakoda and Sirutis (1985), it was pointed out, based on four winter cases, that the geographical configurations of rms error are sensitive to the subgrid physical parameterization and that these rms error patterns resemble remarkably well arithmetic mean error patterns for the respective models. For example, the distribution of errors for the model with E-physics is different from those of the models with other physics. Particularly, the effect of "envelope mountains" (Wallace et al., 1983) exhibits appreciably different geographical distribution of forecast error. Other aspects of the model's bias are summarized in the Appendix, in which the eight January ensemble-mean statistics are displayed. The most pronounced deficiencies are found in the transient eddy kinetic energy and moisture in tropics.

c. *Climate-drift adjustment*

The arithmetic mean error is regarded as the "climate drift," though the value for Day 10–30 has not reached its asymptotic level yet. The previous subsection suggests that the model's systematic error is a major component of the rms error, and that *a large portion of the forecast errors are due to the model's climate drift.* This fact has been increasingly noticed in the community of atmospheric general circulation modeling. The removal of this bias, however, may require decades of hard and tedious work (Miyakoda and Chao, 1982). Pointing out this impending limitation of the GCM approach, Miyakoda and Chao previously proposed to use an "anomaly model" as another option. The final answer for the utility of the anomaly model remains to be seen.

On the other hand, Shukla (1983) proposed a strategy: "The model climatology is subtracted from the predicted field to obtain the predicted anomaly field." Shukla was probably the first to advocate the simple subtraction of drift from numerical results of long time integration. Indeed, the success of sensitivity study in climate simulations or predictability study in weather forecasts has been based on this reasoning. Namely, two parallel runs, i.e., the control and the perturbation, may have the same model bias, and therefore, the pair of runs can provide significant and useful information through subtraction of the two numerical results, no matter how large the basic errors are.

The subtraction of systematic bias will be applied to our forecasts presented in the foregoing sections. Applying this procedure, we shall use the words "the climatic-drift adjustment" or the "empirical adjustment" to refer to the subtraction process. In this paper, the adjustment is made only after all prediction calculations are finished. (ECMWF is also taking this approach—Molteni et al., 1985).

6. Empirical adjustment

To implement the empirical adjustment, we first prepared the model's systematic errors defined by Eq. (3.5), based on eight January cases. The necessary information for the climate adjustment are only the variables of interest every 10 days, for the whole forecast range of 30 days. This set of information is a function of space coordinates and forecast time.

a. *Anomaly maps*

Figure 20 presents five samples of the predicted and observed height anomaly maps. The prediction includes original anomalies, $\delta\langle Z\rangle$, as well as empirically adjusted anomalies, $\delta\langle Z\rangle_{\text{adj}}$. The climate drift used for the adjustments is exactly the same throughout all eight cases.

Among the five samples in this figure, there are typical cases in terms of forecast performance and the degree of positive impact by empirical adjustment, while the sample of 16 January 1979 is the case in which empirical adjustment did not improve, but worsened the original forecasts.

The charts in this figure reveal that unadjusted anomaly patterns resemble each other irrespective of the cases, indicating that the systematic bias was indeed dominant. Therefore, modifications for the adjustment

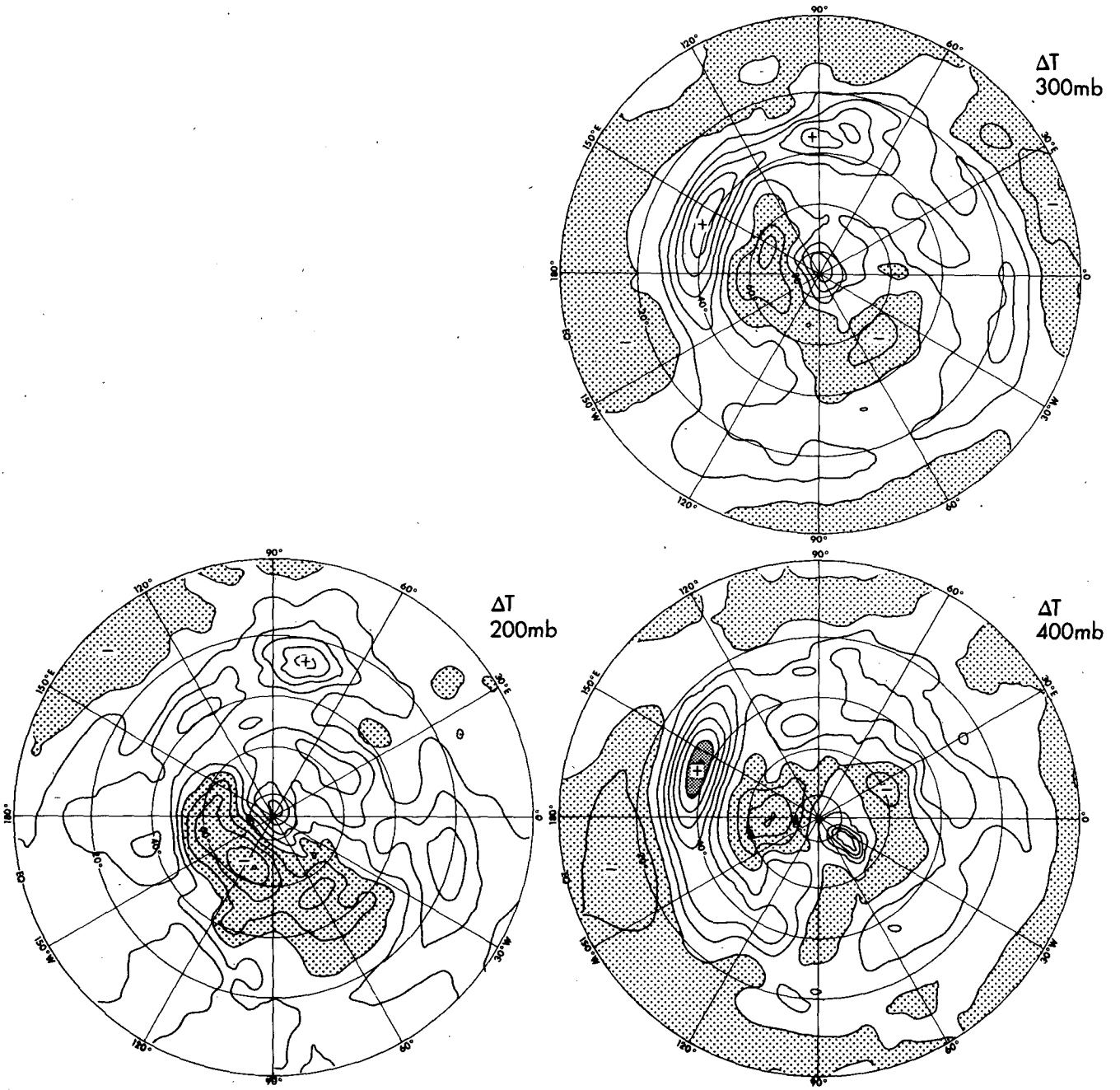

FIG. 18a–b. Eight-case average of arithmetic mean temperature error for Day 10–30. Contour interval is 1°C.

are considerable in magnitude. Detailed investigations reveal that empirical adjustment has improved most of the prediction in terms of teleconnection indices and that the modifications are pronounced for WP index and next PNA, but not for other indicies.

It is interesting to note that the patterns of 1981 and 1983 show a remarkable similarity in both observations and forecasts, though the latter was an extraordinary case in the records.

b. Skill scores

Figures 21 repeat the skill scores for stochastic means in Fig. 8 but also include new scores for the empirically adjusted forecasts. It is worthy to note that in all cases except one, both skill scores, i.e., correlation coefficients and rms error for 500 mb height, have been improved. The exceptional case is for 16 January 1979, in which the starting day is different from the rest of the samples.

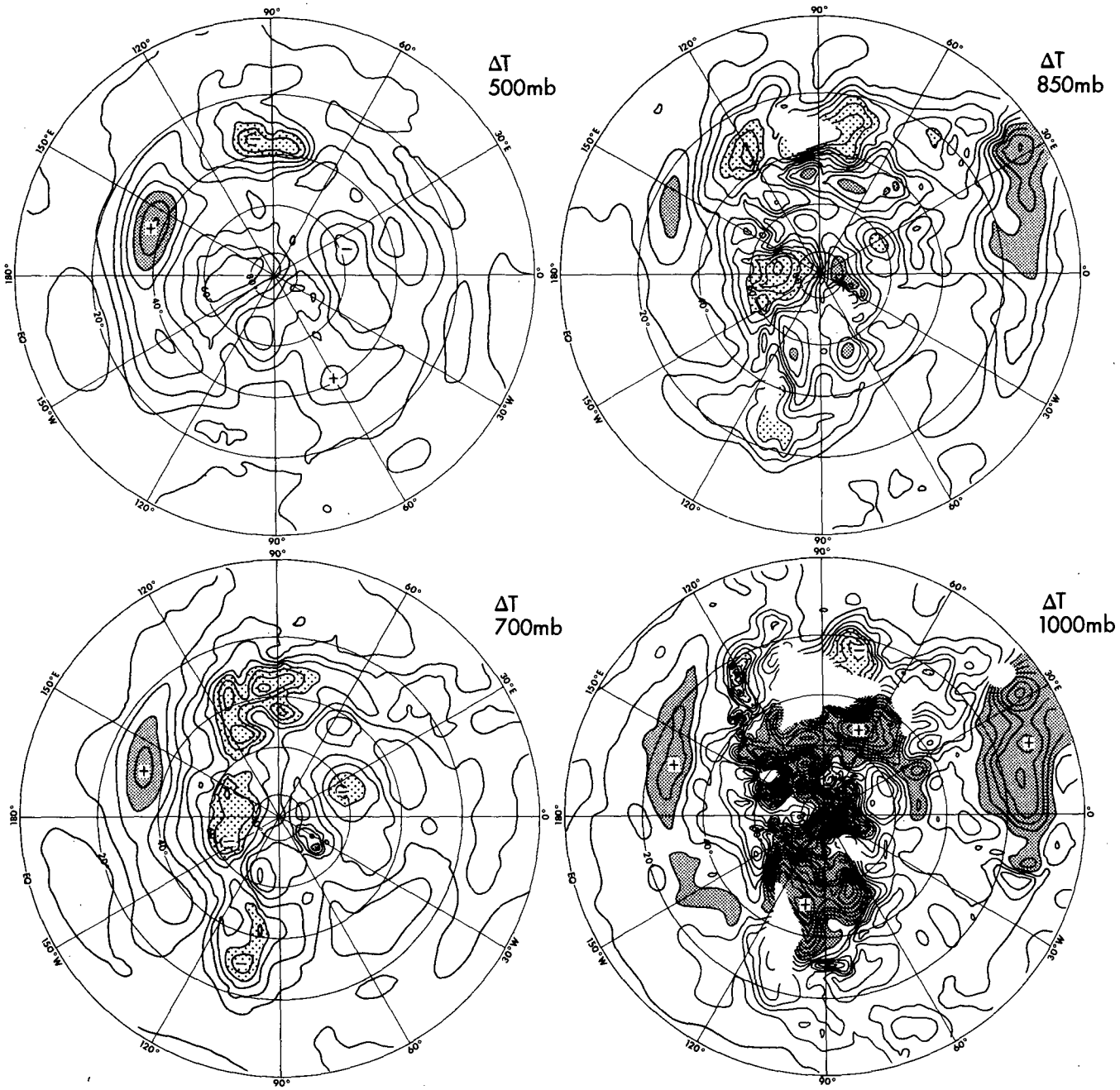

FIG. 18b. (*Continued*)

Figures 22 and 23 are summaries for the eight cases. The rms error diagram is now modified by normalizing the error curves by the persistence level. As a whole, the improvement is remarkable in both scores. The correlation coefficients are well above the persistence level. The rms errors are below the persistence level, and they are also below the climate level until the end of the month.

It is worthy to note that, in the 1000 mb rms error diagram, the 10-day and the 20-day error curves have not yet reached their plateau, or equilibrium, suggesting that the limit of predictability has not yet been reached. It is straightforward to show that the curves of rms error follow a systematic behavior, if the climate drift is adjusted. Namely, the rms errors approach the level of 10- or 20-day mean variability (climate) or somewhat higher with increasing forecast time. Theoretically, the level of *climate departure* should be $2^{-1/2} = 0.71$ relative to the persistence level at sufficiently long time, when the persistence forecast has zero correlation coefficient.

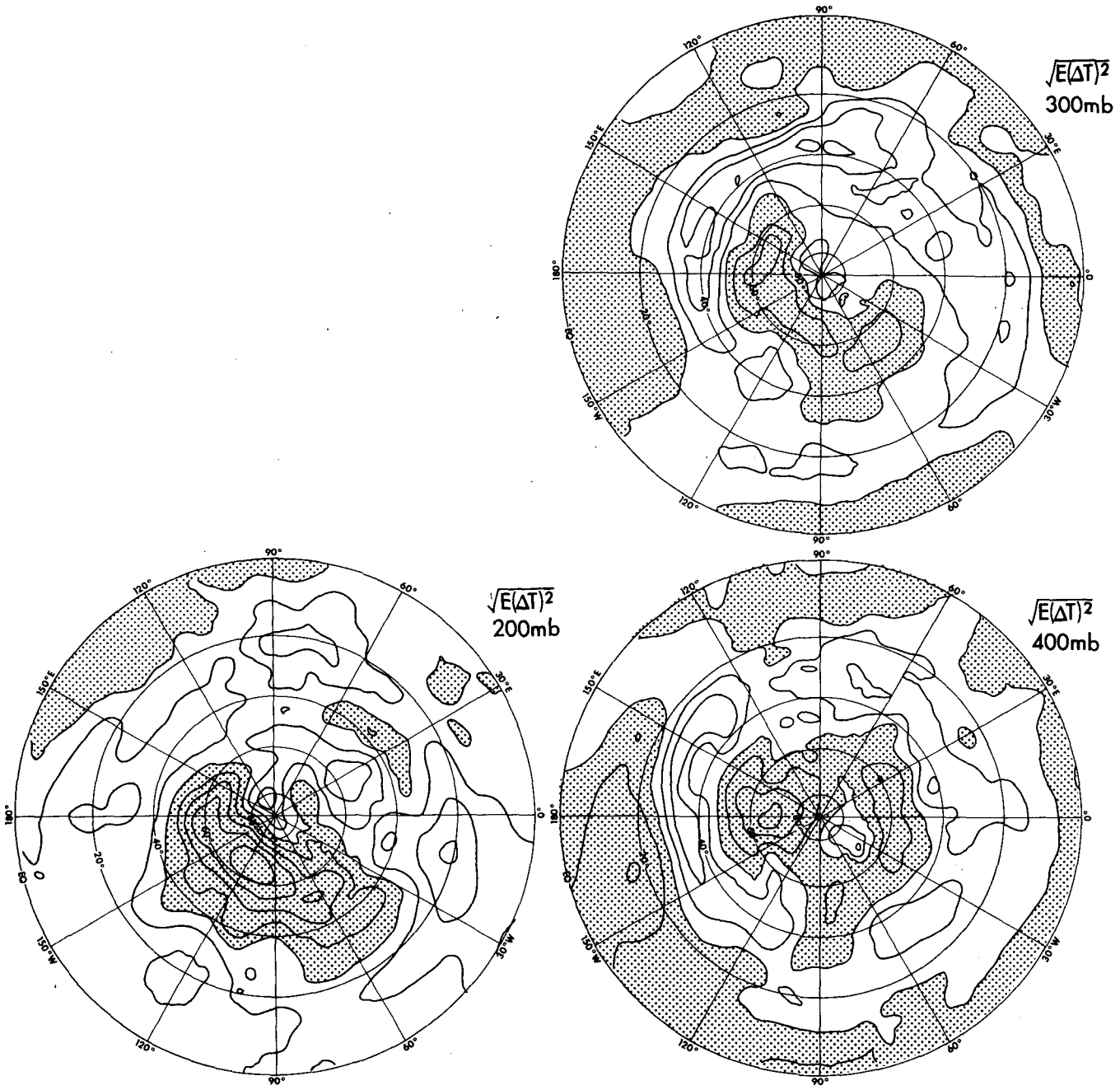

FIG. 19a–b. Eight-case average of rms temperature error for Day 10–30. Contour interval is 1°C.

However, since in our case the persistence correlation is nonzero due to sampling bias, it is possible that the persistence level is underestimated.

Figure 24 is the time evolution of correlation coefficients at various levels, in which the empirical adjustment is included. This figure is to be compared with Fig. 12. A pronounced improvement is evident for all levels.

To summarize this section, it appears that the results are encouraging for one-month forecasts. The potential improvement is a 5 ~ 10 day extension in skill for 10-day means. Note, however, that the eight January cases adopted in this study all include blocking actions. There is a hypothetical view that the blocking situation has a longer memory, and accordingly, that these phenomena tend to be more predictable.

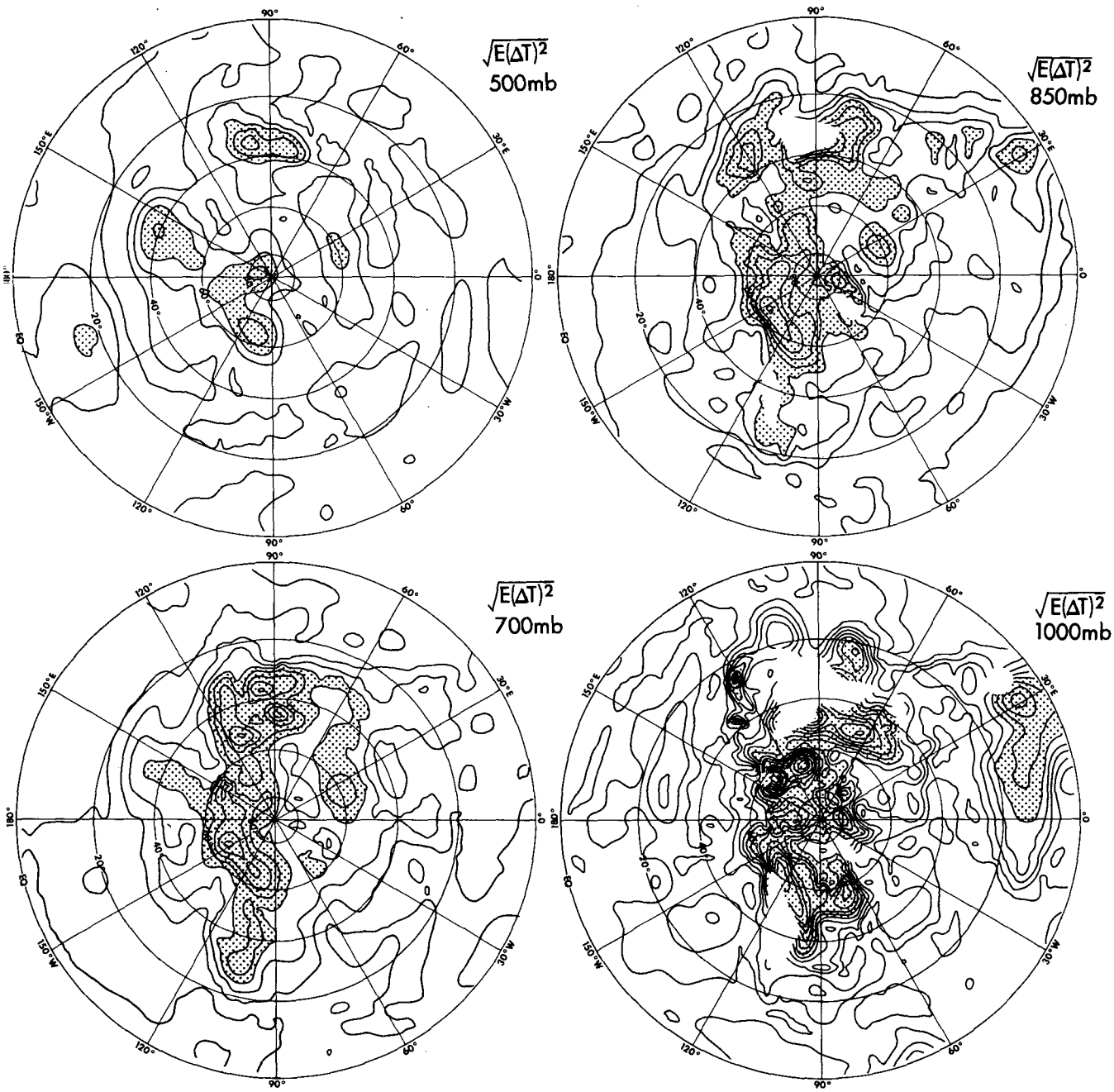

FIG. 19b. (*Continued*)

7. Remarks

a. Application of the empirical adjustment to other cases

There may be a question as to whether the climate drift obtained in the eight January cases can be applied to the case immediately outside the mother sample. In order to investigate this point and also to check the hypothesis on the starting day mismatch in the case of 16 January 1979, we applied the forecast as well as the empirical adjustment to two additional cases, i.e., 1 January 1984 and 16 December 1982. The forecasts were made using the NMC initial conditions. Table 2 shows the results of the test, indicating that the climate drift was applied to both cases; only eight of 12 cases

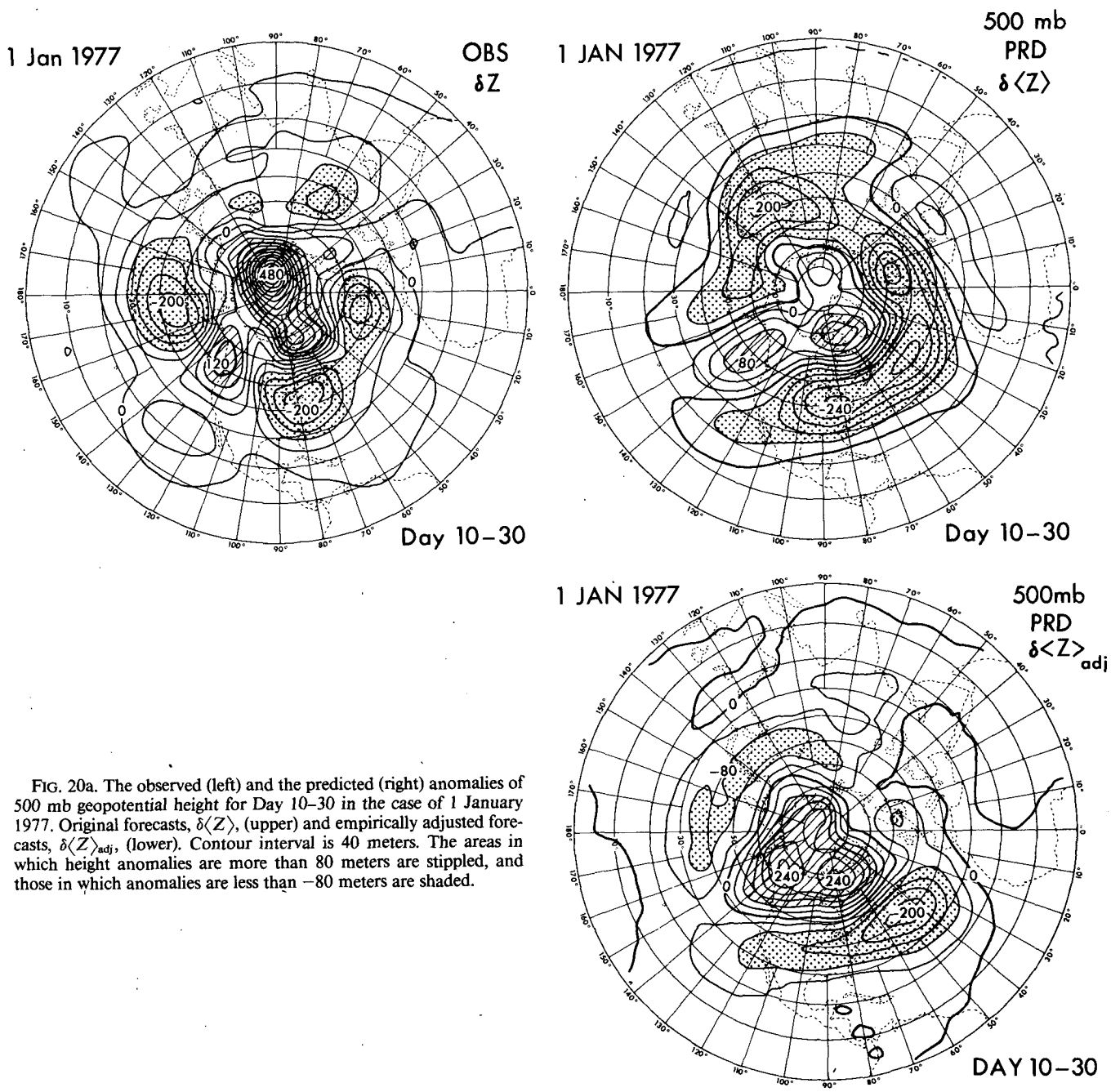

FIG. 20a. The observed (left) and the predicted (right) anomalies of 500 mb geopotential height for Day 10–30 in the case of 1 January 1977. Original forecasts, $\delta\langle Z\rangle$, (upper) and empirically adjusted forecasts, $\delta\langle Z\rangle_{adj}$, (lower). Contour interval is 40 meters. The areas in which height anomalies are more than 80 meters are stippled, and those in which anomalies are less than −80 meters are shaded.

show that the 500 mb correlation coefficients have been increased in the 10-, 20- and 30-day means.

The fact that the case of 16 December 1982 has been improved by the empirical adjustment suggests that the hypothesis of the starting day mismatch was not valid in this case. Concerning the case of 16 January 1979, we have evidence that the forecast was improved by using the Arakawa–Schubert (1974) cumulus parameterization (Miyakoda and Sirutis, 1984) instead of the moist convective adjustment, and that the forecast was also improved by using the 18-level model (*E*-physics) instead of the 9-level model.

b. Adequacy of the GFDL data assimilation scheme

A second point for remarks is concerned with the last statement of section 4. The scores of 500 mb geopotential height with respect to the three initial con-

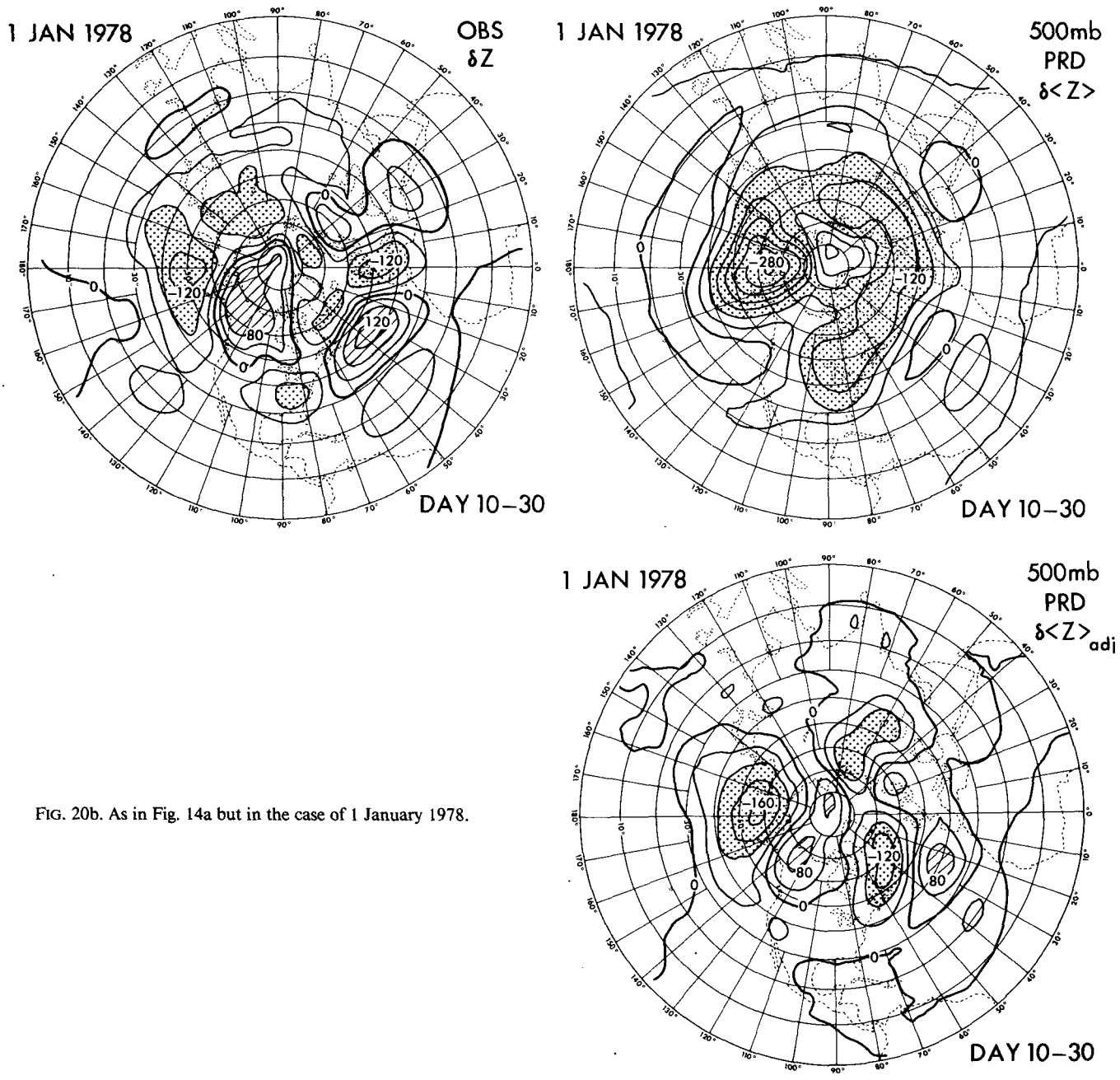

FIG. 20b. As in Fig. 14a but in the case of 1 January 1978.

ditions, i.e., the A, the B and the C, are investigated (not shown here).

The scores show that in the first 10 days the B is best, partially because the verification is the NMC analyses, but that later the A appears slightly better than others. However, the sample number may not be statistically sufficient to draw any definite conclusion. Perhaps this comparison simply shows that all three analyses are adequate for this type of forecasting. (See the similar comparison for the medium-range forecast in Arpe et al., 1985, and Baumhefner, 1984.)

There is an argument that, if the models used in two phases, i.e., data assimilation and prediction, are the same or comparable in the dynamic characteristic, favorable skill scores are expected for the forecast performance. From this point of view, it is suspected that initial condition A is unfairly advantageous in the test described above. The fact is, however, that the models

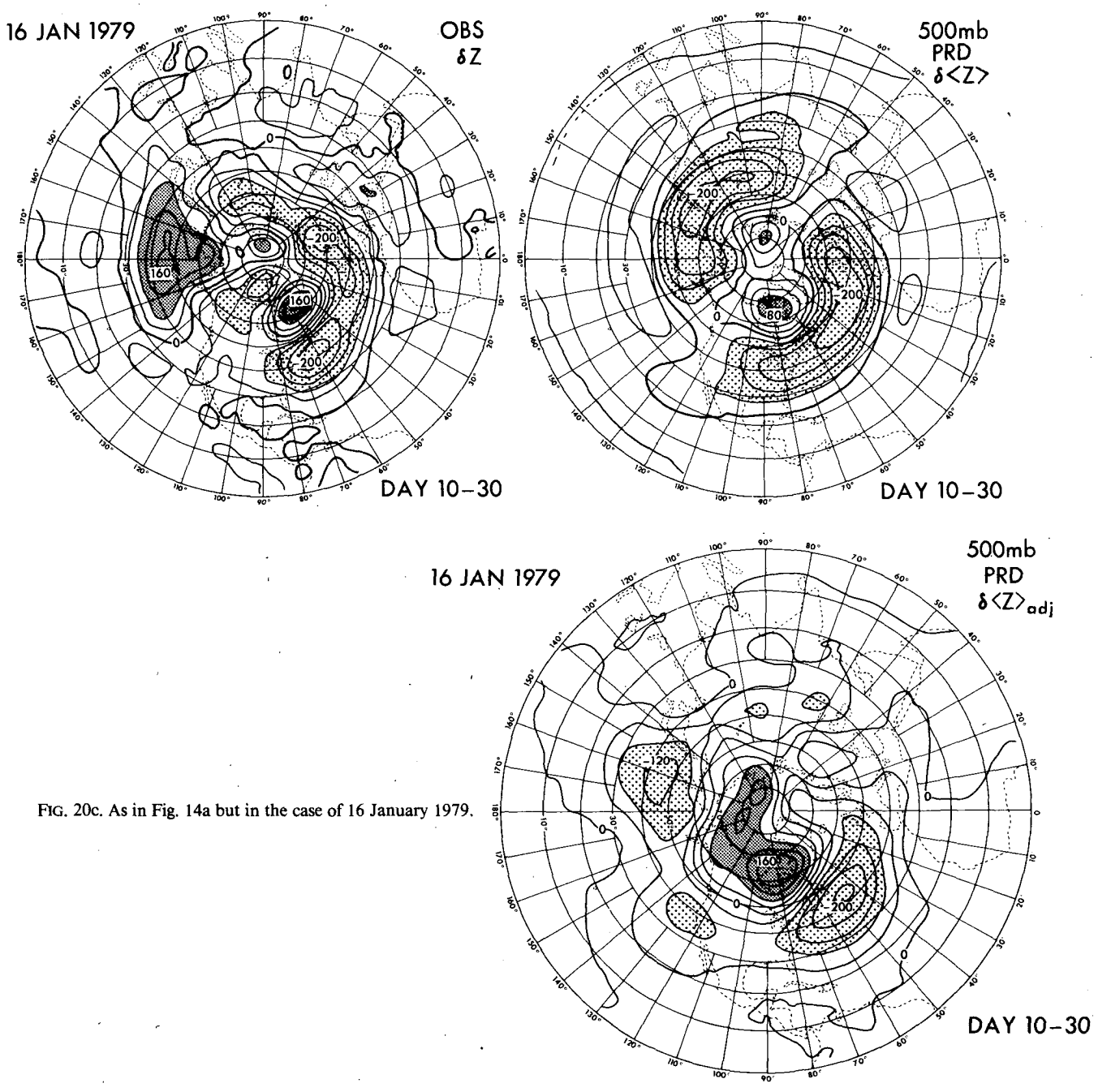

FIG. 20c. As in Fig. 14a but in the case of 16 January 1979.

related to initial condition A are not the same for the two phases (unfortunately). The data assimilation model is an 18 vertical level spectral model without a turbulence closure scheme, whereas the forecast model is a 9 vertical level finite difference model with turbulence closure. Besides, the former includes diurnal variability, while the latter does not, and the radiation packages are different in the two phases.

Yet the concept of the model's consistency in the two phases is valid and, perhaps, important. Dickinson (1982) stated that an ideal system of long-range forecasts would be an optimum combination of a high-quality prediction model and an adequate data assimilation scheme. We concur with Dickinson in this respect. It is envisaged that the future stochastic system is likely to be "lagged average forecasts" proposed by

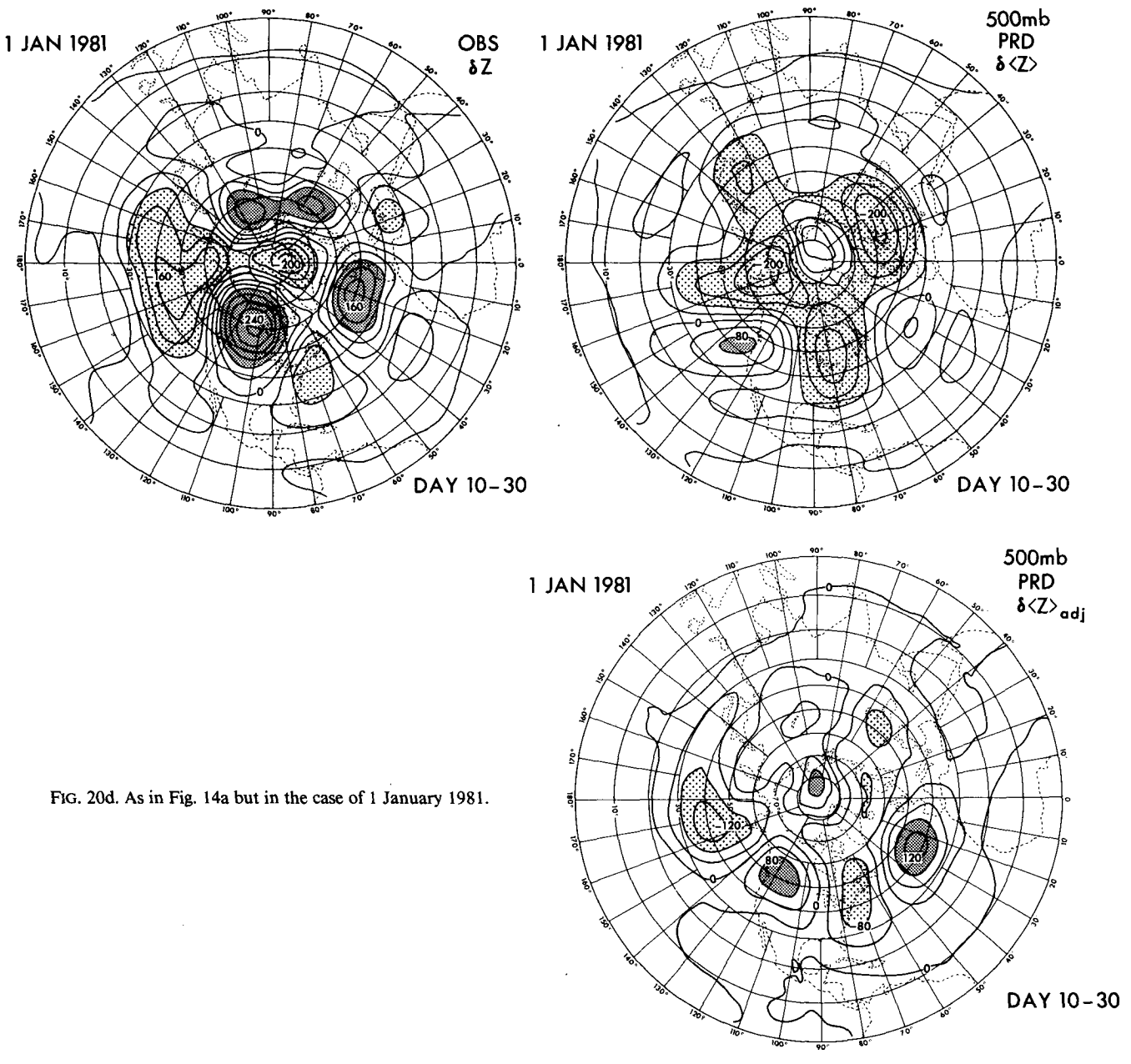

FIG. 20d. As in Fig. 14a but in the case of 1 January 1981.

Hoffman and Kalnay (1983). For these reasons, the role played by the assimilation scheme on long-range forecasts would be increased, and therefore, the adequacy of the assimilation scheme is our great interest.

c. *The limit of predictability*

Monthly predictability of 10-day means discussed in this paper does not contradict the conventional theory of a two-week limit of instantaneous forecast in deterministic predictability. In fact, one-month forecasts of 10-day means are subsets of the conventional forecasts (except for the stochastic aspect) in the sense that only low frequency components are selected.

In the end, the removal of climate drift turns out to be essential. Although the empirical adjustment is a backdoor approach, the results may indicate the minimum level of attainable predictability (the lower bounds in the definition of Lorenz, 1982).

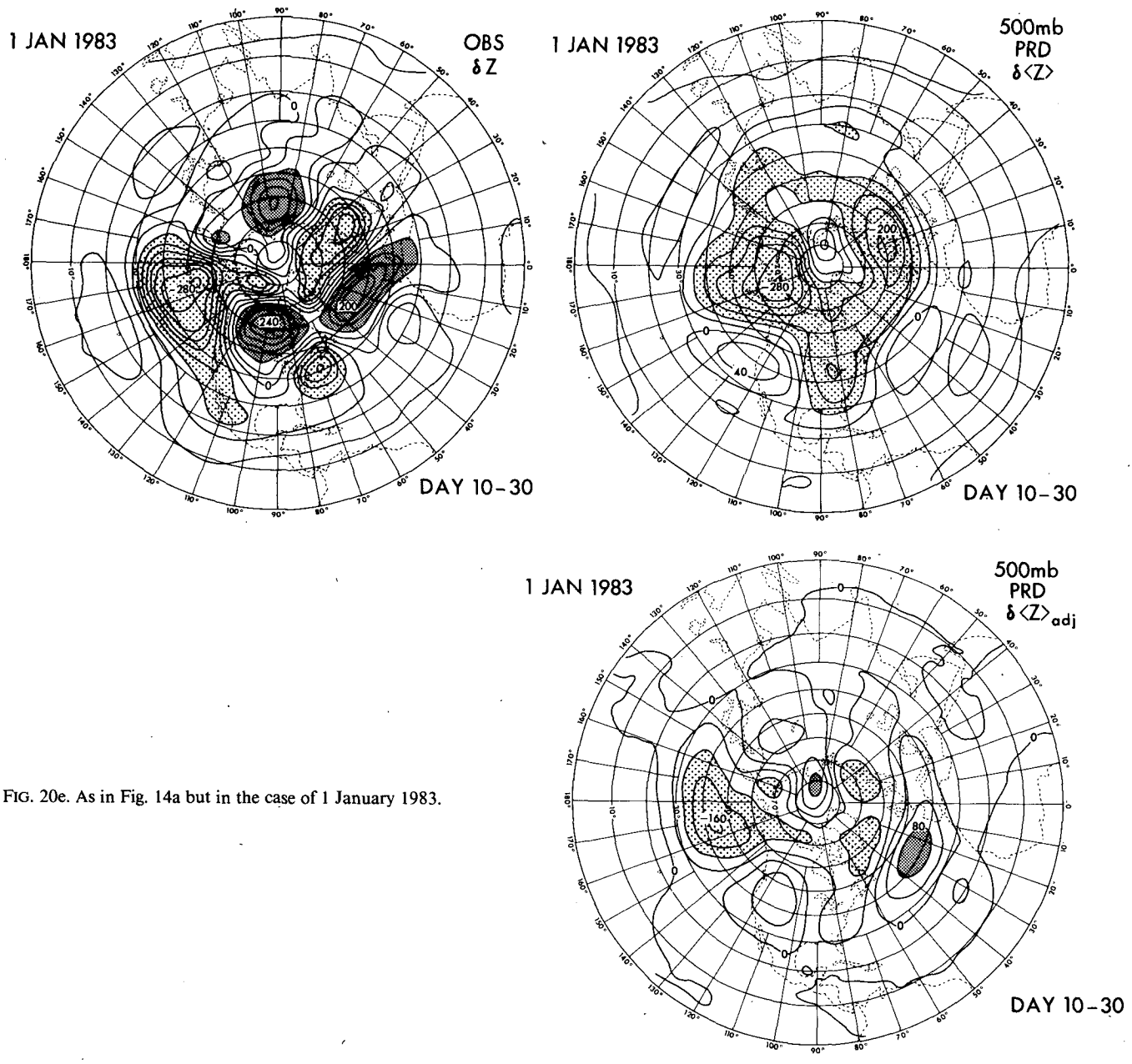

FIG. 20e. As in Fig. 14a but in the case of 1 January 1983.

Another aspect of the predictability is the degree of scatter of three integrations in the framework of stochastic forecasting. The measure is the scatter normalized by the variance over the eight January cases, i.e.,

$$\sigma_f(x)/\tilde{\sigma}_N(x) \tag{7.1}$$

where x is an arbitrary variable. The numerator, $\sigma_f(x)$, is the eight-case average of the scatter as defined by Eq. (3.7); the denominator, $\tilde{\sigma}_N(x)$, is the interannual variance of x, i.e., the standard deviation of x with respect to eight samples of different Januaries, defined by

$$\tilde{\sigma}_N(x) = \left[\frac{1}{n-1}\sum^{n}\{x - E(x)\}^2\right]^{1/2} \tag{7.2}$$

where $n = 8$. For x, the forecast results were employed instead of observations, because of the homogeneous quality of forecast products.

Figures 25 show examples of the interannual vari-

ance, $\tilde{\sigma}_N(x)$, and the normalized scatter, $\sigma_f/\tilde{\sigma}_N$, for the 20-day mean (Day 10–30) of the 500 mb geopotential height, Z_{500}, and the sea level pressure, P_{SL}. In practice, the maps in the lower part of Fig. 25 are the mean of $\sigma_f/\tilde{\sigma}_n$ over three groups based on different initial conditions, A, B and C, i.e.,

$$\langle \sigma_f/\tilde{\sigma}_N \rangle = \frac{1}{3}\sum^{3} \sigma_f/\tilde{\sigma}_N. \tag{7.3}$$

The maps of $\tilde{\sigma}_n$ in the upper part of Fig. 25 are

$$\langle \tilde{\sigma}_N \rangle = \frac{1}{3}\sum^{3} \tilde{\sigma}_N. \tag{7.4}$$

The normalized scatter, exceeding unity, is interpreted as the measure that reaches the limit of predictability in the stochastic forecasting system. Under this assumption, Figs. 25 show the geographical distributions of the predictability for respective variables. Shukla (1981) discussed the significance level based on an F-test, which has some relevance to the present study, at least in intention.

Observations of Figs. 25 may lead to the following conclusions.

- The patterns of $\tilde{\sigma}_N(Z_{500})$ and $\tilde{\sigma}_N(P_{SL})$ are similar to each other, indicating the barotropic nature of variabilities and also showing the approximate agreement of blocking-prone areas. By the way, $\sigma_f(Z_{500})$ and $\sigma_f(P_{SL})$ (not shown here) also resemble each other. It is noted, however, that these samples are slightly different from the observed variance, $\tilde{\sigma}_N$, for a long-term mean (Blackmon et al., 1977), probably because of the climatological bias for this sample, as commented earlier.
- For Z_{500} and P_{SL}, the normalized scatters $\sigma_f/\tilde{\sigma}_N$ are distributed in a similar way. The less-than-unity areas tend to be located in the middle of the Atlantic and the Pacific oceans. A noteworthy point is that the P_{SL} is more predictable than the Z_{500}. It is interesting to note that the pattern for Z_{500} in Fig. 25 has a certain degree of resemblance to Shukla's F-value distributions (Day 1–30 and Day 8–37 in his figures), with respect to the spiral shape of the pattern, and the overall regions of greater-than-unity or of less-than-95% significance level.
- The $\sigma_f/\tilde{\sigma}_N$ pattern of T_{850} (not shown here) is different from the previous quantities. It appears that the region of greater-than-unity for T_{850} is more widely spread, particularly over eastern Europe, the Philippine area and South America.
- The standard deviation of precipitation, σ_f(PRC), (not shown here) and the interannual variability, $\tilde{\sigma}_N$(PRC), show large values over the areas of heavy precipitation (see Fig. A5). For PRC, the greater-than-unity regions are located dominantly over the tropics.

d. The study of systematic errors

According to Epstein (1985), the empirical adjustment has long been used in operational practice for short-range forecasts (Glahn and Lowry, 1972; Glahn, 1980; Hughs, 1982), and the concept was advocated by Lorenz (1977). In fact, the removal of systematic error for short-range prediction is a better approximation because the errors are more linear.

In these empirical adjustments, the drift is simply subtracted only after all prediction calculations are finished, as was done in our study. Another version is to apply adjustment continuously in an interactive mode. A method of tuning the governing equations based on one-time-step prediction was proposed by Faller and Schemme (1977) as a statistical correction. The method has not been practically implemented yet.

This systematic error becomes more pronounced and serious in the medium-range and monthly forecasts (e.g., Hollingsworth et al., 1980; Derome, 1981; Bengtsson and Lange, 1981; Bettge and Baumhefner, 1984). The causes of the drift have been investigated from various angles (Wallace et al., 1983; Burridge and Sadourny, 1982; Arpe and Klinker, 1986), and yet the causes do not appear to be easily unravelled. The problems are related to a model's spatial resolution, physics and even numerical methods, coupled with the fluid mechanics. The dynamics involved in the drift are intrinsically nonlinear in character, as mentioned for Figs. 18 and 19. (See Wallace et al., 1983.)

Perhaps it is safe to state that there are two kinds of error, i.e., the primary and the secondary errors (Arpe and Klinker, 1986). The secondary error is the error development due to the dynamic instability, and it is, therefore, uncontrollable. What one can achieve is to reduce the primary error sources. In this sense, it is speculated that the "statistical correction" in the interactive version is very difficult to apply because of the nonlinear character of the error development.

The most serious deficiency in models is the excessive cooling, particularly during the winter polar night region. Mahlman and Umscheid (1985) mentioned, based on their experience with a high resolution model, i.e., N90L40, that the resolution might contribute significantly to reducing the coldness at high latitudes. The most recent study of Cubasch and Tibaldi (personal communication, 1985) has revealed that the spatial resolution could be relevant in alleviating the climate drift based on a numerical experiment with various resolution spectral models, including the currently highest resolution model of triangular truncation 106 and 15 level, i.e., T106L15. Other possible effects are the adequate treatment of cirrus clouds (Ramanathan et al., 1983), the arrangement of a sufficient supply of water vapor from the earth's surface, and the inclusion of diurnal variability.

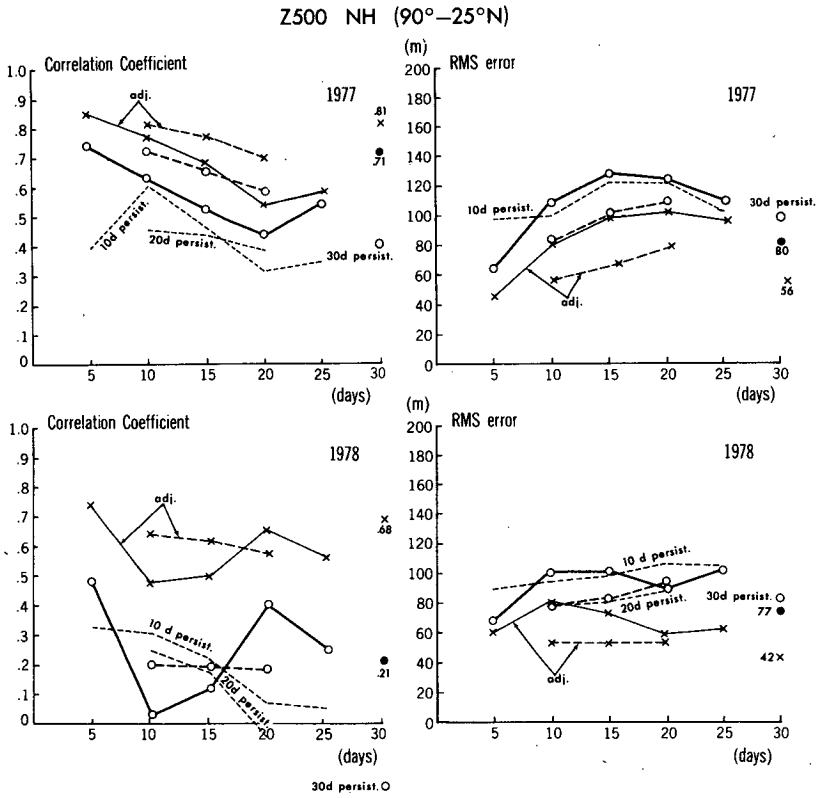

FIG. 21a. Correlation coefficients (left) and rms error (right) for the stochastic mean of 500 mb geopotential height in the cases of January 1977 and 1978. The original forecasts are small circles, and the empirically adjusted forecasts are small crosses. Solid and dashed lines are for the 10-day and the 20-day means, respectively. See Fig. 8a for a further explanation.

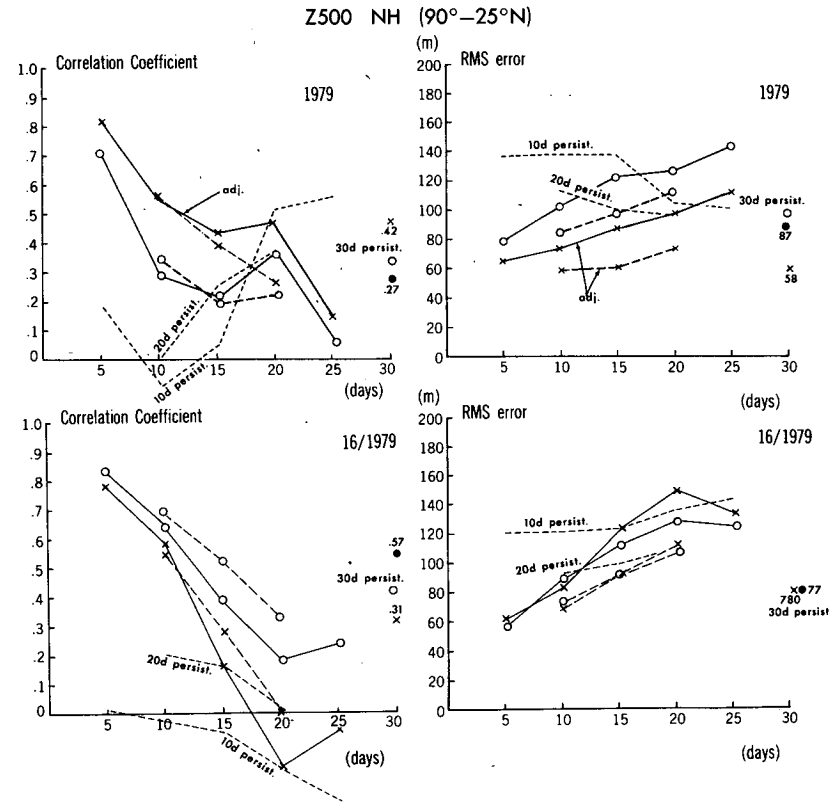

FIG. 21b. As in Fig. 17a but in the cases of 1 January and 16 January 1979.

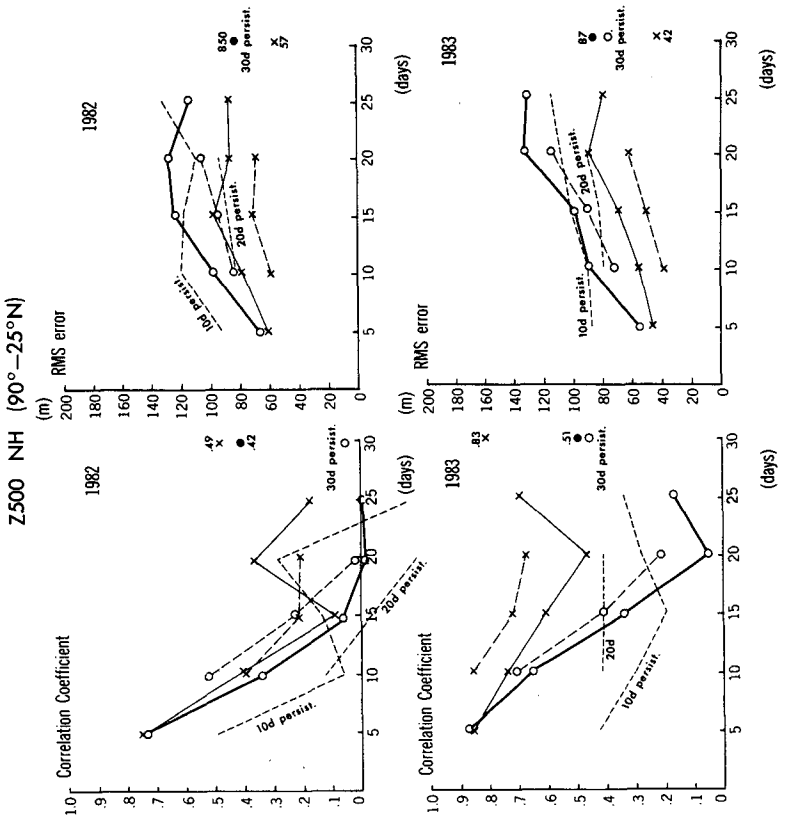

FIG. 21d. As in Fig. 17a but in the cases of January 1982 and 1983.

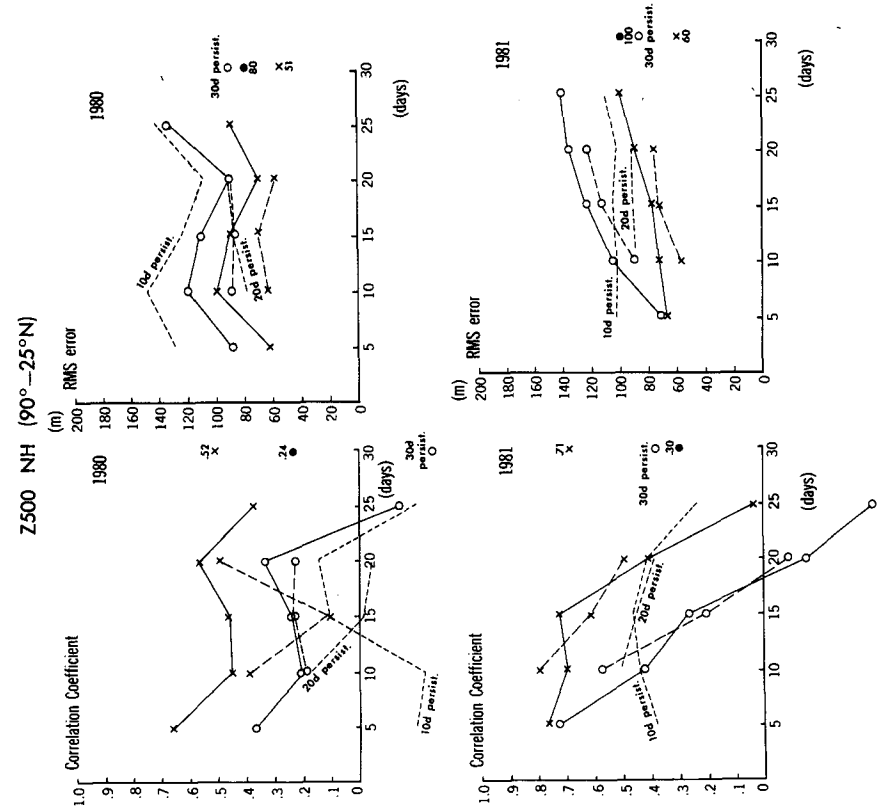

FIG. 21c. As in Fig. 17a but in the cases of January 1980 and 1981.

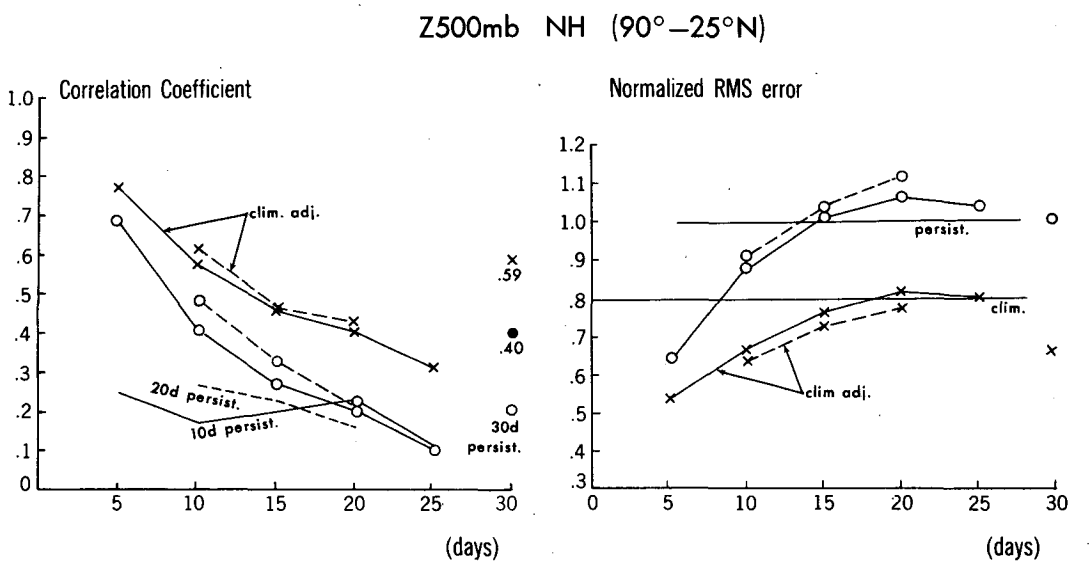

FIG. 22. Ensemble mean skill scores over the eight January cases for 500 mb geopotential height. The rms error is normalized by the persistence. See Figs. 10 and 21a for further explanation.

It should be emphasized that the empirical adjustment does not solve the basic difficulty. The search for primary error sources is far more preferable than the mere application of empirical corrections. The effort in the front-door approach should be encouraged. The reduction of climate drift is fundamental for the improvement of the one-month forecasts and the advancement to the seasonal forecasts, even if the empirical adjustment has to be applied after all.

8. Conclusions

Based on a series of numerical experiments of the eight January cases with a particular GCM and with a climatologically normal sea-surface temperature at the lower boundary, conclusions are drawn as follows.

(i) A marginal skill is found for monthly forecasts, using the 10-day mean filter applied to the prognoses (see Figs. 10 and 11).

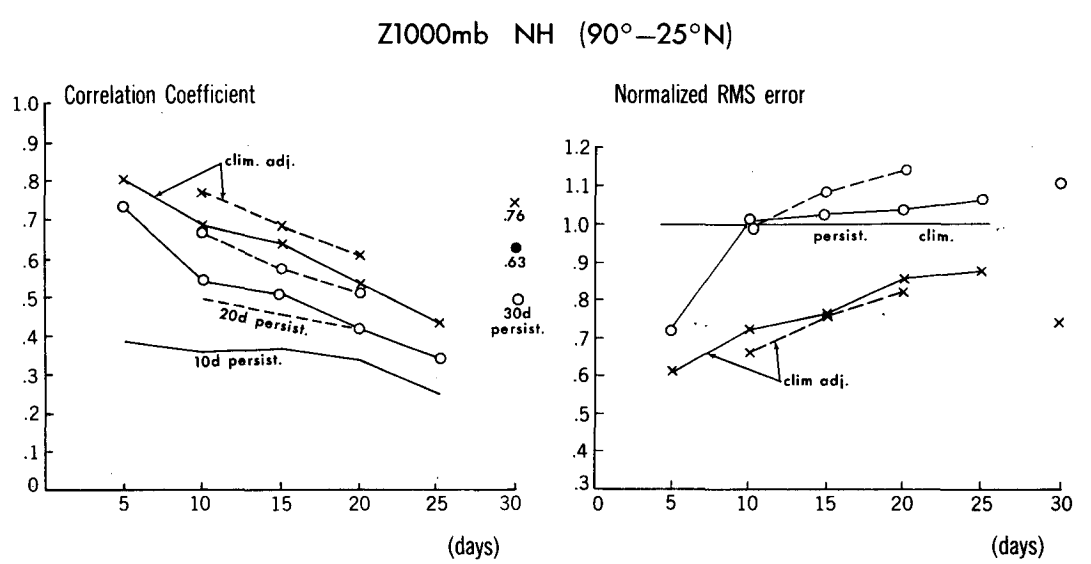

FIG. 23. As in Fig. 22 but for 1000 mb geopotential height.

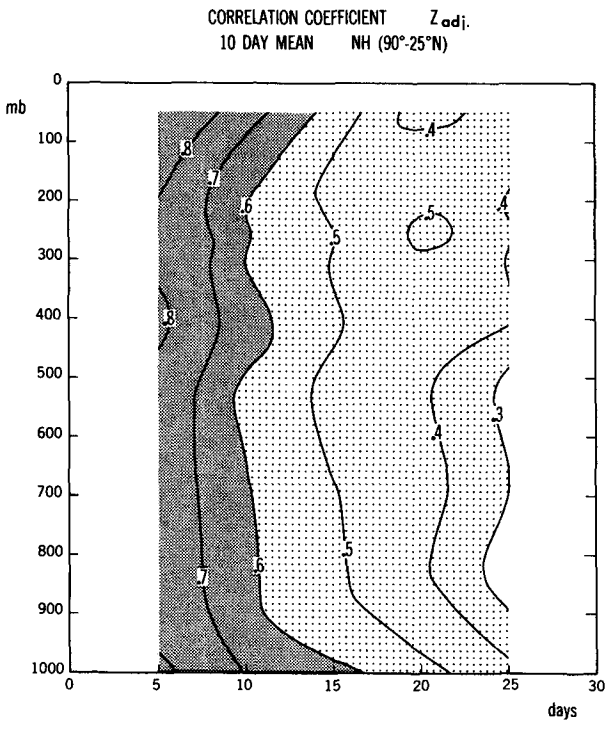

FIG. 24. Time evolution of the forecast skill, based on the empirically adjusted forecasts. See Fig. 12 for further explanation.

(ii) The forecast error pattern in terms of rms error is highly correlated with the arithmetic mean error pattern, suggesting that an appreciable systematic bias (climate drift) is included.

(iii) The predictive skill scores are substantially raised by subtraction of the climate drift from the original prognoses. Out of eight, five were improved substantially. The potential improvement of monthly forecasts is, therefore 5 ~ 10 days.

(iv) In view of these results, it is conceived that one-month forecasts may be feasible. However, the reduction of the climate drift is required for the outright application of a dynamical forecast approach.

(v) The GFDL four-dimensional data assimilation procedure is acceptable for the use of one-month forecasts.

(vi) The geographical distribution of forecast scatter in the stochastic prediction is different with respect to variables. The geopotential height and the sea level pressure, for example, exhibit a statistically distinctive configuration of the spread, roughly agreeing with the distribution of blocking-ridge-prone areas.

(vii) The stochastic scatter is not large in the one-month time scale, but at the end of a month, it grows to a substantial magnitude.

It has been remarked that the eight January cases (1977–83) adopted in this paper belong to a climato-

TABLE 2. Correlation coefficients and rms error in units of meters for the original and the empirically adjusted forecasts of 500 mb geopotential height in the cases of 1 January 1984 and 16 December 1982.

		10 day mean			20 day mean		30 day mean
		0–10	10–20	20–30	0–20	10–30	0–30
		1 January 1984					
Z_{500}							
Correlation coefficient	Original	0.86	0.22	0.36	0.60	0.31	0.53
	Adjusted	0.83	0.24	0.22	0.69	0.33	0.57
rms error in meters	Original	52.2	117.8	126.7	74.1	105.7	79.2
	Adjusted	58.6	106.0	120.7	62.1	85.0	64.4
Z_{1000}							
Correlation coefficient	Original	0.91	0.46	0.38	0.77	0.48	0.72
	Adjusted	0.91	0.57	0.35	0.85	0.57	0.78
rms error in meters	Original	42.2	81.1	87.6	52.7	72.9	55.6
	Adjusted	40.7	62.3	81.7	40.2	57.7	43.4
		16 December 1982					
Z_{500}							
Correlation coefficient	Original	0.81	0.08	0.05	0.02	−0.04	0.26
	Adjusted	0.76	0.09	0.04	0.15	0.01	0.33
rms error in meters	Original	63.2	135.3	148.5	127.8	146.0	108.6
	Adjusted	67.7	125.4	132.7	109.2	127.8	87.6
Z_{1000}							
Correlation coefficient	Original	0.87	0.26	0.32	0.35	0.25	0.51
	Adjusted	0.88	0.23	0.34	0.47	0.28	0.58
rms error in meters	Original	44.6	98.9	110.2	88.4	106.0	80.0
	Adjusted	38.4	92.3	98.4	70.5	95.0	65.3

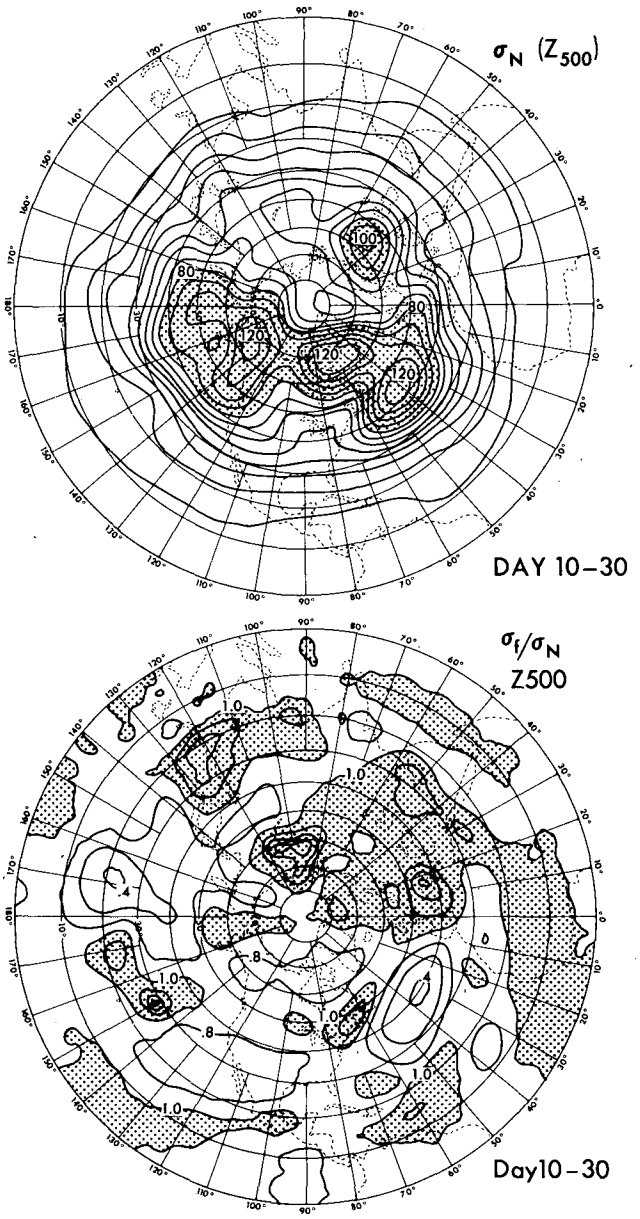

FIG. 25a. The interannual variances, $\tilde{\sigma}_N$, (upper) and the normalized scatter, $\sigma_f/\tilde{\sigma}_N$, (lower) for the 500 mb geopotential height, Z_{500}.

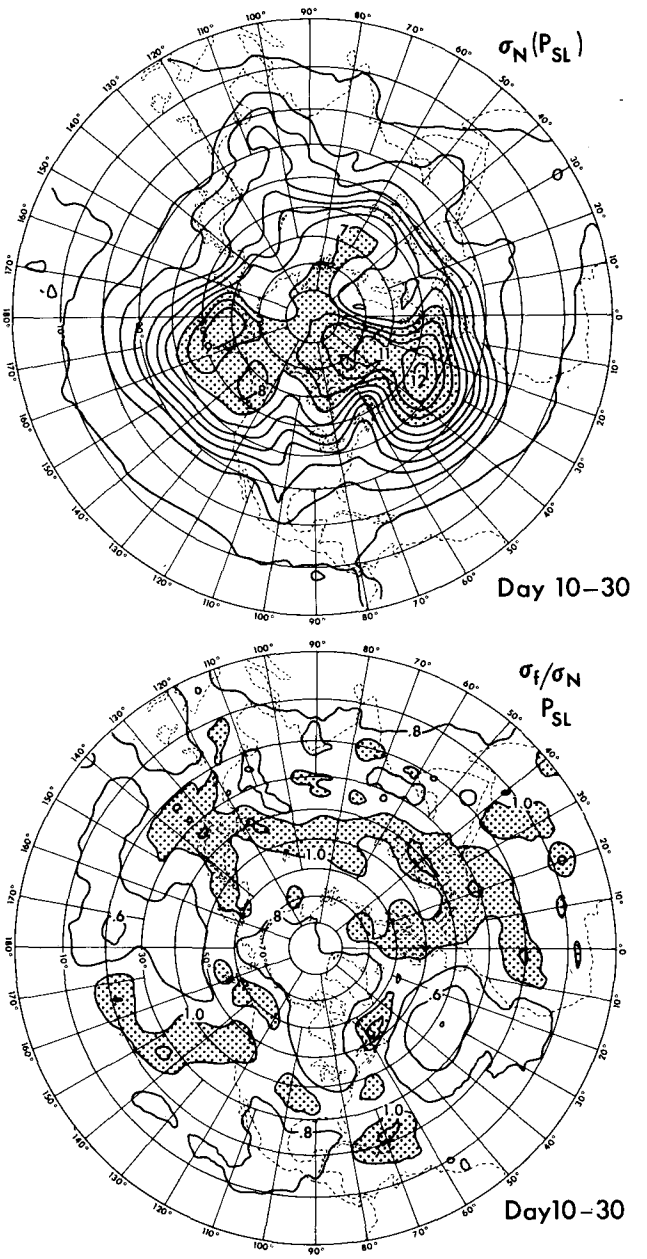

FIG. 25b. As in Fig. 25a but for the sea level pressure, P_{SL}.

logically biased period. For example, these cases all include blocking actions. There is a hypothetical view that the blocking phenomena tend to be more predictable.

Acknowledgments. We are grateful to Drs. L. Bengtsson, D. Gilman, Y. Hayashi, A. Hollingsworth, J. Mahlman, A. Oort, R. Pierrehumbert, A. Simmons, S. Tibaldi and A. Gordon. Many individuals at GFDL assisted in this work. Thanks are expressed to Mr. R. Caverly, L. Umscheid, R. White, P. Baker and D. Fork. We appreciate the contributions from Mrs. J. Pege, Mr. P. Tunison and his team, and Mr. J. Conner.

APPENDIX

General Circulation Feature

To gain a better insight into the model's characteristics and systematic biases, some of the general cir-

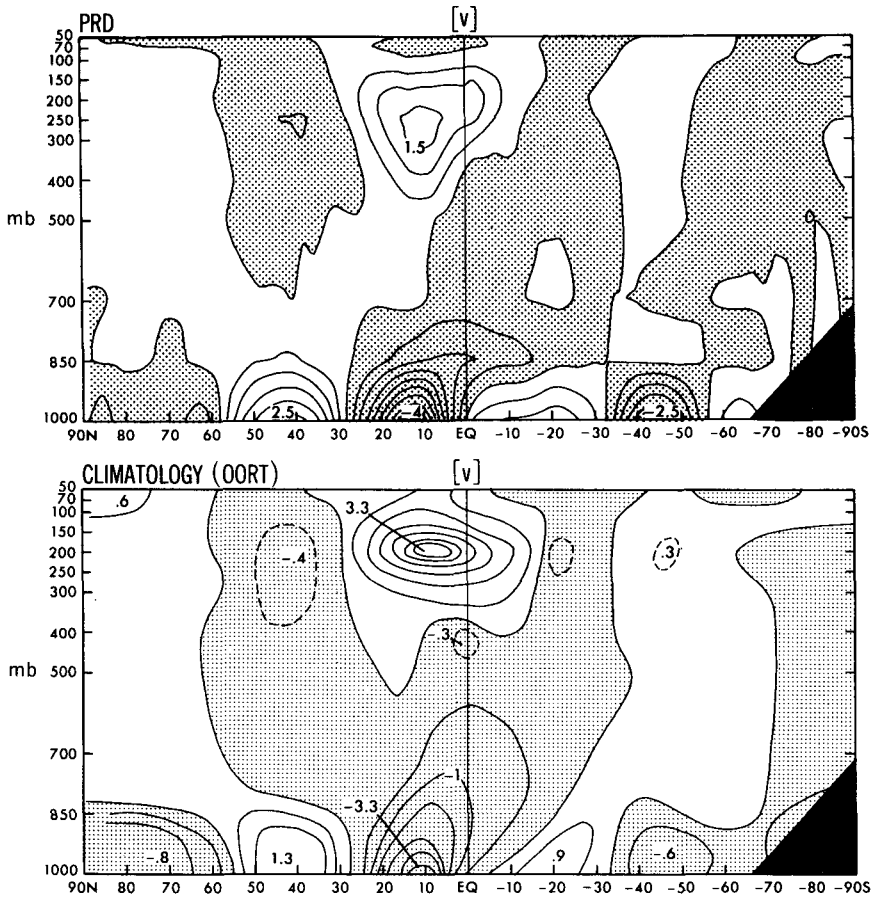

FIG. A1. The monthly and zonal mean of meridional wind for the forecast (upper) and Oort's January climatology (lower) with contour interval of 0.5 m s^{-1} in solid lines and 0.25 m s^{-1} in dashed lines.

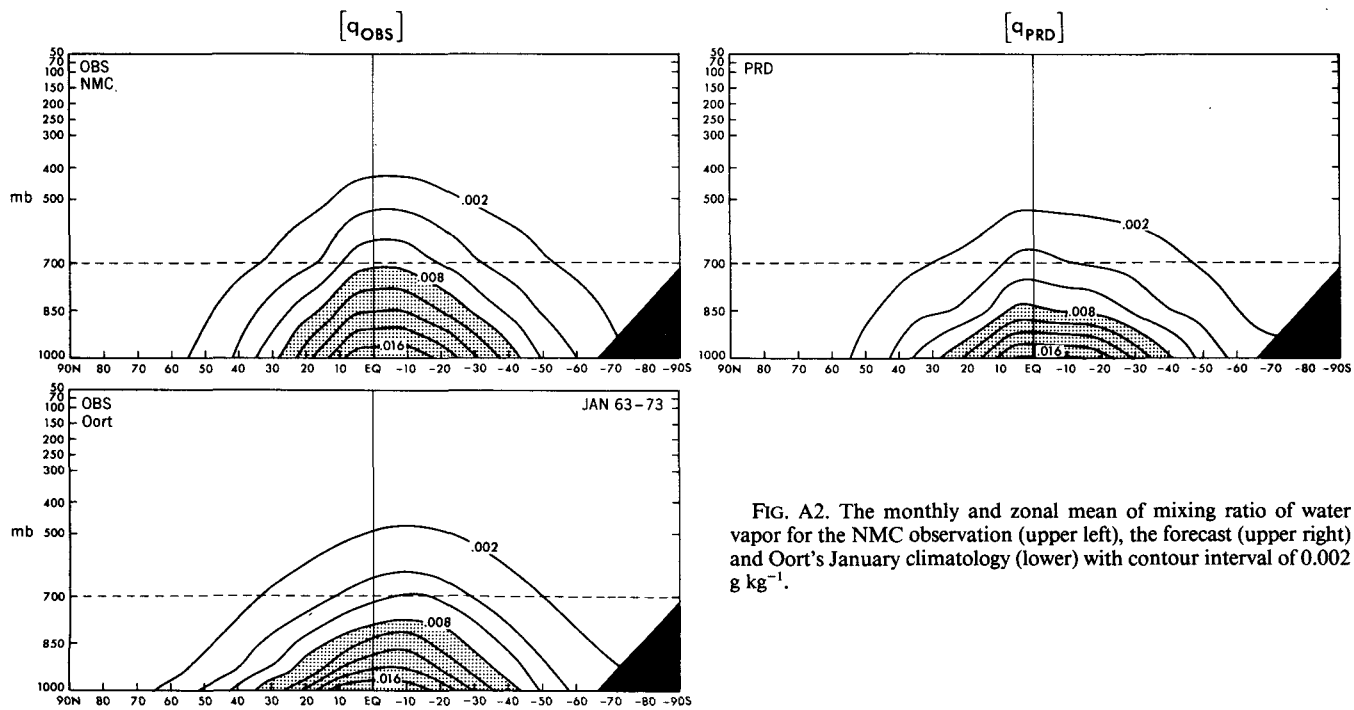

FIG. A2. The monthly and zonal mean of mixing ratio of water vapor for the NMC observation (upper left), the forecast (upper right) and Oort's January climatology (lower) with contour interval of 0.002 g kg^{-1}.

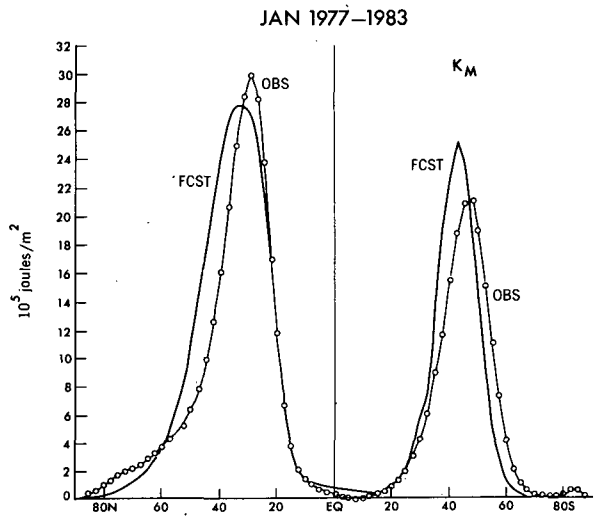

FIG. A3a. Latitudinal distributions of zonal mean kinetic energy, K_M, which are averaged vertically. Small circles are observation.

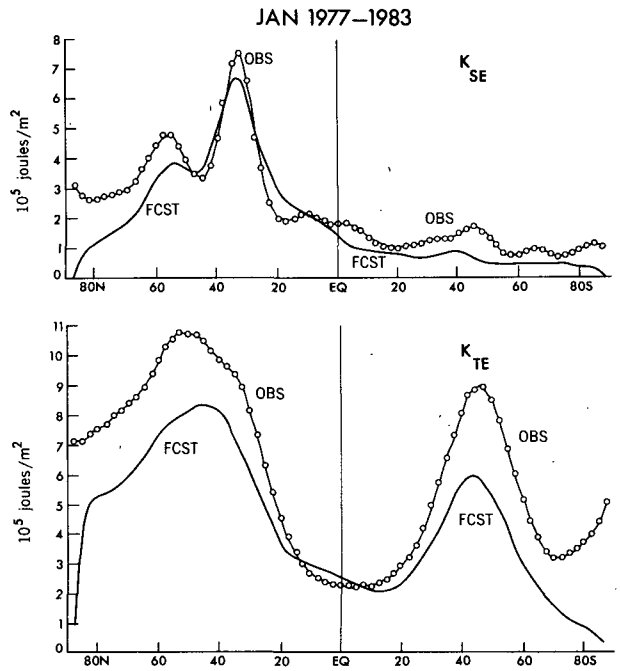

FIG. A3b. Latitudinal distributions of stationary eddy kinetic energy, K_{SE} (upper) and transient eddy kinetic energy, K_{TE} (lower), which are averaged vertically and zonally.

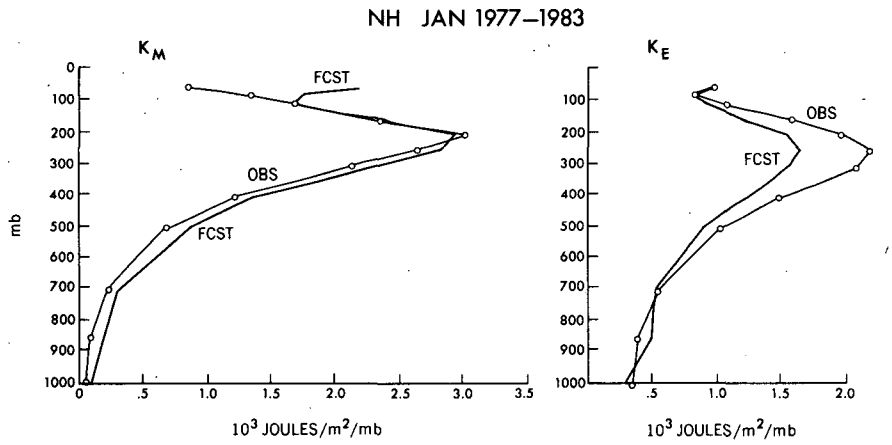

FIG. A4a. Vertical distributions of zonal mean kinetic energy, K_M (left), and eddy kinetic energy, K_E (right), averaged horizontally for the entire globe. Small circles are observations.

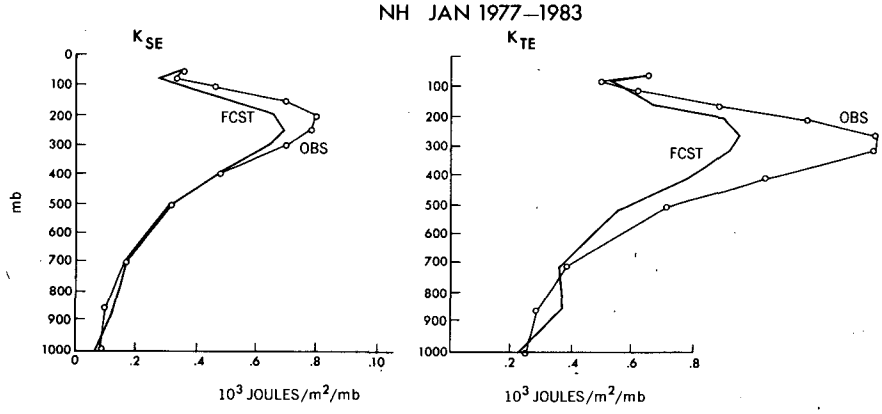

FIG. A4b. As in Fig. A4a but for stationary eddy kinetic energy, K_{SE} (left), and transient eddy kinetic energy, K_{TE} (right).

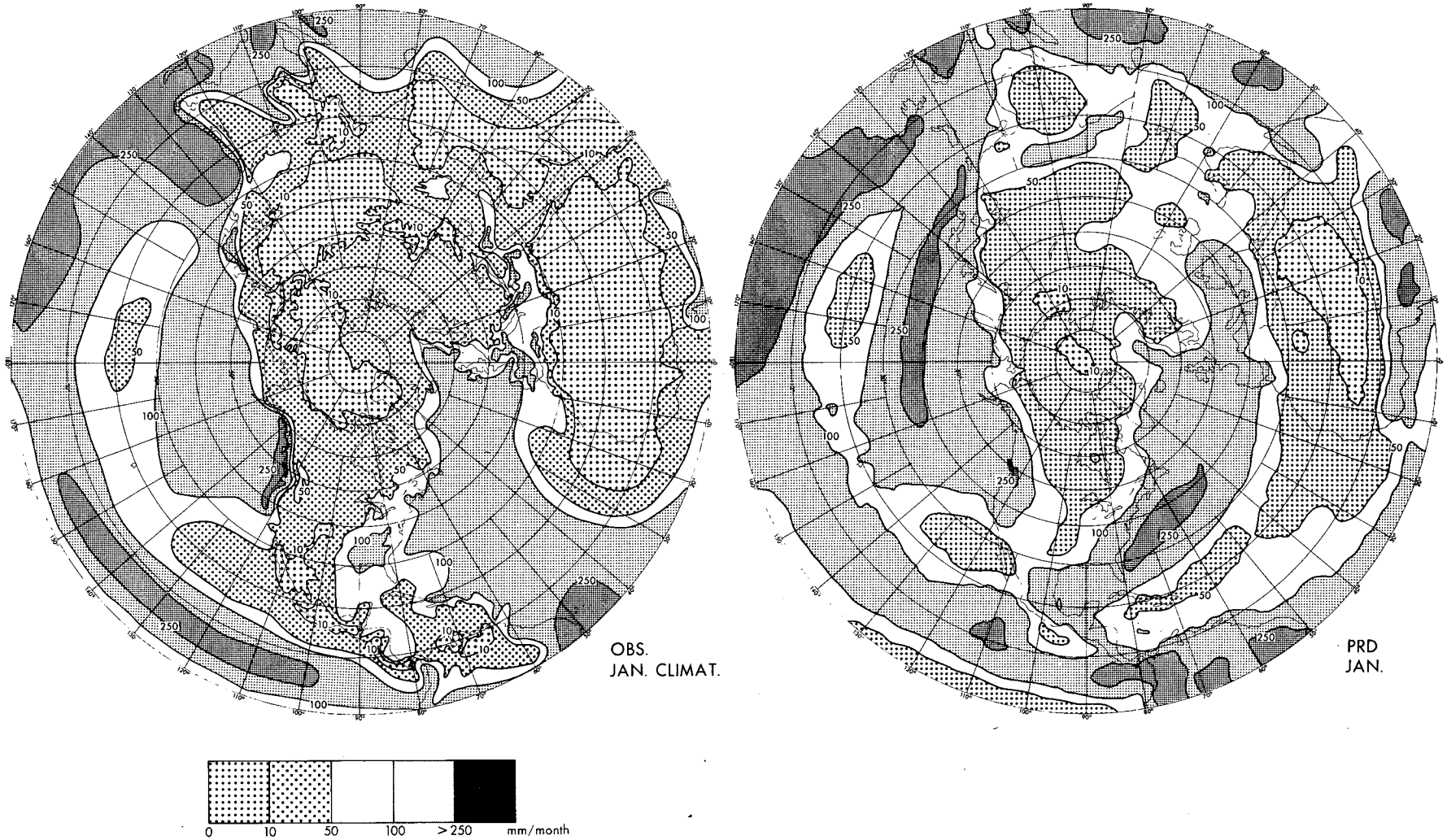

FIG. A5. The rate of precipitation for January climatology (left) and forecasts (right). Contours are 500, 250, 100, 50 and 10 mm mon^{-1} as shown in the legend.

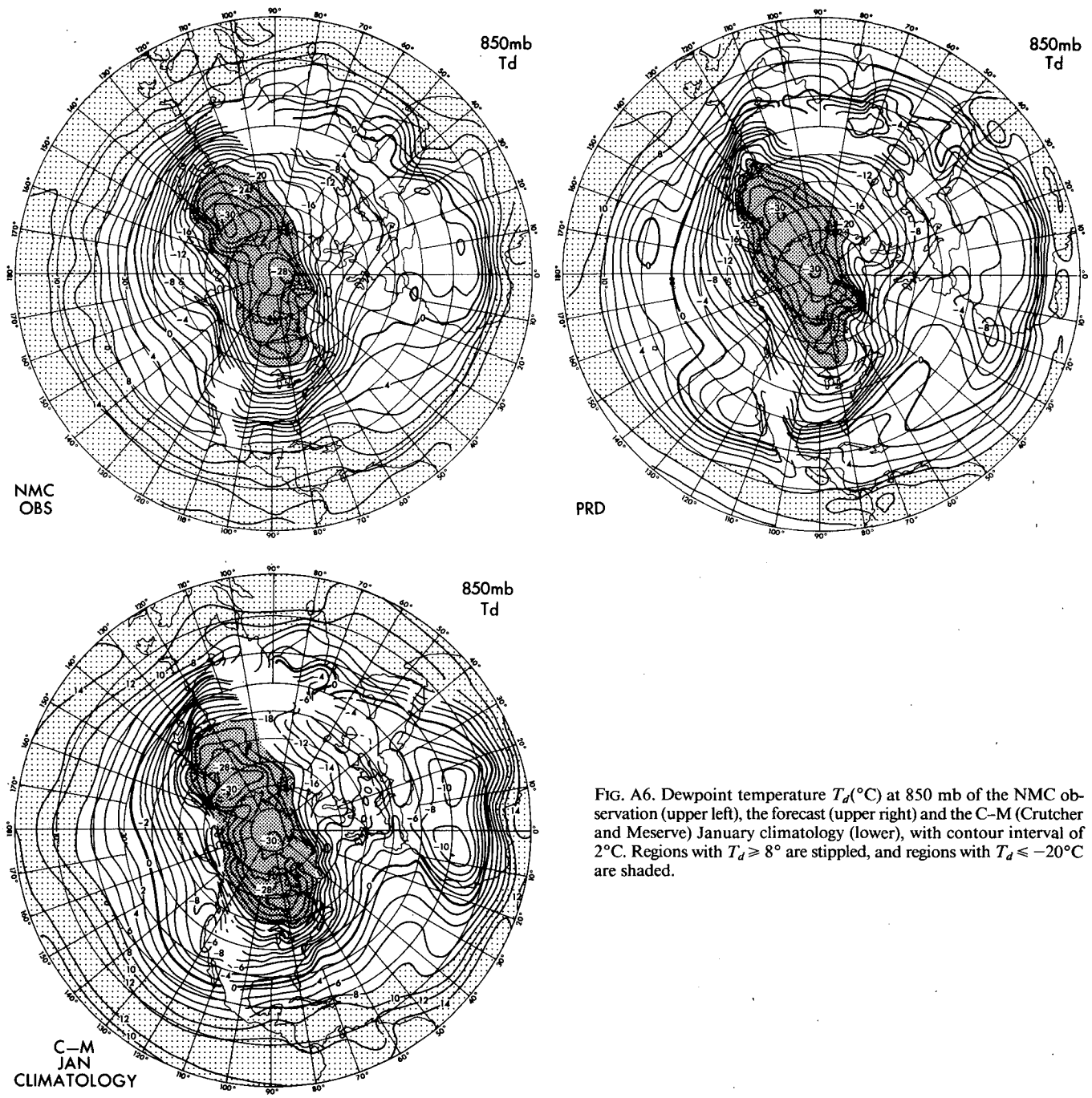

FIG. A6. Dewpoint temperature T_d(°C) at 850 mb of the NMC observation (upper left), the forecast (upper right) and the C–M (Crutcher and Meserve) January climatology (lower), with contour interval of 2°C. Regions with $T_d \geq 8°$ are stippled, and regions with $T_d \leq -20°$C are shaded.

culation features are presented in Figs. A1–6. The variables in the model's results are averaged for 30 days with respect to three realizations over eight January cases, and they are compared with the corresponding observations or climatologies.

Six items are shown. The first two are the zonal means of meridional wind and the mixing ratio of water vapor; the third item is the latitudinal distribution of stationary and transient eddy kinetic energies, which are integrated vertically; the fourth item is the vertical distributions of zonal mean and eddy kinetic energies as well as stationary and transient eddy kinetic energies,

which are averaged horizontally over the globe; the fifth and the sixth items are the Northern Hemispheric distributions of precipitation and dewpoint temperature.

The stationary and transient eddy kinetic energies are based on Oort's (1983) definition. Eddies are the deviations from zonal means, and the stationaries and transiences are referred to the monthly means. Then, the eddy kinetic energies K_E are written by

$$K_E = K_{SE} + K_{TE}$$

where K_{SE} and K_{TE} are the stationary and transient energies.

The observations are taken from the NMC analyses for the corresponding time to that of the forecasts, but the meridional wind and moisture are compared with January climatologies taken from Oort (1983); the precipitation is compared with January climatology from Lvovitch and Ovtchinnikov (1964); and the dewpoint temperature is compared with January climatology from Crutcher and Meserve (1970).

Six salient features are noted below.

- The K_M of forecasts in the Northern Hemisphere is almost comparable with or slightly larger than the observation, but the peak of forecasts is clearly shifted poleward.
- The K_E in the forecast is substantially less than that in observation.
- The observed distribution of K_{SE} in the Northern Hemisphere shows two maxima at latitudes of 35° and 55°. The K_{SE} at 55°N is related to blocking activities (Miyakoda and Sirutis, 1984). The prediction of K_{SE} captures this feature, but the magnitude is somewhat less than the observation. The K_{SE} in the Southern Hemisphere is substantially lower in the forecasts than in the observation.
- The K_{TE} is underestimated considerably (at least 20%).
- The forecasted distribution of precipitation is appreciably improved compared with the previous results (Miyakoda and Strickler, 1981). However, there is still clear disagreement with the climatology over Africa, tropical Pacific, South America, and two storm tracks in the midlatitude oceans.
- The forecasted moisture is appreciably lower than the observed or the climatology in the tropics. The dewpoint temperature pattern has been substantially improved in this paper compared with the previous study in the aforementioned paper.

REFERENCES

Arakawa, A., and W. H. Schubert, 1974: Interaction of cumulus cloud ensemble with the large-scale environment. Part I. *J. Atmos. Sci.*, **31**, 674–701.

Arpe, K., and E. Klinker, 1986: Systematic errors of the ECMWF operational forecasting model in midlatitudes. *Quart. J. Roy. Meteor. Soc.*, **112**, 181–202.

——, L. Bengtsson, A. Hollingsworth and Z. Janjic, 1976: A case study of a ten-day prediction. Tech. Rep. No. 1, European Centre for Medium-Range Weather Forecasts, UK, 105 pp. [Available at ECMWF, Shinfield Park, Reading, Berkshire, RG2 9AX, England.]

——, A. Hollingsworth, M. S. Tracton, A. C. Lorenc, S. Uppala and P. Kallberg, 1985: The response of numerical weather prediction systems to FGGE level II-b data. Part II: Forecast verifications and implications for predictability. *Quart. J. Roy. Meteor. Soc.*, **111**, 67–101.

Barnett, T. P., 1978: Estimating variability of surface air temperature in the Northern Hemisphere. *Mon. Wea. Rev.*, **106**, 1353–1367.

Baumhefner, D. P., 1984: Forecast intercomparisons using the global weather experiment data base. GARP Special Rep. No. 43. Report of the seminar on progress in numerical modelling and the understanding of predictability as a result of the global weather experiment, Sigtuna, Sweden, WMO/ICSU, III-141-154.

——, and T. W. Bettge, 1981: Characteristics of atmospheric planetary circulations and associated model forecast skill during FGGE case studies selected by WGNE. *Proc. Int. Conf. on Early Results of FGGE and Large-Scale Aspects of the Monsoon Experiments.* Tallahassee. [Available at WMO, Case Postale No. 5, CH-1211 Geneva 20, Switzerland.]

Bengtsson, L., 1981: Numerical prediction of atmospheric blocking—a case study. *Tellus*, **33**, 19–42.

——, and A. Lange, 1981: Results of the WMO/CAS NWP data study and intercomparison project for forecasts for the Northern Hemisphere in 1979–1980. PWPR Rep., WMO, 26 pp.

Bettge, T. W., and D. P. Baumhefner, 1984: Total and planetary-scale systematic errors in recent NMC operational model forecasts. *Mon. Wea. Rev.*, **112**, 2317–2325.

Blackmon, M. L., J. M. Wallace, N. C. Lau and S. L. Mullen, 1977: An observational study of the Northern Hemisphere wintertime circulation. *J. Atmos. Sci.*, **34**, 1040–1053.

Burridge, D., and R. Sadourny, 1982: Systematic errors in general circulation models. *WMO/ISCU Study Conf. on Physical Basis for Climate Prediction.* Leningrad World Climate Programme. WMO, 169–188.

Crutcher, H. L., and J. M. Meserve, 1970: Selected level heights, temperatures and dew points for the Northern Hemisphere. Naval Weather Service Command, U.S. Govt. Printing Office, 29 pp.

Daley, R., A. Hollingsworth, J. Ploshay, K. Miyakoda, W. Baker, E. Kalnay, C. Dey, T. Krishnamurti and E. Barker, 1985: Objective analysis and assimilation techniques used for the production of FGGE III-b analyses. *Bull. Amer. Meteor. Soc.*, **66**, 532–538.

Derome, J., 1981: On the average errors of an ensemble of forecasts. *Atmos.-Ocean*, **19**, 103–127.

Dickinson, R. E., 1982: Feasibility of monthly and seasonal forecasting by numerical models. *WMO/ICSU Study Conf. on Physical Basis for Climate Prediction*, Leningrad World Climate Programme. WMO, 235–245.

Druyan, L. M., R. C. J. Somerville and W. J. Quirk, 1975: Extended-range forecasts with the GISS model of the global atmosphere. *Mon. Wea. Rev.*, **103**, 779–795.

Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus*, **21**, 739–759.

——, 1985: Procedure for the statistical correction of medium-range spectral forecasts. NMC Office Note No. 301. [Available at NMC, Washington, DC.]

Esbensen, S. K., 1984: A comparison of intermonthly and interannual teleconnections in the 700 mb geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **112**, 2016–2032.

Faller, A. J., and C. E. Schemm, 1977: Statistical corrections to numerical prediction equations. *Mon. Wea. Rev.*, **105**, 37–56.

Gilchrist, A., 1977: An experiment in extended range prediction using

a general circulation model and including the influence of sea-surface temperature anomalies. *Beitr. Phys. Atmos.*, **50,** 25–40.

——, 1982: Aspects of the simulation of climate and climate variability in middle latitudes. *WMO/ISCU Study Conference on Physical Basis for Climate Prediction,* Leningrad World Climate Programme, WMO, 129–150.

Glahn, H. R., 1980: Methods and accuracy of statistical weather forecasting in the United States. *Collection of papers presented at the WMO Symp. on Probabilistic and Statistical Methods in Weather Forecasting,* Nice, 387–396.

——, and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11,** 1203–1211.

Hayashi, Y., 1986: Statistical interpretations of ensemble-time mean predictability. *J. Meteor. Soc. Japan,* **64,** 167–181.

Heckley, W. A., 1984: Systematic errors of the ECMWF forecasting system in tropical regions. ECMWF Tech. Rep. [Available at ECMWF, Shinfield Park, Reading, Berkshire, RG2 9AX, England.]

Hoffman, R. N., and E. Kalnay, 1983: Lagged-average forecasting, an alternative to Monte Carlo forecasting. *Tellus,* **35,** 100–118.

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

——, A. C. Lorenc, M. S. Tracton, K. Arpe, G. Cats, S. Uppala and P. Kallberg, 1985: The response of numerical weather prediction systems to FGGE level II-b data. Part I: Analyses. *Quart. J. Roy. Meteor. Soc.*, **111,** 1–66.

Hughes, F. D., 1982: An objective method to modify NMC/NWP model mean sea level progs. NMC Office Note No. 259. [Available at NMC, World Weather Building, 5200 Auth Road, Washington, DC.]

Kanamitsu, M., 1985: A study of the predictability of the ECMWF operational forecast model in the tropics. *J. Meteor. Soc. Japan,* **63,** 779–804.

Karl, T. R., R. E. Livezey and E. S. Epstein, 1984: Recent unusual mean winter temperatures across the contiguous United States. *Bull. Amer. Meteor. Soc.*, **65,** 1302–1309.

Lau, N. C., 1984: A comparison of circulation statistics based on FGGE level III-b analyses produced by GFDL and ECMWF for the special observing periods. NOAA Data Rep. ERL GFDL-6237 pp. [Available at Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, NJ.]

Leith, C. E., 1973: The standard error of time-averaged estimates of climatic means. *J. Appl. Meteor.*, **12,** 1066–1069.

——, 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102,** 409–418.

Lorenz, E. N., 1977: An experiment in nonlinear statistical weather forecasting. *Mon. Wea. Rev.*, **105,** 590–602.

——, 1982: Atmospheric predictability experiments with a large numerical model. *Tellus,* **34,** 505–513.

Lvovitch, M. I., and S. P. Ovtchinnikov, 1964: *Physical-Geographical Atlas of the World.* Academy of Science, USSR, and Dept. of Geodesy and Cartography, State Geodetic Commission, Moscow, USSR.

Mahlman, J. D., and L. J. Umscheid, 1985: Current results from the GFDL SKYHI model: Effects of resolution. *Proc. Fifth Conf. on the Meteorology of the Stratosphere and Mesosphere,* Boulder, Amer. Meteor. Soc., 6-6.

Manabe, S., D. G. Hahn and J. L. Holloway, 1974: The seasonal variation of the tropical circulation as simulated by a global model of the atmosphere. *J. Atmos. Sci.*, **31,** 43–83.

Mansfield, D. A., and T. N. Palmer, 1983: Predictability of the quasi-stationary flow in the atmosphere, and its dependence on anomalous boundary forcing. *IAMAP-WMO Symp. on Maintenance of the Quasi-stationary Components of the flow in the Atmosphere and in Atmospheric Models,* Paris, WMO 209–212.

Mellor, G. L., and T. Yamada, 1974: A hierarchy of turbulence-closure models for planetary boundary layer. *J. Atmos. Sci.*, **31,** 1791–1806.

Miyakoda, K., and J. Sirutis, 1977: Comparative integrations of global models with various parameterized processes of subgrid-scale vertical transports: Description of the parameterizations. *Beitr. Phys. Atmos.*, **50,** 445–587.

——, and R. F. Strickler, 1981: Cumulative results of extended forecast experiment. Part III: Precipitation. *Mon. Wea. Rev.*, **109,** 830–842.

——, and J. P. Chao, 1982: Essay on dynamical long-range forecasts of atmospheric circulation. *J. Meteor. Soc. Japan,* **601,** 292–307.

——, and J. Sirutis, 1984: Impact of subgrid-scale parameterizations on monthly forecasts. *Workshop on Cumulus Convection. Reading, European Centre for Medium Range Weather Forecasts.* 231–277. [Available at ECMWF, Shinfield Park, Reading, Berkshire, RG2 9AX, England.]

——, and ——, 1985: Extended range forecasting. *Advances in Geophysics,* Vol. 28, Part B, S. Manabe, Ed., Academic Press, 55–85.

——, G. D. Hembree, R. F. Strickler and I. Shulman, 1972: Cumulative results of extended forecast experiments. I: Model performance for winter cases. *Mon. Wea. Rev.*, **100,** 836–855.

——, T. Gordon, R. Caverly, W. Stern, J. Sirutis and W. Bourke, 1983: Simulation of a blocking event in January 1977. *Mon. Wea. Rev.*, **111,** 846–869.

Molteni, F., U. Cubasch and S. Tibaldi, 1985: Experimental monthly forecasts at ECMWF using the lagged-average forecasting technique. Long-range Forecasting Research Reports Series No. 6, CAS/WMO. *Proc. of the First Workshop on the Diagnosis and Prediction of Monthly and Seasonal Atmospheric Variations over the Globe.* College Park, MD, 598–607.

Namias, J., 1982: Case studies of long period air–sea interaction relating to long-range forecasting. *WMO/ISCU Study Conference on Physical Basis for Climate Prediction.* Leningrad World Climate Programme, WMO, 293–325.

Oort, A. H., 1983: Global atmospheric circulation statistics, 1958–1973. NOAA Prof. Paper 14. U.S. Govt. Printing Office, 180 pp.

——, and E. M. Rasmusson, 1971: Atmospheric circulation statistics. NOAA Prof. Paper 5. U.S. Govt. Printing Office, 323 pp.

Ploshay, J., R. White and K. Miyakoda, 1983: FGGE III-b daily global analyses—Part I. NOAA Data Rep., ERL GFDL-1.

Preisendorfer, R. W., and T. P. Barnett, 1983: Numerical model–reality intercomparison tests using small-sample statistics. *J. Atmos. Sci.*, **40,** 1884–1896.

Puri, K., and W. Stern, 1984: Investigations to reduce noise and improve data acceptance in the GFDL 4-dimensional analysis system. *ECMWF Seminar Proc.—Data Assimilation Systems and Observing System Experiment with Particular Emphasis on FGGE,* 157–190.

Quiroz, R. S., 1983: The climate of the "El Niño" winter of 1982–83: A season of extraordinary climate anomalies. *Mon. Wea. Rev.*, **111,** 1685–1706.

Ramanathan, V., E. J. Pitcher, R. C. Malone and M. L. Blackmon, 1983: The response of a spectral general circulation model to refinements in radiative processes. *J. Atmos. Sci.*, **40,** 605–630.

Shukla, J., 1981: Dynamical predictability of monthly means. *J. Atmos. Sci.*, **38,** 2547–2572.

——, 1983: On physical basis and feasibility of monthly and seasonal prediction with a large GCM. *Long-range Forecasting Research Publications Series,* No. 11, WMO, 142–153.

——, and D. S. Gutzler, 1983: Interannual variability and predictability of 500 mb geopotential heights over the Northern Hemisphere. *Mon. Wea. Rev.*, **111,** 1273–1279.

——, and K. C. Mo, 1983: Seasonal and geographical variation of blocking. *Mon. Wea. Rev.*, **111,** 388–402.

Smagorinsky, J., 1969: Problems and promises of deterministic extended range forecasting. *Bull. Amer. Meteor. Soc.*, **50,** 285–311.

Spar, J., J. J. Notario and W. J. Quirk, 1978: An initial state perturbation experiment with the GISS model. *Mon. Wea. Rev.*, **106,** 89–100.

Stern, W. R., J. J. Ploshay and K. Miyakoda, 1984: Continuous data assimilation at GFDL during FGGE. *ECMWF Seminar Proc.—Data Assimilation Systems and Observing System Experiment with Particular Emphasis on FGGE,* 125–156.

Wagner, A. J., 1977: Weather and circulation of January 1977. The coldest month on record in the Ohio Valley. *Mon. Wea. Rev.*, **105,** 553–560.

——, 1978: Weather and circulation of January 1978. Cold with record snowfall in the midwest and northeast, mild and wet in the west. *Mon. Wea. Rev.*, **106,** 579–585.

——, 1979: Weather and circulation of January 1979. Widespread record cold with heavy snowfall in the midwest. *Mon. Wea. Rev.*, **107,** 489–506.

——, 1980: Weather and circulation of January 1980: Commencement of a major index cycle. *Mon. Wea. Rev.*, **108,** 531–538.

——, 1981: Weather and circulation of January 1981: Record warmth in the west, record cold in the southeast and widespread severe drought. *Mon. Wea. Rev.*, **109,** 920–928.

——, 1982: Weather and circulation of January 1982: A stormy month with two record cold waves. *Mon. Wea. Rev.*, **110,** 310–319.

Wallace, J. M., and D. S. Gutzler, 1980: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109,** 784–812.

——, S. Tibaldi and A. J. Simmons, 1983: Reduction of systematic forecast errors in the ECMWF model through the introduction of an envelope orography. *Quart. J. Roy. Meteor. Soc.*, **109,** 683–717.