Dynamical Predictability of Monthly Means

J. SHUKLA

Goddard Laboratory for Atmospheric Sciences, NASA/GSFC, Greenbelt, MD 20771

(Manuscript received 14 January 1981, in final form 10 July 1981)

ABSTRACT

We have attempted to determine the theoretical upper limit of dynamical predictability of monthly means for prescribed nonfluctuating external forcings. We have extended the concept of "classical" predictability, which primarily refers to the lack of predictability due mainly to the instabilities of synoptic-scale disturbances, to the predictability of time averages, which are determined by the predictability of low-frequency planetary waves. We have carried out 60-day integrations of a global general circulation model with nine different initial conditions but identical boundary conditions of sea surface temperature, snow, sea ice and soil moisture. Three of these initial conditions are the observed atmospheric conditions on 1 January of 1975, 1976 and 1977. The other six initial conditions are obtained by superimposing over the observed initial conditions a random perturbation comparable to the errors of observation. The root-mean-square (rms) error of random perturbations at all the grid points and all the model levels is 3 m s^{-1} in u and v components of wind. The rms vector wind error between the observed initial conditions is >15 m s^{-1}.

It is hypothesized that for a given averaging period, if the rms error among the time averages predicted from largely different initial conditions becomes comparable to the rms error among the time averages predicted from randomly perturbed initial conditions, the time averages are dynamically unpredictable. We have carried out the analysis of variance to compare the variability, among the three groups, due to largely different initial conditions, and within each group due to random perturbations.

It is found that the variances among the first 30-day means, predicted from largely different initial conditions, are significantly different from the variances due to random perturbations in the initial conditions, whereas the variances among 30-day means for days 31–60 are not distinguishable from the variances due to random initial perturbations. The 30-day means for days 16–46 over certain areas are also significantly different from the variances due to random perturbations.

These results suggest that the evolution of long waves remains sufficiently predictable at least up to one month, and possibly up to 45 days, so that the combined effects of their own nonpredictability and their depredictabilization by synoptic-scale instabilities is not large enough to degrade the dynamical prediction of monthly means. The Northern Hemisphere appears to be more predictable than the Southern Hemisphere.

It is noteworthy that the lack of predictability for the second month is *not* because the model simulations relax to the same model state but because of very large departures in the simulated model states. This suggests that, with improvements in model resolution and physical parameterizations, there is potential for extending the predictability of time averages even beyond one month.

Here, we have examined only the dynamical predictability, because the boundary conditions are identical in all the integrations. Based on these results, and the possibility of additional predictability due to the influence of persistent anomalies of sea surface temperature, sea ice, snow and soil moisture, it is suggested that there is sufficient physical basis to undertake a systematic program to establish the feasibility of predicting monthly means by numerical integrations of realistic dynamical models.

1. Introduction

The deterministic prediction of subsequent evolution of atmospheric states is limited to a few days due to the presence of dynamical instabilities and nonlinear interactions. Two predictions made from the same initial conditions, except for small and random differences in the initial state, begin to differ from each other (Lorenz, 1965; Charney *et al.*, 1966; Smagorinsky, 1969). The rate and degree to which the two predictions diverge depends on the growth rates of the hydrodynamic instabilities, the nature of nonlinear interactions, and the structure of the differences between the two initial states. Since the observed state of the atmosphere, due to errors in observations and their interpolation to data-void areas, always contains some uncertainties, and since the formulation of the dynamical equations and parameterizations of the physical processes are only approximate, there is an upper limit on the range of deterministic prediction. This upper limit is mainly determined by the error growth rates associated with the instabilities of the mean flows with respect to the synoptic-scale disturbances and therefore it

strongly depends on the structure of the initial conditions. Some initial conditions are more predictable than others. This upper limit may be referred to as the predictability limit for synoptic scales. Moreover, since the limits of predictability of synoptic scales are considered under fixed external forcing [including slowly varying sea surface temperature (SST), soil moisture, snow and ice, etc.], and since the evolution of the two initial states is determined completely by dynamical instabilities and their interactions, we propose to refer to it as the *dynamical predictability of synoptic scales*. In subsequent discussions we have distinguished between the *dynamical predictability* and *predictability due to external forcings*.

The dynamical predictability of synoptic scales is of interest for short-range prediction and considerable literature already exists on the different aspects of this problem. In this study we propose to investigate the predictability of space and time averages. We ask the following questions: While the detailed structure of instantaneous flow patterns cannot be predicted beyond a few days, is it possible that space and time averages can be predicted over longer averaging periods?

In an earlier paper, Charney (1960) had raised the question of predictability of space and time averages. It is perhaps appropriate to quote Charney:

> "The crucial question is 'What is really remembered?' If, for example, the system only remembers a spatially or temporally averaged mean state, then it should at least be possible to predict this mean state. It seems to me that this is just the problem of long-range prediction."

The degree of dynamical predictability of space–time averages will naturally depend upon the space–time spectra of the atmospheric states and the nature of interaction among different scales. For example, if there were no stationary forcings at the earth's surface and if the day-to-day fluctuations were determined solely by the baroclinic instability of radiatively maintained zonal flows, the variability of time averages would mainly depend on the length of the averaging period and, therefore, the time averages will not be any more predictable than the amplitudes and phases of individual disturbances. In this case, most of the energy would be contained in the most unstable scales and larger scales will grow either through the cascade of energy from the fast growing unstable scales or due to their self amplification.

Recent studies by Charney and Devore (1979), Charney and Straus (1980) and Charney *et el.* (1981) have shown that for a given external forcing, thermally and orographically forced circulations can possess multiple equilibrium states and that some of these states are more stable than others. This suggests that in the presence of thermal and orographic forcings, self interaction and nonlinearities can be important mechanisms for fluctuations of longer periods. It is quite likely that large-amplitude synoptic instabilities play an important role in destabilizing the existing equilibria and thus the predictability of transitions from one equilibrium to another may not be any more promising than the predictability of intense cyclone waves.

Fig. 1a shows the space-time spectra for the observed geopotential height field at 500 mb for 15 winter seasons. It is seen that most of the variance is contained in the planetary-scale (wavenumbers 1–4) low-frequency (10–90 day) components of the circulation. Fig. 1b shows the variances in different wavenumber and frequency domains for 15 different winter seasons. Most of the interannual variability is contributed by the low-frequency planetary waves. Since the variability of space–time averages is mainly dominated by the planetary-scale low-frequency components, it can be anticipated that the prospects of predictability of space–time averages may not be as hopeless as those of the synoptic scales.

In the earlier classical studies of deterministic predictability, the theoretical upper limit was determined by the rate of growth and the magnitude of error between the two model evolutions for which the initial conditions differed by only a small random perturbation. For a study of the predictability of time averages, a more appropriate question would be: If numerical predictions are made from initial conditions which are as different as two randomly chosen years (for example, as different as observations on the same calendar date for different years), how long would it take before the time-averaged predictions become indistinguishable from the predictions made with small random perturbations in the initial conditions? The differences between two atmospheric states for the same calendar day in two different years can be very large because each dynamical state has evolved through complex nonlinear interactions under the influence of different boundary forcings, and therefore the amplitudes and phases of the main energy-bearing planetary waves will be much more different than those which can be expected by adding a small random perturbation. Thus it is natural to ask: Is it possible that a given configuration of planetary waves remembers itself much longer than the limit of deterministic prediction, which mainly refers to the predictability of synoptic scales?

The main difference between the previous studies and this study is that the earlier studies examined the growth rate of the errors, and the upper limit of synoptic-scale predictability was determined by the

dominant instabilities. In the present study, we have carried out the analysis of variance to compare the variability among very different initial conditions and among randomly perturbed initial conditions. The limit of predictability is not determined solely by the growth rates of synoptic-scale instabilities, but by the relative magnitudes of the deterministic growth rates of the planetary waves and the degradation of planetary waves by the rapidly amplifying synoptic-scale instabilities. In other words, the lack of predictability of a high-frequency small-scale system itself is not very significant because its effect will be minimized due to averaging; however, its effects on the longer waves are important. The earlier studies of predictability considered the rate of error growth under fixed boundary conditions, i.e., they examined only the dynamical predictability of the initial conditions. There is the possibility that the predictability can be different for different boundary forcings. For example, time averages may be more predictable for certain anomalous structures of sea surface temperature, soil moisture, sea ice and snow.

The present study and the earlier studies share a common deficiency: both introduce only a random perturbation in the initial conditions. There is no evidence that the errors of observations are randomly distributed. Observational errors have systematic biases over the data-void areas (*viz.*, oceans) and it is not clear that the growth rates of systematic errors will be the same as the growth rates of random errors. However, just as the earlier predictability studies with random perturbations were intended mainly to illustrate the theoretical upper limit of synoptic-scale predictability, the present study also is an idealized study to establish the theoretical upper limit of predictability of time averages.

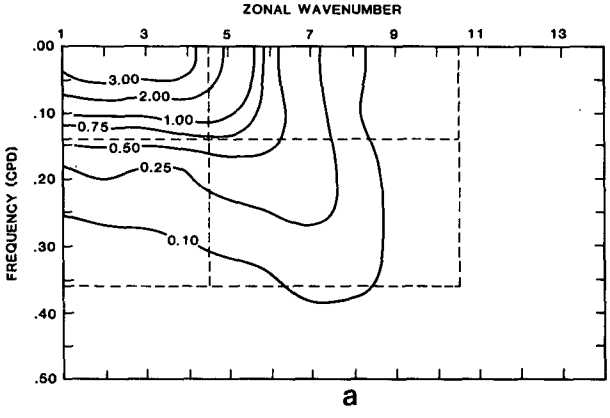


FIG. 1a. Wavenumber–frequency decomposition of transient variances of 500 mb geopotential height field along 50°N for 14 winter seasons.

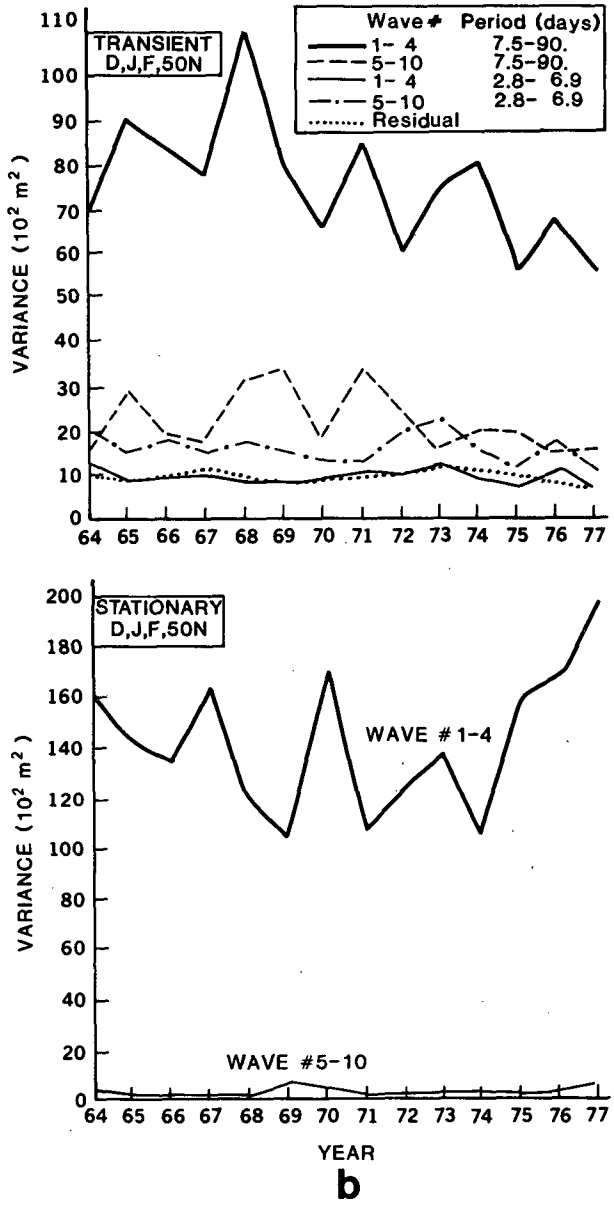


FIG. 1b. Interannual variability of transient variances (upper panel) and stationary variances (lower panel) of 500 mb geopotential height field along 50°N.

2. Mechanisms for the interannual variability of monthly means

Let us assume that $\Psi(\mathbf{X}, t)$ represents the state of the atmosphere at any time t (schematic Fig. 2), where $\mathbf{X}$ is the three-dimensional space vector.

At the earth's surface, which is the lower boundary of the atmosphere, sea surface temperature, sea ice and snow, soil moisture and vegetation, etc., act like slowly varying external forcings. Although solar variability is the only forcing truly external to the

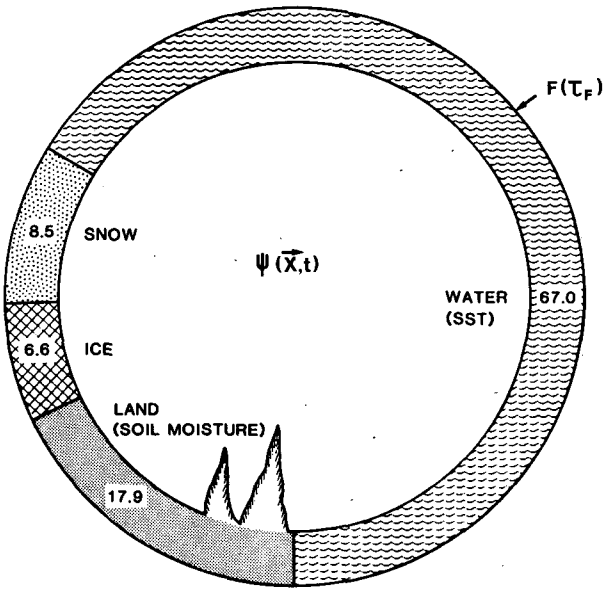


FIG. 2. Schematic illustration of the roles of internal dynamics and slowly varying boundary conditions.

atmosphere, since the time scale of the change of such parameters as large-scale sea surface temperature, soil moisture and sea ice cover, etc., is much larger than the time scale of synoptic-scale instabilities, we would refer to these slowly varying boundary conditions as external forcings to the atmospheric system described by $\Psi(\mathbf{X}, t)$. The time evolution of Ψ can now be treated as the combined effects of the initial conditions (because they will determine the nature of instabilities and their interactions) and the effects of external forcing as defined by the boundary conditions. There are three distinct time scales which enter any consideration of the predictability of time averages. First is the time scale of day-to-day fluctuations of Ψ(to be denoted by τ_Ψ) due to inherent instability and nonlinearity of the system, and this determines the level of uncertainty in estimating a time average (Leith, 1973), second is the time scale of external forcing (to be denoted by τ_F); and third is the averaging time (T) for which we wish to examine the predictability. For the particular case of predictability of monthly means, it can be stated that $\tau_\Psi < T < \tau_F$, because τ_Ψ for the atmosphere is a few days and τ_F for large-scale sea surface temperature and sea ice anomalies is a few months (we do not have a good observational basis to estimate τ_F for soil moisture). Since these boundary conditions change slowly, they can be assumed to be constant (or prescribed) for a period of one month. However, even under fixed external forcings, internal dynamics can change the evolution of initial states to such an extent that time averages will be different from each other. The mechanisms responsible for the interannual variability of monthly means can be broadly categorized as follows:

(i) *Internal dynamics* Due to combined effects of instabilities, nonlinear interactions, thermal and orographic forcings and fluctuating zonal winds, tropical–extratropical interactions, etc. (Since orography and land–sea contrast at the earth's surface is fixed, thermally and topographically forced motions and their interactions are considered as a part of the internal dynamics.)

(ii) *Boundary forcings* Due to fluctuations of sea surface temperature, sea ice/snow, soil moisture and other slowly varying boundary conditions and their effects on the amplitudes and phases of planetary waves which in turn may determine the tracks and intensity of cyclone-scale disturbances. Fluctuations in solar or other extraterrestrial energy sources are not considered in the present study.

An understanding of the relative contributions of the two factors is essential to determine the predictability of climate in general and of monthly means in particular. We recognize that the dynamical properties of the small-scale instabilities can be very different under extremely different boundary conditions. However, the observed anomalies in the boundary conditions are not found to be so large that they could change the basic dynamics of the synoptic disturbances. The changes in the boundary conditions can affect the amplitudes and phases of the planetary waves in extratropical latitudes and large-scale Hadley and Walker circulations in low latitudes. Changes in the structure and persistence of the midlatitude planetary waves, in turn, can change the frequency, intensity and propagation properties of synoptic-scale instabilities. Similarly, in low latitudes, SST and soil moisture anomalies of moderate magnitude can change the growth and amplitude of synoptic-scale tropical disturbances, whereas anomalies of large magnitude can drastically alter the geographical locations of convergence zones and precipitation. Anomalous heat sources in low latitudes can also produce, under favorable conditions, large responses in middle latitudes. In this paper, however, we propose to study the predictability of monthly means under fixed boundary conditions. This we would refer to as the *dynamical predictability* of monthly means. Examination of the role of slowly varying boundary conditions in determining the changes of monthly means is a topic for separate investigation.

General circulation models allow us to examine the dynamical predictability and boundary-forced predictability separately. It is not possible to address these questions by analysis of observed data alone because the observations reflect the combined effects of internal dynamics and fluctuating boundary

conditions. Even in the absence of fluctuating external forcings, internal dynamics can generate interannual variations of monthly and seasonal means. Sometimes, this component of the variability has been referred to as the "climate noise" (Leith, 1975; Straus and Halem, 1981). This terminology can be misleading if the word "noise" refers implicitly to those components of the circulation which are considered to be unpredictable. Since it has not yet been established that the changes due to internal dynamics are unpredictable, it is not appropriate to refer to the variability due to internal dynamics as climate noise. We propose to distinguish between the fast growing synoptic instabilities and slowly varying planetary scales which possess different predictability characteristics. Lack of synoptic-scale predictability beyond two weeks is not sufficient to assume that certain statistical properties (e.g., monthly means) of the internal dynamics are not predictable. It is quite likely that some components of the atmospheric flows may retain a certain degree of nonlinear memory for a period beyond the limits of deterministic prediction for synoptic scales and therefore a nonlinear dynamical prediction scheme could predict the evolution of large-scale components over a period longer than that for synoptic scales.

3. Numerical experiments

We have carried out 60-day integrations of the GLAS climate model starting from three different initial conditions corresponding to 1 January of 1975, 1976 and 1977. These are referred to as the control runs. The differences among the initial conditions for 1 January of different years are larger than those due to errors of observations because they reflect the multitude of effects of varying boundary conditions and dynamical interactions during each preceding year. Amplitudes and phases of planetary waves were significantly different from each other. The rms vector wind differences at 500 mb for observed initial conditions over Northern Hemisphere were 16.9 m s^{-1} between 1975 and 1976, 17.6 m s^{-1} between 1976 and 1977, and 18.5 m s^{-1} between 1975 and 1977; the differences were $\sim$10 m s^{-1} in the lowest tropospheric levels and $\sim$20 m s^{-1} in the upper troposphere. The rms differences for the initial conditions of 500 mb geopotential height between 30 and 70°N were 160.4 m between 1975 and 1976, 187.7 m between 1976 and 1977, and 186.7 m between 1975 and 1977. In comparison, the rms difference between maps from two randomly chosen years, calculated from 15 years of daily values, was found to be 178.8 m.

Each of these initial conditions was then randomly perturbed such that the spatial structure of the random perturbations in u and v components at all the nine levels of the model had a Gaussian distribution with zero mean and standard deviation of 3 m s^{-1}. The amplitudes of the random perturbation in u and v were not allowed to exceed 12 m s^{-1} at any grid point. These are referred to as the perturbation runs. For each control run there are several perturbations runs. The general circulation model used in the present study has been described by Halem *et al.* (1980).

TABLE 1. Schematic summary of the control and the perturbation runs (C_{ji}).

	Initial condition	1 Jan 1975 $j = 1$	1 Jan 1976 $j = 2$	1 Jan 1977 $j = 3$
Control	$i = 1$	C_{11}	C_{21}	C_{31}
Random perturbation	$i = 2$	C_{12}	C_{22}	C_{32}
Random perturbation	$i = 3$	C_{13}	C_{23}	
Random perturbation	$i = 4$		C_{24}	

Since all the integrations are made with fixed boundary conditions, variability among the predictions of monthly means made from the control runs gives a measure of the long-range memory of the initial conditions characterized by different configurations of the planetary waves. Variability among different perturbation runs for the same control run gives a measure of the degradation of the predictability of those initial conditions. If for a given averaging period T, the variability among the control runs is not statistically different from the variability among the perturbation runs, we can conclude that there is no dynamical predictability for the time averages over period T. It should be pointed out, however, that this statement applies only to the model being used for the study and not necessarily to the real atmosphere. However, since ultimately one or the other model has to be used for any prediction, lack of predictability for the models will imply our inability to make actual predictions of the atmosphere. Absence of *dynamical predictability* would not imply, however, the absence of boundary-forced predictability due to anomalous structures of sea surface temperature, soil moisture and sea ice, etc.

Table 1 gives a summary of the numerical integrations carried out for this study. C_{11}, C_{21}, and C_{31} refer to the three 60-day control runs made from the initial conditions of 1 January 1975, 1976 and 1977, respectively. C_{12}, C_{13}, are the perturbation runs for which the initial conditions of C_{11} were randomly perturbed. Likewise, C_{22}, C_{23} and C_{24} are the perturbation runs for the control run C_{21}, and C_{32} is the perturbation run for the control run C_{31}. The statistical properties of the random perturbations were the same in all the perturbation cases; however, the actual values at any grid point were different in each case.

4. Predictability of planetary and synoptic scales

We have examined the scale dependence of the theoretical upper limit of deterministic predictability. The errors for planetary waves and synoptic-scale waves are very different. We show that the planetary waves which have the largest contribution to monthly means have a much longer predictability time compared to the synoptic-scale waves.

The 500 mb geopotential height (ϕ) for each day and each latitude for control (c) and perturbation (p) runs can be expressed as

$$\phi^c(\lambda, t) = \phi_0^c + \sum A_k^c \cos\theta_k + B_k^c \sin\theta_k,$$

$$\phi^p(\lambda, t) = \phi_0^p + \sum A_k^p \cos\theta_k + B_k^p \sin\theta_k,$$

where λ is the longitude and θ_k the phase for the wavenumber k. The mean value of 72 grid points along the latitude circle is denoted by $k = 0$. The total mean-square error E along the latitude circle can be expressed as

$$\{E(t)\}^2 = \frac{1}{2\pi}\int_0^{2\pi} [\phi^c(\lambda, t) - \phi^p(\lambda, t)]^2 d\lambda$$

$$= (A_0^c - A_0^p)^2 + \frac{1}{2}\sum_{k=1}^{36} (A_k^c - A_k^p)^2 + (B_k^c - B_k^p)^2,$$

where A_k and B_k are functions of time.

The right-hand side gives the contribution of each wavenumber to the total rms error. We have examined the errors, as defined below, in the planetary scales (wavenumbers 0–4), synoptic scales (wavenumbers 5–12) and short scales (wavenumbers 13–36) separately:

Planetary-scale error =

$$[(A_0^c - A_0^p)^2 + \sum_{k=1}^{4} (A_k^c - A_k^p)^2 + (B_k^c - B_k^p)^2]^{1/2}$$

Synoptic-scale error =

$$[\sum_{k=5}^{12} (A_k^c - A_k^p)^2 + (B_k^c - B_k^p)^2]^{1/2}$$

Short-scale error =

$$[\sum_{k=13}^{36} (A_k^c - A_k^p)^2 + (B_k^c - B_k^p)^2]^{1/2}.$$

We have calculated the daily values of errors for the six pairs (see Table 1) of control and perturbation runs: (C_{11}, C_{12}), (C_{11}, C_{13}), (C_{21}, C_{22}), (C_{21}, C_{23}), (C_{21}, C_{24}), (C_{31}, C_{32}). Figs. 3a and 3b show the mean of the six error values for 500 mb geopotential height for wavenumbers 0–4 and 5–12 respectively, averaged over the latitude belt 40–60°N. The vertical bars on each curve give the standard deviation among six error values. Mean and standard deviations of the persistence error in the corresponding wavenumber range for the three control runs, C_{11}, C_{21}, C_{31} also are shown on each figure. Errors in wavenumbers 13–36 are very small and are not shown here.

The differences in the two figures are rather remarkable. If the crossover point between the ran-

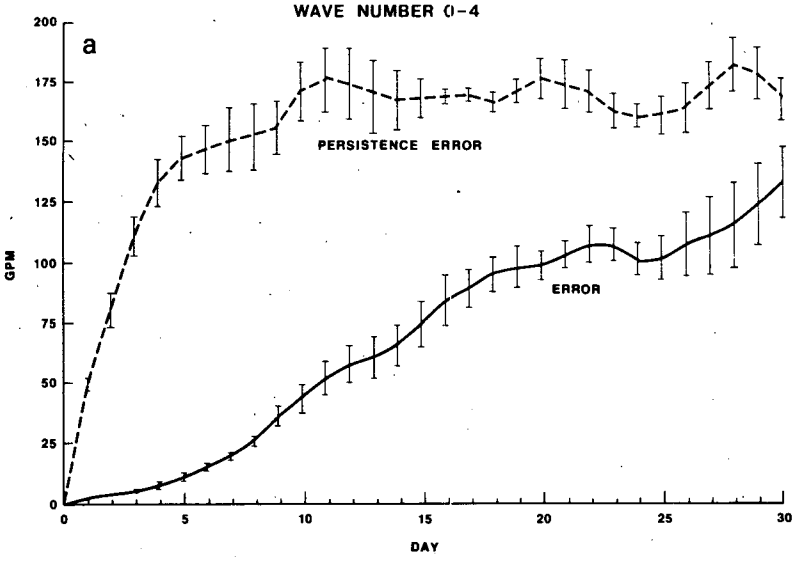


FIG. 3a. Root mean square error, averaged for six pairs of control and perturbation runs and averaged for latitude belt 40–60°N for 500 mb geopotential height (gpm), for wavenumbers 0–4. Dashed line is the persistence error averaged for the three control runs. Vertical bars denote the standard deviation of the error values.

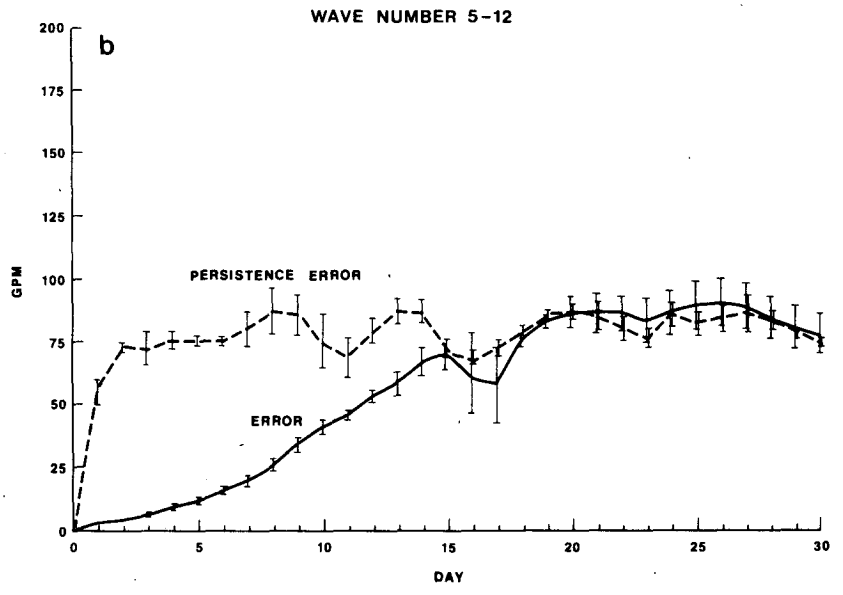


FIG. 3b. As in Fig. 3a except for wavenumbers 5–12.

dom error growth curve and the persistence error curve is considered to be the theoretical upper limit of the deterministic predictability, the synoptic scales are found to lose complete predictability after two weeks. The persistence error is a measure of error between two randomly chosen charts and is $\sqrt{2}$ times larger than the rms of daily fluctuations (climatology error). The planetary scales, on the other hand, seem to show theoretical predictability even beyond one month. This further suggests that the space and time averages, which are mainly determined by the planetary-scale motions, have a potential for predictability at least up to or beyond one month.

We also have examined the day-to-day changes of sea level pressure and 500 mb temperature for control run and perturbation runs. Figs. 4a–4f show the plots of daily sea level pressure and temperature averaged over some of the areas shown in Fig. 7. It can be seen that the initial conditions do not persist in the course of integration and even large spatial averages show large day-to-day fluctuations. During the first month, although the day-to-day fluctuations of either the control or the perturbation runs are large, the differences between the control run and the random perturbation run are not as large. However, for the second month, the departures between the control and the perturbation runs are so large that even the monthly means are indistinguishable from the monthly means for a completely different initial condition.

Relatively small differences between spatially averaged atmospheric variables for the first 20–25 days of control and perturbation runs would suggest that the time averages for the first month should be predictable. We have shown this to be the case by a more systematic analysis of variance of model simulations described in the following sections.

5. Results

Figs. 5a–5c show the plots of monthly mean sea level pressure for January (day 1–31) for the model runs (C_{11}, C_{12}), (C_{21}, C_{22}) and (C_{31}, C_{32}), respectively. The upper panel shows the monthly mean for the control run and the lower panel shows the monthly mean for the random perturbations over that control run. To reduce the number of figures and to emphasize the large changes in winter, we have only shown the Northern Hemisphere maps. For January, any control run is much more similar to its own perturbation runs than to any other control run or any other perturbation run. The upper and lower panels of Figs. 5a–5c have large similarities compared to any two upper panels or any two lower panels. For example, both the upper and the lower panels of Fig. 5c show a very deep Aleutian low which is neither as strong nor located at the same place in other maps for January. The same is true for other major circulation features. This suggests that the random perturbations in the initial conditions have not changed the 31-day evolution of the flow so drastically that the monthly means may look very different.

The results are very different for the month of February (days 32–60), which are shown in Figs. 6a–6c. Although some of the large-scale features retain their general configuration in the upper and lower panels, the displacements in the centers of highs and lows is large enough to give very signifi-

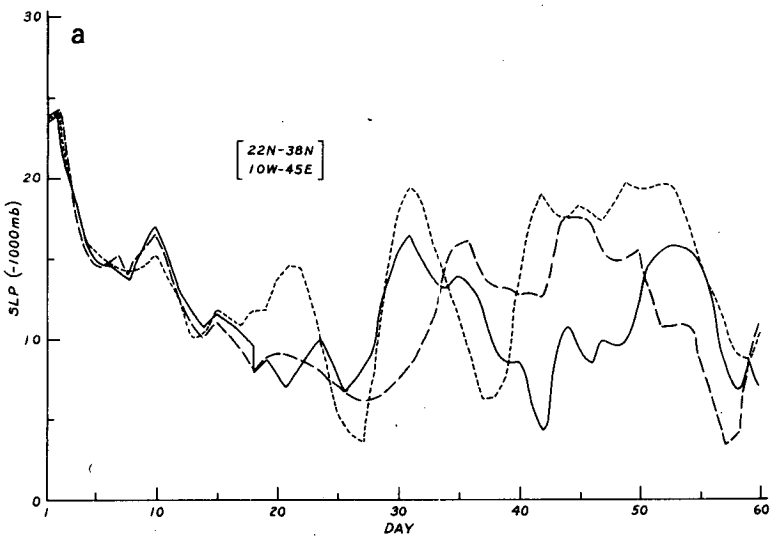


FIG. 4a. Daily values of sea level pressure (mb) averaged for grid points between 22 and 38°N, 10°W and 45°E, for control run (solid line) and two perturbation runs (dashed lines) for initial conditions of 1 January 1975.

cant quantitative differences between the upper and the lower panels. The differences between the upper and the lower panels are comparable to the differences between any two figures. In the following section, we present a more quantitative description of the differences between the control and the perturbation runs.

a. Space and time averaging

We have examined the predictability of monthly means for several space averaging domains. Let us first ask a simple question: Why do we examine monthly means rather than examining 23-day or 39-day means? We have not carried out, and we are not aware of, any study which examines the most appropriate averaging periods for time-averaged predictions. Naturally the answer would depend on the frequency spectra of the different components of the flow. Since the space and time scales of atmospheric motions are dynamically related, it also should depend on the appropriate combination of space and time averaging domains. Our choice of monthly time period is based simply on a qualita-

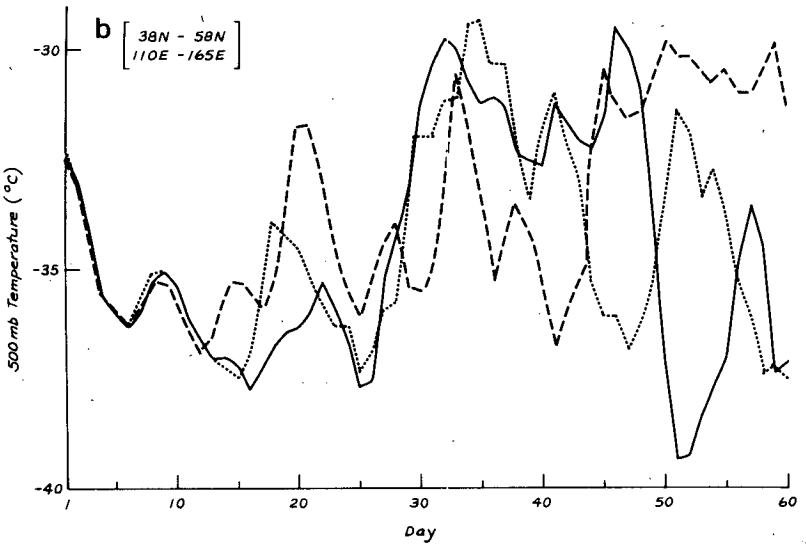


FIG. 4b. Daily values of 500 mb temperature (°C) averaged for grid points between 38 and 58°N, 110 and 165°E, for control run (solid line) and two perturbation runs (dashed lines) for initial conditions of 1 January 1976.

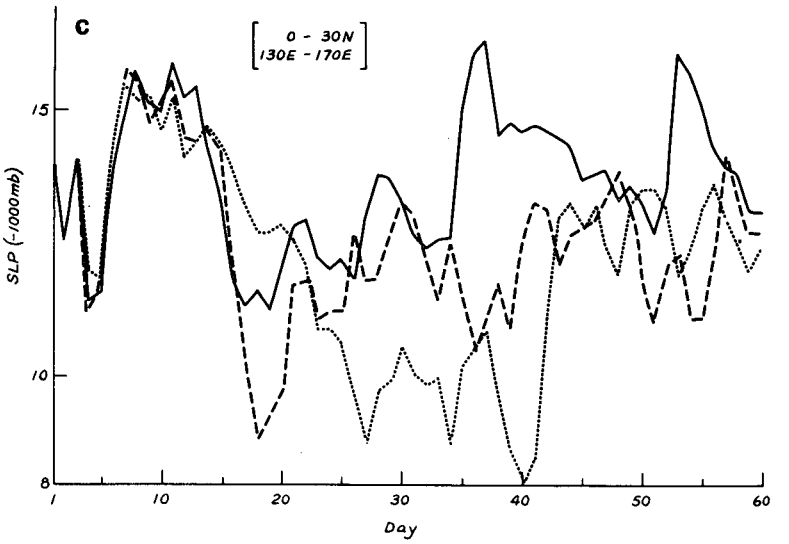


FIG. 4c. Daily values of sea level pressure (mb) averaged for grid points between 0 and 30°N, 130 and 170°E, for control run (solid line) and two perturbation runs (dashed lines) for initial condition of 1 January 1975.

tive reasoning that it is more than two weeks, which is the upper limit of deterministic prediction, and less than a season which is an appropriate time scale for the external solar forcing. Moreover, for social consumption, forecasts are normally given in terms of monthly and seasonal means.

Similarly, there is arbitrariness in choosing the space scales for averaging. From an examination of the past data over United States one can see that the space scales of the monthly mean anomalies for the large-scale dynamical variables (e.g., pressure, temperature, wind, etc.) is always larger than 500 × 500 km, which is the resolution of the model, but not as large as the whole of continental United States. In fact, typically, different quadrants of United States show different "signs" of monthly and seasonal anomaly field. It would, therefore, be meaningless to predict a monthly mean for the whole United States. In this study we have examined the monthly means averaged over different areas shown in Fig. 7. While choosing the averaging area we have tried to isolate the oceanic and continental areas, the monsoonal and nonmonsoonal areas and the areas of active midlatitude cyclogenesis.

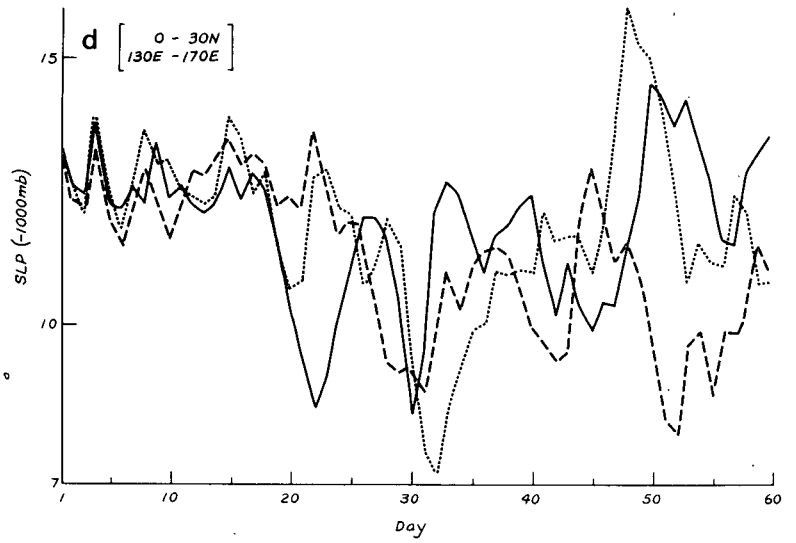


FIG. 4d. Daily values of sea level pressure (mb) averaged for grid points between 0 and 30°N, 130 and 170°E, for control run (solid line) and two perturbation runs (dashed lines) for initial conditions of 1 January 1976.

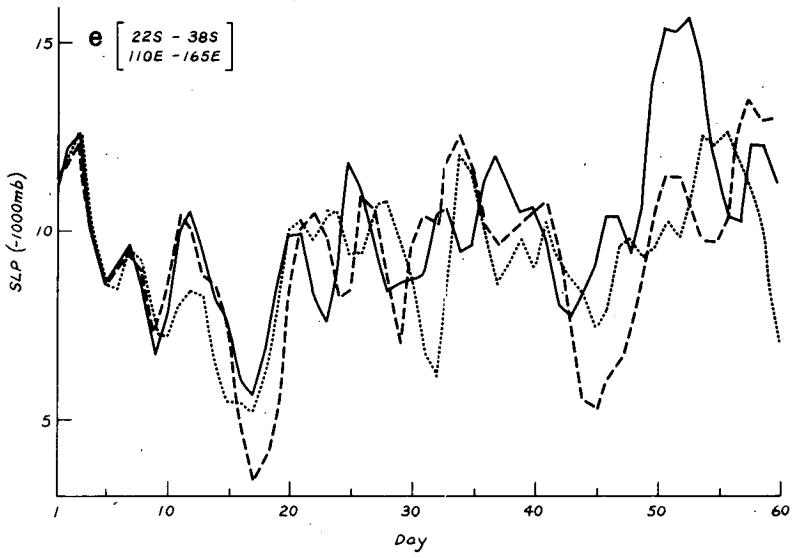


FIG. 4e. Daily values of sea level pressure (mb) averaged for grid points between 22 and 38°S and 110 and 165°E, for control run (solid line) and two perturbation runs (dashed lines) for initial conditions of 1 January 1975.

b. Analysis of variance

We have carried out the analysis of variance (F test) to determine the statistical significance of the differences in the variances among different control runs and among different perturbation runs. This technique is described in statistics text books (e.g., Hays, 1963).

In Table 1 the three boxes labeled as $j = 1, 2, 3$ refer to the three different initial conditions. For each box j, index i refers to the perturbed runs made with the initial condition for box j. If M_{ji} denotes the monthly mean value of any variable for the model run C_{ji}, and J is the number of boxes, the expression for F can be written as

$$F(\nu_1, \nu_2) = \frac{\sum_j \frac{(\sum_i M_{ji})^2}{n_j} - \frac{(\sum_j \sum_i M_{ji})^2}{N}}{\sum_j \sum_i M_{ji}^2 - \sum_j \frac{(\sum_i M_{ji})^2}{n_j}} \frac{(N-J)}{(J-1)} = \frac{F_1}{F_2} \frac{(N-J)}{(J-1)}.$$

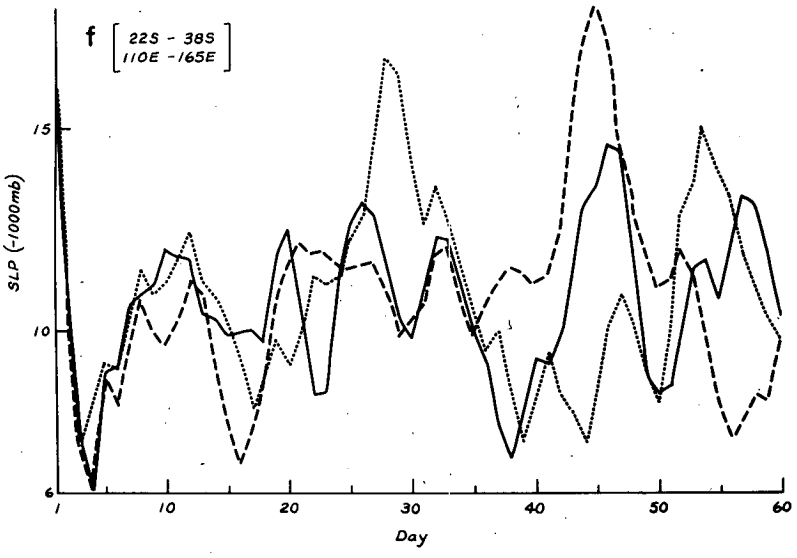


FIG. 4f. Daily values of sea level pressure (mb) averaged for grid points between 22 and 38°S and 110 and 165°E, for control run (solid line) and two perturbation runs (dashed lines) for initial conditions of 1 January 1976.

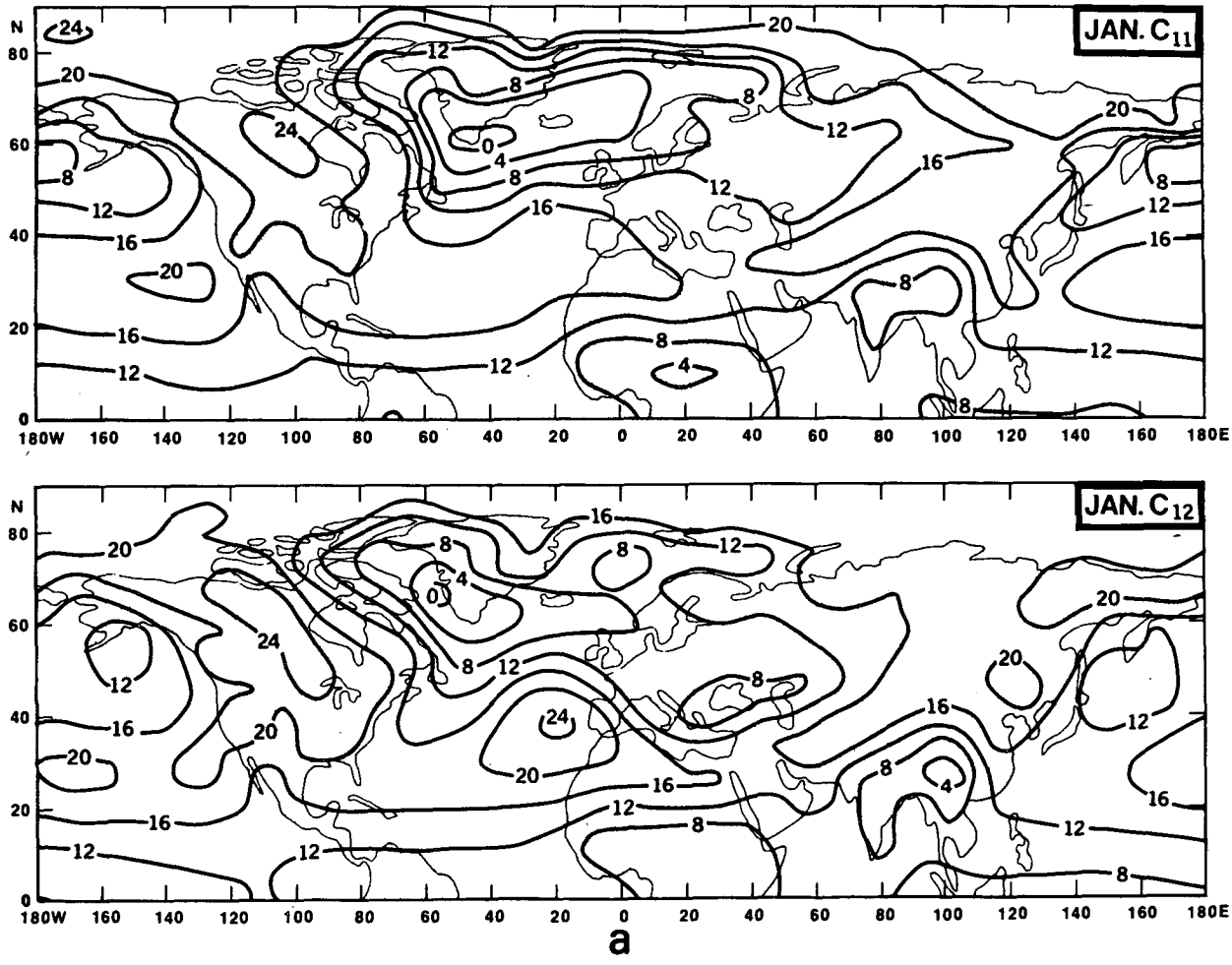


FIG. 5a. 31-day mean sea level pressure for days 1–31 for the control run from the initial conditions of 1 January 1975 (upper panel: JAN C_{11}) and its perturbation run (lower panel: JAN C_{12}).

In the present case (see Table 1) $n_1 = 3$, $n_2 = 4$, $n_3 = 2$, $N = n_1 + n_2 + n_3 = 9$, $\nu_1 = J - 1 = 2$, $\nu_2 = N - J = 6$.

The numerator (F_1) is a measure of variability among the monthly means of largely different initial conditions and the denominator (F_2) is a measure of variability among the monthly means for randomly perturbed initial conditions. By carrying out an analysis of variance, we are examining the significance of the fluctuations of the monthly means and not the monthly means themselves.

The algebraic mean of all the runs in any one box differs from the algebraic mean of all the runs in another box. The numerator F_1 is a measure of the variability among such box means, and the denominator F_2 is a measure of variability within the boxes. If the three different initial conditions are indeed very different, and if the random perturbations do not change them substantially, the numerator will be large and remain large, whereas the denominator will be small and remain small. However, if the random perturbations can produce large changes in the time averages, the denominator will increase and reduce the value of F. A reduction in the value of the numerator will also reduce the value of F. This will occur if the algebraic mean of all the runs in one box is not very different from the algebraic mean of all the runs in other boxes.

The numerical value of F determines the level of significance for the differences between the variability among the boxes and the variability within the boxes. From F tables, $F(\nu_1, \nu_2) = 5.1$ for 95% and 10.9 for 99% and 14.5 for 99.5% level of significance. Therefore, if the value of F, calculated from Eq. (1) for all the integrations in Table 1, is larger than 5.1, we can conclude that the variability among the monthly means of the control runs with very different initial conditions is larger (at the 95% significance level) than the variability among the monthly means due to random perturbations in the initial conditions.

For any initial condition, the day-to-day evolution

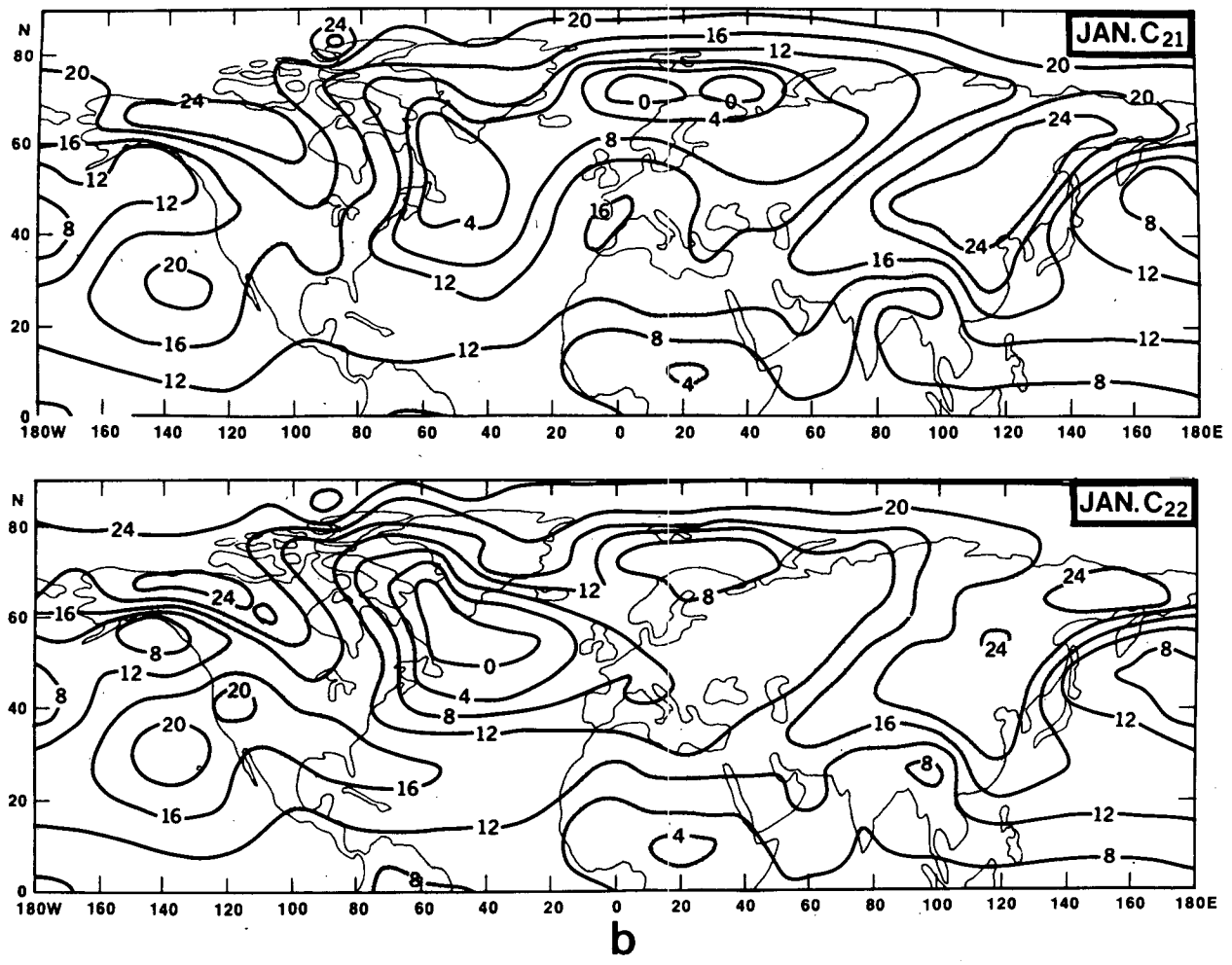


FIG. 5b. 31-day mean sea level pressure for days 1–31 for the control run from the initial conditions of 1 January 1976 (upper panel: JAN C_{21}) and its perturbation run (lower panel: JAN C_{22}).

of the flow will be different between the control run and its perturbation run. This will lead to different values of monthly means. The magnitude of this difference will depend, among other things, on the number of days the flow is allowed to evolve before calculating the monthly means. For example, the differences between the monthly means of the control run and the perturbation run for days 31–60 will be different from those for days 1–30. Similarly, the differences between the monthly means of two control runs will be different from the differences between the monthly means of a control run and its perturbation run. The purpose of this analysis is to introduce a quantitative measure to examine these differences and compare them with each other. If for days 31–60, the variability among the control runs is not significantly larger than the variability among the perturbation runs, it will suggest that during a 60-day integration of the global GCM, the evolution of the flow was so modified by the presence of random errors in the initial condition that even a 30-day mean (for the last 30 days) was indistinguishable from a similar 30-day mean for very different initial conditions. Since the boundary conditions are identical for all the integrations, this will imply that the different initial conditions of the control runs had no bearing on the time averages for days 31–60.

In this study we have carried out 60-day numerical integrations for all the nine cases. We have first examined the predictability of time averages for days 1–31, 16–46 and 32–60. For convenience we refer to these as January (J), January–February (J/F) and February (F), respectively.

It is recognized that the observed initial conditions on any day are not unrelated to the observed boundary conditions for the same day and therefore certain inconsistencies might occur through using climatological mean boundary conditions. Presumably, this inconsistency is common to all the model runs with different initial conditions. The

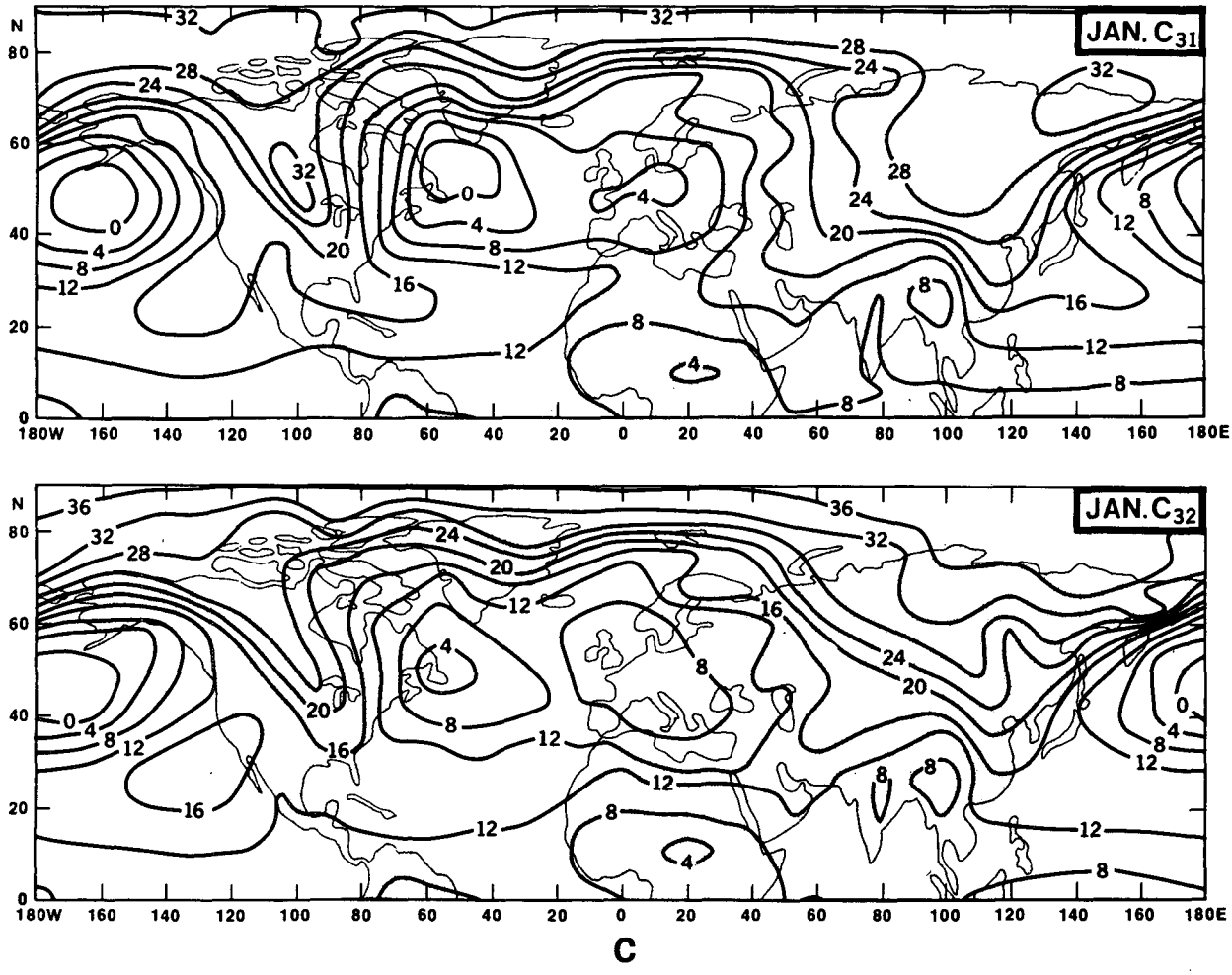


FIG. 5c. 31-day mean sea level pressure for days 1–31 for the control run from the initial conditions of 1 January 1977 (upper panel: JAN C_{31}) and its perturbation run (lower panel: JAN C_{32}).

error growth rates were very similar for all the perturbation runs. Since we propose to study the dynamical predictability under identical boundary conditions, we had no better alternative than choosing the climatological mean boundary conditions. A more appropriate procedure would have been to choose the control initial conditions from a long (several years) GCM integration carried out with constant (or seasonally varying) boundary conditions. Due to limitations on available computer time, we could not carry out such long integrations. However, such long integrations already have been carried out by other modeling groups, (S. Manabe and M. Schlesinger, personal communication) and it may be useful to carry out similar studies with these models.

c. Results of F calculation

Figs. 8a–8c give the F values for sea level pressure averaged over each area for the averaging periods of days 1–31, 16–46 and 32–60, respectively. Areas with F values of 5 or more are shaded. For the monthly means of days 1–31, 29 out of 38 areas show F values of 5 or more. The number of such significant areas drops down to 11 for monthly means of days 16–46 and drops further down to 5 for monthly means of days 32–60. Since two of these five areas did not show significance for averaging periods of days 16–46, it is reasonable to conclude that only three of the areas show significance for the averaging period of days 32–60. Since we are considering the significance level above 95%, two of 38 areas may be significant by chance. It can be concluded that there is a complete loss of dynamical predictability for averaging periods of days 32–60.

If the number of significant areas are segregated according to the hemisphere, it is seen that only three out of 18 areas in the Southern Hemisphere are significant above 95% for the averaging period of days 16–46. This suggests that the Northern Hemi-

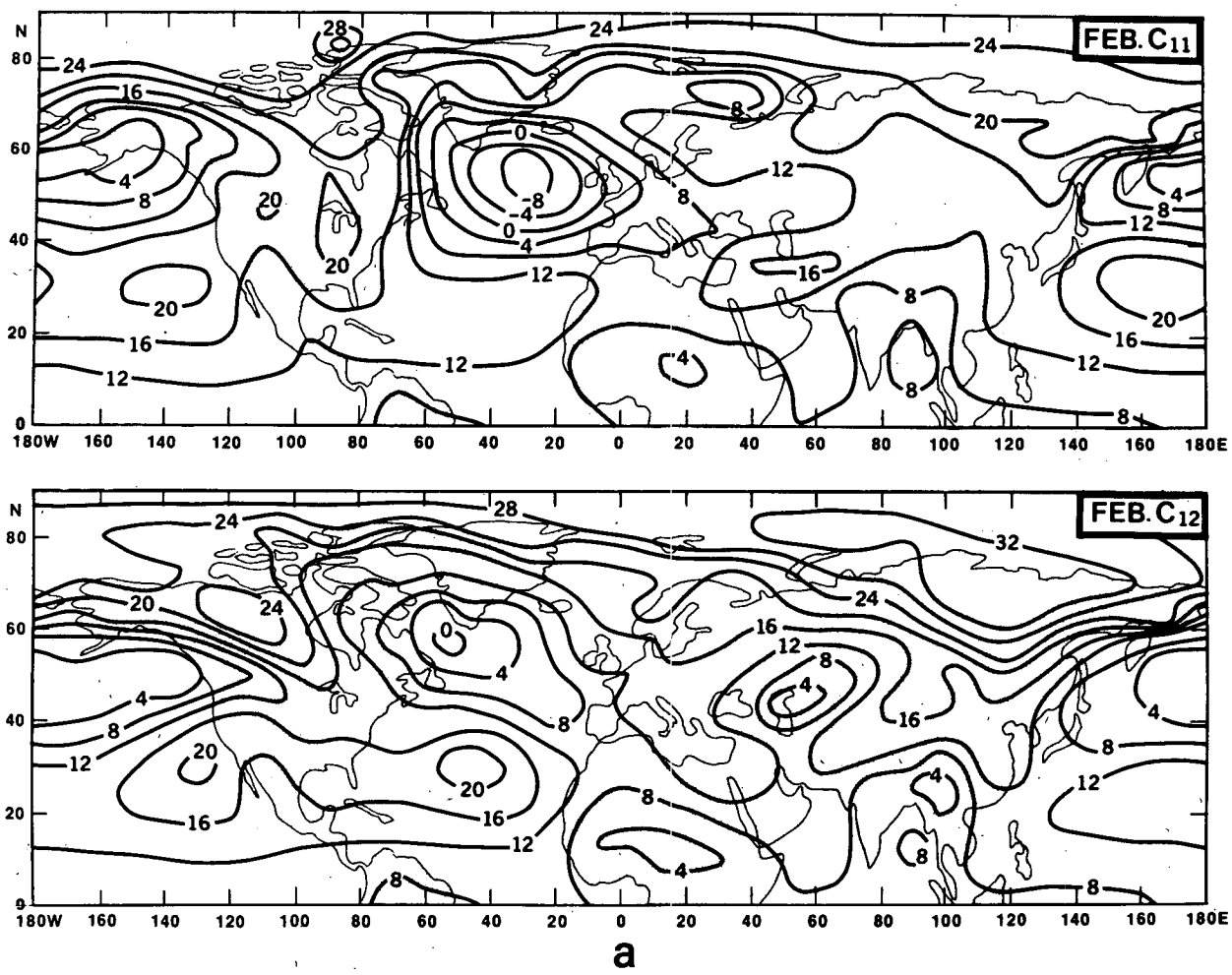


FIG. 6a. 29-day mean sea level pressure for days 32–60 for the control run from the initial conditions of 1 January 1975 (upper panel: FEB C_{11}) and its perturbation run (lower panel: FEB C_{12}).

sphere monthly means are potentially more predictable than the Southern Hemispheric monthly means. It is quite likely that the presence of orographic and thermal forcings due to mountains and continental–oceanic heat sources in the Northern Hemisphere winter establishes planetary-scale motions which are sufficiently stable to allow a dynamical prediction for a longer range compared to the Southern Hemisphere. It may be conjectured likewise that the prospects of longer term predictability for the Northern Hemispheric summer may not be as favorable.

We have examined the sensitivity of these results to the size of the averaging area. We have repeated the calculations of F for areas smaller and larger than those shown in Fig. 7 but centered at the same 38 areas. We refer to the modified areas as Area − 1, Area + 1, Area + 2, and Area + 4, respectively. Area + n (Area − n) refers to an area which is increased (decreased) by n grid points on either side of the area shown in Fig. 7. One grid length in latitude and longitude is 4 and 5°, respectively. Tables 2a–2c give the values of F for the averaging periods of days 1–31, 16–46 and 32–60, respectively. For January, the number of boxes significant at more than the 95% confidence level are maximum (=29) for the areas shown in Fig. 7 and Area + 1. There is a slight reduction in the number of significant areas for smaller and larger areas. The effect seems to be more clear for Southern Hemisphere areas. The number of significant areas decreases from 14 to 10 with an increase in the spatial averaging size from the area shown in Fig. 7 to the largest area, referred to as Area + 4. This can be due to the absence of strong planetary-scale stationary waves in the Southern Hemisphere summer. The number of significant areas for predictability of averages for days 16–46 and 32–60 is too small to investigate its dependence on the size of the averaging area.

From these results it can be conjectured that at

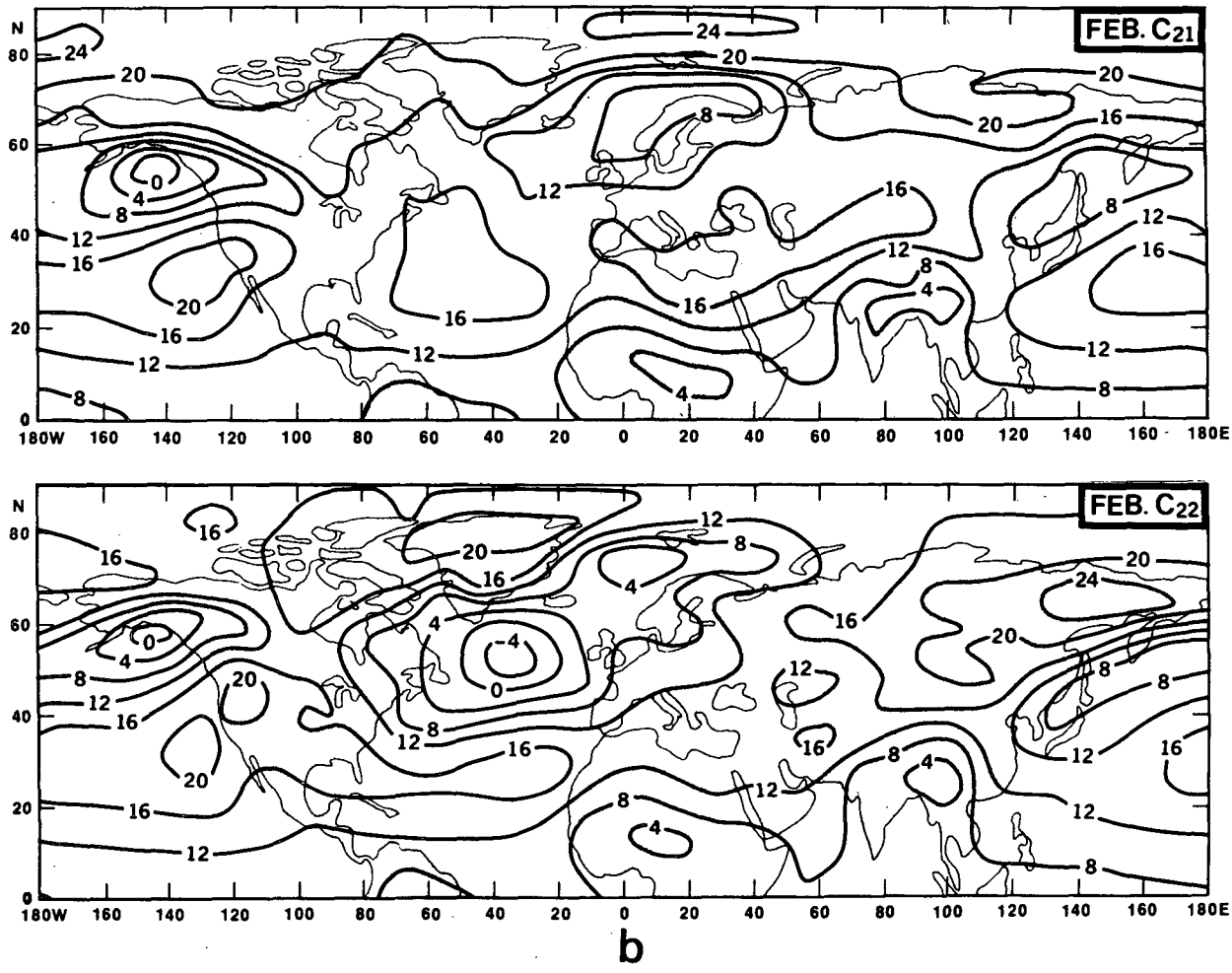


FIG. 6b. 29-day mean sea level pressure for days 32–60 for the control run from the initial conditions of 1 January 1976 (upper panel: FEB C_{21}) and its perturbation run (lower panel: FEB C_{22}).

the 95% confidence level there is almost no predictability for the averaging period of days 32–60; there is a substantial degree of predictability for days 1–31 but there is only partial predictability for days 16–46. This suggests that the observed initial conditions, together with their instabilities and nonlinear interactions, remember themselves, at least up to a month, to such an extent that the monthly means for different integrations are significantly different from the monthly means of randomly perturbed initial conditions.

We also have calculated the values of F at each grid point for five different 30-day means corresponding to days 1–30, 8–37, 15–44, 22–51 and 29–58. Figs. 9a–9e give the plots of F for the 500 mb geopotential height over the Northern Hemisphere. The dotted areas are significant at the 95% level and the cross-hatched areas are significant at the 99.5% level. A systematic degradation of predictability is found to occur from the first to the fifth map. Most of the grid points are significant at the more than 95% confidence level for the predictability of the first 30-day mean. This is a clear indication of the dominance of initial conditions which were different, and it also indicates that the initial planetary-wave configurations could not be drastically changed by the random perturbations. A more encouraging conclusion can be drawn from the distribution of F values for the 30-day mean of days 8–37 and 15–44. Even after ignoring the first 7 and 14 days of integration, which are considered to be the "useful" and "theoretical" upper limits of synoptic-scale deterministic prediction, there are large areas for which the significance level is more than 95%. Fig. 10 gives the number of grid points between 2 and 78°N which are significant at the 95 and 99% confidence levels for the five averaging periods. The number of grid points for the averaging period of days 29–58 is too small to be statistically significant and it may be reasonable to conclude that the model used in this

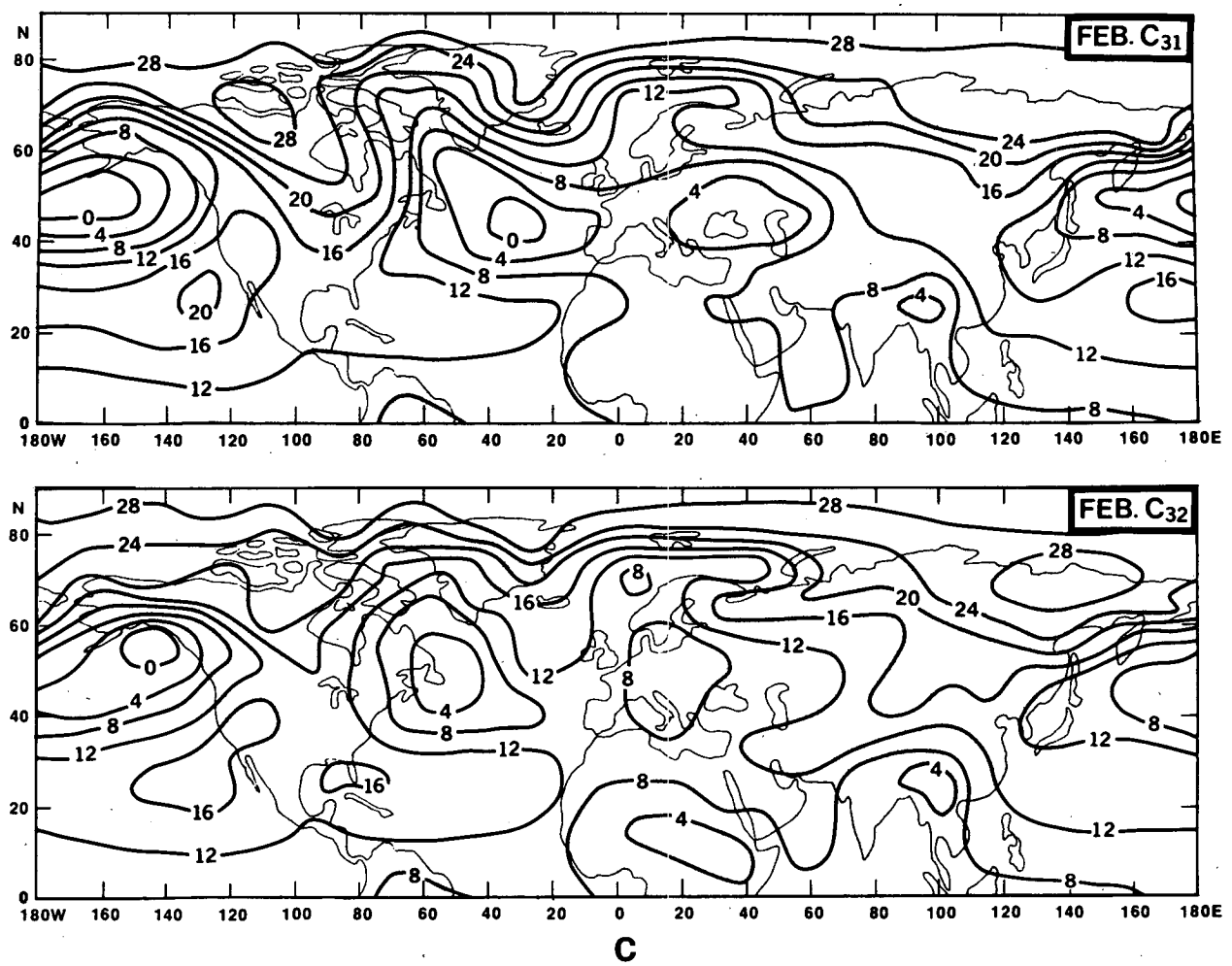


FIG. 6c. 29-day mean sea level pressure for days 32–60 for the control run from the initial conditions of 1 January 1977 (upper panel: FEB C_{31}) and its perturbation run (lower panel: FEB C_{32}).

study, *with climatological mean boundary conditions*, may not have any success in predicting the monthly mean of the second month.

6. Discussion of results

Since the results of any such study are bound to be model dependent, it is necessary to examine especially those characteristics of the model which may have a bearing on the results and their interpretation. For example, if integrations were started with largely different initial conditions for a hypothetical defective model, in which initial conditions persist throughout the course of the integration, one can get a very large value of F. This may lead to a false conclusion about the existence of predictability. In this section we will show that the day-to-day fluctuations simulated by the model are realistic and, except near the poles, comparable to the observed day-to-day fluctuations in the atmosphere.

Similarly, the reduction in the values of F for the second month could be caused by the model property, or more appropriately the model deficiency, that the model simulations for the second month converged to the same state and therefore their variance was reduced. We will show that this is not the case for the present model. The reduction in F occurred mainly due to very large departures between the control run and the perturbation runs for the second month of integration and not due to the relaxation of all the simulations to the same model state.

a. Day-to-day fluctuations

In order to show that the large values of F for January were not due to the persistence of different initial conditions, we have examined the day-to-day fluctuations of sea level pressure, 500 mb geopotential height, and temperature for the nine models runs.

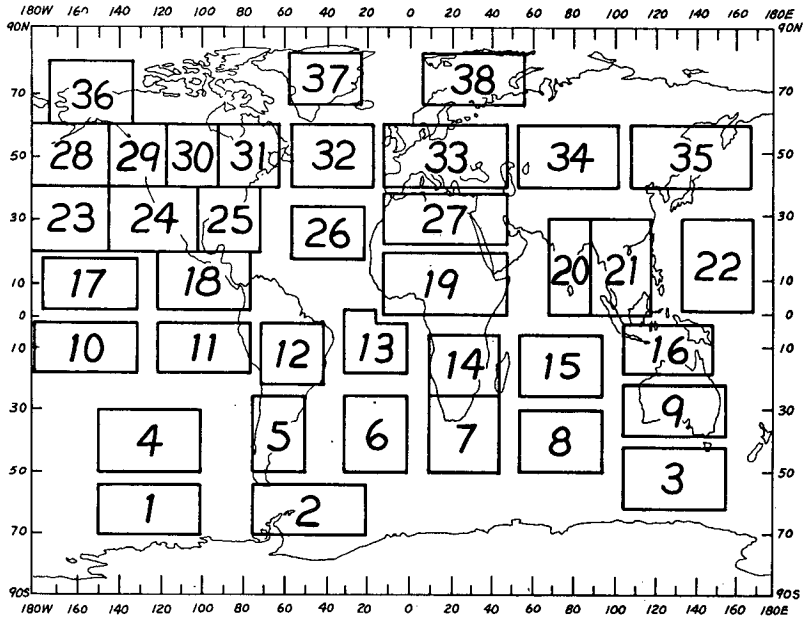
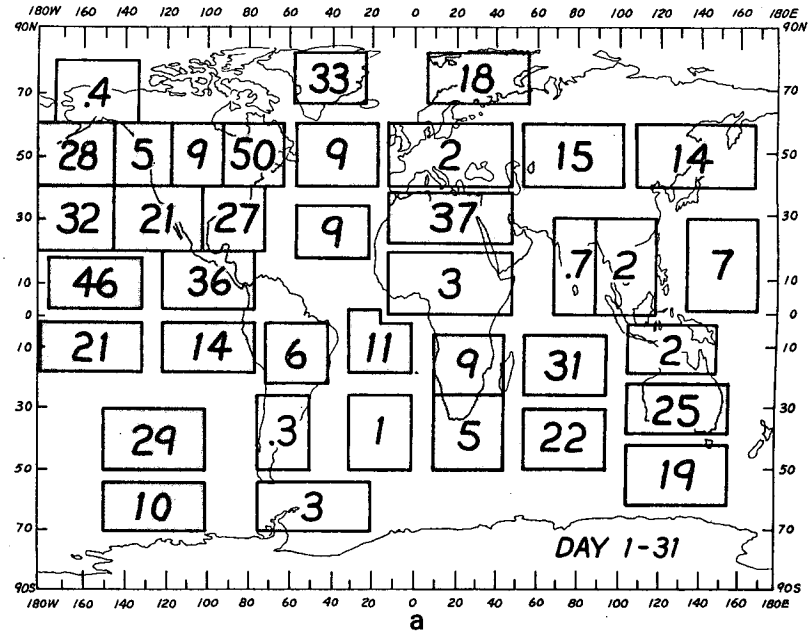


FIG. 7. Locations and sizes for the 38 areas over which sea level pressure is averaged for the analysis of variance.

In order to determine the model's ability to simulate the day-to-day fluctuations, we have also calculated, at each grid point, the standard deviation of daily values for nine model runs for January and February.

If P_{ijkt} denotes the daily value of any meteorological variable (*viz.*, sea level pressure or 500 mb geopotential height) at grid point i, j for model run k on day t, we calculate a measure of day-to-day variability σ_{ij} at grid point i, j as

$$(\sigma_{ij})^2 = \sum_k \sum_t (P_{ijkt} - \bar{P}_{ijk})^2/(K*T)$$



a

FIG. 8a. F values for 31-day mean (days 1–31) sea level pressure averaged over different areas.

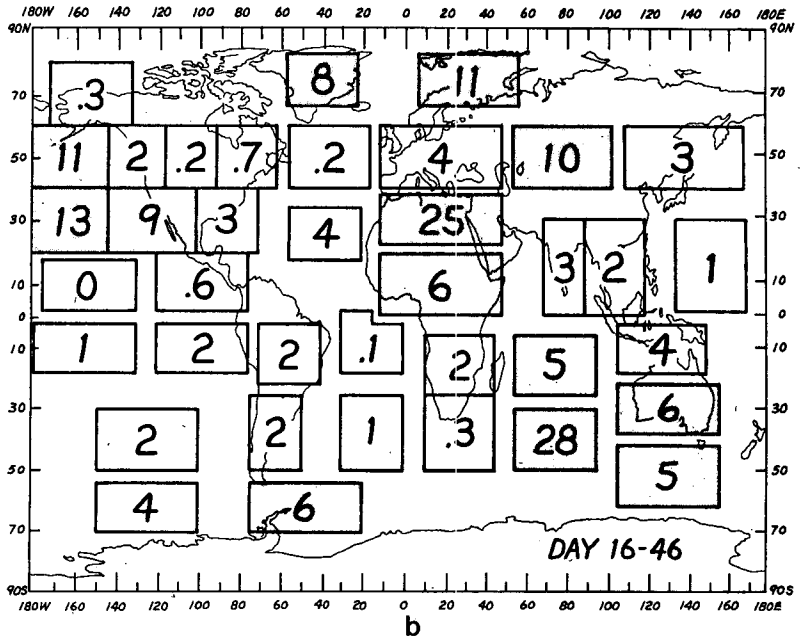


FIG. 8b. F values for 31-day mean (days 16–46) sea level pressure averaged over different areas.

where

$$\bar{P}_{ijk} = T^{-1} \sum_{t} P_{ijkt},$$

and

$$k = 1, 2, \ldots, K \ (K = 9)$$

and

$$t = 1, 2, \ldots, T$$

(T = 31 for January and 29 for February).

For comparison, we have also calculated σ_{ij} for 15 years of observation (K = 15). The observational data set consisted of the NMC (National Meteorological Center) analyses which were received from NCAR (National Center for Atmospheric Research). Fig. 11 shows the zonal averages of σ_{ij} for nine model runs and 15 months of observations for sea level pressure for January and February. Results for 500 mb geopotential height field (not shown) also

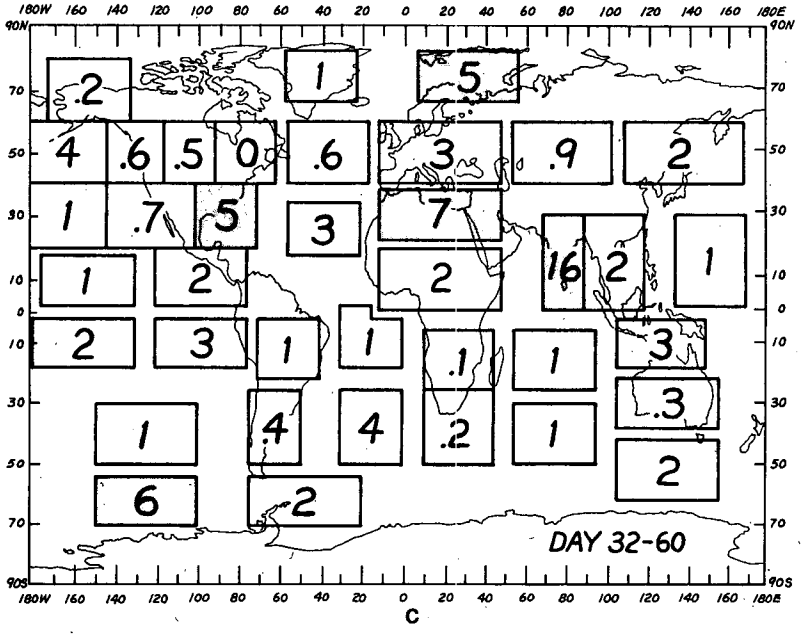


FIG. 8c. F values for 29-day mean (days 32–60) sea level pressure averaged over different areas.

were similar except that the differences near the poles were even larger. The model tends to underestimate the day-to-day variability near the winter pole. This is not unrelated to a major model deficiency that gives an unrealistically cold upper troposphere and strong zonal wind. In the tropics and the midlatitudes, the day-to-day variability of the model is found to be realistic and quite comparable to the day-to-day fluctuations observed in the real atmosphere.

We have not compared the day-to-day predictions with the corresponding observations because numerical weather prediction was not the main intent of this study. Moreover, we have not used the observed boundary conditions for the corresponding initial conditions. The primary aim here was to determine the time limit for which the random perturbations can make the space–time averages completely unpredictable.

b. Does the model relax to the same state for the second month?

Since the external forcings (solar heating and climatological mean boundary conditions) for all the model runs are identical, it is very likely that all the simulated model states will relax to the same mean climate. The relaxation time, in general, will be determined by the time scale of the external forcing and the nature of dynamical system which might have its own internal feedbacks which, in turn, may interact with the external forcing (*viz.*, cloud-radiation interaction). A comprehensive examination of this question is beyond the scope of the present study, especially because the prospects of predictability of a few days to a month are largely, and sometimes solely, determined by the dynamical instabilities and their interactions. In the present study we investigate a limited aspect of the question. Does the reduction in the value of F for February occur due to similarity between all the model simulated states or is it due to very large differences between the control run and its perturbation run? In order to provide a quantitative answer to this question, we have examined the following three aspects of the model simulations.

TABLE 2a. Values of F for January (days 1–31) for different sizes of the area of averaging.

Region no.	(Area − 1)	Area (Fig. 7)	(Area + 1)	(Area + 2)	(Area + 4)
1	9.0	10.3	14.9	19.6	8.9
2	3.7	3.4	3.1	3.4	3.6
3	25.6	18.6	17.4	19.6	39.3
4	40.3	29.0	22.2	12.4	1.9
5	0.1	0.2	0.3	0.2	0.2
6	1.4	0.9	0.5	0.3	0.0
7	5.7	5.1	4.1	2.5	1.3
8	23.9	22.2	19.5	16.0	8.4
9	16.0	25.4	31.1	32.9	30.5
10	32.8	20.8	28.5	39.7	19.5
11	9.6	14.1	18.6	21.1	41.8
12	4.4	6.1	6.8	6.2	9.9
13	11.2	10.7	9.0	6.1	5.5
14	12.1	9.0	8.3	4.2	0.8
15	34.8	30.6	15.9	4.4	1.6
16	2.6	1.7	1.6	1.7	3.5
17	33.0	46.1	63.9	20.4	12.6
18	25.6	36.1	25.2	27.2	32.3
19	2.4	3.1	5.1	10.3	39.7
20	0.9	0.7	0.1	0.0	0.8
21	1.0	1.7	2.2	2.0	2.1
22	8.0	7.0	6.8	2.5	4.1
23	28.6	32.0	19.8	16.0	2.3
24	18.9	20.9	22.9	31.7	36.3
25	16.9	26.6	44.0	62.2	53.1
26	6.7	8.7	13.1	20.0	30.8
27	42.8	36.8	29.4	24.6	24.9
28	24.4	27.7	6.9	2.2	0.5
29	2.4	5.4	9.5	18.1	28.5
30	9.9	8.7	7.9	9.1	13.4
31	50.4	50.1	47.5	49.3	46.1
32	8.6	9.0	11.0	14.8	26.8
33	2.1	1.8	1.8	4.2	17.8
34	7.6	15.4	26.6	28.9	40.0
35	8.4	14.0	16.5	11.3	10.5
36	0.5	0.4	0.1	0.6	0.4
37	30.7	33.2	17.2	14.3	7.3
38	18.4	17.9	17.1	15.0	21.4

TABLE 2b. Values of F for mid-January and mid-February (days 16–46) for different sizes of the area of averaging.

Region no.	(Area − 1)	Area (Fig. 7)	(Area + 1)	(Area + 2)	(Area + 4)
1	4.3	3.9	5.5	5.5	4.9
2	4.3	5.9	5.5	5.3	2.9
3	5.9	4.6	2.4	2.3	2.2
4	2.2	1.7	1.5	0.8	0.5
5	1.5	1.6	1.4	0.8	0.7
6	1.7	1.2	1.0	0.8	0.6
7	0.2	0.2	0.3	0.2	0.1
8	29.3	28.3	24.7	16.5	7.9
9	5.9	5.5	4.9	4.6	3.0
10	1.2	1.0	0.5	0.0	0.4
11	1.6	1.8	1.5	1.7	2.9
12	1.2	1.8	2.9	4.2	4.4
13	0.2	0.1	0.3	0.9	4.6
14	1.4	2.4	1.8	1.5	0.9
15	4.2	4.5	2.5	1.3	2.2
16	3.0	3.8	3.9	3.1	1.9
17	0.2	0.0	0.1	1.4	1.2
18	0.7	0.6	0.4	1.0	3.4
19	3.3	5.7	7.8	11.5	16.3
20	3.1	3.2	2.8	2.8	0.3
21	2.0	2.0	2.8	3.2	0.2
22	1.8	1.3	1.0	0.1	1.2
23	11.3	12.8	2.0	1.2	0.4
24	4.4	9.0	15.8	19.1	7.7
25	3.0	3.4	3.6	3.6	2.6
26	3.5	4.0	5.5	7.2	6.7
27	27.4	25.0	22.1	19.6	10.8
28	10.7	11.0	0.8	0.1	0.3
29	1.8	2.1	4.2	5.6	4.3
30	0.5	0.1	0.0	0.0	0.2
31	0.5	0.7	1.0	1.2	1.9
32	0.1	0.1	0.4	0.7	1.8
33	4.0	3.8	3.6	4.0	4.6
34	4.9	9.6	16.1	16.3	17.2
35	1.8	2.5	3.4	6.9	2.9
36	0.0	0.2	0.5	0.7	1.1
37	7.9	8.0	10.0	8.4	4.5
38	9.1	11.0	11.1	9.0	7.2

1) CHANGE IN F_1 AND F_2 FROM JANUARY TO FEBRUARY

Table 3 gives the values of F_1 and F_2 for January and February. If all the model simulations relaxed to

TABLE 2c. Values of F for February (days 32–60) for different sizes of the area averaging.

Region no.	(Area − 1)	Area (Fig. 7)	(Area + 1)	(Area + 2)	(Area + 4)
1	5.2	6.3	7.9	9.3	10.6
2	1.8	2.0	2.4	3.0	5.0
3	1.4	1.3	2.2	2.4	2.2
4	0.8	0.9	1.3	2.7	12.0
5	0.7	0.4	0.5	1.4	4.0
6	3.7	4.2	5.3	6.0	6.3
7	0.3	0.1	0.1	0.1	0.1
8	1.7	1.2	0.8	0.5	1.0
9	0.3	0.2	0.4	0.4	1.5
10	2.1	2.4	2.9	2.0	1.3
11	3.2	3.2	3.3	3.4	3.4
12	0.8	1.0	1.2	1.6	1.5
13	1.3	1.2	1.1	0.8	0.5
14	0.5	0.1	0.1	0.3	0.6
15	2.1	1.1	0.8	0.9	0.7
16	2.3	2.9	2.9	2.7	2.0
17	1.3	1.1	0.8	0.8	0.6
18	4.4	2.0	2.4	1.8	1.1
19	0.5	1.5	2.9	4.7	8.1
20	7.6	16.1	15.7	5.0	0.6
21	1.5	1.6	2.7	3.1	5.8
22	1.0	0.8	0.7	6.2	2.2
23	0.8	1.0	0.7	0.5	0.3
24	0.5	0.7	1.2	1.3	0.9
25	4.6	5.1	3.6	2.5	1.1
26	2.2	3.0	4.0	3.9	2.1
27	7.5	7.3	7.3	7.2	5.1
28	4.1	4.3	1.3	0.4	0.1
29	0.8	0.6	1.0	1.1	1.0
30	0.3	0.4	0.4	0.3	0.3
31	0.0	0.0	0.1	0.1	0.3
32	0.6	0.6	0.5	0.4	0.2
33	4.0	3.2	2.0	1.5	0.6
34	1.6	0.8	0.3	0.4	1.5
35	2.3	1.6	1.6	0.2	0.4
36	0.0	0.1	0.2	0.1	3.3
37	1.1	0.8	1.5	1.6	1.1
38	3.7	5.0	5.3	5.3	3.1

the same state for the second month, the value of F_2, which measures variability among the control and their perturbation runs, would be very small. However, it can be seen that, from January to February, the value of F_2 has increased for 32 out of 38 areas. Most of the reduction in F from January to February is due to increase in value of the denominator (F_2). This supports the contention that the loss of predictability for the second month is not due to relaxation of the model to a mean state but due to large differences in the evolution of the dynamical system represented by the model.

Table 3 also shows that the values of F_1 have decreased for 26 out of 38 areas (although the decrease for most of the areas is very small). A possible reason for the decrease in F_1 could be a model deficiency, as mentioned above, that all model states were similar. We have shown above that this is not the case because F_2 has increased. Since F_1 is a measure of variability among the different box means, large deviations of control and perturbation runs will lead to reduction in the value of F_1 from January to February. It is important to note that this indeed is the case for the present model. If all the model simulations for the second month relaxed to the same state, the prospects for dynamical predictability with this particular model would have been quite hopeless. However, since the lack of predictability for the second month is due to the large deviations between the dynamically evolving flow fields, it is conceivable that, with the use of

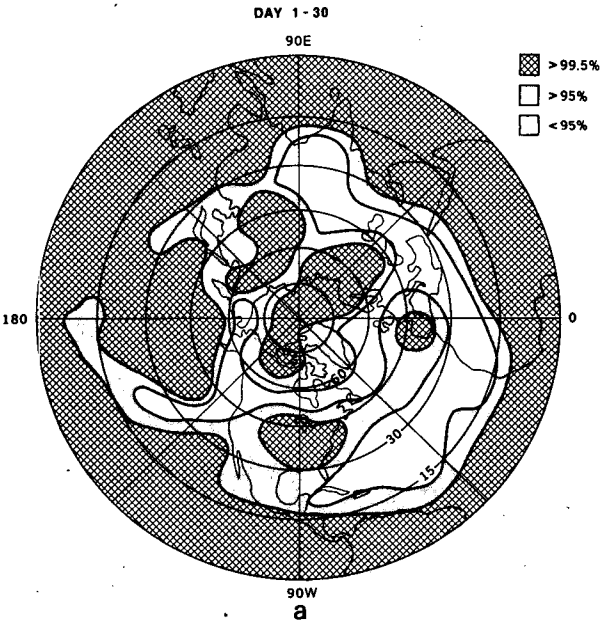


FIG. 9a. F values for 500 mb geopotential height field, calculated at each grid point, for 30-day mean (days 1–30). Light-shaded areas are significant at 95%, and cross-hatched areas are significant at 99% confidence level. Significance level for blank areas is <95%.

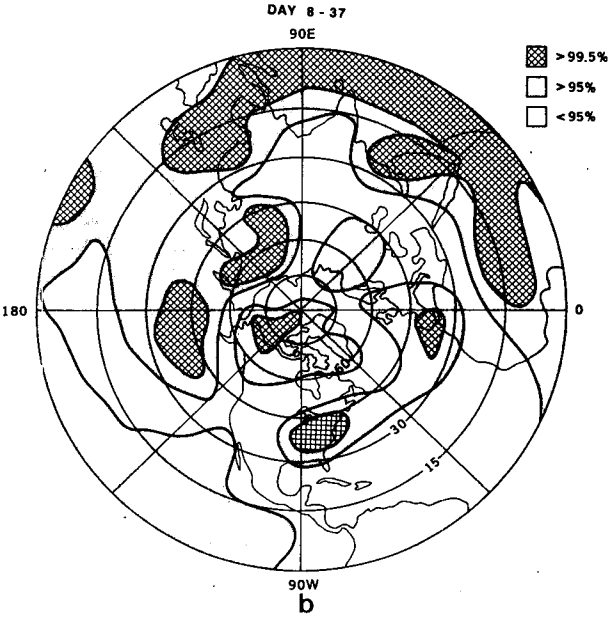


FIG. 9b. As in Fig. 9a except for the 30-day mean (days 8–37).

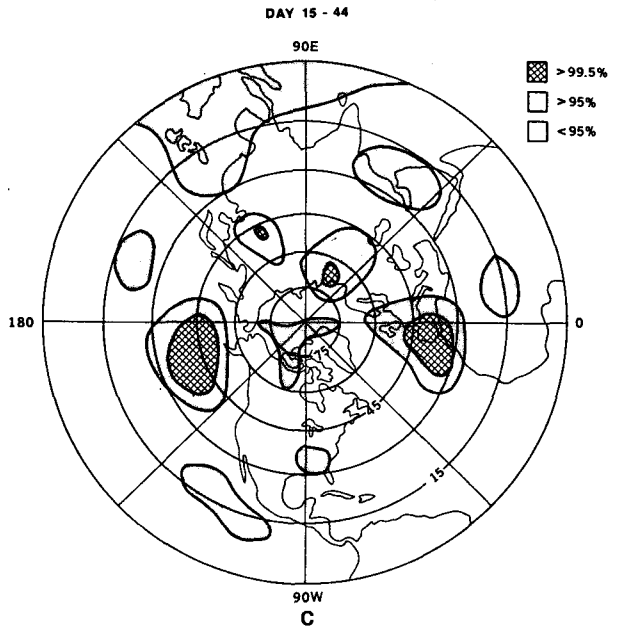


FIG. 9c. As in Fig. 9a except for the 30-day mean (days 15–44).

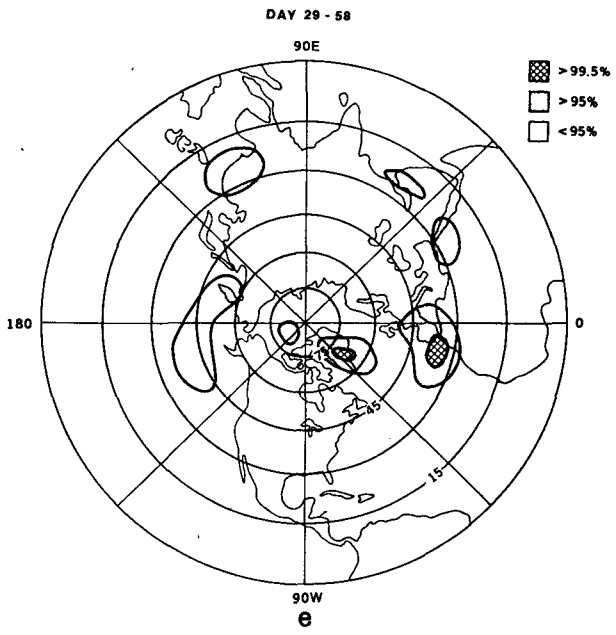


FIG. 9e. As in Fig. 9a except for the 30-day mean (days 29–58).

more realistic dynamical models and accurate physical parameterizations, the limit of predictability of time averages may be extended beyond one month.

2) VARIABILITY AMONG THE CONTROL RUNS FOR JANUARY AND FEBRUARY

We know that the initial conditions for the three control runs were as different as three randomly chosen years and, therefore, standard deviation among them gave a measure of interannual variability. If these initial conditions relaxed to a very similar model state for the second month, the standard deviation for the second month (February) should be smaller than the interannual variability. Fig. 12 shows the zonally averaged standard deviation among the three control runs for January and February for 500 mb geopotential height and sea-level pressure. It is seen that the variability among the three control runs for the second month is indistinguishable from the variability for the first month. This provides an additional support to the contention that the reduction in F and the loss of predictability of the second month is not due to the relaxation of the model to the same state, but due to

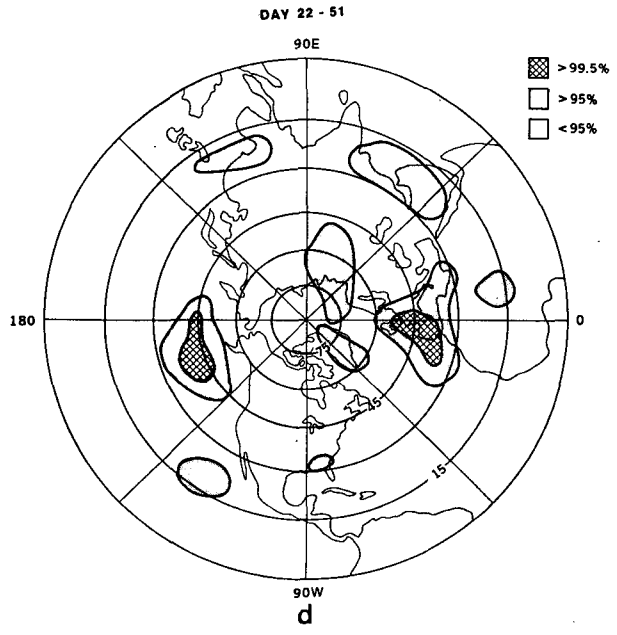


FIG. 9d. As in Fig. 9a except for the 30-day mean (days 22–51).

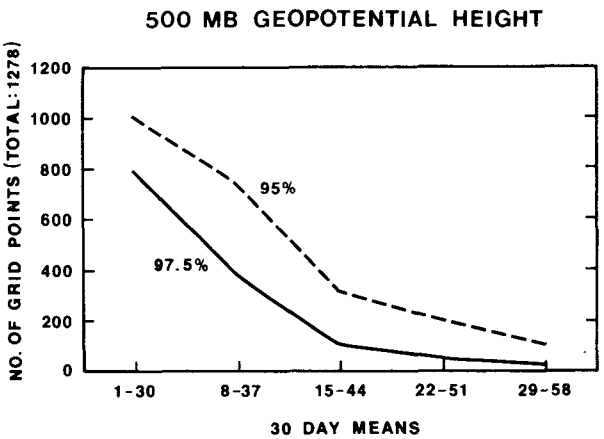


FIG. 10. Number of grid points for which F value is significant at more than 95 and 97.5%.

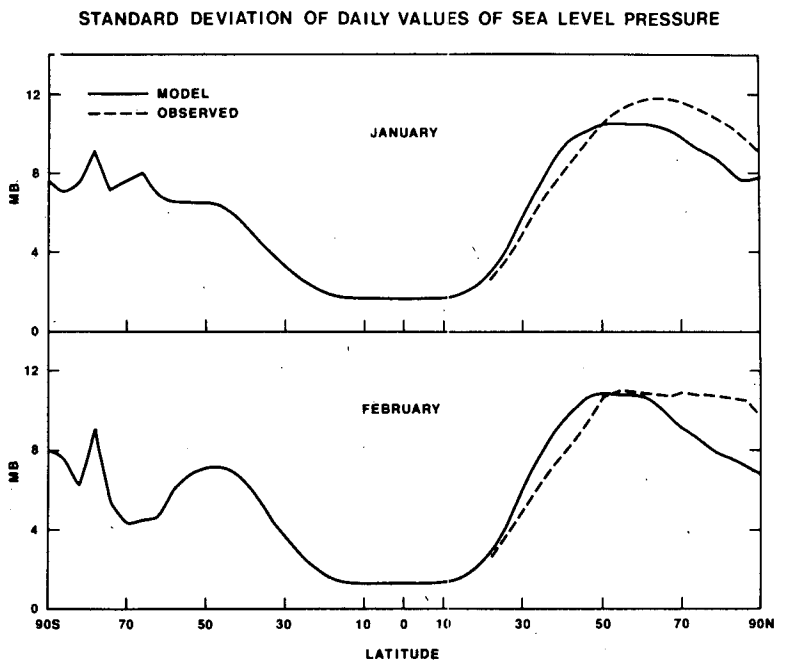


FIG. 11. Zonally averaged standard deviation of daily grid point values for sea level pressure for January (upper panel) and February (lower panel), for model (solid line) and observations (dashed line).

large errors between the control and the perturbation runs.

3) ERROR BETWEEN CONTROL AND PERTURBATION RUNS

We also have calculated the magnitude of the error between a control run and its perturbation runs. The differences between the control and its perturbation increase steadily with an increase in the length of integration, and we have shown that the magnitude of this difference for the second month is large enough to be comparable to the differences between the observed monthly means. The error $E(\phi)$, which is a measure of the rms error among the monthly means due to random perturbations, for any latitude ϕ, is defined as

$$\{E(\phi)\}^2 = \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{10}\left\{\begin{array}{r}\{(M_{11}-M_{12})^2+(M_{11}-M_{13})^2+(M_{12}-M_{13})^2+(M_{21}-M_{22})^2\\ +(M_{21}-M_{23})^2+(M_{21}-M_{24})^2+(M_{22}-M_{23})^2\\ +(M_{22}-M_{24})^2+(M_{23}-M_{24})^2+(M_{31}-M_{32})^2\}d\lambda\end{array}\right\},$$

where M_{ji} refers to the grid-point value of monthly mean sea level pressure for run number C_{ji} (refer to Table 1) and integral over λ denotes integral over all longitudes (72 grid points for $d\lambda = 5°$) along the latitude ϕ.

Fig. 13 shows the zonally averaged values of $E(\phi)$ for three different averaging periods. Dotted, dashed and solid lines, labeled as J, J/F and F refer to the monthly means for days 1–31, 16–46 and 32–60 respectively. The corresponding value of $E(\phi)$ among all possible combinations of 16 years (1962–77) of observed monthly mean sea level pressure over the Northern Hemisphere are shown by a thin solid line. Except for the low latitudes, the average error among the control and the perturbations runs for days 32–60 is comparable to the observed differences between Januaries of different years. This clearly demonstrates that the model-simulated monthly means for the second month are not relaxing to the same value but they are perhaps as different as possible for the prescribed external forcing.

The dash–dot line in Fig. 13 gives the ratio of observed $E(\phi)$ and the model $E(\phi)$ for February. In agreement with the results of Charney and Shukla (1980), it is found that the ratio is maximum in low latitudes. This supports their hypothesis that a

large fraction of the interannual variability of the monthly means in the low latitudes may be related to the changes in the slowly varying boundary conditions.

c. Model limitations

As mentioned earlier, a major deficiency of the present model is the underestimation of day-to-day variability near the poles. Variability of monthly means for days 32–60 also is smaller than the observed variability near the poles. Fig. 14 shows the zonally averaged standard deviation for 16 years of observed January means, and monthly mean (days 32–60) for nine model runs. The largest discrepancy is found to occur near the winter pole. North of 60°N, the model variability is only ~60% of the observed variability. This is related to a major model deficiency that the upper tropospheric–lower stratospheric temperatures are very cold compared to the observations and the zonal winds at the upper levels are quite large. Removal of such systematic and serious deficiencies of the general circulation models will be a prerequisite for their utility as forecast models for dynamical prediction of monthly means.

It should be emphasized that we have subjected this model to a rather rigorous and perhaps even unfair test of its predictability by imposing climatological mean boundary conditions for integrations starting from observed initial conditions. Since the initial conditions are not quite consistent with the imposed stationary forcings, and since the model does not simulate the stationary circulation realistically, spurious traveling wave components are generated and tend to make the model unpredictable from the very beginning of the integration. If all the integrations were started with the corresponding observed boundary conditions, it would reduce the amount of initial inconsistency between the initial dynamical state and the underlying forcing and therefore it might reduce the errors due to spurious traveling waves.

We also should summarize other known model deficiencies which may have a bearing on the present results and their interpretation. Straus and Shukla (1981) have shown that although the total variance simulated by the GLAS model is not too unrealistic, the cyclone wave variances are overestimated and the low-frequency planetary wave variances and the stationary variances are significantly underestimated. This will tend to reduce the variance among the monthly means. Since we have no reason to believe that the reduction in the variance among control runs will be smaller than the reduction in the variance among the perturbation runs, we conclude that this particular deficiency of the model may not alter our conclusions. On the other hand, since the cyclone-scale variances are overestimated in the GLAS model and since these scales are known to grow rapidly and contribute toward the unpredictability of larger scales, it will tend to overestimate the variance among the perturbation runs. Since the differences among the control runs is very large at the initial time itself, dynamic instabilities cannot make them any more different from each other. It is, therefore, reasonable to conclude that this particular deficiency of the GLAS model tends to underestimate the predictability, and if the model was more realistic in simulating the cyclone-scale variances, the theoretical upper limit of dynamical predictability would have been larger. It should be pointed out that one of the major motivations of carrying out this study, in contradistinction to a study in which model generated forecasts are compared with

TABLE 3. Values of F_1 and F_2 for different regions for January and February.

	F_1		F_2	
Region no.	January days 1–31	February days 32–60	January days 1–31	February days 32–60
1	41.7	31.8	12.1	14.9
2	11.6	16.4	10.0	23.9
3	8.9	8.2	1.4	13.5
4	11.3	2.1	1.1	6.6
5	0.4	0.2	5.1	1.5
6	1.8	10.2	5.8	7.2
7	14.6	0.6	8.5	10.1
8	58.8	4.9	7.9	11.6
9	9.2	0.7	1.0	7.4
10	0.7	0.9	0.1	1.1
11	0.3	1.0	0.0	0.9
12	0.5	0.8	0.2	2.3
13	0.3	0.2	0.1	0.6
14	1.1	0.0	0.3	1.7
15	2.5	0.2	0.2	0.7
16	0.1	0.7	0.3	0.7
17	1.0	0.5	0.0	1.6
18	1.3	0.4	0.1	0.7
19	1.0	0.2	0.9	0.5
20	0.2	1.4	1.1	0.2
21	1.2	1.3	2.1	2.4
22	1.3	1.7	0.5	5.9
23	77.8	9.0	7.2	26.3
24	6.7	1.7	0.9	7.2
25	26.8	3.8	3.0	2.2
26	41.5	13.3	14.3	13.3
27	19.9	26.3	1.6	10.7
28	195.8	47.5	21.1	32.6
29	21.4	10.5	11.9	50.5
30	30.8	3.7	10.6	25.0
31	85.6	0.3	5.1	64.1
32	103.0	34.6	34.3	174.3
33	5.7	27.4	9.2	25.7
34	79.1	4.3	15.3	14.8
35	26.0	13.7	5.5	24.5
36	1.7	8.7	12.3	138.7
37	125.0	29.8	11.2	111.4
38	226.7	122.7	37.8	72.4

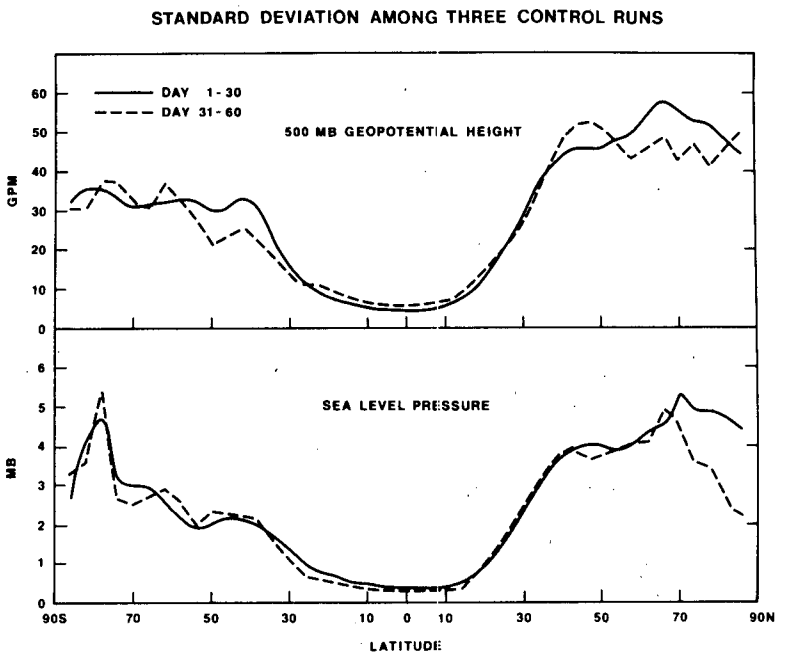


FIG. 12. Zonally averaged standard deviation among monthly values for three control runs for first 30-day mean (solid line) and next 30-day mean (dashed line), for 500 mb geopotential height (upper panel) and sea level pressure (lower panel).

the observations, was to compare two different properties of the same model rather than comparing the model results with the observations. Model results are generally compared with the observations to determine the skill of prediction and its operational usefulness. This is not the objective of this study. Here we are addressing a different question: Is there a basis to make dynamical prediction of monthly means?

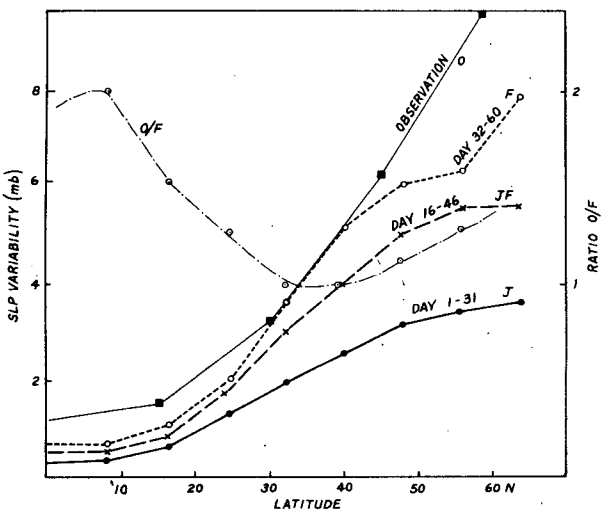


FIG. 13. Zonally averaged rms error between control and perturbation runs averaged for days 1–31 (solid line), days 16–46 (long dashed line) and days 32–60 (short dashed line). Thin solid line is for the observed January sea level pressure for 16 years (1962–77). Dashed-dotted line is the ratio of observed and model (days 32–60) variances.

Although we conclude from this study that monthly means are potentially predictable with dynamical models, it remains to be seen whether the skill of such a prediction will be any better than that of a statistical prediction scheme. It is likely that a combined statistical–dynamical approach may have more skill than either a pure dynamical or a pure statistical approach.

7. Summary and conclusions

The GLAS general circulation model is integrated for 60 days with nine different initial conditions. Three of the initial conditions are the observations on 1 January of 1975, 1976 and 1977. Six other initial conditions are obtained by adding random perturbations, comparable to the errors of observations, to the three observed initial conditions. For all the integrations, the boundary conditions of sea surface temperature, sea ice, snow and soil moisture are the same as their climatological values changing with season.

We have carried out an analysis of variance to compare the differences of monthly means predicted from largely different observed initial conditions to the monthly means predicted from randomly perturbed initial conditions. If the variability among the monthly means predicted from very different

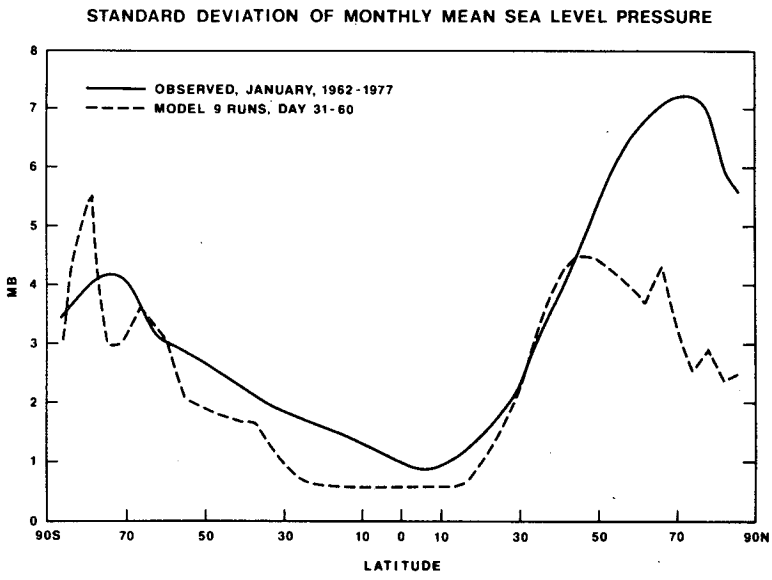


FIG. 14. Zonally averaged standard deviation among monthly (days 31–60) sea level pressure for nine model runs (dashed line) and 16 observed Januaries (solid line).

initial conditions becomes comparable to the variability for randomly perturbed initial conditions, it is concluded that the model does not have any predictability beyond that period.

It is found that the observed initial conditions in different years have large differences compared to the errors of observation, and different planetary-scale configurations retain their identity at least up to one month or beyond. The predictability limit for the synoptic-scale waves (wavenumbers 5–12) is only about two weeks, but for the planetary waves (wavenumbers 0–4) it is more than a month. We have shown that there is a physical basis to make dynamical prediction of monthly means at least up to one month.

For climatological mean boundary conditions, there is no potential for predictability in the second month (days 31–60) because the monthly means predicted from largely different initial conditions become indistinguishable from the monthly means predicted from random perturbation. It is encouraging to note, however, that the lack of predictability for the second month is not because all the model integrations converge to the same mean state but because of very large differences due to random perturbations in the initial conditions. This suggests that it is possible, at least in principle, to extend the predictability limit even beyond one month by improving the model, the initial conditions, and the parameterizations of physical processes.

In this paper we have examined only the dynamical predictability with climatological mean boundary conditions. There could be additional predictability due to fluctuations of the slowly varying boundary conditions of sea surface temperature, soil moisture, sea ice and snow. There is sufficient observational and GCM-experimental evidence to suggest that the fluctuations of sea surface temperature and soil moisture in low latitudes produce significant changes in the monthly and seasonal mean circulation. Likewise, if the anomaly of sea surface temperature, sea ice or snow is of large magnitude and of large-scale, it can produce significant changes in the mid-latitude circulation. Due to the presence of strong instabilities in the middle latitudes, boundary effects have to be sufficiently large to be significant. However, there seems to be some evidence that the fluctuations of the heat sources in the tropics can produce significant changes in the extratropical latitudes, and, therefore, there is potential for additional midlatitude predictability because of its interaction with low latitudes.

8. Suggestions and further considerations

One of the serious limitations of the present general circulation models seems to be their inability to simulate the stationary circulation. This problem seems to be common to dynamical weather prediction models also, because predicted fields show systematic geographically fixed error structures. The following factors may be cited as possible reasons for incorrect simulation of monthly mean fields:

1) *Inadequate horizontal resolution*: Errors in small scales may affect the large scales, and higher-order accuracy is needed for long waves.

2) *Inadequate vertical resolution, location of the upper boundary, and upper boundary condition*: Most of the models do not treat adequately the vertical propagation and damping of stationary waves.

3) *Inadequate treatment of orography*: Either due to coarse resolution or due to improper finite-difference treatment, flow over and around mountains is not well simulated.

4) *Diabatic heating processes over land and ocean*: Due to inadequate parameterization of boundary-layer, convection, cloud-radiation interaction and physical processes at ocean and land surfaces, the vertical structure of the diabatic heating field is not realistic and therefore the forced stationary waves are not realistic. That, in turn, causes the amplitudes and locations of the storms and their tracks to be unrealistic. If the stationary component of the circulation is not simulated correctly, the discrepancy also will manifest itself into the propagating components of the circulation.

The results of the present study and the results of several other long-range prediction studies of Miyakoda (personal communication) suggest that the prospects of useful long-term dynamical prediction will largely depend upon our ability to develop highly accurate dynamical models which can realistically simulate the stationary and the transient components of the circulation.

The present generation of dynamical models already have demonstrated a degree of verisimilitude which justifies a systematic program to establish the feasibility of long-term dynamical prediction. What is needed first is a very systematic evaluation and intercomparison of dynamical models to determine the precise nature of their weaknesses in simulating the structure and amplitudes of stationary and transient circulations. This is beyond the scope of any individual research scientist and it requires an institutional and organizational framework to begin such a massive undertaking. It is perhaps fair to say that the physical basis and justification to initiate a program of dynamical long-range prediction today is at least as strong as was the basis to start numerical weather prediction 25 years ago, and likewise, a systematic program of long-range dynamical prediction will undoubtedly increase our understanding of the dynamics of the atmosphere and ocean.

Acknowledgments. I am grateful to Professors J. G. Charney and E. N. Lorenz for the benefit of many useful discussions and suggestions during the course of this study. I wish to express my deep appreciation for discussions with and comments by Drs. M. Halem, R. Hoffman, Y. Mintz, K. Miyakoda, D. Randall, D. Straus, D. Gutzler and Prof. M. Wallace. I am thankful to Dr. D. Straus for providing the data on the observed interannual variability of variances. It is a pleasure to express my appreciation to Tom Warlan and Ricky Sabatino for numerical calculations, Karen DeHenzel and Debbie Boyer for typing the manuscripts, and Laura Rumburg for drafting the figures.

REFERENCES

Charney, J. G., 1960: Numerical prediction and the general circulation. *Dynamics of Climate*, R. L. Pfeffer, Ed., Pergamon, 12–17.

——, and J. G. Devore, 1979: Multiple flow equilibria in the atmosphere and blocking. *J. Atmos. Sci.*, **36**, 1205–1216.

——, and J. Shukla, 1980: Predictability of monsoons. *Monsoon Dynamics*, Sir James Lighthill and R. P. Pearce, Eds., Cambridge University Press.

——, and D. M. Straus, 1980: Form-drag instability, multiple equilibria and propagating planetary waves in baroclinic, orographically forced, planetary wave system. *J. Atmos. Sci.*, **37**, 1157–1176.

——, R. G. Fleagle, V. E. Lally, H. Riehl and D. Q. Wark, 1966: The feasibility of a global observation and analysis experiment. *Bull. Amer. Meteor. Soc.*, **47**, 200–220.

——, J. Shukla and K. C. Mo, 1981: Comparison of a barotropic blocking theory with observation. *J. Atmos. Sci.*, **38**, 762–779.

Halem, M., J. Shukla, Y. Mintz, M. L. Wu, R. Godbole, G. Herman and Y. Sud, 1980: Climate comparisons of a winter and summer numerical simulation with the GLAS general circulation model. GARP Publ. Series, **22**, 207–253.

Hays, W. L., 1963: *Statistics*. Holt, Rinehart and Winston, 719 pp.

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

——, 1975: The design of a statistical-dynamical climate model and statistical constraints on the predictability of climate. The Physical Basis of Climate and Climate Modeling, GARP Publ. Ser. No. 16, Appendix 2.2.

Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321–333.

Smagorinsky, J., 1969: Problems and promises of deterministic extended range forecasting. *Bull. Amer. Meteor. Soc.*, **50**, 286–311.

Strauss, D. M., and M. Halem, 1981: A stochastic-dynamical approach to the study of the natural variability of the climate. *Mon. Wea. Rev.*, **109**, 407–421.

——, and J. Shukla, 1981: Space-time spectral structure of a GLAS general circulation model and a comparison with observations. *J. Atmos. Sci.*, **38**, 902–917.