PREDICTABILITY

J. SHUKLA

Center for Ocean–Land–Atmosphere Interactions
Department of Meteorology
University of Maryland
College Park, Maryland

1. INTRODUCTION

The mathematical equations governing the dynamics of the atmospheric flows are nonlinear, and the observed structure of the atmosphere is characterized by horizontal and vertical gradients of wind, temperature, and moisture that permit hydrodynamical and thermodynamical instabilities to grow. These characteristics of atmospheric motion are the primary reason for an upper limit on the deterministic predictability of atmospheric flows. In addition, the equations and the physical parameterizations used for prediction are not exact, and they introduce a source of error in predictions made with a model. Even if the models were perfect, small uncertainties in the initial state can grow due to the inherent instability of the flow and nonlinear interactions among motions of different space and time scales. The quantitative upper limit for deterministic prediction, even for an exact model, is determined by the growth rates and equilibration of the most dominant instabilities.

During the past three decades there have been several attempts to estimate the upper limit for deterministic prediction of the instantaneous state of the atmosphere, to be referred to as the weather. There are some conceptual difficulties in arriving at an estimate of upper limit for deterministic predictability. For example, how does one decide that an upper limit has been reached? The most common answer to this question has been that if the differences between two predictions made with identical models but slightly different initial conditions become equal to the differences between two randomly chosen weather maps (for the same time of year), then the upper limit of prediction has been reached. Even if we accept this as a working definition for the upper limit, we have yet to define a way of measuring the differences. For example, should it be the root-mean-square difference between the two maps averaged over all the appropriately scaled variables or correlation coefficients or some other measure of space time variability? It should be pointed out that the aforesaid working definition of the upper limit of predictability implicitly assumes that we already have a good knowledge of the climatology of the atmosphere and therefore we do not consider a prediction to be of any value if it does not improve on our pre-existing knowledge of the climatology. The concept of usefulness is, therefore, implicit in this definition, although it does not follow that the prediction for a range of time equal to the upper limit as defined above has any practical utility. There have been some attempts to distinguish between the upper limits of theoretical and practical (useful) predictability.

In some of the earlier works, predictability was defined by a single parameter, the growth rate for root-mean-square error over large areas (one hemisphere or the globe). For example, if the doubling time for the error in temperature is about 3 days and if the initial error is about 1°C, assuming a constant growth rate with time, the error will become about 8°C in 9 days. If 8°C is the root-mean-square error between randomly chosen weather maps, the upper limit will be considered to be 9 days. The difficulty in defining predictability in terms of growth rate arises mainly because the growth rate depends on the structure of the initial large-scale flow and the value of the maximum permissible error strongly depends on the latitude and season. The growth rate and equilibration also depend on the variable under consideration.

In this chapter, we shall first review the earlier attempts to determine the limits of atmospheric predictability. We shall describe the results from simple models and turbulence theories in Section 2.1, the analog method in Section 2.2, and results from global general circulation models in Section 2.3. In Section 3 we shall review the work on predictability of space–time averages, and in Section 4 we shall review some of the outstanding problems of predictability. In Section 5 we shall present the concluding remarks.

2. Classical Predictability Studies

We shall use the words *classical predictability studies* to refer to those works during the past 30 years in which an attempt was made to arrive at a quantitative estimate of the growth rate of an initial error and to determine the limits of predictability. Such studies have used either simple models of atmospheric flow or turbulence models or complex models of the general circulation of the atmosphere with explicit treatments of mechanical and thermal forcings. Historical records of the atmospheric flows have also been examined to find naturally occurring analogs and the growth rate of their initial differences.

2.1. Simple Models

The first such study reported in the literature is the one by Thompson (1957), who showed, using a simple barotropic model, that the initial errors tend to grow with time and that the atmospheric flow is not predictable beyond a week. He emphasized the effects of the scale of initial error on predictability and commented on the relative merits of expanding the observational network and improving the models for weather prediction. It was implicit in his work that the instability of the atmospheric flow is the main reason for limits on predictability. He also introduced the concept of the error between two randomly chosen maps as a convenient upper limit of the error beyond which the flow is competely unpredictable. He also showed that the zonally averaged flow is more predictable than the unaveraged flow.

Simple predictability experiments were carried out, serendipitously, by Lorenz (1963) in connection with his work on numerical integration of a simplified nonlinear baroclinic model. The motivation for integrating the model and producing a long-time series was to test the ability of a linear statistical model to predict the behavior of a hydrodynamical flow using the data generated by the nonlinear governing equations. However, since the computing facility available to Lorenz at that time was far inferior to the present-day personal computers, from time to time he had to re-enter the solution printed by the machine to continue the integration further. Since the machine did the computations with an accuracy of six significant digits, but printed out only three digits, solutions were changed in the last three digits every time new values were punched in. Lorenz noticed that if calculations were repeated with such rounding off, the solutions began to diverge and that for longer integrations they became quite different.

Soon thereafter, Lorenz (1965) wrote a comprehensive paper on the predictability of a 28-variable atmospheric model. He used a two-layer quasi-geostrophic model having a zonal flow with two north–south modes and perturbations with three east–west wavelengths each with two north–south modes in each layer. He showed that the doubling time for the initial errors strongly depended on the structure of the flow [a conclusion repeatedly confirmed by the general circulation model (GCM) experiments and the experience of operational weather prediction] and that for synoptic-scale observation errors the doubling time may range from a few days to a few weeks. On the average, the doubling time was about 4 days. He also noticed that although instantaneous flow patterns become completely unpredictable after a few days, some properties of the flow remain predictable much beyond that time. This will suggest some possibility for predicting space–time averages.

Several investigators (Robinson, 1967; Lilly, 1969; Lorenz, 1969a; Leith, 1971; Leith and Kraichnan, 1972; Lorenz, 1984) have also used turbulence models to determine the predictability of an idealized hydrodynamical flow. These models do not include spherical geometry and the rotation of the Earth, nor do they include thermal and mechanical forcing functions and the physical processes of radiation and condensation. They do, however, provide useful insight into the error growth characteristics due simply to nonlinear interactions among the various scales of the fluid motion. These studies necessarily require *a priori* assumptions about the spectra of the kinetic energy of the atmosphere, and the results are quite sensitive to such assumptions.

Robinson (1967) proposed the idea of a virtual viscosity that would dissipate eddies of all sizes, but with the time taken for dissipation of a particular scale (which is a measure of the predictability time for that scale) depending on the rate at which energy from that scale is transferred to the scales at which "true" dissipation takes place. He used a $-\frac{5}{3}$ power law for the large-scale atmospheric energy spectrum. Simply stated, Robinson's concept of limited predictability is based on an assumption that eddies get "diffused" or "dissipated away" in a finite time and therefore there is no hope for predictions beyond a few days. This line of reasoning is not consistent with our intuition (based on observations of the atmosphere) that eddies do not get dissipated but are maintained by well-defined physical processes, and hence the problem of predictability is not the nonexistence of eddies, but rather their growth, movement, and decay, as well as their interactions with other scales of motions. Robinson's concept of dissipative time scales seems more appropriate for determining the time steps for numerical integrations of atmospheric models rather than for determining the predictability of the atmosphere.

Lorenz (1969a) (also using a $-\frac{5}{3}$ power law for the energy spectrum) calculated the time taken for each scale of the motion to be totally unpredictable, defined as the state at which error energy in a given scale becomes equal to the energy at that scale in the initial prescribed spectrum. He found that the interactions take place only among the adjacent scales; however, the error in the smaller scales gets saturated rather quickly. Even if the synoptic scales were free of any error initially, errors from the neighboring smaller scales produced errors in synoptic scales within a day or so. Lorenz further suggested that the predictability would be increased if the energy spectra had a -3 power law rather than $-\frac{5}{3}$. Leith (1971) and Leith and Kraichnan (1972) used improved turbulence closure approximations and showed that for the two-dimensional eddy kinetic energy spectrum similar to the one observed in the atmosphere, the doubling time for error was about 2 days.

It is rather interesting that these estimates of error-doubling time are quite close to the estimates made by current state-of-the-art GCMs and also the estimates made by Lorenz (1969b) by using analogs in the past observations of the atmosphere.

In his paper, Lorenz (1984) has shown that the presence of a moderately strong spectral gap in the mesoscale range of the assumed energy spectrum will increase the predictability by about 3 more days. Lorenz's calculations suggest that the error level at day 1 without the spectral gap will be about the same as the error at day 4 in the presence of the spectral gap.

2.2. *Observations (Analogs)*

Dr. J. Namias once remarked that the analog method of weather forecasting is as old as the second weather chart. Before the advent of the statistical and dynamical models for weather prediction, the analog method was perhaps the most commonly used technique for weather forecasting. Even now analogs are commonly used to make extended-range predictions.

Lorenz (1969b, 1973) proposed an ingenious method of studying classical atmospheric predictability using naturally occurring analogs from past records of atmospheric observations. He proposed that if it were possible to find two rather closely resembling atmospheric states, the rate with which the differences between the two states grow would give a measure of the classical predictability error growth. He used five years (1963–1967) of twice-daily height data over the Northern Hemisphere for the 200-, 500-, and 850-mb surfaces to carry out his search for good analogs. He could not find good analogs: the difference (rms error) between the two states corresponding to his best analog pair was 62% of the error between two randomly chosen states. Lorenz found that the doubling time for error between these "mediocre" analogs was about 8 days. Since the main objective was to find the growth rate of small errors, Lorenz extrapolated the growth rate for small errors from the knowledge of the growth rate for large errors. For this he proposed a quadratic hypothesis for error growth rate that gave a doubling time of 2.5 days for small errors. It is rather remarkable that this estimate of doubling time is very close to the estimates from the state-of-the-art global general circulation models. Considering the crudeness of the technique utilized, the quantitative exactness of this result should be considered as a combination of Lorenz's brilliance and serendipity. As Lorenz pointed out, a cubic hypothesis would have given a doubling time of 5 days, but the data did not show as good a fit. Lorenz further suggested that the chances of obtaining really good analogs does not seem to be good even for larger data sets. However, it may still be worthwhile to process large data sets once they are available.

Gutzler and Shukla (1984) have analyzed 15 years (1963–1977) of winter season daily, 500-mb height observations for the Northern Hemisphere to search for analogs. They restricted their search to 500 mb only because of the equivalent barotropic nature of a significant part of the atmospheric variability. They examined the analogs for the planetary waves and the synoptic waves separately and also for limited spatial domains. They also looked for analogs for 5-day-mean maps.

They found that by considering the 15-year data, but only at 500 mb, the root-mean-square (rms) error between the best pair of analogs was only about 50% of the rms error between randomly chosen maps. This percentage error was reduced to 40% if rms error was calculated for the planetary waves only (zonal wave numbers 0–4), and further reduced to about 32% for limited regions over the North American and European sectors. The error doubling time was also reduced for reduced errors between the analog pairs, indicating thereby a faster growth for smaller errors. They also examined the accuracy of short-range predictions based on the best analogs, and in each case, with the exception of synoptic waves (wave numbers 5–36), such short-range predictions were inferior to the corresponding persistence forecasts. This was true even for the 5-day-mean circulation maps. Surprisingly, the rms error between the best pair of 5-day-mean analogs was close to 69% of the error between two randomly chosen 5-day-mean maps.

In summary, the work of Gutzler and Shukla using natural analogs supports the last statement of Lorenz's (1969b) paper, "Probably we can gain some additional insight into our problem by processing the largest sample of data which we can assemble, but we must not expect miracles."

2.3. General Circulation Models

Charney *et al.* (1966) were the first to apply general circulation models (GCMs) to the study of the classical predictability of model-simulated atmospheric circulations. They utilized the three GCMs described by Smagorinsky (1963), Mintz (1964), and Leith (1965) to examine the growth rate of initial sinusoidal and random temperature error fields. The results were highly model dependent; the error growth characteristics were quite different for each of the models. The Leith model showed a rapid decay of the initial error for the first 4 days, followed by an error increase to half of its initial value up to day 7. After day 10, the error began to level off. Thus, the expected exponential growth of the initial error was not manifested by the Leith model. The Mintz–Arakawa model showed a near-exponential growth of error after an initial drop for a few days. The doubling time of the error was estimated to be about 5 days. The Mintz–Arakawa results were considered to be the most realistic because this was the only model that exhibited strong aperiodic behavior during a long-term (about 300 days) integration of the model and also because the error growth characteristics were similar to what one could expect from theoretical considerations. The Smagorinsky model was integrated with initial random and sinusoidal temperature error fields of various amplitudes (0.02, 0.1, 0.5, and 2.0 K). For small initial error amplitudes, the error grew very slowly for the first 30 days, after which it showed a doubling time of 6 to 7 days. However, the actual flow patterns showed a primarily periodic behavior. It should be remarked that this large divergence among the results of various models is merely a reflection of the fact that none of the models were realistic. Predictability experiments with state-of-the-art models available today will show more convergent results.

Later papers on classical predictability studies with GCMs were reported by Smagorinsky (1969), Jastrow and Halem (1970), and Williamson and Kasahara (1971). The model used by Smagorinsky was the improved version of the earlier model described by Miyakoda *et al.* (1969) and Manabe *et al.* (1965). The model used by Jastrow and Halem was the improved and modified version of the earlier Mintz–Arakawa model. Williamson and Kasahara (1971) used the model developed at the National Center for Atmospheric Research (NCAR). One of the important conclusions of the Jastrow–Halem and Williamson–Kasahara papers was that the growth rate of the initial error depended on the resolution of the model; error growth was slower for coarse resolution models.

The most comprehensive study of the classical predictability at that time was reported by Smagorinsky (1969), presented as the Wexler memorial lecture at the 49th Annual Meeting of the American Meteorological Society. For the first time, he raised the question of predictability of different spectral modes. In fact, perhaps because this study was so comprehensive (and perhaps because of the reputation of the author), no major work on classical predictability was published for the next 12 years, although a large effort was devoted to actual weather prediction. Smagorinsky presented results of error growth for various initial error amplitudes using two different model resolutions. For an initial random error of about 0.25°C, the doubling time was about 2.5 days, but it took about 7 days for the error of 1°C to double to 2°C. Although the nature of the error growth with time was consistent with the theoretical concepts of hydrodynamical instabilities, the quantitative estimate of the growth rate of the error was quite different from the one suggested by Lorenz, which used simple models or observed analogs. Lorenz's estimates for the doubling time of large-scale error fields (as expected to be for a typical observational network) was about 2 to 3 days, whereas Smagorinsky's model results suggested a doubling time of about 5 to 7 days. As evidenced by a 800-word footnote in Smagorinsky's paper, this difference in the result produced very useful and sharp discussion in the field. It is now generally agreed that Lorenz's estimates were quite close to the present estimates using the current state-of-the-art GCMs.

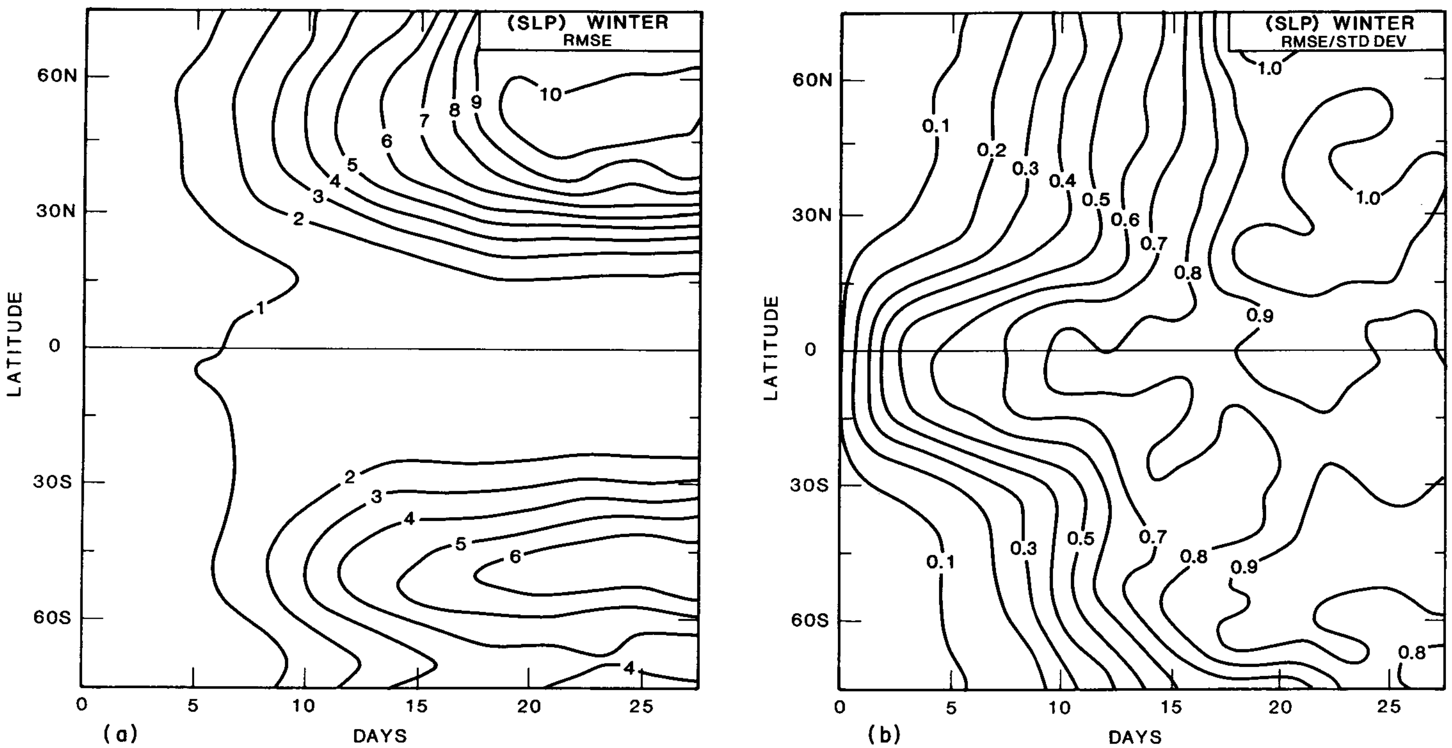
Recently there has been a renewed interest in classical predictability studies. Some examples of the recent works are found in papers by Shukla (1981a), Lorenz (1982), Baumhefner (1984), and Shukla (1984a). Error growth rates have now been examined separately for the tropics and mid-latitudes, for winter and summer, for different space scales, for different hemispheres, and for different initial conditions. In this section, we shall summarize some conclusions from these studies. In particular, we shall show the actual results from some of the integrations of the GLAS climate model. These results are, naturally, model dependent, and it is quite likely that the results from a different GCM with different treatments of numerics and physics will give different quantitative results. We believe, however, that our results on the relative predictability of winter and summer seasons, tropics and mid-latitudes, Northern and Southern Hemispheres, and large and small scales and for random and systematic initial errors will remain unchanged for any GCM that produces a reasonable simulation of climate and its variability.

The model used for these predictability studies has been described by Shukla *et al.* (1981). It is a global primitive-equation model with nine levels in the vertical and a horizontal resolution of 4° latitude by 5° longitude. The model includes parameterizations of radiation, convection, and fluxes at the Earth–atmosphere interface. The observed annual cycle of sea-surface temperature, soil moisture, snow, and sea ice is prescribed at the model grid points.

We shall present results of 30-day integrations using nine winter initial conditions and four summer initial conditions. The model was first integrated with the observed initial conditions of 1 January 1975 (control run), and then two additional integrations were carried out with random perturbations in the initial conditions. The rms error between the control run and the first perturbation run will be referred to as E_{11}, and the rms error between the control run and the second perturbation run will be referred to as E_{12}. Similarly, observed initial conditions of 1 January 1977 were integrated along with three perturbations, and rms error between this control and three perturbations will be referred to as E_{21}, E_{22}, and E_{23}, respectively. The rms error between a control run starting from the observed initial conditions of 1 January 1978 and one perturbation run will be referred to as E_{31}. For each perturbation run, the statistical properties of the random perturbation to the initial conditions were the same (a spatially Gaussian distribution with zero mean and standard deviation of 3 m s^{-1} in u and v components at all the model grid points and at all the levels), but the actual grid-point values of the random perturbations were different for different cases. These integrations were earlier used by Shukla (1981a) to study the dynamical predictability of monthly means. Similar integrations were carried out for the summer season by using the observed initial conditions in the middle of June as control run and three perturbation runs.

In the past, most attention was paid to the error growth rate (or doubling time) as the key predictability parameter. This is not a very useful parameter, partly because it varies greatly for different values of the error and partly because the ultimate limit of predictability is not only determined by the growth rate, but also by the saturation value of the error. This becomes an important consideration when we are examining the predictability of different seasons and different parts of the globe. We have, therefore, presented the results for error growth with time, as well as the ratio of error to the standard deviation of daily fluctuations. A larger error growth does not necessarily mean lesser predictability because it will also depend on the equilibration value of the error that depends on the magnitude of the day-to-day fluctuations.

Figures 1–3 show the results for sea-level pressure, geopotential height at 500 mb, and wind at 300 mb, respectively. Each figure has four panels,


(SLP) WINTER
RMSE
LATITUDE
60N
30N
0
30S
60S
DAYS
0
5
10
15
20
25
(a)
1
2
3
4
5
6
7
8
9
10
(SLP) WINTER
RMSE/STD DEV
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
(b)

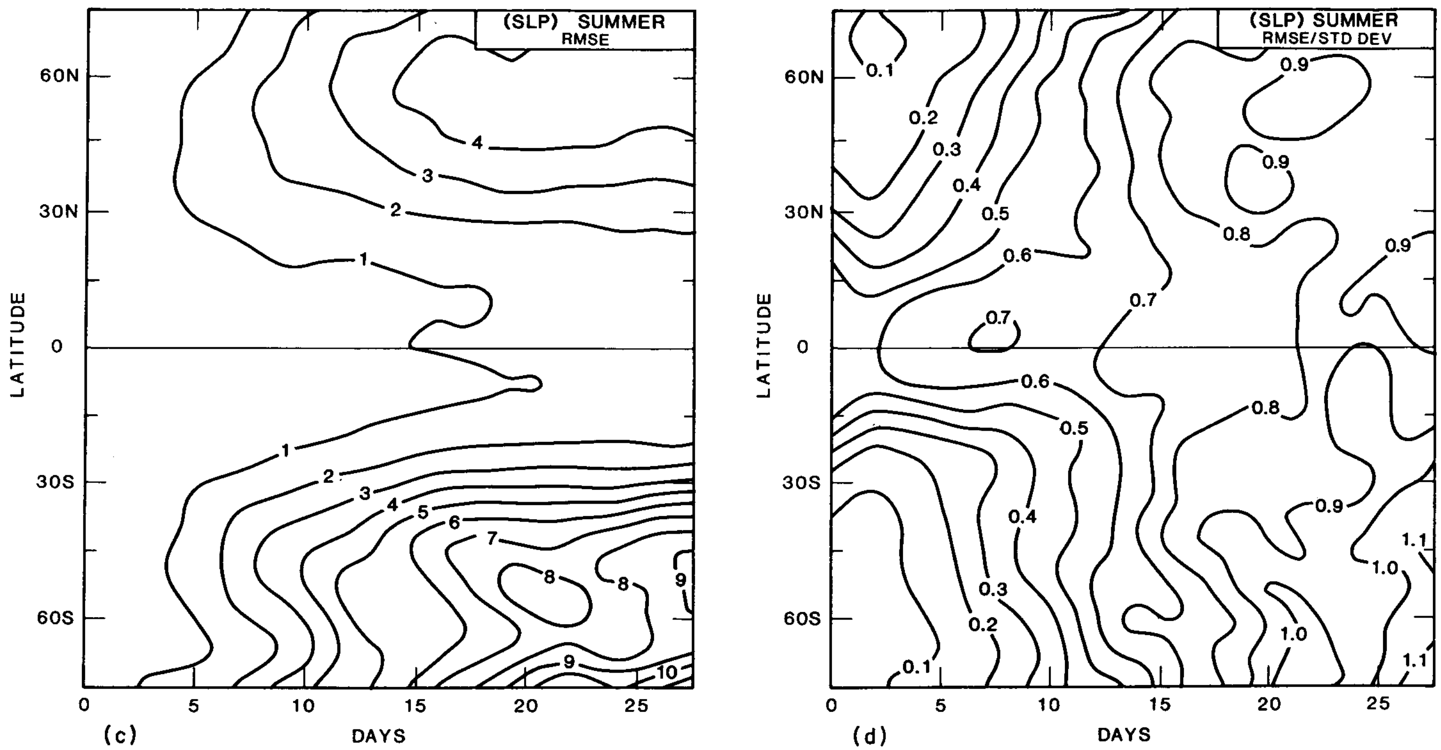
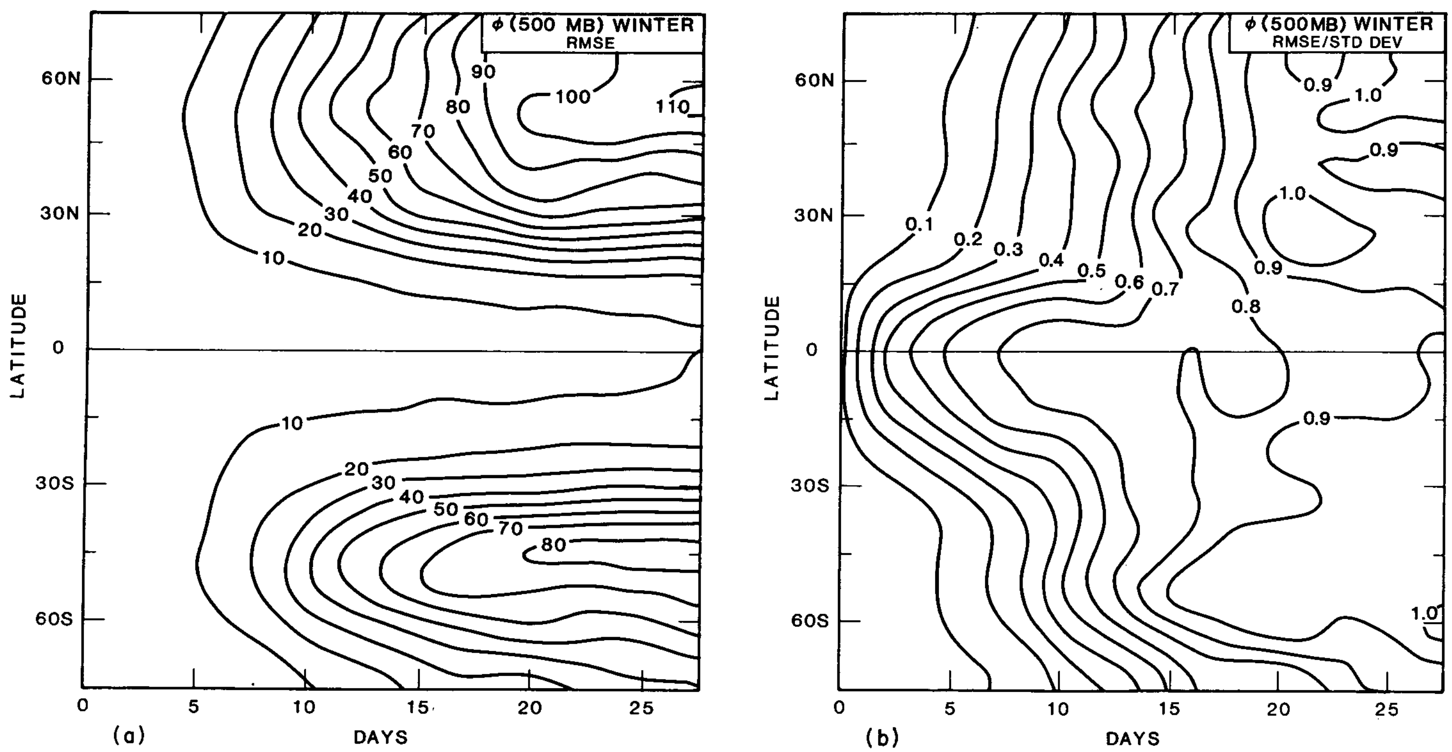


FIG. 1. Zonal average of geometric mean error (RMSE) and ratio (RMSE/STD) of root mean square and standard deviation of daily values for sea-level pressure. (a) RMSE and (b) RMSE/STD for six pairs of control and perturbation runs during winter; (c) RMSE and (d) RMSE/STD for three pairs of runs during summer.


φ (500 MB) WINTER
RMSE
LATITUDE
DAYS
(a)
φ (500MB) WINTER
RMSE/STD DEV
LATITUDE
DAYS
(b)

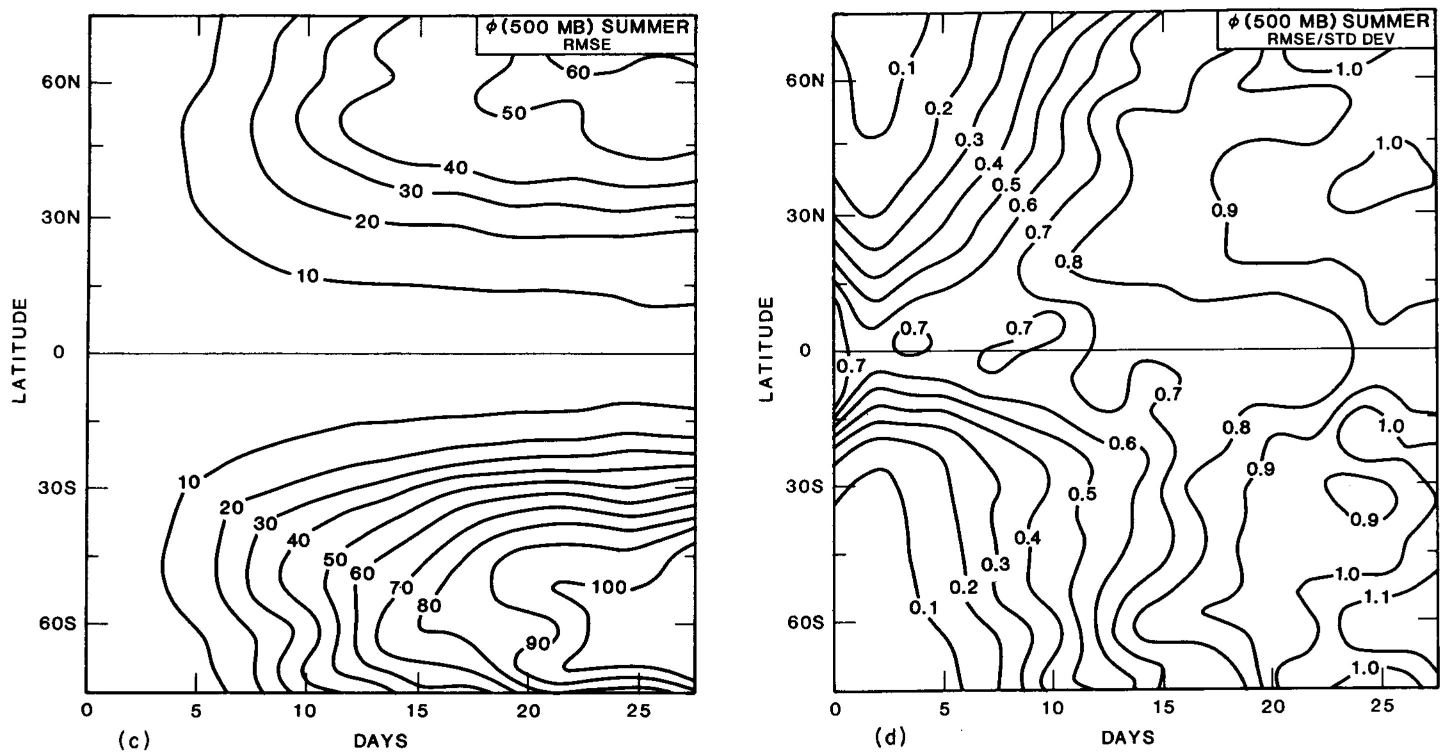
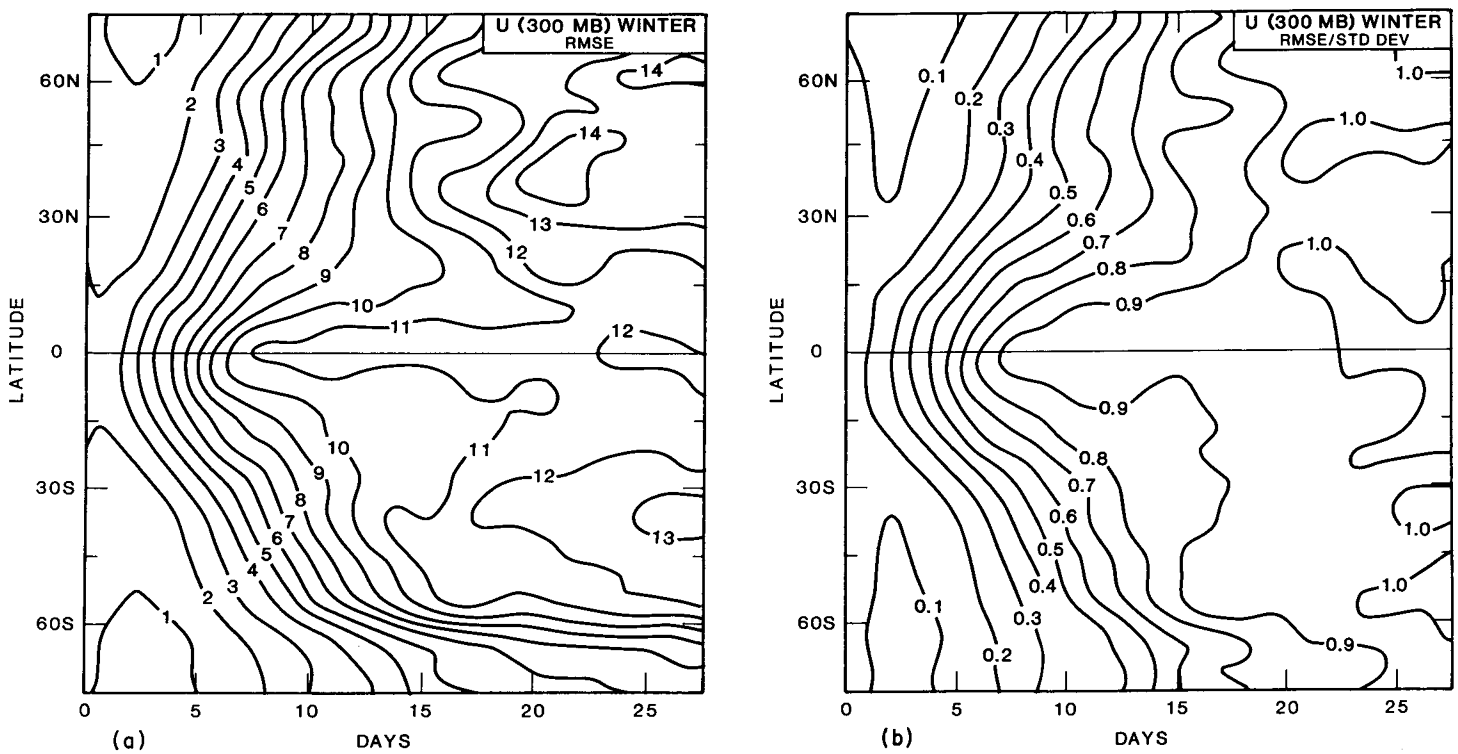


FIG. 2. Zonal average of geometric mean error (RMSE) and ratio (RMSE/STD) of root mean square and standard deviation of daily values for 500-mb height. (a) RMSE and (b) RMSE/STD for six pairs of control and perturbation runs during winter; (c) RMSE and (d) RMSE/STD for three pairs of runs during summer.


U (300 MB) WINTER
RMSE
U (300 MB) WINTER
RMSE/STD DEV
LATITUDE
DAYS
60N
30N
0
30S
60S
0 5 10 15 20 25
(a)
(b)

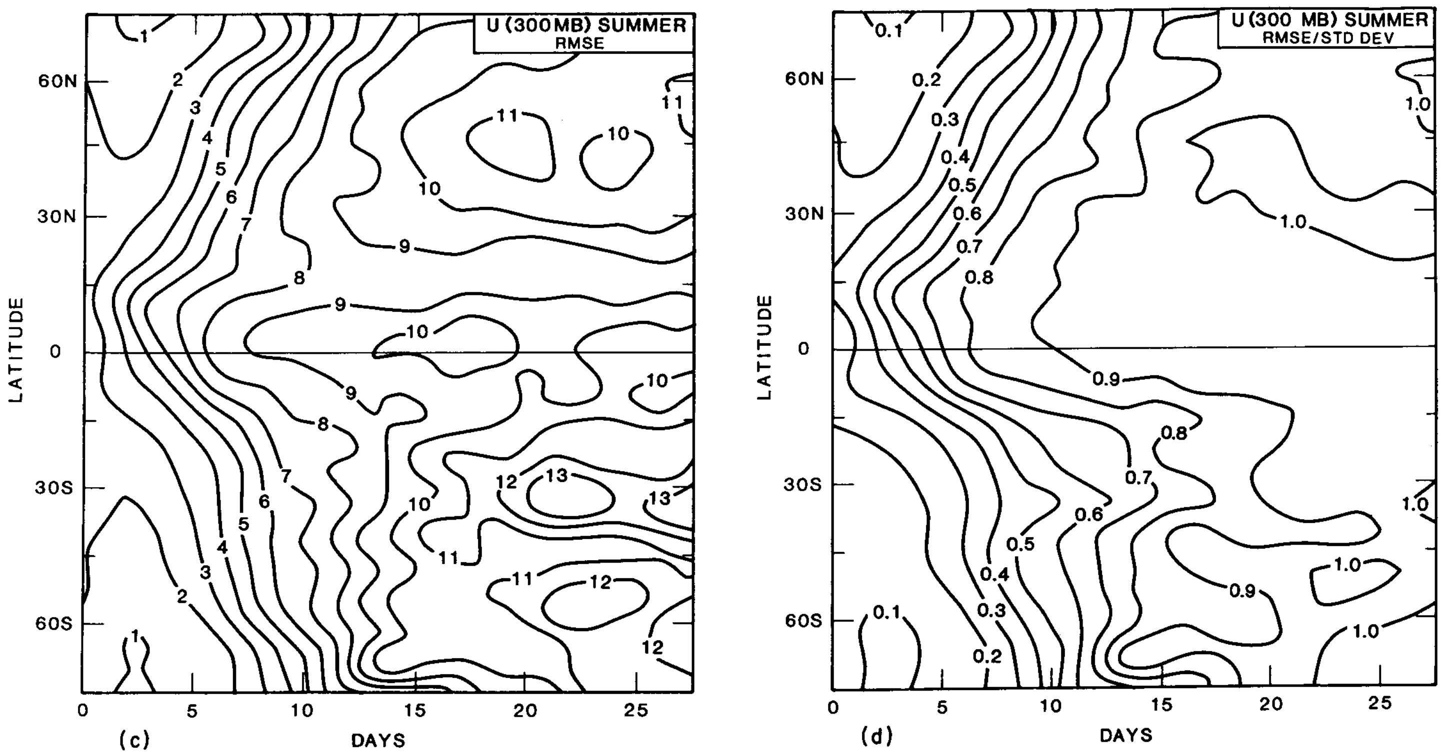


FIG. 3. Zonal average of geometric mean error (RMSE) and ratio (RMSE/STD) of RMS and standard deviation of daily values for 300-mb zonal wind. (a) RMSE and (b) RMSE/STD for six pairs of control and perturbation runs during winter; (c) RMSE and (d) RMSE/STD for three pairs of runs during summer.

which represent the rms error (a) and the ratio of error to standard deviation (b) for the winter season and similarly (c) and (d) for the summer season. The error in panel (a) represents the geometric mean error, the square root of average value of squares of E_{11}, E_{12}, E_{21}, E_{22}, E_{23}, and E_{31}; and the standard deviation (STD) used in panel (b) represents the square root of the sum of squares for all daily values of deviations for all integrations for that season. These figures show the dependence of predictability on latitude, season, and the weather variable in question.

The main conclusions from the results of the classical predictability studies described in the preceding section and presented by several other investigators are summarized in the following subsections.

2.3.1. Predictability of the Tropics and Extratropics. Since the growth rate and equilibration mechanisms for the tropical and mid-latitude instabilities are quite different, it is desirable to examine their predictability separately. It is recognized that there is considerable interaction between the tropics and mid-latitudes. However, since the time scales of growth and equilibration of synoptic-scale tropical disturbances is much smaller than that of the tropical–extratropical interactions, we are justified in examining their predictability separately.

In an earlier paper (Shukla, 1981b), the author has shown that the limit of deterministic predictability for the tropics is only 3 to 5 days compared to 2 to 3 weeks for the mid-latitudes. This is because the standard deviation of the day-to-day fluctuations (which is the saturation value of errors) is much smaller in the tropics and because the instabilities associated with the growth of the tropical disturbances are driven by moist convection, leading to larger growth rates than those of the dynamical instabilities of the mid-latitudes, which are driven by horizontal or vertical wind shear. Some of the earlier studies on predictability examined only the global or the hemispheric average rms error, and since the tropical errors are small in magnitude, results were dominated by the mid-latitude errors. Thus the results on tropical predictability were overlooked. Figure 4 from Shukla (1981b) shows the rms error averaged over 10° latitude belts centered at 6, 30, and 58°N for sea-level pressure. The equilibration value of the error is largest for 58°N and smallest for 6°N, reflecting the latitudinal variability of the daily standard deviation. It is also seen that the initial growth of error is the largest for the tropics. Thus, a combination of faster growth rate and smaller equilibration value makes the tropical regions of the globe the least predictable for day-to-day weather forecasting. This conclusion is not inconsistent with the current experiences in operational weather forecasting, where it has been found that the skill of tropical forecasts is not better than that of a persistence forecast, even at day 2 or 3.

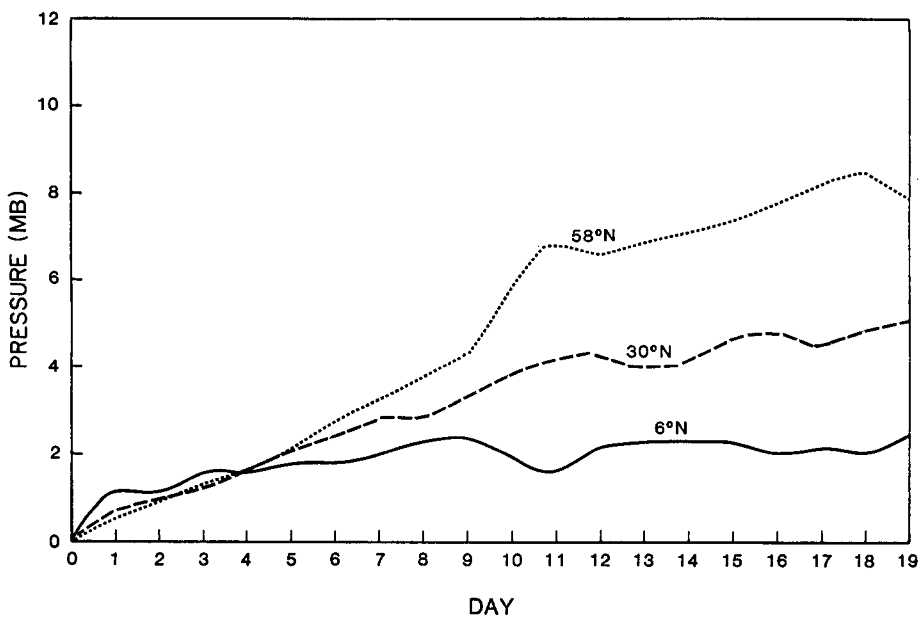


FIG. 4. Root mean square as a function of time between a summer control and a predictability run for sea-level pressure (mb). Solid line, dashed line, and dotted line refer to an average over 10° latitude belt centered at 6°N, 30°N, and 58°N, respectively. [From Shukla (1981b).]

This result is further supported by Figs. 1–3, in which it is seen that for each variable and each season, although the rms error is quite small in the tropics, the ratio of rms error to standard deviation is more than 0.5 within 1 to 5 days, whereas it takes about 5 to 12 days to reach that value for the mid-latitudes. These conclusions are based on an assumption of an idealized, initial error field over the whole globe. In reality, the observational network over the tropics is worse than that over the mid-latitudes, and the prospects for deterministic prediction of day-to-day weather in the tropics appear to be rather dim. The prospects for predicting space–time averages, on the other hand, are good. It will be shown in a later section that partly because of smaller day-to-day variability in the tropics, and partly because of strong influence of boundary conditions, the space–time averages are more predictable in the tropics than in the mid-latitudes.

2.3.2. Predictability of the Northern and Southern Hemispheres. In an article by Louis Purret (1976) in the NOAA magazine Smagorinsky conjectured, "I would suspect that there is a little less predictability in the Southern Hemisphere than there is in the Northern Hemisphere." Our results support this conjecture. The equilibration value for the error is higher in the Northern Hemisphere than in the Southern Hemisphere. The absence of large, stationary asymmetric boundary forcings in the Southern Hemisphere reduced the amplitude and variability of planetary scales, which in turn reduces the equilibration value of the errors. The possibility of enhancement of predictability due to the presence of stationary forcings was supported by another study, in which we examined the predictability of the idealized atmosphere of an ocean-covered Earth, and it was found to be smaller than the predictability of the atmosphere with mountains at the Earth's surface.

2.3.4. Predictability during the Winter and Summer Seasons. As shown in Figs. 1–3 and discussed earlier with Shukla (1984a), circulations during the Northern Hemisphere winter are more predictable than those during the summer. This provides a good illustration for the point that the error growth rate alone is not an adequate parameter to describe predictability. Although the error growth rate is higher during winter than in summer, the day-to-day variability during winter is also considerably larger than during summer, so that it takes longer for the initial error to reach its saturation value during the winter season. It should be pointed out that on the basis of the values of error growth rate alone, Charney *et al.* (1966) had erroneously concluded that circulations during the summer season might be more predictable than those during the winter season. Results of operational numerical weather prediction are not inconsistent with the conclusion arrived at by our predictability studies.

2.3.5. Predictability of Planetary and Synoptic Scales. Smagorinsky (1969) was the first to examine the predictability of different scales separately, and he correctly concluded that the larger scales are more predictable than the smaller scales. However, it was not clear whether a short sample of one case could resolve the predictability of various scales. We examined this question again (Shukla, 1981a) by using six pairs of control and predictability integrations, and the results are reproduced in Fig. 5. It is seen that for the latitude belt 40–60°N for 500 mb, predictability of planetary scales (wave numbers 0–4) is more than 4 weeks compared to about 2 weeks for synoptic scales (wave numbers 5–12). For wave numbers 13–36, predictability was only a few days. It is interesting to note, however, that the initial growth rate for both the planetary and synoptic scales is nearly the same. The doubling time of initial small errors is about 2.5 days and once the error has reached a value of about 25 m, the doubling time is close to 3 days. The higher predictability of the planetary waves is due to higher values of their amplitude and variability. If the error growth for the synoptic scales were attributed to the fast-growing dynamical instabilities at those scales, it will be of interest to determine

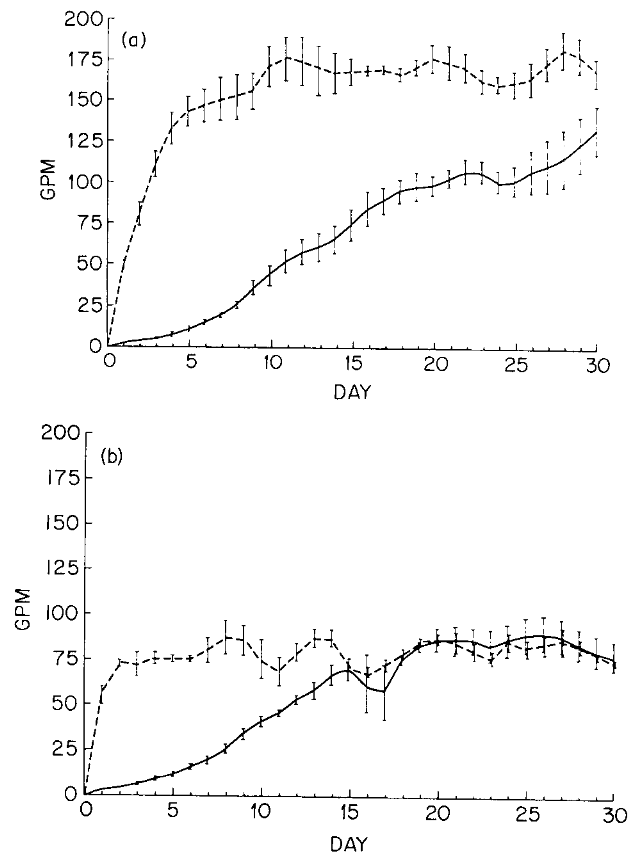


FIG. 5. Root-mean-square error (solid line) averaged for six pairs of control and perturbation runs and averaged for latitude belt 40–60°N for 500-mb height for (a) wave numbers 0–4 and (b) wave numbers 5–12. Dashed line is the persistence error averaged for the three control run. Vertical bars denote the standard deviation of the error values. [From Shukla (1981a). Reproduced with permission from the Journal of Atmospheric Sciences, a publication of the American Meteorological Society.]

the relative importance of planetary scale instability itself and the influence of synoptic scales in making the planetary scales unpredictable.

2.3.6. Predictability of High- and Low-Resolution Models. Smagorinsky (1969) compared the error growth for models with two different resolutions. The low-resolution model had 20 grid points between the equator and pole, with a grid size of 640 km at the pole and 320 km at the equator. The high-resolution model had 40 grid points between the equator and pole. He did not find any significant difference in the doubling time for initial errors for high- and low-resolution models, although the persistence error for the high-resolution model was clearly larger than that for the low-resolution model. Subsequent papers by Jastrow and Halem (1970) and Williamson and Kasahara (1971) showed that the doubling time of the error decreased for higher spatial resolution of the model and that the doubling time for the synoptic-scale errors was closer to 3 days rather than 5 days as reported by Smagorinsky.

2.3.7. Predictability of "Balanced" and "Unbalanced" Initial States. Daley (1980) has shown that the error growth in the rotational component of the flow could be rather small if the initial error was only in the gravitational component. This might explain, at least partially, the faster growth rate of spatially coherent initial errors compared to purely random errors. In a GCM, through convection and other diabatic processes, errors in gravitational component will also soon feed back to the rotational components.

Besides the question of dynamical balance between the mass and motion fields, there is also the question of consistency between the observed initial conditions used as input for a GCM integration and the boundary conditions of sea-surface temperature (SST), soil moisture, sea ice, snow, etc., used in the model. It is quite conceivable that this inconsistency can be an additional source for error growth. However, it can be argued that in classical predictability studies with GCMs, the errors due to imbalances in the initial and boundary conditions should not be too large because of the "perfect model" assumption in which time growth of a small initial perturbation is examined.

We have attempted to address this question by repeating our predictability error growth calculations for an initial condition that was obtained after integrating a GCM for 30 days. Figure 6 shows the 500-mb rms error averaged for 40–60°N for six pairs of integrations during the winter season. The curves labeled E_{11}, E_{21}, and E_{31} refer to the rms error between a control run that started from an *observed* initial condition and a perturbation run. In each case, the control run was extended up to 60 days. At the end of the 30 days, similar random perturbations were introduced and the rms error calculated for 15 days. The curves labeled E'_{11}, E'_{21}, and E'_{31} in the lower part of Fig. 6 show the time growth of rms error for the three cases corresponding to E_{11}, E_{21}, and E_{31}, respectively. It is quite clear that the growth rates of error for integrations starting from day 30 of the control run are smaller than those from the observed initial conditions. This result suggests that improvements in methods for analysis and initialization hold promise for improvements in short-range weather prediction. It is difficult to determine whether the above reduction in the error growth was due to a better balance between the mass and motion fields (i.e.,

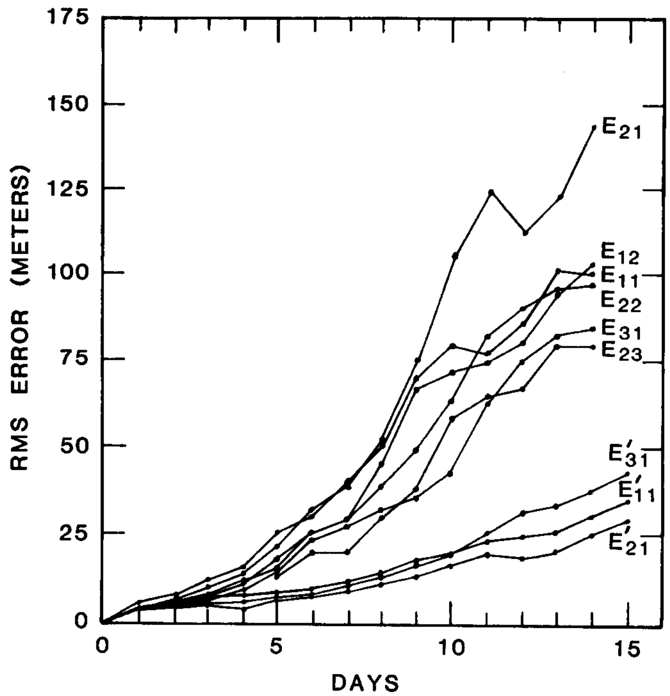


FIG. 6. Root-mean-square error between control and perturbation runs averaged for latitude belt 40–60°N for 500-mb height. E_{11} and E_{12} are for the initial condition of 1 January 1975 two perturbations; E_{21}, E_{22}, and E_{23} are for 1 January 1977 three perturbations; and E_{31} is for 1 January 1978 one perturbation. E'_{11}, E'_{21}, and E'_{31} are for the perturbation being superposed on day 30 of control runs used in E_{11}, E_{21}, and E_{31}.

reductions in the gravitational component of the flow) or due to reductions in the inconsistency between the atmospheric flow and the underlying boundary conditions.

2.3.8. Dependence of Predictability on the Structure of Large-Scale Flow. Since the nature of error growth is determined by the nature of the dominant instabilities, which in turn depends on the dynamical structure of the flow fields, it is natural to expect that different initial conditions will show different predictability characteristics. Figure 6 shows the time series for the error field for six different pairs of control and predictability integrations. The random perturbation in the initial conditions had the same statistics (zero mean and standard deviation of 3 m s^{-1} for u and v components) for all the cases. The curves labeled E_{11}, E_{21}, and E_{31} refer to three different initial conditions. It is clearly seen that the error growth rate depends on the initial conditions. It is therefore natural to expect that, for this reason alone, the skill of numerical weather prediction will not be the same on each day, although in operational forecasting the quality and quantity of input observations could also vary from one day to the other.

2.3.9. Dependence of Predictability on the Structure and Magnitude of the Initial Error. Figure 6 also shows the effects of structure of the initial error field on predictability. For example, the curves labeled E_{21}, E_{22}, and E_{23} not only had identical planetary- and synoptic-scale flow, but even the statistics of the random perturbation to the initial conditions were the same. The only differences were the actual grid-point values of the random error, and that was sufficient to produce differences in the rate of error growth. This again provides at least a partial explantation for the day-to-day changes in the skill of operational numerical weather prediction.

In another set of experiments, we smoothed the initial random error over the oceans, corresponding to the assumption that oceanic errors are more systematic than those over land, which can be assumed to be mostly instrumental and therefore random. The magnitudes of the individual grid-point values were adjusted to keep the standard deviation of the error field the same as that for the globally random error. We found that the growth rate for spatially coherent initial error was larger than that for the random initial error. These results raise some interesting questions about the relative virtues of possible observing systems with uniform but large error over the whole globe compared to the existing system of relatively smaller error over the land and larger errors over the oceans.

The dependence of error growth rate on the size of the error itself was well recognized from the earlier pioneering works of Lorenz (1969a) and Smagorinsky (1969). Lorenz had shown that very small errors confined to very small scales grow much faster than larger errors at larger scales. Smagorinsky had also shown that the smaller the error, the faster the growth rate. There is a smallest scale that can be resolved by a GCM, and GCM calculations can address the question of error growth only for scales larger than that. In earlier predictability studies by Shukla (1981b, 1984a), it was found that the doubling time for tropical errors was quite large. This was merely a manifestation of the fact that the error over the tropics had almost reached their saturation value, and there was no possibility for the error to double again. For example, an initial rms error of 1°C in temperature, of 3 mb in sea-level pressure (Jastrow and Halem, 1970), was already comparable to the saturation value of the error in the tropics.

2.3.10. Predictability of Different Variables. Figures 1–3 show the results for predictability of sea-level pressure, geopotential height, and wind field, respectively. Since these fields are dynamically coupled, it is reasonable to see small differences in predictability characteristics for different variables, especially in the mid-latitudes where the dynamical coupling is quite strong. However, in the tropics the mass and motion fields are not strongly coupled, and in relation to the mid-latitudes, the day-to-day changes in sea-level pressure and temperature are quite small compared to the wind field. Following the criteria of error growth and error equilibration, the wind field in the tropics seems to be a little more predictable than the pressure or temperature field.

We had carried out similar predictability calculations for rainfall for the four summer runs reported by Charney and Shukla (1977) and Shukla (1981b), and we found that deterministic predictability for rainfall was even smaller than that for the circulation variables.

Lorenz (1982) has examined the 10-day forecasts produced by the European Centre model for 100 consecutive days and has calculated error growth between model integrations starting from consecutive days. Assuming that the analyzed fields on two consecutive days do not differ greatly, the rms error between two integrations will give estimates of error growth similar to the ones obtained in the classical predictability studies. While the doubling time for the smallest observed error of 25 m was about 3.5 days, Lorenz estimated the doubling time for small errors to be about 2.5 days. He further introduced empirical methods for improving the forecast error, but the improvement in forecast was also accompanied with a decrease in the doubling time for the small errors.

In this paper, Lorenz introduced the concept of lower and upper bounds on predictability. The lower bound on predictability refers to the *minimum accuracy* with which forecasts can be made, and the upper bound on predictability refers to the *maximum error* for forecasts at a given range. The performance of the current operational numerical weather prediction (NWP) models, therefore, gives an estimate of the lower bound of predictability (i.e., there is a possibility of doing better than that), while classical predictability studies give an estimate of the upper bound on predictability (i.e., we cannot do better than that). Improvements in NWP models and observing systems will lead to a larger lower bound and a smaller upper bound on predictability.

Lorenz estimated the lower and upper bounds on predictability for the European Centre model and estimated that even without further improvements in 1-day forecast, 10-day forecasts as good as the present 7-day forecasts can be made. The range of predictability could be further extended by 2 more days by halving the 1-day forecast error.

3. Predictability of Space–Time Averages

It has generally been recognized that although the upper limit for prediction of instantaneous weather lies somewhere between 1 and 3 weeks, space–time averages of weather elements could be predicted for periods beyond this limit. During the last 5 to 10 years, a large body of observational and numerical modeling works have been reported that have collectively established a physical basis for dynamical prediction of monthly and seasonal averages. In the following sections, we shall present a brief review of the recent work and several remaining outstanding problems that need to be addressed.

We shall first describe the mechanisms for the variability of monthly and seasonal averages and then examine the potential for their predictability. A convenient conceptual framework to describe the mechanisms of variability is afforded by the following two categories (Shukla, 1981a): (1) internal dynamics and (2) boundary forcings.

(1) *Internal dynamics:* Even if there were no changes in the external forcing functions and even if the boundary conditions at the Earth's surface were constant, there will be changes in day-to-day weather and in monthly and seasonal averages. These will occur due to the combined effects of dynamical instabilities and nonlinear interactions among different scales of motions, the interaction of fluctuating zonal winds with quasi-stationary mechanical and thermal forcings, etc. Monthly and seasonal averages can be made different by sampling different segments of this evolving nonperiodic flow. We shall further describe the predictability of internal dynamics in Subsection 3.1.

(2) *Boundary forcings:* Slowly varying boundary forcings due to anomalies of sea-surface temperature, soil moisture, sea ice, snow, etc., can produce anomalous sources and sinks of heating and moisture that can influence the amplitudes and phases of planetary-scale waves, which, in turn, can change the location, intensity, and frequency of synoptic-scale disturbances. Since we shall confine our discussion only to the monthly and seasonal time scales, we shall not consider the external forcings due to fluctuations in solar or other extraterrestrial energy sources.

There has been considerable interest in determining the relative importance of the internal dynamics and boundary forcings for the observed interannual variability of monthly or seasonal averages. Due to the strong coupling between the internal dynamics and boundary forcing mechanisms, it is not possible to determine their relative roles by analyzing observed data without making some drastic assumptions about the role of one or the other [see, for example, the paper and correspondence by Madden, 1976; Shukla, 1983a; Madden, 1983]. One possible way to gain some insight into the problem is by idealized numerical experiments with GCMs, which can be integrated with and without boundary condition anomalies. Intercomparison between such integrations can suggest the possible role of the boundary conditions. Some attempts have been made in this direction (Charney and Shukla, 1977, 1981; Lau, 1981), but the conclusions remain questionable because they were based on comparison of numerical simulations with actual observations of the atmosphere, rather than on the comparison of two simulations (with and without the anomalous boundary forcing) from the same model. If model simulations without the changing boundary conditions can produce variances comparable to the observations, there is no justification to conclude that the boundary conditions are not important because internal dynamics can be overemphasized in such a hypothetical simulation.

Numerical experiments with several GCMs have been carried out to determine the influence of boundary forcings due to regional anomalies of SST or soil moisture, etc., and they suggest an important role of boundary forcings for interannual variability of monthly and seasonal averages.

3.1. Dynamical Predictability

Since monthly and seasonal average anomalies are primarily determined by low-frequency, planetary-scale flow patterns, predictability of planetary waves will crucially determine the predictability of space–time averages. Predictability of planetary scales can be limited either by the instabilities at their own scale or by their interactions with highly unstable synoptic scales. If the growth and decay of the planetary waves were completely determined by their interactions with the synoptic scales, there will be no real hope for predictions beyond the limits of deterministic predictability. However, there is no evidence that that is the case for the atmospheric flows. A quantitative determination of the role of synoptic scales in the evolution and equilibration of planetary scales is quite essential to realize the potentials of dynamical predictability.

Dynamical predictability of monthly means was investigated by the author using the GLAS climate model (Shukla, 1981a). The model was integrated for 60 days with three different observed initial conditions during three different years. These were supposed to represent large differences in the initial conditions. Six additional 60-day integrations were made after changing the observed initial conditions by superposition of random perturbations with root mean square of 3 m s^{-1} in u and v components. These were supposed to represent small differences in the initial conditions. It was hypothesized that for a given averaging period, if the rms error among the time averages predicted from largely different (observed) initial conditions became comparable to the rms error among time averages predicted from small differences (random perturbations) in the initial conditions, the time averages would be considered to be unpredictable. It was found that the variances among the first 30-day means for largely different initial conditions were significantly different from the variances due to random perturbations, and it was concluded that the first 30-day means were dynamically predictable. It was also found that the next 30-day means (days 31–60) were not dynamically predictable. It has been pointed out by Dr. Y. Hayashi of the Geophysical Fluid Dynamics Laboratory (GFDL) (personal communications) that based on analysis of variance presented in our paper, it is not appropriate to declare the lack of predictability for second 30-day means, and the possibility remains that even the second 30-day means could be dynamically predictable. This was an idealized study, in the spirit of classical predictability studies for day-to-day weather prediction, and actual forecast experiments must be carried out to determine the predictability of monthly or seasonal averages.

Miyakoda *et al.* (1983) have presented an example of a dynamical prediction for 30 days. This is an excellent illustration for potential of extended-range, dynamical predictability using advanced models for dynamics and physics. It is natural to expect that all initial conditions will not be equally predictable, but even one good example provides encouragement to pursue it further. There is some indication that the blocking situations have relatively greater predictability (Bengtsson, 1981).

3.2. Boundary-Forced Predictability

If the changes in the boundary conditions at the Earth's surface were able to produce changes in the atmospheric circulation that were large and coherent enough to be distinguishable from the natural variability of the internal dynamics, boundary forcings would provide additional potential for predictability of space–time averages. Based on the correlations between the observed changes in boundary conditions, atmospheric circulation, and rainfall and also on GCM sensitivity studies with prescribed changes in the boundary conditions, it has been suggested that under favorable structures of the large-scale flow and appropriate locations of the boundary anomaly, significant and predictable changes in the atmospheric flow can indeed be produced. Changes in the boundary conditions produce local changes in the surface heat flux and moisture convergence, which in turn produce deep heat sources that can influence the remote, as well as the local, circulation.

A summary of several such experiments carried out with the GLAS climate model has been presented in Shukla (1982, 1984b). Similar experiments have been, and are being, carried out at several other GCM groups around the world. However, to our knowledge not a single case of model integration has been reported in which observed global boundary conditions of all the slowly varying fields (SST, soil moisture, sea ice, snow, etc.) were used to integrate the observed initial conditions. We hope that the encouragement provided by the results from regional boundary anomalies will lead to study of predictability for global boundary conditions.

We shall present here, as an illustration, one example of a sensitivity study with the GLAS climate model using the observed SST anomalies over tropical Pacific during the winter of 1982–1983 (Fennessy *et al.*, 1985). In Fig. 7a, the observed SST anomaly during January 1983 was added to the climatological SST to integrate the model for 60 days. This integration is referred to as the "anomaly run," and a similar integration with climatological SST is referred to as the "control run." Such pairs of integrations were made for three different initial conditions. The difference between the anomaly and control runs averaged for three pairs for the period days 11–60 for precipitation (b) and the rainfall anomaly calculated from the observed outgoing long-wave radiation for 1982–1983 winter (c) is also shown in Fig. 7. The outgoing long-wave radiation anomalies are changed to rainfall anomalies by using empirical relations developed by Arkin (1983). It is gratifying that the model calculations have been able to simulate rather well the location as well as intensity of rainfall anomaly. Similar results were obtained by several other modeling groups (Liege Colloquium on Hydrodynamics, May 1984) who used similar SST anomalies, although quantitative differences were found for different models with different parameterizations of boundary layer and moist convection.

A comprehensive study of the role of tropical SST anomalies has been carried out by Lau and Oort (1985) in which they have examined a 15-year integration of the GFDL model, with the observed SST anomalies over the equatorial Pacific. The results are most remarkable, especially for the tropics. In the simulation without the SST anomalies (Lau, 1981), there was no evidence for the planetary-scale seesaw of surface pressure

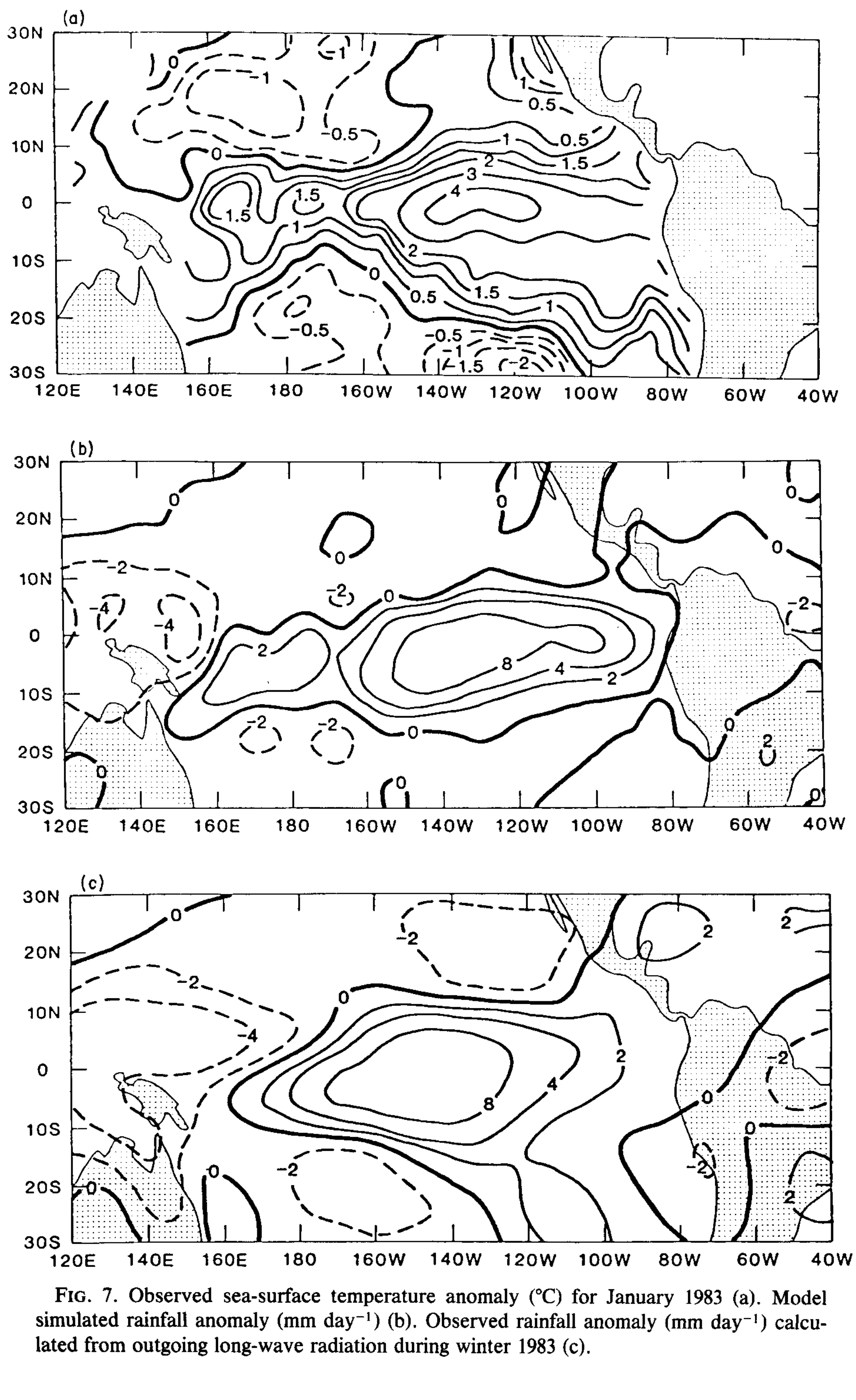


FIG. 7. Observed sea-surface temperature anomaly (°C) for January 1983 (a). Model simulated rainfall anomaly (mm day^{-1}) (b). Observed rainfall anomaly (mm day^{-1}) calculated from outgoing long-wave radiation during winter 1983 (c).

referred to as the Southern Oscillation, which is one of the most dominant modes of the tropical variability; whereas in the simulation with the SST anomalies, the Southern Oscillation is simulated remarkably well. The observed correlations between the tropical SST anomalies and mid-latitude circulation is also simulated remarkably well in the 15-year simulation with SST anomalies. These results, combined with some of Philander and Seigel (1985) and other ocean modeling groups on simulation of the oceanic circulation and SST that use the prescribed atmospheric wind stress forcing from the observations, suggest that the predictability of the coupled ocean–atmosphere system could be larger than the predictability of the atmosphere alone.

3.3. Prospects for Dynamical Extended-Range Forecasting

A large body of observational, theoretical, and GCM results collectively suggests that there are good prospects for dynamical prediction of monthly and seasonal averages. Miyakoda *et al.*, (1983) have already demonstrated the existence of extended-range predictability. Recent works by Miyakoda and his group at GFDL has shown that limits of predictability of time-averaged flow can be further extended by improving the dynamical model, especially the physical parameterizations of the model.

The factors that have provided hope for dynamical extended-range forecasting (DERF) can be summarized as follows:

(1) The planetary scales are more predictable than the synoptic scales.

(2) Slowly varying boundary conditions at the Earth's surface can produce significant and predictable changes in the time-averaged atmospheric circulation.

(3) There exists a reasonable conceptual framework to understand the structure and evolution of atmospheric variability at medium- and long-time scales.

(4) Global atmospheric GCMs are now able to simulate well the important features of the mean and transient components of atmospheric circulation.

(5) Tropical ocean GCMs are also able to simulate the response of the prescribed atmospheric forcing reasonably well.

(6) Advances in computer technology, space observations, communication, and data-processing techniques make it feasible to carry out a large number of integrations of atmospheric and oceanic GCMs by using real-time observations of global initial and boundary conditions.

It is recognized that extended-range integrations with the atmospheric GCMs still show systematic errors (climate drift), and models must be continuously improved to reduce the climate drift. However, as previously suggested by the author (Shukla, 1983b), an appropriate mean climate drift can be subtracted from the predicted time averages to reduce the forecasting error.

4. Some Outstanding Problems

Although it is possible to list a very large number of problems that need to be understood in the general area of predictability, we have chosen to comment only on the following three problems, which require further discussion.

4.1. Mean (Climate Drift) and Transient Predictability

Operational numerical weather prediction centers are constantly trying to improve their day-to-day forecasts by improving the parameterizations of model physics, increasing the resolution of the model, and improving the quality of data and data-analysis techniques. One of the important sources of error has been referred to as the "systematic error," which is considered to arise mainly due to the "climate drift" of the model. The systematic error is generally defined as the average error for a large number of forecasts. Wallace *et al.* (1983) have shown that changes in the mountain heights (envelope orography) produced clear reductions in the systematic error. It is quite likely that systematic errors can be further reduced by changing the diabatic heating and dissipation mechanisms in the models. Climate drift for NWP models can be diagnosed by making long-term integrations of the forecast model with appropriate boundary conditions and comparing the simulated climate with observations.

Systematic errors can also be reduced by statistical corrections to the forecasts (Faller and Lee, 1975; Lorenz, 1977). These methods have not been too popular with the operational NWP centers because they do not provide any physical insight, and although it might reduce the forecast error, it is not possible to understand either the cause of the error or reasons for improvement. The general assumption has been that improving the forecasts by changing the model adds to our understanding, whereas changing the forecast empirically does not.

Systematic errors of NWP models have been examined mainly for flow parameters because it is possible to verify against observations. Similar analysis of systematic errors for rainfall, cloudiness, and vertical distributions of heating will be useful for determining the sources of systematic errors. Based on predictability studies for the tropics and a limited number of tropical forecasts, it was suggested (Shukla, 1981b) that systematic errors in tropical heat sources are so large and develop so fast that short- and medium-range forecasts could be improved by prescribing the large-scale tropical heat sources rather than calculating them by using model dynamics and physics. Some recent work at the European Centre (M. Tiedtke, personal communication) seems to support this conjecture.

The extensive experience of operational numerical weather prediction and the predictability studies described in an earlier section suggest that some initial conditions are clearly more predictable than the others. Assuming that the quality and quantity of data do not change from one day to the other, and assuming that the model remains the same, the only factor that remains to be considered to explain the transient behavior of predictability is the dynamical structure of the initial state. An outstanding problem in weather forecasting is to identify the important features of the initial state that might reveal the predictability properties of the flow. For example, if it were true that in a large number of cases highly amplified presistent blocking ridges were predictable for longer periods [as has been suggested by Bengtsson (1981)], it will be possible to attach a higher confidence to a prediction that maintained the initial amplified blocking ridge. These considerations suggest a need for detailed synoptic study of predictability. We are not aware of any comprehensive study of predictability as a function of the synoptic structure of flow. For example, is it likely that location and intensity of the jet streams, intertropical convergent zones (ITCZ), or Walker cells could affect predictability? Although forecasts ultimately degrade either due to inadequacies of the models or the initial data, if there were significant relationships among the large-scale features of the initial flow and its associated predictability, such relationships can be exploited to improve the operational predictions.

It should be noted that any statistical correction to the forecast based on past records will be helpful only in reducing the systematic errors and will not affect the errors that depend on the structure of the initial state.

4.2. Observational Errors and Model Errors

It is rather interesting that the very first paper on predictability (Thompson, 1957) discussed here was motivated, at least in part, by the considerations of relative importance of good initial data and good models for weather forecasting. This question is equally, if not more, valid today as it was about three decades ago. During the past 10 years, there has been a large increase in the use of observations from satellites to define the initial state for NWP. However, major improvements in short- and medium-range forecasting appear to have come from better models. Arpe *et al.* (1985) have suggested that the observational errors are the dominant factor only for first 1 to 2 days of the forecast, after which the errors are dominated by model errors. This is a rather tricky question because model errors also contribute to the amplification of the initial observational errors, and we would have no way to know the time taken for the model-produced errors to become large if there were *no* observational errors to start with. Determination of relative roles of observational and model errors for short- and medium-range forecasting needs further work.

4.3. Predictability of Predictability

It is well established that there is an upper limit on predictability of weather. However, it is also evident that within that limit there can be changes in predictability that depend on the structure of the initial state. Earlier we discussed the possibility of identifying the important synoptic features of the flow that might provide some clue to the accuracy of the forecasts. Some formal procedures have also been suggested (Epstein, 1969; Leith, 1974; Hoffman and Kalnay, 1983) to make a quantitative determination of the reliability of the forecast. Leith (1974) suggested that instead of one forecast from a given initial condition, several (say about eight) forecasts from the same initial conditions can be prepared by integrating the model with slightly perturbed initial conditions. Divergence among these various forecasts will be a measure of the instability of the initial state and therefore a possible measure of the reliability of the forecast. Hoffman and Kalnay (1983) suggested that rather than perturbing the initial state (either randomly or systematically), forecasts from successively observed initial states can be combined, with suitable weighting functions, to produce a better forecast and to estimate the reliability of the forecasts. This suggestion eliminates the need for additional model integrations because such forecasts are produced routinely.

If divergence among predictions from slightly different initial conditions were a good measure of the reliability of the predictions, these methods would provide not only a prediction (average of all integrations) of the flow, but also a prediction of the predictability of the flow. Although these methods have not yet been tried operationally, chances of their success will remain limited for forecast models with large systematic errors and climate drift. Persistence of atmospheric flows for several days may also reduce the advantages of using observed initial conditions during that period. For example, forecasts from several initial states within a 2- to 3-day period could be very similar, but each could be very different from the observations.

An examination of the forecast errors for operational NWP models suggests a tendency for the persistence of correlation between forecasts from consecutive days. Since the decorrelation time for the atmospheric flows is about 5 days, this further suggests that predictability depends on circulation regime. This also suggests that errors of predictions from initial conditions in the immediate past can be a useful guide for predictability on a given day. It should be noted that the statistical correction techniques that use the error history for a large number of forecasts in the past will not be able to take into account the transient nature of predictability. For producing forecasts from a given day, the current operational NWP methods do not use any information from either the analyses or the forecasts during the last several days. The only exception is the use of a short-term forecast as a first guess for the analysis. Optimal interpolation techniques require the use of spatial structure functions, which are derived from past data over a much longer period than the decorrelation time of the atmosphere. Since weather forecasting is considered to be an initial value problem and since the prediction equations are highly nonlinear, there is no compelling reason to use the past history of forecast errors. However, considering the inadequacies of the models as well as of the observations, it should be possible to use information on the deficiencies of the forecasts from initial conditions in the recent past to improve short- and medium-range forecasts.

5. Concluding Remarks

There is a complete agreement among scientists that the instantaneous weather is not predictable at infinite range. In fact, there is no serious challenge to the statement that the instantaneous weather is not predictable even beyond 2 to 3 weeks. As implied by the work of Lorenz (1982), it may be convenient to discuss the predictability at day 1 and at days beyond day 1 separately. Lorenz has shown that even if we could not improve predictions at day 1, there is potential for improving predictions beyond day 1. There is little room for disagreement on this point. However, the question of predictability at day 1 needs more discussion. Is it possible to make significant reductions in the current 1-day forecast errors? Lorenz's work implies that it is highly unlikely. The argument is as follows: There are, and there always will be, scales of motion unresolved by the NWP models, and even if there were no errors in the resolved scales, the errors of the unresolved scales would quickly make the resolved scales unpredictable. The underlying assumption is that by a reduction in the grid size of the model and by the introduction of more sophisticated and complex physical processes, the growth rates of errors will be increased, and therefore forecast errors at day 1 will remain nearly same as that for the current models. We are not quite sure about the validity of these conjectures because there is no evidence that the current 1-day forecast errors are mainly due to the unresolved scales. In fact, there is some evidence to the contrary, viz., that the current 1-day forecast errors are also due to errors in the observations at the synoptic scales. It does not appear to be unreasonable to expect that by improving the current NWP models and the description of the initial state at the current resolution, 1-day forecast errors could be reduced without increasing the growth rate of the error. Estimates of the rates at which the unresolved scales influence the synoptic scales, and thereby the planetary scales, have been made only for simple models that do not have forcing and dissipation mechanisms, and there is no guarantee that these estimates will hold good for more realistic models of the atmosphere with well-defined forcing functions.

The lower curves labeled E'_{11}, E'_{21}, and E'_{31} in Fig. 6 suggest that the growth rate of the initial error can be reduced considerably by improving the initial conditions and reducing the inconsistency between the initial conditions and the boundary conditions. In the opinion of this author, we are not yet at a stage where the problems due to unresolved scales and the intrinsic instability of the flow are the primary factors contributing to the forecast error at day 1 or 2. It is not unlikely that the errors in defining the synoptic and planetary scale itself and in parameterizing the diabatic forcings at the synoptic and planetary scale are the primary reasons for the short-range operational NWP forecast errors. It is also of interest to note that the standard deviation of short-range forecast errors does not show any preferred areas of maxima in the storm track regions, which would have been expected from the classical predictability arguments of fast growth rates in those regions.

Recent works on the predictability of space–time averages (Miyakoda, personal communication) indicate that the prospects for dynamical long-range forecasting of monthly and seasonal averages are quite good.

However, this needs to be substantiated by a reasonably large number of actual forecasts.

REFERENCES

Arkin, P. A. (1983). Ph.D. Thesis, Univ. of Maryland, College Park.

Arpe, K., Hollingsworth, A., Lorenc, A. C., Tracton, M. S., Uppala, S., and Kallberg, P. (1985). *Q.J.R.Meteorol. Soc.* **111,** 67–101.

Baumhefner, D. P. (1984). *In* "Predictability of Fluid Motions" (G. Holloway and B. J. West, eds.) pp. 169–180. American Institute of Physics, New York.

Bengtsson, L. (1981). *Tellus* **33,** 19–42.

Charney, J. G., and Shukla, J. (1977). *Symposium on Monsoon Dynamics, New Delhi, 1977.*

Charney, J. G., and Shukla, J. (1981). *In* "Monsoon Dynamics" (J. Lighthill and R. Pearce, eds.) pp. 99–109. Cambridge Univ. Press, London and New York.

Charney, J. G., Fleagle, R. G., Riehl, H., Lally, V. E., and Wark, D. Q. (1966). *Bull. Am. Meteorol. Soc.* **47,** 200–220.

Daley, R. (1980). *Mon. Weather Rev.* **108,** 1719–1735.

Epstein, E. S. (1969). *Tellus* **21,** 739–759.

Faller, A. J., and Lee, D. K. (1975). *Mon. Weather Rev.* **103,** 845–855.

Fennessy, M., Marx, L., and Shukla, J. (1985). *Mon. Weather Rev.* **113,** 858–864.

Gutzler, D. S., and Shukla, J. (1984). *J. Atmos. Sci.* **41,** 177–189.

Hoffman, R. N., and Kalnay, E. (1983). *Tellus* **35A,** 100–118.

Jastrow, R., and Halem, M. (1970). *Bull. Am. Meteorol. Soc.* **51,** 490–513.

Lau, N. C. (1981). *Mon. Weather Rev.* **109,** 2287–2311.

Lau, N. C., and Oort, A. (1985). *In* "Coupled Atmosphere–Ocean Models," Elsevier., Amsterdam (in press).

Leith, C. E. (1965). *In* "Methods in Computation Physics," (B. Alder and S. Fernbach, eds.) Vol. 4, pp. 1–28. Academic Press, New York.

Leith, C. E. (1971). *J. Atmos. Sci.* **28,** 148–161.

Leith, C. E. (1974). *Mon. Weather Rev.* **102,** 409–418.

Leith, C. E., and Kraichnan, R. H. (1972). *J. Atmos. Sci.* **29,** 1041–1058.

Lilly, D. K. (1969). *Phys. Fluids, Suppl. II,* 24–249.

Lorenz, E. N. (1963). "*Trans N.Y. Acad. Sci.,*" **25,** 409–432.

Lorenz, E. N. (1965). *Tellus* **17,** 321–333.

Lorenz, E. N. (1969a). *Tellus* **21,** 289–307.

Lorenz, E. N. (1969b). *J. Atmos. Sci.* **26,** 636–646.

Lorenz, E. N. (1973). *J. Appl. Meteorol.* **12,** 543–546.

Lorenz, E. N. (1977). *Mon. Weather Rev.* **105,** 590–602.

Lorenz, E. N. (1982). *Tellus* **34,** 505–513.

Lorenz, E. N. (1984). *In* "Predictability of Fluid Motions" (G. Holloway and B. J. West, eds.) pp.133–139. American Institute of Physics, New York.

Madden, R. A. (1976). *Mon. Weather Rev.* **104,** 942–952.

Madden, R. A. (1983). *Mon. Weather Rev.* **111,** 586–589.

Manabe, S., Smagorinsky, J., and Strickler, R. J. (1965). *Mon. Weather Rev.* **93,** 769–798.

Mintz, Y. (1964). *WMO-IUGG Symp. Res. Dev. Aspects Long-Range Forecasting* **66,** 141–155.

Miyakoda, K., Smagorinsky, J., Strickler, R. F., and Hembree, G. D. (1969). *Mon. Weather Rev.* **97,** 1–76.

Miyakoda, K., Gordon, C. T., Caverly, R., Stern, W. F., Sirutis, J., and Bourke, W. (1983). *Mon. Weather Rev.* **111,** 846–869.

Philander, S. G. H., and Seigel, A. D. (1985). *In* "Coupled Atmosphere–Ocean Models," Elsevier, Amsterdam (in press).

Purrett, L. (1976). *NOAA Mag.* 16–17.

Robinson, G. D. (1967). *Q. J. R. Meteorol. Soc.* **43,** 409–418.

Shukla, J. (1981a). *J. Atmos. Sci.* **38,** 2547–2572.

Shukla, J. (1981b). *NASA Tech. Memo. 83829.* Goddard Space Flight Center, Greenbelt, Maryland.

Shukla, J. (1982). *NASA Tech. Memo. 85092.* Goddard Space Flight Center, Greenbelt, Maryland.

Shukla, J. (1983a). *Mon. Weather Rev.* **111,** 581–585.

Shukla, J. (1983b). *Proc. WMO-CAS/NSC Expert Study Conf. Long-Range Forecasting, Princeton,* **1,** 142–153.

Shukla, J. (1984a). *In* "Predictability of Fluid Motions" (G. Holloway and B. J. West, eds.) pp. 449–456. American Institute of Physics, New York.

Shukla, J. (1984b). *In* "Problems and Prospects in Long and Medium Range Weather Forecasting" (D. M. Burridge and E. Kallen, eds.) pp.155–206. Springer-Verlag, New York.

Shukla, J., Straus, D., Randall, D., Sud, Y., and Marx, L. (1981). *NASA Tech. Memo. 83866.* Goddard Space Flight Center, Greenbelt, Maryland.

Smagorinsky, J. (1963). *Mon. Weather Rev.* **91,** 99–164.

Smagorinsky, J. (1969). *Bull. Am. Meteorol. Soc.* **50,** 286–311.

Thompson, P. D. (1957). *Tellus* **9,** 275–295.

Wallace, J. M., Tibaldi, S., and Simmons, A. J. (1983). *Q. J. R. Meteorol. Soc.* **109,** 683–718.

Williamson, D. L., and Kasahara, A. (1971). *J. Atmos. Sci.* **28,** 1313–1324.