Random Error Growth in NMC's Global Forecasts

CAROLYN A. REYNOLDS* AND PETER J. WEBSTER

Program in Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado

EUGENIA KALNAY

Development Division, National Meteorological Center, NWS/NOAA, Washington, D.C.

(Manuscript received 24 May 1993, in final form 18 November 1993)

ABSTRACT

The three-dimensional structure of random error growth in the National Meteorological Center's Medium-Range Forecast Model is investigated in an effort to identify the sources of error growth. The random error growth is partitioned into two types: external error growth, which is due to model deficiencies, and internal error growth, which is the self-growth of errors in the initial conditions. Forecasts from winter 1987, summer 1990, and winter 1992 are compared to assess seasonal variations in regional error growth as well as forecast model improvement. The following is found:

- In the tropics, large external error growth at the 200-mb level is closely associated with deep convection. There is evidence of significant model improvements in the tropics at the 850-mb level between 1987 and 1992.
- The spatial structure of the external error growth in the midlatitudes suggests that the representation of orography in the model, especially over Antarctica and the Rockies, is a significant source of errors.
- Internal error growth in the midlatitudes is greater over the Atlantic and European regions than over the Pacific region and appears to be associated with blocking phenomena, especially over the North Atlantic and Europe. The Northern Hemisphere exhibits a seasonal cycle in the magnitude of error growth, but the Southern Hemisphere does not.

The results for the external and internal error growth rates were obtained using a parameterization of the correlation between forecasts and the verifying analyses. The parameterization is based on the assumption that linear random error growth is caused primarily by model deficiencies, and the validity of this assumption is examined. The results suggest that, in the tropics, significant increases in forecast skill may be obtainable through both model and analysis improvement. In the midlatitudes, however, there is less potential for increases in forecast skill through model improvement, and decreasing the analysis error becomes more important. The parameterization yields results that are physically meaningful and in agreement with previous predictability studies, and that provide quantitative estimates of the spatial and temporal distribution of the sources of forecast errors.

1. Introduction

Random error growth in the National Meteorological Center's (NMC) Medium-Range Forecast Model (MRF) is studied in an effort to estimate the limits of predictability as well as what skill might be obtainable through model improvements or from better observations on a regional basis. This is done by parameterizing the random error growth as a function of two sources: internal and external. The internal error source represents the self-growth of errors in the initial conditions. For example, internal error growth may be large in regions of high baroclinic instability or in convergence regions of advected errors. The external error source represents the random error growth due to model deficiencies. Large external error growth will indicate where the model is most deficient in simulating the real atmosphere. An understanding of the spatial structure of this type of error should aid in determining the mechanisms responsible for the error and ultimately lead to their reduction and to model improvement.

a. Internal errors

It is expected that a significant part of the internal error growth will be due to instabilities in the atmosphere. As Lorenz (1965) pointed out, the self-growth of small-scale errors will eventually contaminate larger scales, and this, in turn, results in an upper bound on predictability. Many studies have been done using turbulence theory as well as simple and complex models

* Current affiliation: Marine Meteorology Division, Naval Research Laboratory, Monterey, California.

Corresponding author address: Dr. Carolyn Reynolds, Naval Research Laboratory, 7 Grace Hopper Avenue, Stop 2, Monterey, CA 93943-5502.

in an effort to estimate limits of predictability (Leith 1971; Leith and Kraichnan 1972; Shukla 1985).

Lorenz (1982) used random forecast error growth rates and random forecast difference growth rates (i.e., differences between forecast runs) to estimate the current, as well as the upper bound on, predictability. He found that if model predictability is the same as atmospheric predictability (an admittedly overoptimistic assumption), then model improvements could result in 10-day forecasts as good as the then present 7-day forecasts. This skill could be extended by 2 days by halving the 1-day forecast error. He also pointed out that as the model improves the actual forecast skill will increase, but the approximation of the upper bound on predictability will decrease. These results are consistent with the results of a more recent study (Chen 1989) that estimated that the 1986 NMC MRF had skill for the first week of forecast time and a limit of dynamic predictability of about two weeks.

Obtaining a single value for the error doubling time or growth rate for the entire globe or for a hemisphere may be misleading. The limit of predictability is a function not only of the error growth rate but also of the error saturation value, whose magnitude is roughly double that of the variance of the field (Miyakoda et al. 1972). Both the variance and the error growth rates have been shown to be functions of latitude (Shukla 1985; Kalnay et al. 1988), season (Shukla 1985), and scale (Lorenz 1969; Smagorinsky 1969; Dalcher and Kalnay 1987). Palmer (1988) and Chen (1990) have found relationships between forecast skill and the phase of the Pacific–North America pattern (Horel and Wallace 1981). Ferranti et al. (1990) document links between the strength of the intraseasonal oscillation and medium-range forecast skill.

Schubert and Suarez (1989) investigated the spatial distribution of error growth in a two-level GCM. An eigenvector analysis of the error correlation matrix suggested that the fastest large-scale error growth is due to large-scale–small-scale interaction found over the Atlantic. By examining the root-mean-square error tendencies, they found preferential error growth over the eastern oceans, which they suggest is due to large-scale–small-scale interactions associated with the mature and decaying phases of baroclinic waves. Consistent with their eigenvector analyses, the error growth rates over the North Atlantic are larger than those over the North Pacific for the first 10 days, although growth rates over the North Pacific become larger than those over the North Atlantic later in the integration.

b. Model deficiencies

Model deficiencies can cause both systematic errors (i.e., the climate drift of the model) and random errors that do not result in a time mean signal. Since it is not possible to determine the exact errors in the initial conditions of a forecast, it is difficult to separate the random error growth caused by model deficiencies (differences between the model and the atmosphere) from the unstable self-growth of the initial conditions errors, except by estimation and parameterization. Leith (1978) approached this problem by assuming that the exponential part of the error growth is due to the self-growth of errors, while the linear portion of the random error growth results from model deficiencies. Sources of model deficiencies such as inadequate cumulus parameterization and orographic forcing have since been shown to result in linear error growth (Wergen 1984; Wallace et al. 1983), supporting Leith's hypothesis.

Dalcher and Kalnay (1987) modified Leith's approach by including the effects of error saturation in their parameterization. They found that model deficiencies dominated error growth in the zonal tropical belt, while the self-growth of errors was the dominating process in the midlatitudes. In this study, we modify the approach of Dalcher and Kalnay by using the correlation between forecasts and analyses instead of random mean-square error growth, and apply the analyses to smoothed spatial fields. We also perform the same parameterization on correlations between forecasts started 1 day apart to evaluate the validity of the assumption that linear error growth is due primarily to model deficiencies.

Arpe et al. (1985) compared model forecasts run from different analyses and used an error budget equation to separate the contributions to the forecast error arising from model error and analysis error. They found that analysis error dominates error growth for the first 2 days and after day 5, while model error dominates error growth between days 2 and 5. They also found that while, initially, forecast errors increased linearly, forecast differences increased exponentially, supporting Leith's (1978) hypothesis that external error growth is linear in nature.

The method employed to investigate the nature and structure of random forecast error growth, along with the forecast datasets used, will be described in the next section. In section 3 the results of the parameterization of the random error growth for the zonally averaged fields are presented. Section 4 contains the results of the error parameterization for the full spatial fields. The results of the application of the parameterization to the random differences between forecasts are presented in section 5 in order to assess the validity of some of the underlying assumptions that the parameterization is based on. The results of this study are summarized and conclusions are made in section 6.

2. Experimental design

a. Data

All analyses and forecasts used are output from NMC's MRF. For two winter seasons and one summer season, January through March of 1987 and 1992

(JFM87 and JFM92), and June through August of 1990 (JJA90), 90 consecutive daily analyses and forecasts are analyzed. The first two periods are part of the Dynamic Extended-Range Forecasting (DERF) project being conducted at the National Meteorological Center, described by Tracton and Kistler (1988).

The DERFII dataset (Schubert et al. 1988) is a series of 30-day forecasts made during the winter of 1986/87 using the then operational MRF (MRF86), which had a horizontal resolution of rhomboidal 40 (R40) and 18 layers in the vertical. A detailed description of the DERFII dataset and its characteristics, as well as the 1986 version of the MRF, are given by Tracton et al. (1988). Moisture is considered explicitly in the lowest 12 layers, or approximately up to 300 mb.

The DERF90 experiment took place in the summer and fall of 1990, while a window of opportunity existed on NMC's then new Cray computer (Saha et al. 1990). Many important changes had taken place in the model formulation between 1987 and 1990 including the introduction of gravity-wave drag, increased model resolution from R40 to triangular 80 (T80), and the extension of moisture into the upper layers of the model (Alpert et al. 1988; Caplan and White 1989; Kalnay et al. 1990). The DERF90 dataset is a series of consecutive 90-day forecasts running from May to October 1990. The model used was similar to the operational T80 model described above except for the reduced (T40) horizontal truncation, and the relaxation of certain surface fields toward climatology to account for the increased length of the forecasts.

Although it is useful to compare the differences between the results for a winter and a summer period, a second winter period is needed to assess the improvements made in the model between 1987 and the present. The last forecast set used is the operational 10-day output of the MRF for the winter of 1992. Between 1990 and 1992, the horizontal resolution of the model was increased from T80 to T126, mean orography replaced enhanced orography, a new marine stratus parameterization was introduced, the horizontal diffusion in the medium scales was reduced, and an improved SST analysis was implemented (Kanamitsu et al. 1991).

A major change to the global forecasting system at NMC was the introduction of the spectral statistical interpolation (SSI) analysis scheme in June 1991 (Derber et al. 1991; Parrish and Derber 1992) that replaced the previously operational optimal interpolation (OI) scheme. The new SSI produces a smoother correction to the 6-h first-guess forecast, while the root-mean-square fit to the observations remains as good or better than for the previous OI analysis scheme. Thus, the new analyses are more consistent with the MRF climatology, especially in the tropics.

In addition to the recent changes to the model, the last forecast set is produced from a model run at a significantly higher resolution than the first two datasets. The statistical analyses were preformed on 3° latitude × 5° longitude gridded fields transformed from reduced (T21) versions of the DERF datasets and the full T126 resolution for the operational output.

b. *Development of the random error parameterization*

We consider the works of Leith (1978) and Dalcher and Kalnay (1987) in developing a method to determine if random error growth is caused by the internal growth of errors in initial conditions or by model deficiencies. Leith (1978) proposed representing random error growth as a function of both internal and external sources such that

$$\frac{\delta \mathrm{RE}}{\delta t} = a\mathrm{RE} + z, \tag{1}$$

where a would represent the internal (exponential) random error growth. External (linear) random error growth due to model deficiencies is represented by z. Random error (RE) is represented by error variance rather than root-mean-square error because of the useful additive property of squared quantities (Leith 1978). The total mean-square error (TE) is equal to the sum of systematic mean-square error (SE) and random mean-square error (RE):

$$\overline{(X_f - X_a)^2} = (\overline{X_f} - \overline{X_a})^2 + \overline{(X_f' - X_a')^2}, \tag{2}$$

where the overbar represents a time mean and the prime represents a deviation from that mean. The subscript f refers to the forecast value and the subscript a refers to the value from the verifying analysis. These quantities are calculated in gridpoint space for the zonal wind field U and the geopotential height field Z at the 850-, 500-, and 200-mb levels. Random difference (RD) is calculated in the same manner as RE except it involves the differences between two model forecasts started 1 day apart verifying on the same day as opposed to the differences between the forecasts and the verifying analyses.

A parameterization of the temporal correlation between forecasts and verifying analyses is developed based closely on the parameterizations of Dalcher and Kalnay (1987) and Kalnay et al. (1988). The decay of the temporal correlation C, calculated as a function of forecast lead time t, between the forecast and the verifying analysis is parameterized as

$$\frac{\delta C}{\delta t} = -[a(1 - C) + z]C \tag{3}$$

and has the solution

$$C(t) = \frac{1 + za^{-1}}{1 + za^{-1}\exp[(a + z)t]}. \tag{4}$$

Here, as in (1), a represents the exponential loss of correlation due to the internal growth of initial errors, and z represents the linear loss of correlation due to

model deficiencies (external errors). The validity of the assumption that linear error growth is due to model deficiencies is investigated by comparing results for forecast–analyses correlations with forecast–forecast correlations and is discussed in section 5.

The temporal correlation between the forecasts and the analyses, to which the above parameterization is fit, is computed as

$$C(t) = \frac{\overline{(X_f - \bar{X}_f)(X_a - \bar{X}_a)}}{[\overline{(X_f - \bar{X}_f)^2}\,\overline{(X_a - \bar{X}_a)^2}]^{1/2}}, \qquad (5)$$

where the overbar represents a time mean average over the length of the dataset (90 days); $C(t)$ is calculated at each grid point for all forecast lead times.

There are a few reasons for choosing the temporal correlation between forecasts and analyses instead of the random mean-square error as the representation of forecast skill. The RE growth rates are a function of RE_{sat}, the saturation value of RE, which is often poorly defined, and would have to be estimated for the 10-day operational forecast dataset from 1992. Also, the inclusion of the extra parameter (RE_{sat}) does not result in a better fit of the data.

Use of the correlation also accounts for the decrease in variance of the model with forecast time. Figure 1a shows the RE and RD curves for the JFM87 200-mb U field for the 10° latitude band centered on 5°N. The RE saturation value is approximately 85 $m^2\ s^{-2}$ and the RD saturation value is less than 60 $m^2\ s^{-2}$, owing to the large decrease in model variance with forecast time. Figure 1b shows the correlations between forecasts and verifying analysis (CE) and correlations between forecasts starting 1 day apart (CD) as a function of forecast time, also computed for the zonally averaged latitude band centered on 5°N for the JFM87 200-mb U field. When the correlation has been calculated using a high-pass filter to remove the positive signal left by the seasonal cycle, all correlation curves should approach zero asymptotically. Thus, it is not necessary to have forecasts long enough so that the saturation value can be determined empirically. Therefore, the correlation curves are especially convenient to use when one has forecasts of lengths of only 5 or 10 days with which to work.

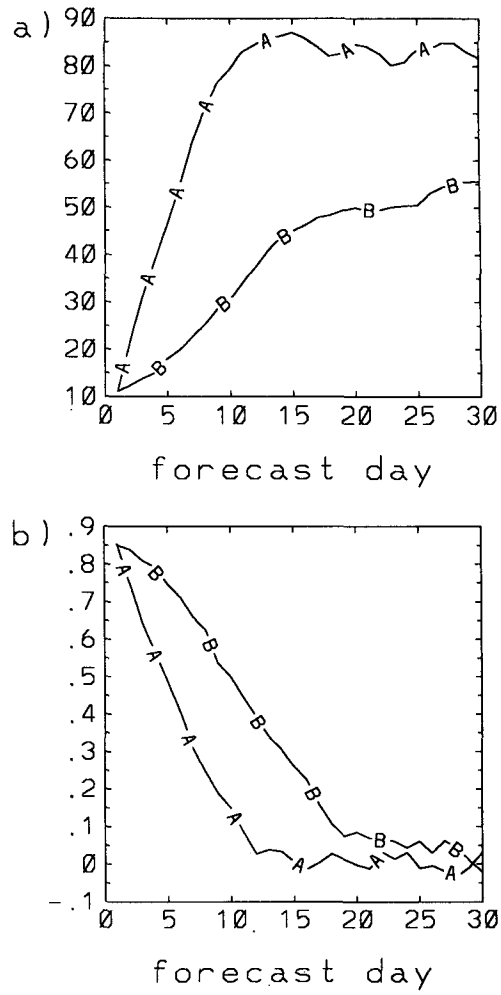

FIG. 1. (a) Random mean-square error ($m^2\ s^{-2}$, curve A) and random mean-square difference ($m\ s^{-2}$, curve B) as a function of forecast time for the 10° latitude band centered on 5°N for the JFM87 200-mb U field. (b) Correlation error (curve A) and correlation difference (curve B) as a function of forecast time for the 10° latitude band centered on 5°N for the JFM87 200-mb U field.

3. Parameterization results: Zonal averages

In previous studies, random error growth has been shown to vary significantly with latitude (Dalcher and Kalnay 1987; Kalnay et al. 1988). The results for zonally averaged fields are a convenient way to compare the differences in error growth between tropical and extratropical regimes.

a. Comparison of the tropics and the extratropics

Visual inspection of the shape of the correlation curve can yield qualitative information about whether model deficiencies or unstable internal errors are dominating the random error growth. Inspection of (3) indicates that the slope of the correlation curve when the correlation is close to 1 is approximately $-z$. Therefore, correlation curves that have the largest initial decreases have the largest values of z, indicating the dominance of model deficiencies in the random error growth. From the second derivative of (3), the time at which the correlation curve has an inflection point t_{infl} is given by

$$t_{infl} = \frac{\ln(a/z)}{a + z}. \qquad (6)$$

From this equation, it is clear that larger values of $z + a$ will result in inflection points closer to $t = 0$. If $z > a$, and model deficiencies are more significant than

the internal growth of initial errors, then t_{infl} becomes negative, and the correlation curve is concave-up, instead of "S"-shaped (initially concave-down).

Figure 2 shows the 10-day correlation error (CE) and correlation difference (CD) curves for both the 0°–10°N band (curves A and B) and the 40°–50°N band (curves C and D) for the JFM87 200-mb height field. In the tropics, the CE (curve A) drops off rapidly initially, resulting in a concave-up curve, illustrating how quickly tropical forecasts diverge from the verifying analyses and how model deficiencies are dominating the random error growth. The CD (curve B) decreases very slowly and in fact increases slightly for the first few days, representing an initial forecast convergence. Although this increase appears slight for the zonal averages for this field, there are regions of the tropics—for instance, the western Pacific—where the CE increases from 0.65 to 0.8 during the first few days. Chen (1989) noticed a decrease in the root-mean-square difference in the tropical height field. The use of the correlations here shows that this convergence of forecasts is due to other factors besides the severe decrease in variance that the model undergoes during the first few forecast days.

In the tropics, even though the systematic error, or the climate drift, is removed, the random parts of the model forecasts not associated with this systematic error behave in a very different manner from the atmosphere, and correlations between the forecasts and the verifying analyses decrease quickly. However, forecasts started 1 day apart become more alike for the first few days; that is, the correlation between them increases. Forecasts become more uniform, more like the model climatology, during the initial "spindown" phase. After the first few days, the forecasts diverge as expected. Thus, it appears that while the model produces poor forecasts in the tropics, the model tropics themselves are relatively predictable. Unfortunately, as the large difference between the CE and CD in the tropics illustrates (Lorenz 1982), the model is most unlike the real atmosphere in this region, and the increase in CD for the first few days is a by-product of the large inconsistencies between the tropical analyses and model climatology.

In the extratropics, the CE (curve C) and CD (curve D) both decrease in a similar fashion, resulting in concave-down curves, with an inflection point around day 6 for the CE, indicating small model deficiencies. The smaller differences between these two curves as compared to the differences between curves A and B also illustrate that model deficiencies are less severe in the midlatitudes than in the tropics. We would expect these aspects to be reflected in the results of our parameterization.

The initial correlation for the extratropical curve (corresponding to a 1-day forecast) is greater than 0.95, while the initial correlation for the tropical curve is only approximately 0.8. This is most probably due in part to

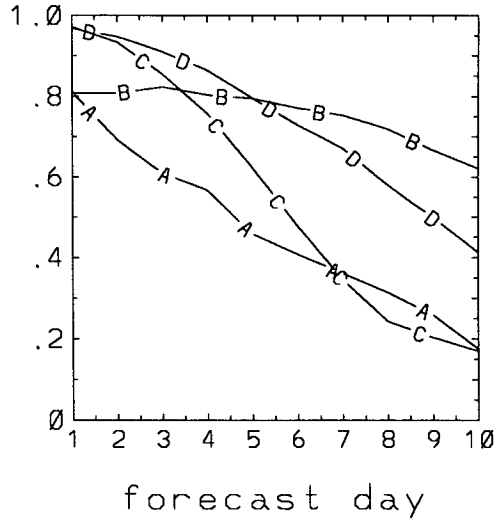

FIG. 2. JFM87 200-mb height field correlation error and correlation difference for the 10° latitude bands centered on 5°N (curves A and B, respectively) and 45°N (curves C and D, respectively).

the large difference in variance between the different regions, so that the same absolute error would be a much higher percent of the saturation error (and likewise a smaller CE) in the region with low variance. However, this large day-1 error in the tropics may also be due to analysis errors. The uneven distribution of observations over the globe results in varying degrees of uncertainty in the analysis errors, which may effect initial error growth as well as the validity of what we consider as verification. The large inconsistencies between the tropical analyses and model climatology mentioned above result in the severe initial spindown and negative difference growth rates discussed above, while the large uncertainties in analyses inhibit the accuracy of estimated error growth rates. The degree of uncertainty in knowing the true atmospheric state, which is particularly large in the tropical and polar regions, should be kept in mind when examining results for which analyses are used as truth.

b. External error results

We now consider the results of the parameterization fit to the correlation error for the first five forecasts days for zonally averaged bands of 10° width, for both the U and geopotential height fields at three atmospheric levels. The period of five forecast days is chosen because of its applicability to operational forecasts. The sensitivity of the results to the number of forecast days included is discussed at the end of this section. We first present the results for the parameter z from (3), which represents error growth due to model deficiencies. For all the figures in this section, curves A, B, and C represent the 850-, 500-, and 200-mb levels, respectively.

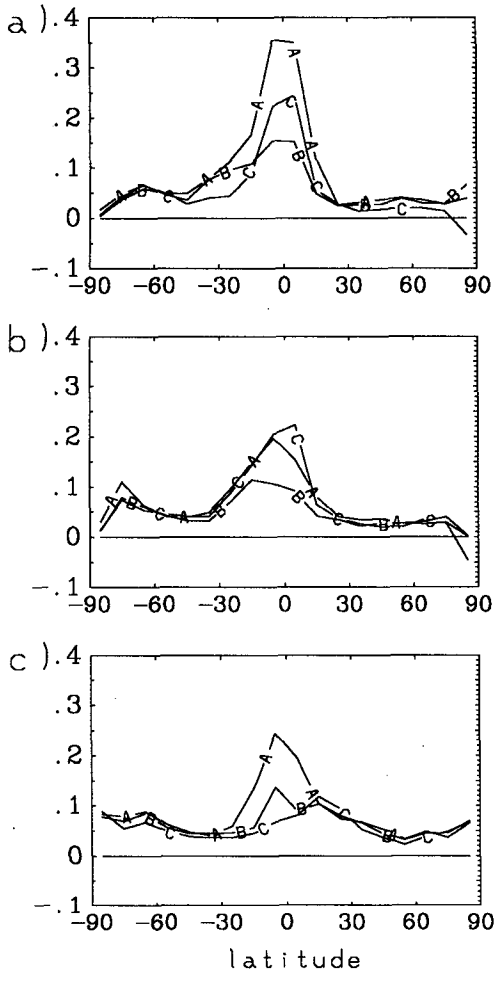

FIG. 3. External error parameter z (day^{-1}) for zonally averaged correlation errors for the height field for (a) JFM87, (b) JFM92, and (c) JJA90 at 850 mb (curve A), 500 mb (curve B), and 200 mb (curve C).

Figure 3 shows the values of z derived for the JFM87, JFM92, and JJA90 periods for the height fields. The most conspicuous feature for all three cases is the large values of z in the tropics. For the JFM87 case, z reaches maximum correlation decay rates of 35% day^{-1} at the 850-mb level. The decay rates at the 200- and 500-mb levels are almost 25% day^{-1} and 15% day^{-1}, respectively.

The results for JFM92 show decreases in the value of z in the tropics, especially at the 850-mb level where z drops to 15% day^{-1} from 35% day^{-1} for the JFM87 period. The improvements suggested by these results are most probably due to the implementation of the new analysis scheme in the MRF, which results in analyses more consistent with the model climatology and significantly reduces the spindown effect. The impact of changes made to the model in 1988 in the vertical diffusion and surface flux schemes (Caplan and White 1989) may also be important, as well as the fact that the JFM92 simulations were performed at a much higher resolution than the JFM87 simulations. In the midlatitudes, the values of z for the two winter periods are similar and are very consistent among all three levels of the atmosphere. The decay rates for the Southern Hemisphere midlatitudes are approximately 5% day^{-1} and are slightly lower for the Northern Hemisphere. It is curious to note that for the JJA90 case the 200-mb level has very low values of z (about 10% day^{-1}) in the tropics. However, because of the change in season, it is not possible to determine if this decrease is due to model changes or changes in the location and intensity of the convection.

The values for z for the U field, shown in Fig. 4, are qualitatively similar to those of the height field, with a few significant differences. The maximum in z in the tropics is still obvious but another equally conspicuous

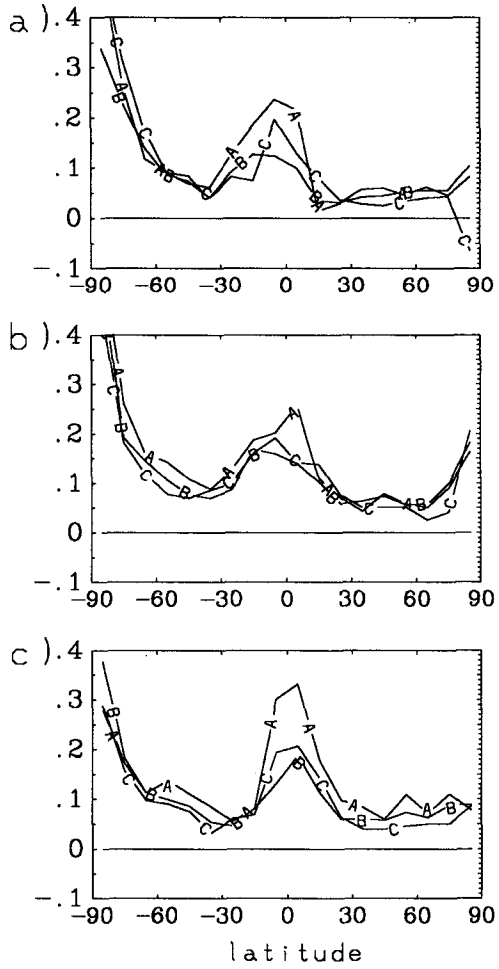

FIG. 4. External error parameter z (day^{-1}) for zonally averaged correlation errors for the U field for (a) JFM87, (b) JFM92, and (c) JJA90 at 850 mb (curve A), 500 mb (curve B), and 200 mb (curve C).

feature is the large values of z in the region over Antarctica, reaching 60% day^{-1} for some fields. These large model deficiencies are in part due to the high terrain (much of Antarctica has surface pressures lower than 600 mb). However, one must consider when interpreting these results that in the polar regions the parameterization does not converge onto the data using the same error tolerance. Figure 5 shows the average absolute difference between the calculated correlation using the parameterized results and the observed correlation for the height and U fields for JFM87. The errors for JFM92 and JJA90 (not shown) are comparable to or smaller than those for JFM87. For most of the midlatitudes, the fitting errors are less than 0.005, except for the tropical height field, where they reach 0.008, and the polar regions, which have much higher errors than the rest of the globe. These large errors, as well as the fact that the accuracy of the analyses in the polar regions is highly suspect, mean that not much confidence can be placed in the parameterization results in the polar regions.

In the tropics, Fig. 4 shows that the JFM92 case exhibits the smallest decay rates of all three periods at the 200- and 850-mb levels. The results for the 500-mb level remain about the same for all three periods. Curiously, the values of z are higher at the 850-mb level for the JJA90 than for the JFM87 case. In the midlatitudes, z is slightly greater for the U field than it is for the height field, averaging around 5%–10% day^{-1}. Higher values of z at the lowest levels indicate that the model deficiencies in the midlatitudes are tied to surface problems rather than to problems in the middle or upper troposphere.

The parameterization was also fit to the correlation difference curves in an effort to test the validity of the assumption that linear error growth is primarily a result of model deficiencies or differences between the model and the atmosphere. For these cases, it is found that z is everywhere within ± 0.05 day^{-1} except for the polar regions, where values of 0.2 day^{-1} and higher are found—yet another reason that the results for these regions should be interpreted with caution. The validity of this assumption on a regional basis is discussed in more detail in section 5, where z calculated using the correlation differences for the full spatial fields is shown.

c. *Internal error results*

In this section we present the results for the parameter a from (3), representing the internal growth of initial errors, which, for the zonally averaged results, is expected to be dominated by instabilities. Figure 6 shows the internal error for the height fields for all three periods. As expected, the latitudinal structures of the values for a are markedly different from that of z. Instead of pronounced maxima in the tropics, there is now a negative maximum in the height field, accurately reflecting areas with the largest spindown effect (i.e., the random differences between two forecasts initially decrease). It also reflects the fact that dynamic instabilities on synoptic time scales are much less energetic in the tropics than in the midlatitudes. At the 850-mb level, the largest negative values in the tropics occur for the JFM87 case, indicating that the model spindown problem has decreased since then. The 200-mb U fields (not shown) exhibit less model improvement. Overall, the JFM92 tropical results show that the spurious effects due to spindown have been reduced since 1987. For this period, the parameterization suggests that, in the tropics, both internal and external errors are of the order of 10% day^{-1}. The JJA90 case seems to have less of a spindown problem than the JFM92 case (around 5% day^{-1}), although, as stated before, it is difficult to separate the effects due to model changes from the effects due to seasonal changes in the location and intensity of convection.

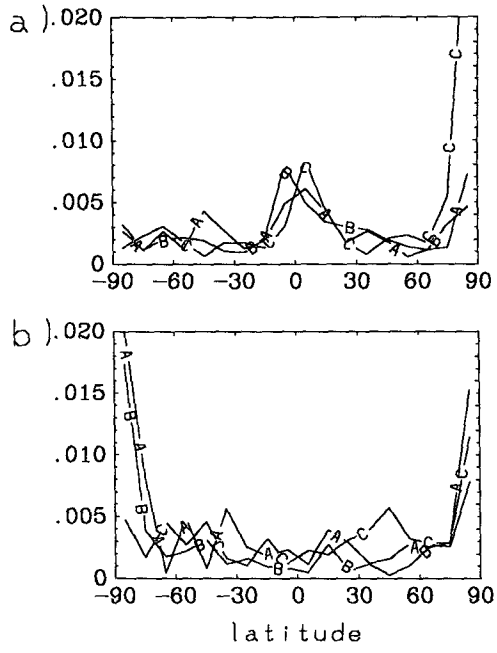

FIG. 5. Average absolute difference between the calculated and actual correlations for the zonally averaged JFM87 (a) height field and (b) U field at 850 mb (curve A), 500 mb (curve B), and 200 mb (curve C).

For the Northern Hemisphere midlatitudes, there is a broad latitude band of high values between 0.4 and 0.5 day^{-1} north of 20°N for JFM87 and north of 30°N for JFM92. These values are equivalent to doubling times for root-mean-square errors of about 2.7–3.5 days and are consistent with those found by Dalcher and Kalnay (1987) and Kalnay et al. (1988). The values of a at all three levels for the two winter cases are similar, with the largest values of a observed at the 200-mb level. For the summer case (JJA90) the 200-mb level values are smaller (about 0.25 day^{-1}, or a dou-

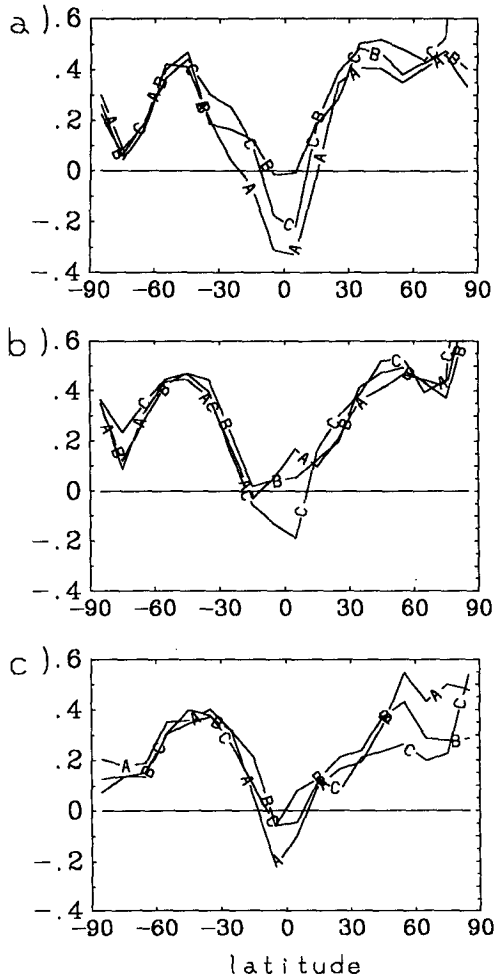

FIG. 6. Internal error parameter a (day^{-1}) for zonally averaged correlation errors for the height field for (a) JFM87, (b) JFM92, and (c) JJA90 at 850 mb (curve A), 500 mb (curve B), and 200 mb (curve C).

bling time for root-mean-square errors of about 5.5 days) except in the extreme north, while the values of a at the 500-mb level vary from about 0.3–0.4 day^{-1}, and the values at 850 mb reach 0.5 day^{-1} north of 50°N.

The processes causing internal error growth in the Northern Hemisphere summer affect the different levels somewhat differently and seem most dominant in the lower troposphere. Likewise, the similarities between the three levels during the winter months indicate that deep processes involving the whole troposphere, perhaps associated with baroclinic instability, are dominating the internal error growth.

For the Southern Hemisphere midlatitudes, there is a very high degree of agreement in a for all three periods. For the two JFM periods (i.e., Southern Hemisphere summer), there is a peak in a of about 0.45 day^{-1} at around 50°S. In the Southern Hemisphere winter case, the peak values are slightly less (around 0.4 day^{-1}) but they exist over a broader latitudinal band and extend farther equatorward (50°–20°S). These findings are consistent with the findings of Trenberth (1982), who shows that in the Southern Hemisphere summer there is still much variability observed in the strong baroclinic zone equatorward of the cold Antarctic continent. Overall, the internal error growth rates for the U fields (not shown) are very similar to those for the height fields although the improvements noted in the decrease in spindown between 1987 and 1992 in the tropics seem to be less pronounced.

The aforementioned results were presented for the parameterization fit to the first five forecast days. Figure 7 shows the external and internal errors for the JFM87 200-mb height field fit to the first 5, 7, and 10 days. Note that while the external errors are relatively insensitive to the number of forecast days included, the internal errors show some differences, especially for the day-10 case. This increased sensitivity to the number of days included in the parameterization after about day 7 or 8 may be related to the increased importance of the advection of errors after this period, which is discussed in section 4b.

4. The spatial distribution of external and internal errors

There is reason to believe that there is much longitudinal structure in the random error growth and that the study of this structure will provide information

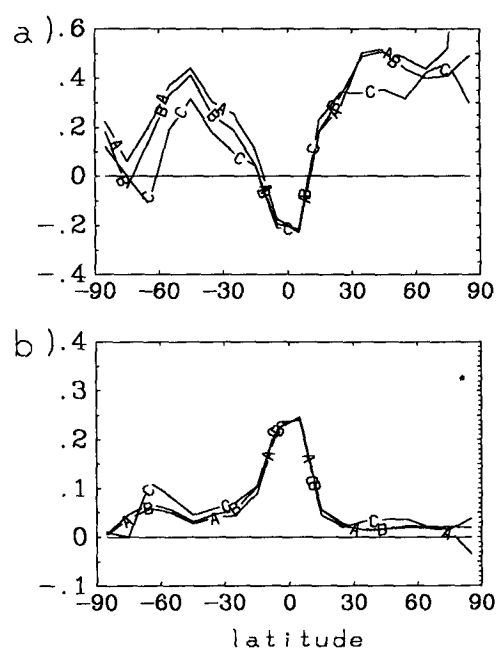

FIG. 7. (a) Internal error parameter a (day^{-1}) and (b) external parameter z (day^{-1}) calculated for the zonally averaged JFM87 200-mb height field by fitting the parameterization to the first 5 (curve A), 7 (curve B), and 10 (curve C) forecast days.

about the processes that are responsible for this error growth. The correlation curves at each grid point will be subjected to much larger sampling errors than the zonally averaged curves, as well as being much more sensitive to the advection of errors. Therefore, to reduce these effects, a 20° latitude × 20° longitude running mean filter was applied to the correlation fields at each forecast time. The results for the full spatial fields are more sensitive to the number of days included in the parameterization than the results for the zonally averaged fields, with local maxima in a shifting eastward as the number of forecast days are increased, but in general the large-scale patterns remain similar. As with the zonally averaged cases, the results of the parameterization fit to the first five forecast days are shown.

The 5-day-averaged absolute difference between the observed correlations and the correlations calculated using the obtained values of a and z for the JFM87 200-mb height field is shown in Fig. 8. The corresponding differences for U (not shown) are quite similar. These errors are in general less than 0.01 for almost all of the midlatitudes, as well as a large portion of the tropics, with higher errors in some regions of the tropics, as well as over the North Atlantic. Particularly high values (up to 0.08) are found for the polar wind fields. The fits for the JFM92 and JJA90 200-mb fields (not shown) were better than those shown for the JFM87 case, while the errors for the lower levels were in general slightly worse. Since the parameterization results are considerably less reliable near the poles and the assumption that linear error growth is caused by model deficiencies may not be applicable in this region, only the results for the area between 75°S and 75°N are presented. The fact that the parameterization does not fit the tropical data as well as the midlatitude data should be kept in mind when interpreting the results.

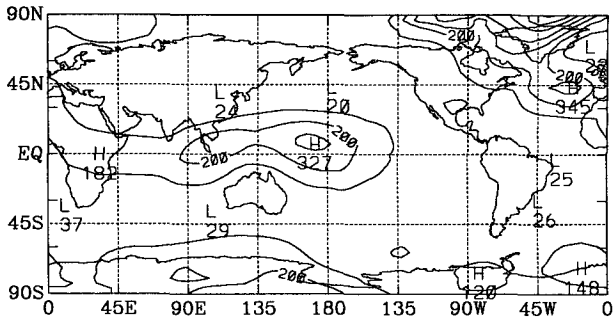

FIG. 8. Averaged 5-day absolute difference between the calculated and actual correlation for the JFM87 200-mb height field. Contour interval is 0.01. Labels are scaled by 10 000.

a. External errors

In this subsection the longitudinal structure of linear random error growth (or correlation decay), which is attributed to model deficiencies, is presented. The linear rate of decay of the correlation between forecasts, as in section 3, is represented by the parameter z.

1) MODEL DEFICIENCIES IN THE MIDLATITUDES

The values of z calculated for the U field for JFM87 are presented in Fig. 9. As demonstrated in section 3 for the zonal averages, the external errors are quite small in the midlatitudes. However, there does seem to be larger values of z in regions associated with high topography.

Antarctica is the region in which there is the largest external error growth. At all levels for all three periods, there are serious model deficiencies south of 60°S, with values of z greater than 0.3 day^{-1} in some regions. These large errors, particularly at the lower levels, are not surprising in light of the fact that a large part of the Antarctic surface is higher than the 600-mb level.

After Antarctica, the topographic feature that seems to cause the most problems in the model is the Rocky Mountains. In the region around western North America extending from the subtropics all the way up to Alaska there are larger values of external error. For both the JFM87 period and JFM92 period (not shown), the values of z are between 0.08 and 0.16 day^{-1}. The values are largest at the 850-mb level and decrease with height, which is to be expected if the problem is related to the topography. For JJA90 (not shown), the larger values appear to be significant only at the 850-mb level, indicating that perhaps deficient orography in this region has a significant effect on forecast error throughout the depth of the troposphere only in wintertime. The region of the Tibetan Plateau is characterized by high values of z between 0.08 and 0.16 day^{-1} for all three periods at the 850- and 500-mb levels but not at the 200-mb level. The relationship between external errors and the Andes are more difficult to discern because of the proximity to Amazon Basin convection.

Even though the Tibetan Plateau is the highest of the features, it seems to be the Rockies that consistently have the largest effect on the upper levels of the atmosphere. This may be because the Rockies are less adequately resolved than the Asian mountains. It might also be related to their north–south orientation and large latitudinal extent, which will have a strong impact on the zonal wind field. This result is consistent with Klinker (1990), who found that the error forcing from the Rockies had a much stronger damping effect on the baroclinic waves in the Atlantic storm track than the Himalayas had on the Pacific storm track.

Errors in the height field resulting from model deficiencies are not as clearly tied to the topography as errors in the U field. This may be because the magnitude of the external errors in the height field seem to be smaller than those in the U field for most regions of the extratropics. However, as in the U field results,

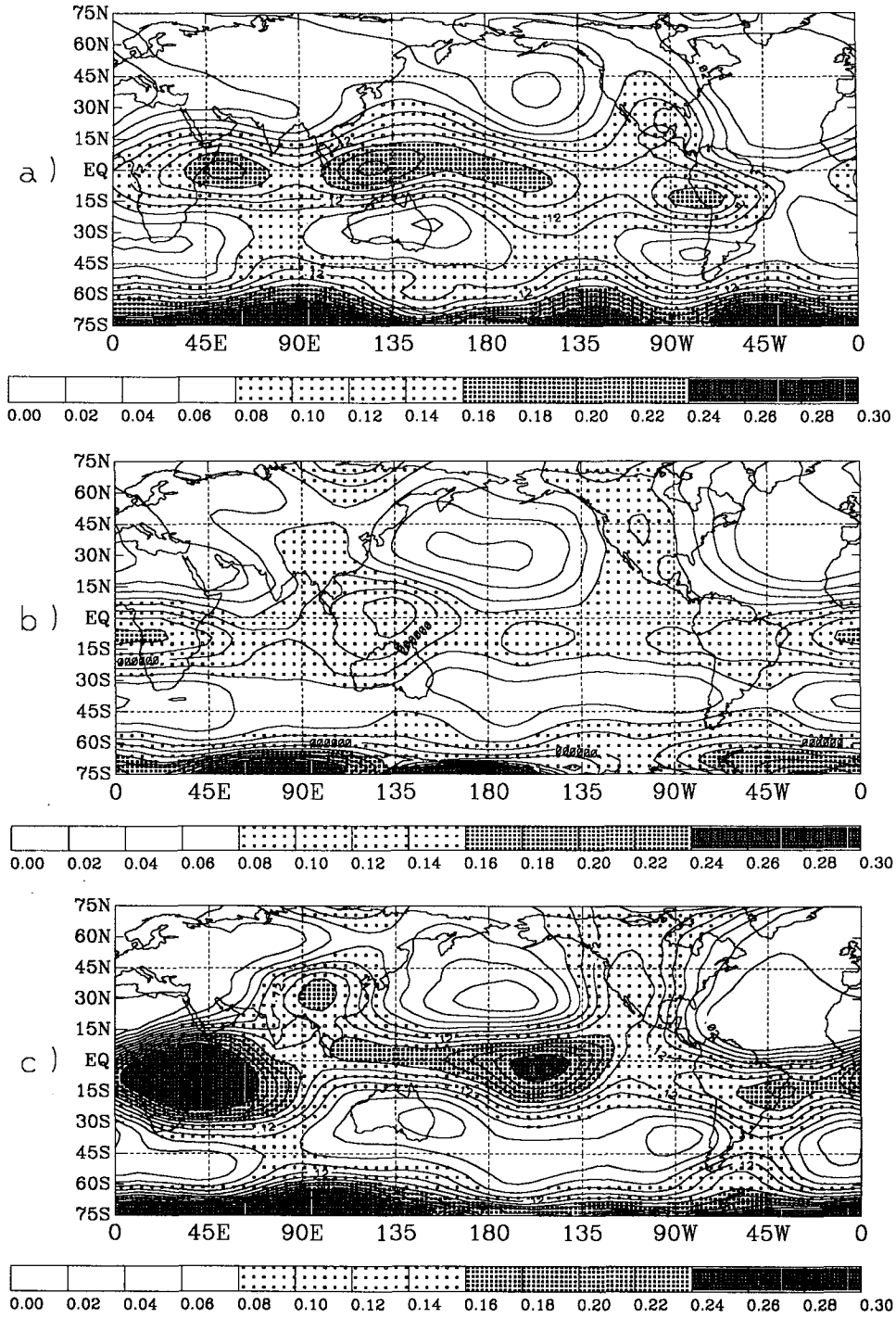

FIG. 9. External error parameter z (day^{-1}) for the JFM87 U field for (a) 200 mb, (b) 500 mb, and (c) 850 mb.

there is a noticeable regional maximum over western North America at all three levels for the winter cases. This maximum is clearest for JFM87, with values above 0.08 day^{-1} at the 850-mb level (not shown) and 200-mb level (Fig. 10).

2) MODEL DEFICIENCIES IN THE TROPICS

It is well known that GCMs have difficulty adequately simulating convective processes and their environmental effects. This is especially troublesome in

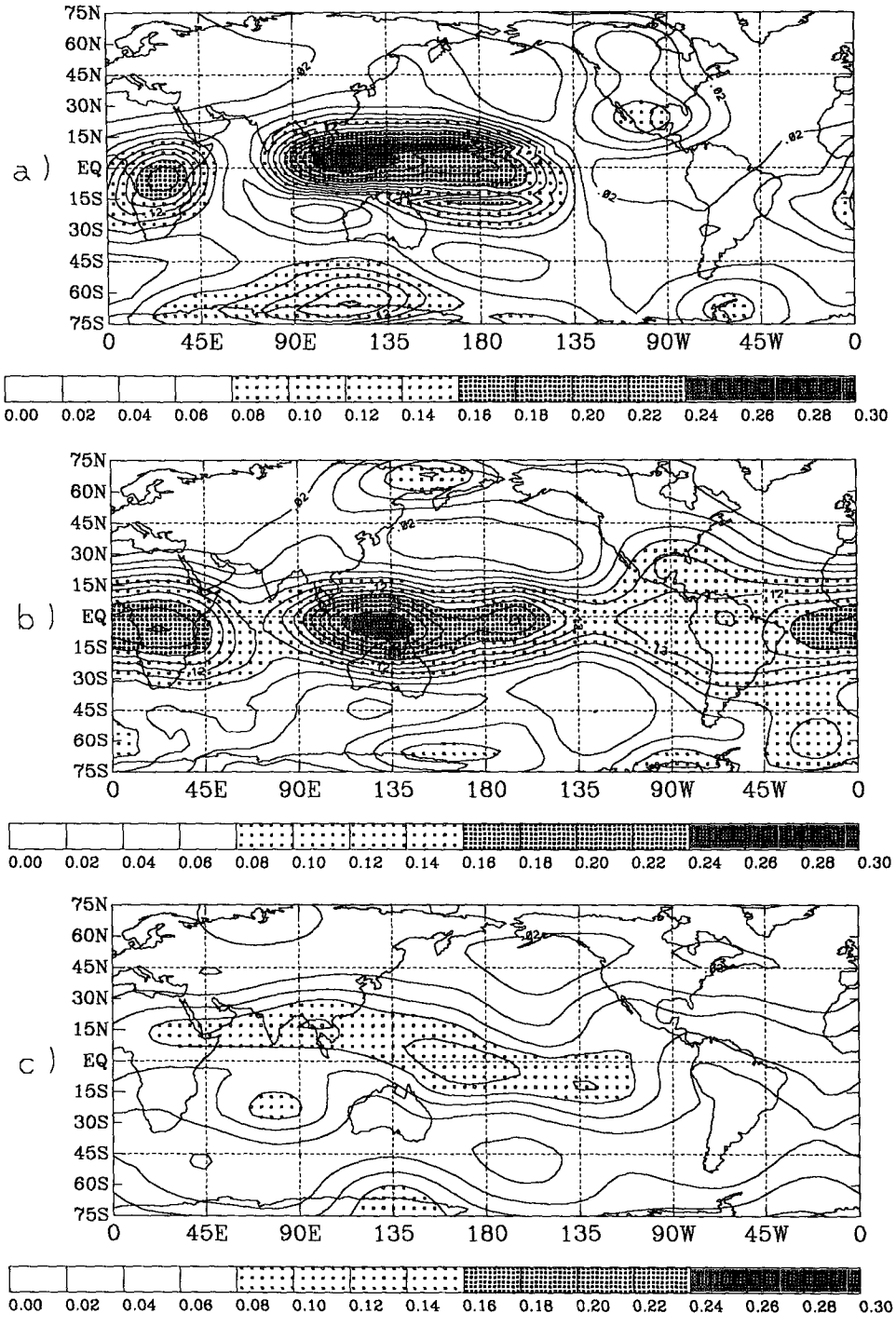

FIG. 10. External error parameter z (day^{-1}) for the 200-mb height field for (a) JFM87, (b) JFM92, and (c) JJA90.

the tropics owing to the extensive areas of active and very deep convection. If the model deficiencies are primarily a result of inadequate cumulus parameterization, then corresponding shifts should be observed in the maxima in external error with the change in season.

Figure 10 shows the external error for all three forecast periods for the 200-mb height field. Figure 11 shows the outgoing longwave radiation (OLR) for the three corresponding time periods, with the values less than 220 W m^{-2} shaded, from which regions of deep

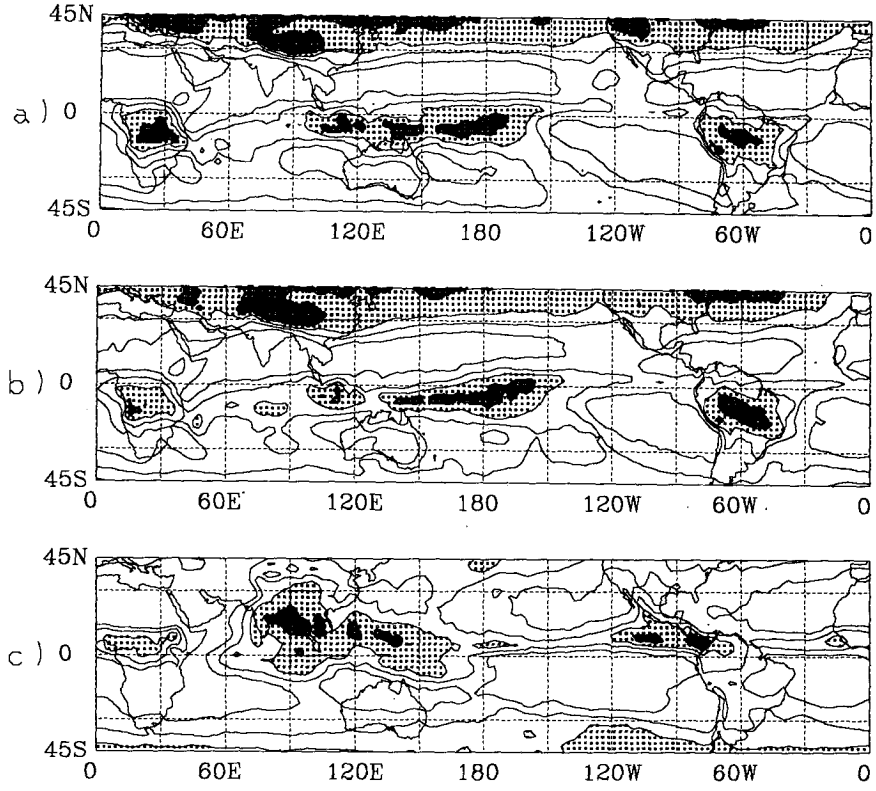

FIG. 11. OLR (W m^{-2}) for (a) JFM87, (b) JFM92, and (c) JJA90. Contour interval is 20 W m^{-2}. Light shading indicates values less than 220 W m^{-2}, and dark shading indicates values less than 200 W m^{-2}.

convection are inferred. There is a good correlation between OLR minima and z maxima in the 200-mb height fields for all three periods. For both JFM87 and JFM92 there is a large maximum in z over the western Pacific and maritime continent, with values reaching 0.24 day^{-1} for the first period. There is a maximum in z with values greater than 0.16 day^{-1} at the 200-mb level corresponding to the African monsoon for both periods. Over the convection in the Amazon Basin there is a maximum in z for the JFM92 case but curiously not for the JFM87 case. For the JJA90 case, the tropical band of large values of z at the 200-mb level follows the northward shift in OLR well, with the latitudinal z maximum over Africa now occurring north of the equator, etc. As with the zonally averaged results, the errors are smallest for the JJA90 case. It is not possible to determine if this is due to seasonal differences in the amount and location of convection, or differences in the model. At the 850-mb level, the values for the JFM87 period are much larger than for the JFM92 period (Fig. 12). This is most probably caused by improvements resulting from the new analysis scheme, as well as improvements in the physical parameterization schemes involving surface fluxes, or the increased resolution at which the JFM92 simulation was run. The external error fields at the two lower levels do not exhibit as coherent a relationship with the convective regions as the external error fields at 200 mb.

The intraseasonal differences in the external error and OLR fields for the JFM87 case illustrate even more clearly the strong relationship between z and deep tropical convection at the 200-mb level. Figure 13 shows the monthly averaged OLR fields for January, February, and March 1987. Convection inferred from OLR over Africa and South America remains fairly constant throughout the period, with the maximum over Africa strengthening slightly during March. However, the convection in the western Pacific changes significantly during the period. It is very strong during the first two months and then weakens considerably during the last month, almost disappearing to the west of the maritime continent.

Figure 14 shows z calculated for the first and second halves of the 1987 winter season for the 200-mb height field. The first half (JF) extends through the month of January and the first half of February. The second period (FM) extends from the middle of February through March. One sees that the external errors over the east and central Pacific are much greater for JF, as is the convection inferred from the OLR (see Fig. 13). For FM, the values of z in this region are reduced, with a smaller maximum centered east of the date line again

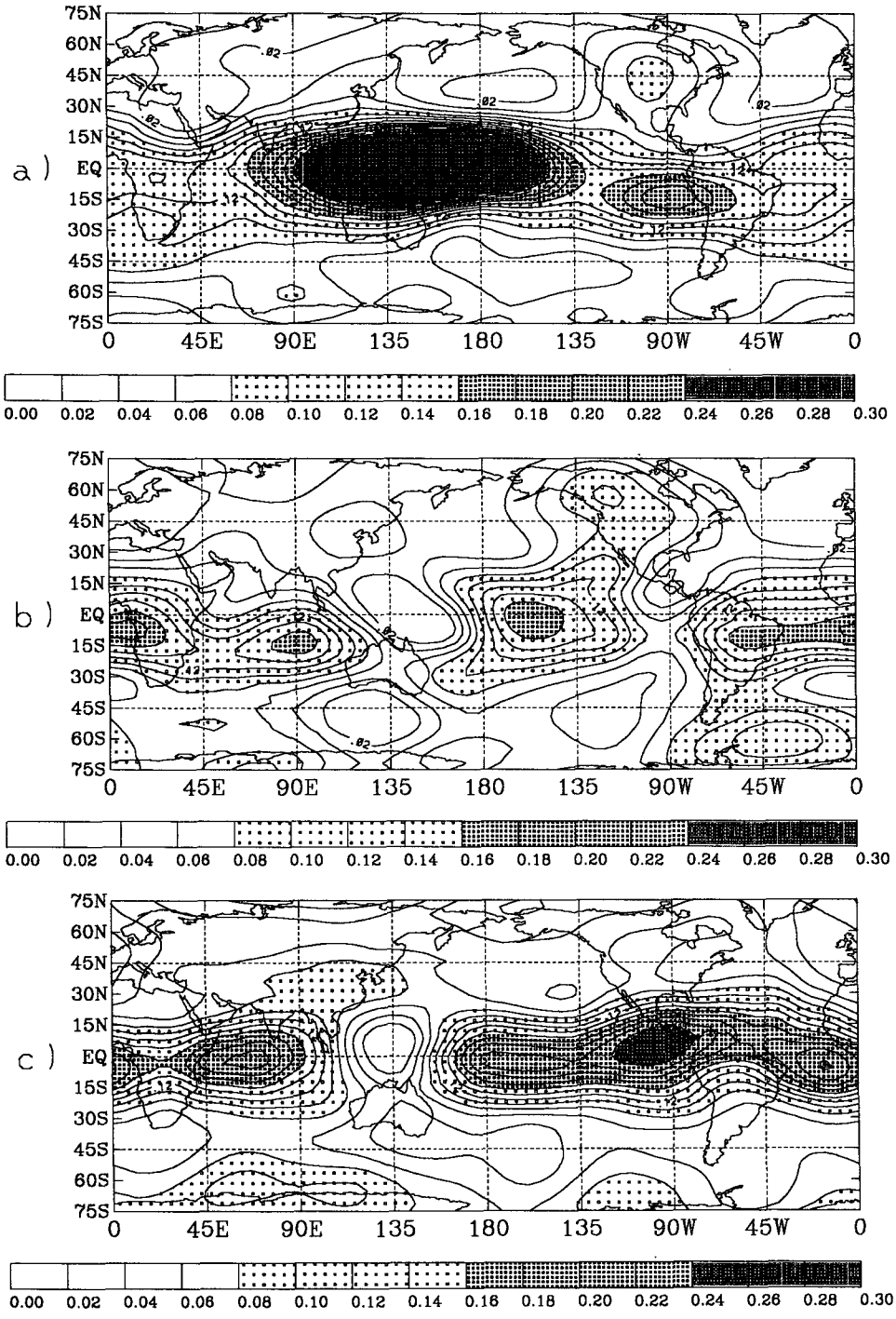

FIG. 12. External error parameter z (day^{-1}) for the 850-mb height field for (a) JFM87, (b) JFM92, and (c) JJA90.

reflecting the shifts in convection. The external error values over Africa are slightly higher during the second period, which is also consistent with the observed intraseasonal shifts in the OLR fields. Similar patterns are observed for the U field (not shown).

b. *Internal errors*

Schubert and Suarez (1989) found large error growth rates in a general circulation model over the eastern oceans, which they associated with the mature

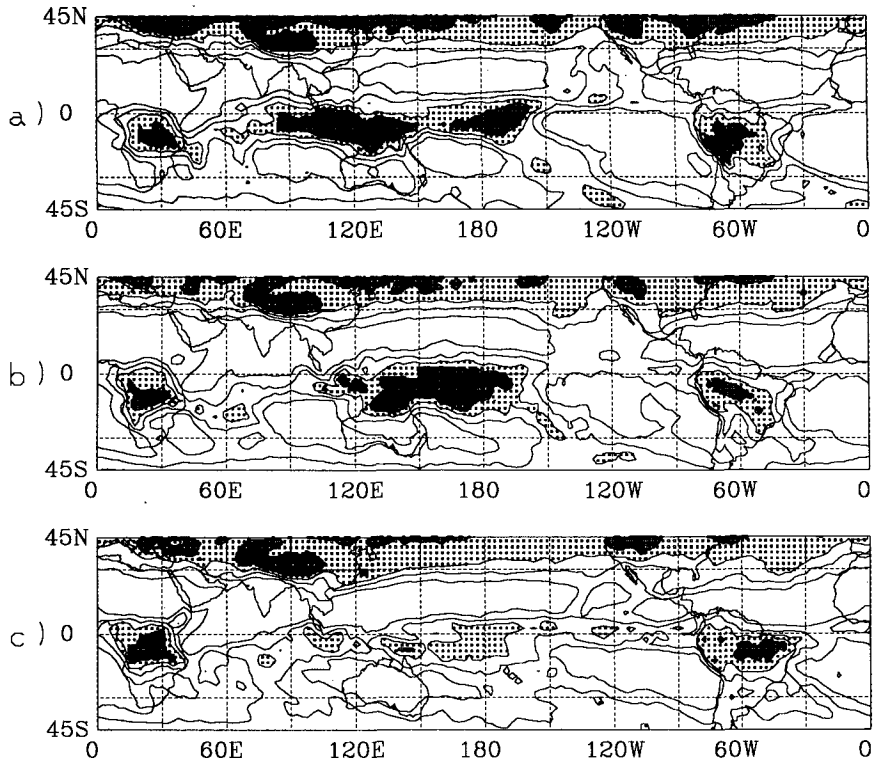

FIG. 13. OLR (W m^{-2}) for (a) January, (b) February, and (c) March 1987. The contour interval is 20 W m^{-2}. Light shading indicates values less than 220 W m^{-2}, and dark shading indicates values less than 200 W m^{-2}.

and decaying phases of baroclinic waves. Tracton et al. (1989) and Tracton (1990) have shown that forecast skill for NMC's MRF is lower when a block forms 3 days or later into the forecast. To investigate the possibilities that predictability may be related to either the intensity of the storm tracks or to blocking, filtered data are used to identify the regions in which storm tracks or blocking dominate. The approach used by Blackmon (1976) and Blackmon et al. (1977) is followed. Time filters are used to identify areas where certain mechanisms dominate the variability of the 500-mb height field. Blackmon used a low-pass filter (10–90 days) of the 500-mb height field to identify regions where blocking ridges and other low-frequency phenomena dominate the variability. A bandpass filter of 2.5–6 days was used to determine the location of storm tracks.

1) SEASONAL VARIATIONS

Figure 15 shows the internal error growth rate a for the 200-mb height fields for all three forecast periods. For both winter periods, one notices regions over the North Atlantic, northern Europe, and western Siberia with very high growth rates, with some regions having growth rates greater than 1.0 day^{-1}, or, alternatively, a root-mean-square doubling time of 1.4 days. In the summertime the growth rates over the North Atlantic are considerably reduced. However, the growth rates over the North Pacific do not exhibit much of a fluctuation in magnitude, with the growth rate for JJA90 slightly higher than for JFM92. Note also that in the Southern Hemisphere there is no noticeable seasonal variation in the magnitude of a, as was the case with the zonally averaged fields. The results for the 500- and 850-mb levels (not shown) are very similar in structure to the 200-mb level in the midlatitudes during winter, indicating that the processes dominating error growth affect all three levels similarly.

2) RELATIONSHIP TO STORM TRACKS

The bandpass-filtered 500-mb height fields are presented in Fig. 16. The positions of the storm tracks, inferred from these fields, are similar to the results found by Blackmon (1976). In the Northern Hemisphere there are two dominant storm tracks along the northern flank of the deceleration regions of the Pacific and Atlantic jets. The maxima are similar in amplitude and orientation for the two winter periods, except for a large secondary maximum over the Norwegian Sea present in the JFM92 case but absent in the JFM87 case. In the JJA90 case, these maxima are still present

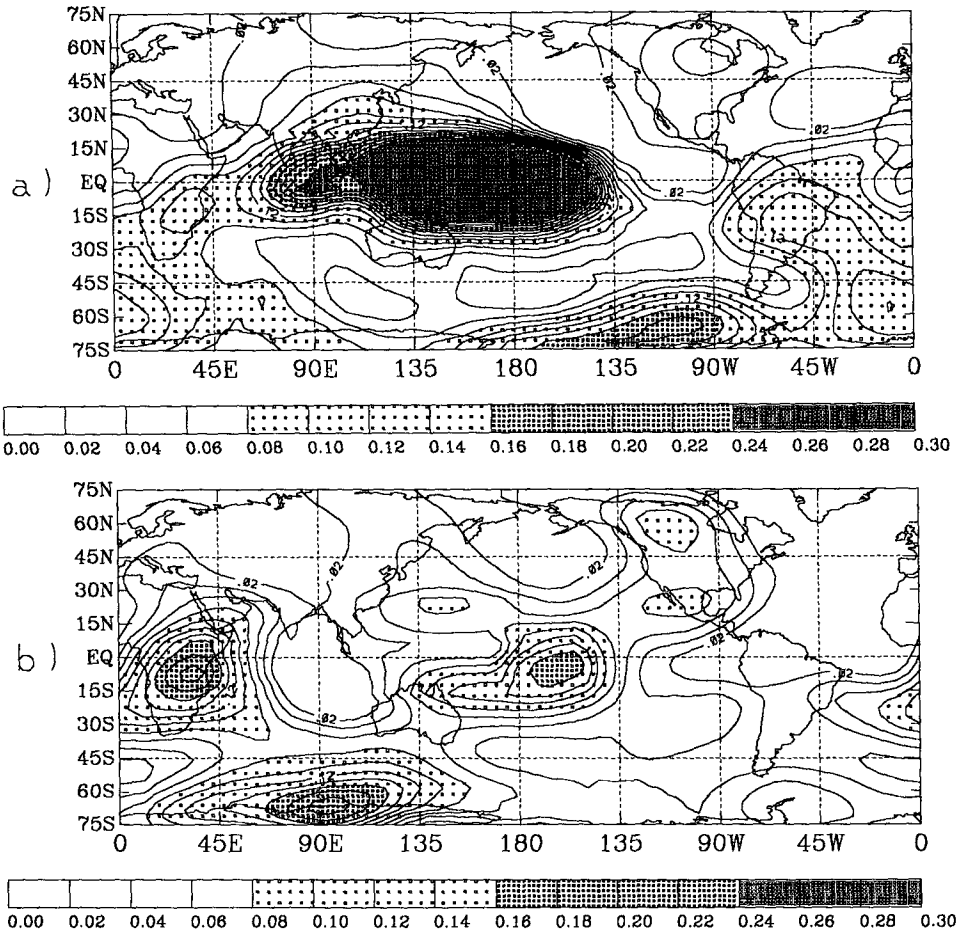

FIG. 14. External error parameter z (day^{-1}) for the 200-mb height field for the (a) first half and (b) second half of the JFM87 period.

but are shifted farther northward and have only one-third to one-fourth the variance of the winter storm tracks. As in Trenberth (1992), the Southern Hemisphere exhibits a much smaller seasonal difference in the magnitude of the variance than the Northern Hemisphere does.

A relationship between a (Fig. 15) and the storm tracks (Fig. 16) in the Northern Hemisphere is not obvious. The maxima in a over the Atlantic and Europe seem to extend eastward beyond the bandpass variance maxima. In the northern Pacific, there is a maximum in a to the east of the storm track in JFM87 and an elongated maximum in a that aligns well with the storm track in JFM92.

3) RELATIONSHIP TO BLOCKING

The low-pass-filtered variance for the 500-mb height fields are shown in Fig. 17. Instead of the 10–90-day window used by Blackmon (1976), a 10–30-day window is chosen in an effort to focus more exclusively on the blocking phenomenon. The Northern Hemisphere shows large differences in variability between the summer and winter periods. There are large maxima over the North Atlantic and eastern Canada, and northern Europe for both winter cases. There are no such corresponding maxima in the summer period. The Southern Hemisphere changes much less with the change of season, and the variance fields are more zonally uniform than the variance fields for the Northern Hemisphere. So, for both the bandpass and low-pass results, the Southern Hemisphere shows almost no seasonal differences in the magnitude of the variance, while the Northern Hemisphere shows a pronounced cycle. If the magnitudes of these variance maxima are related to the internal error growth rates, then this is consistent with the fact that the Southern Hemisphere shows almost no seasonal difference in the magnitude of a maxima, while the Northern Hemisphere shows a pronounced difference.

In the Northern Hemisphere the relationship between the maxima in the low-pass variance and the maxima

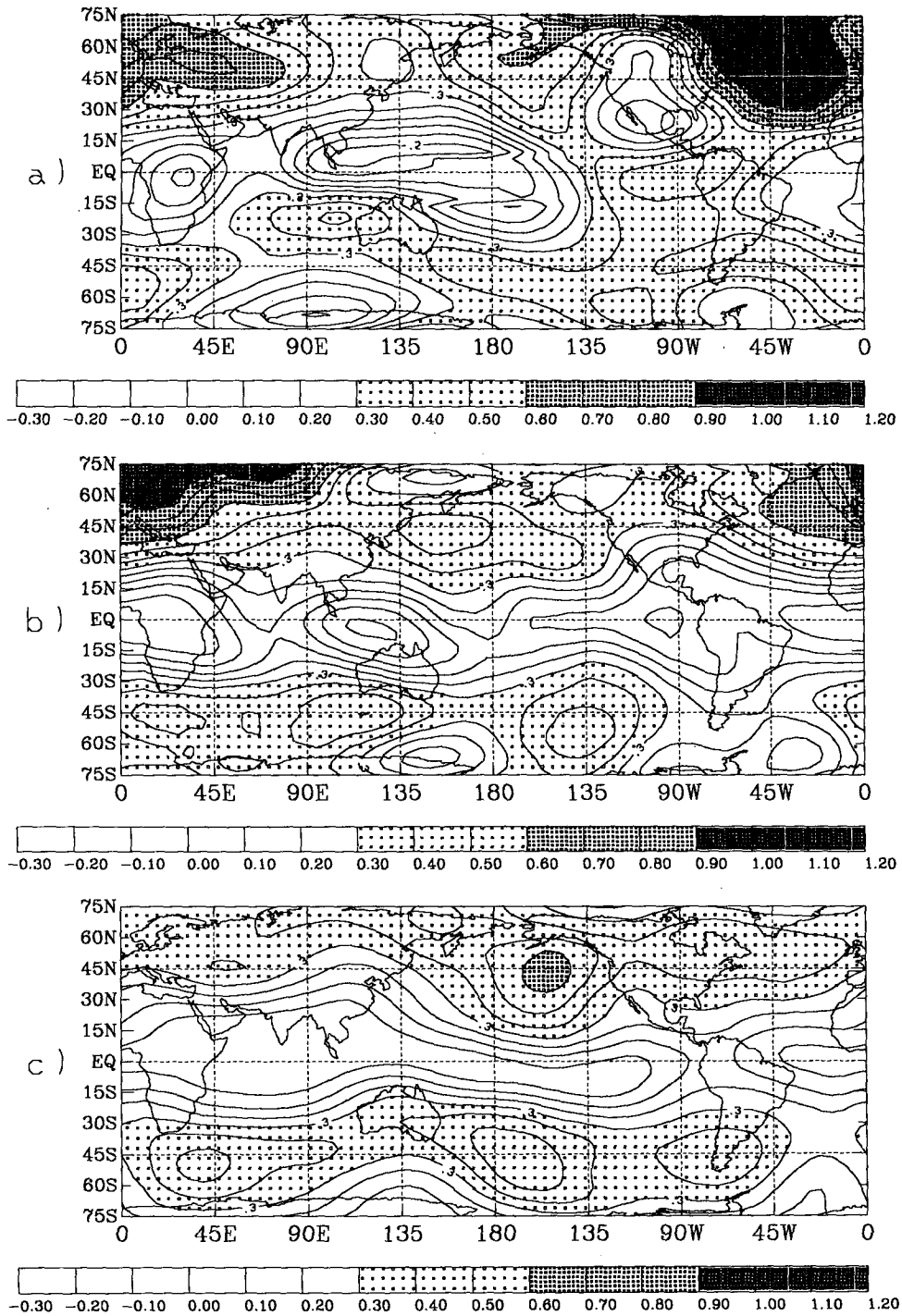

FIG. 15. Internal error parameter a (day^{-1}) for the 200-mb height field for (a) JFM87, (b) JFM92, and (c) JJA90.

in a seems much stronger than that for the bandpass variance. The best correspondence between these two values can be found in the JFM87 fields. The large maximum in a over the North Atlantic corresponds well in size and orientation to the maximum in the low-pass variance found in this region. A good correspondence can also be found for the secondary maximum in both fields found over central Europe and with the maximum found south of Alaska. For this period, these figures suggest that blocking, or at least the onset of

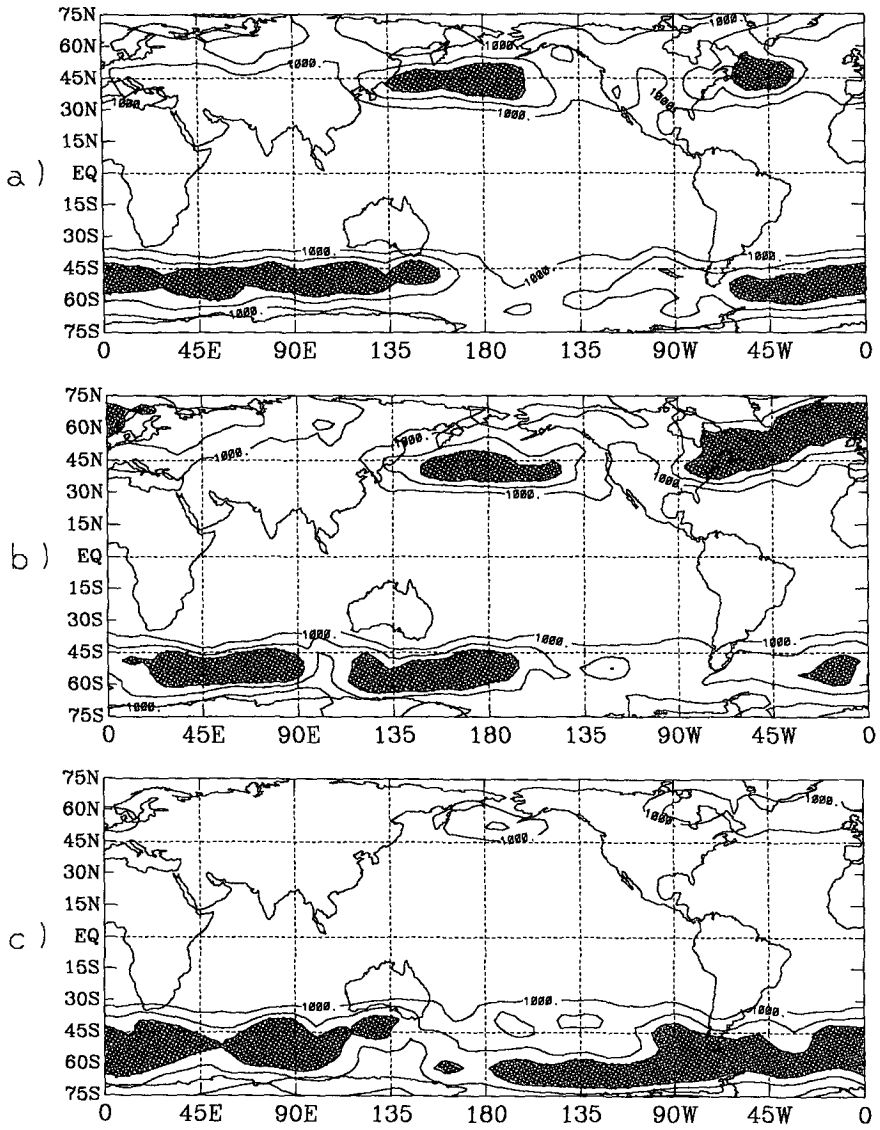

FIG. 16. Variance (m^2) for the bandpass-filtered 500-mb height field for (a) JFM87, (b) JFM92, and (c) JJA90. Contour interval is 1000 m^2. Values greater than 3000 m^2 are shaded.

blocks, have an impact on the 1–5-day forecasts. For the JFM92 case, the correlation between *a* and low-pass variance is not as striking but is still quite good. The large values in *a* that extend from northeastern Canada to northern Siberia correspond roughly to regions with large values of variance. However, the maximum values in variance extend considerably farther west into Canada than do the maximum values in *a*, suggesting that the eastward advection of errors may be important in these regions.

The results discussed above, which appear to illustrate a relationship between blocking and internal error growth in the Northern Hemisphere, are consistent with the findings of Tracton et al. (1989), Tracton (1990), and Chen (1989), who show that forecast skill is below average when a block forms 3 or more days after the initial forecast time.

4) INTRASEASONAL VARIATIONS

Figure 18 shows the internal error growth rates for the first (JF) and second (FM) halves of JFM87 for the 200-mb height field. In the Southern Hemisphere, the *a* fields are quite different for the two periods. In the Northern Hemisphere, there are common characteristics in both periods. For example, there is a large maximum over the Atlantic in *a* in both sets, although this maximum is considerably larger and extends farther northwest for FM than for JF. There is also a large maximum over Europe for both periods, although this

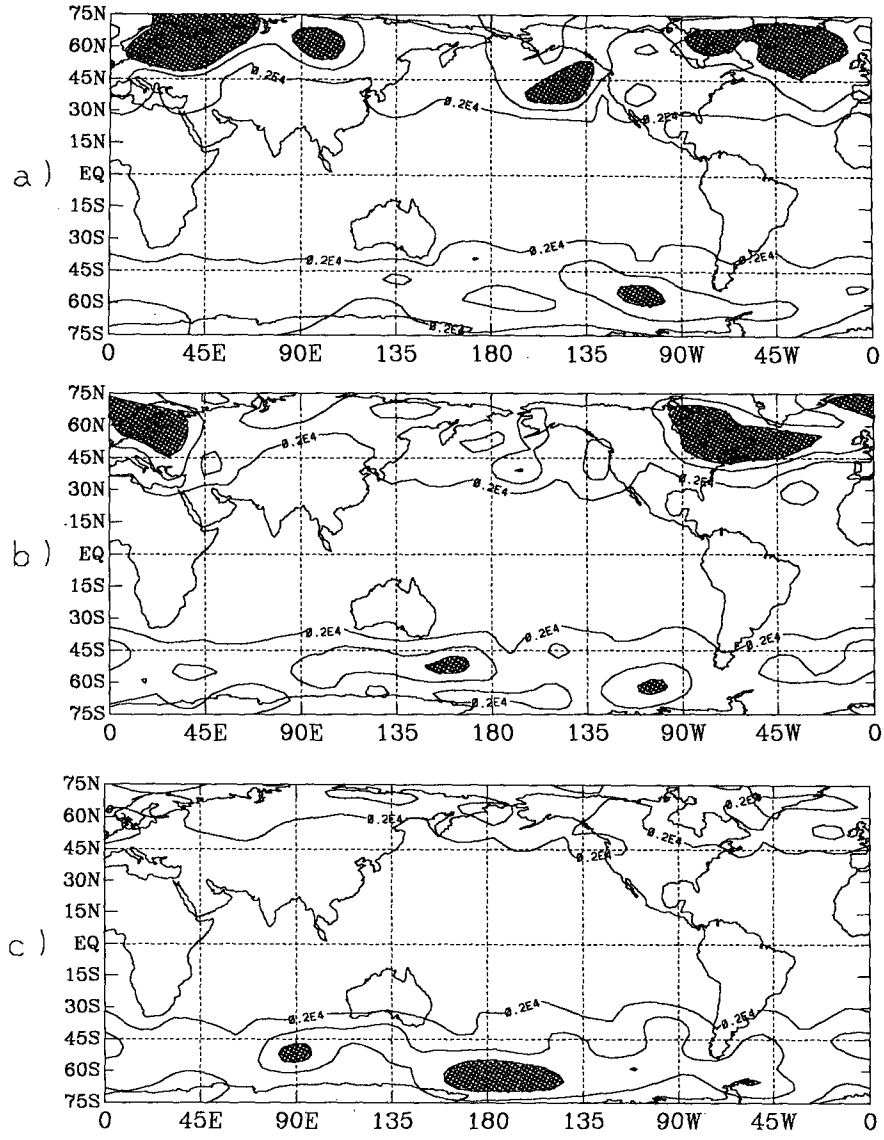

FIG. 17. Variance (m^2) for the low-pass-filtered 500-mb height field for (a) JFM87, (b) JFM92, and (c) JJA90. The contour interval is 4000 m^2. Values greater than 10 000 m^2 are shaded.

maximum is about 30° farther east for the FM case. In the North Pacific, the maximum for the FM time period is located over the Bering Sea, while for JF there are high values of *a* north of 70°N.

Figures 19 and 20 show the 500-mb height field bandpass and low-pass variances for JF and FM. For the bandpass variance fields the values are much higher for the first part of the period than for the second part in the Northern Hemisphere. This is very different from the behavior of the *a* fields, where values are higher for the second part of the period.

For the low-pass variance, there is a clearer link to the change in the *a* fields. The low-pass variance maximum is larger during FM than it is during JF in the North Atlantic and northeast Pacific. The position of the maxima in the low-pass variance is also broadly consistent with the positions and frequencies of the blocks during the DERFII experiment (JFM87) (see Tracton et al. 1989). This increase in low-pass variance for the second part of the period corresponds well with the higher values of *a* found in these blocking regions for FM. In the Southern Hemisphere, it is difficult to find any relationship between the variance fields and the *a* fields.

Local error growth is not only a function of local instabilities but also is affected by the advection of errors from other regions (Epstein 1969). This source of error growth may become particularly important in regions of differential advection. In an effort to estimate the importance of the advection of errors, the

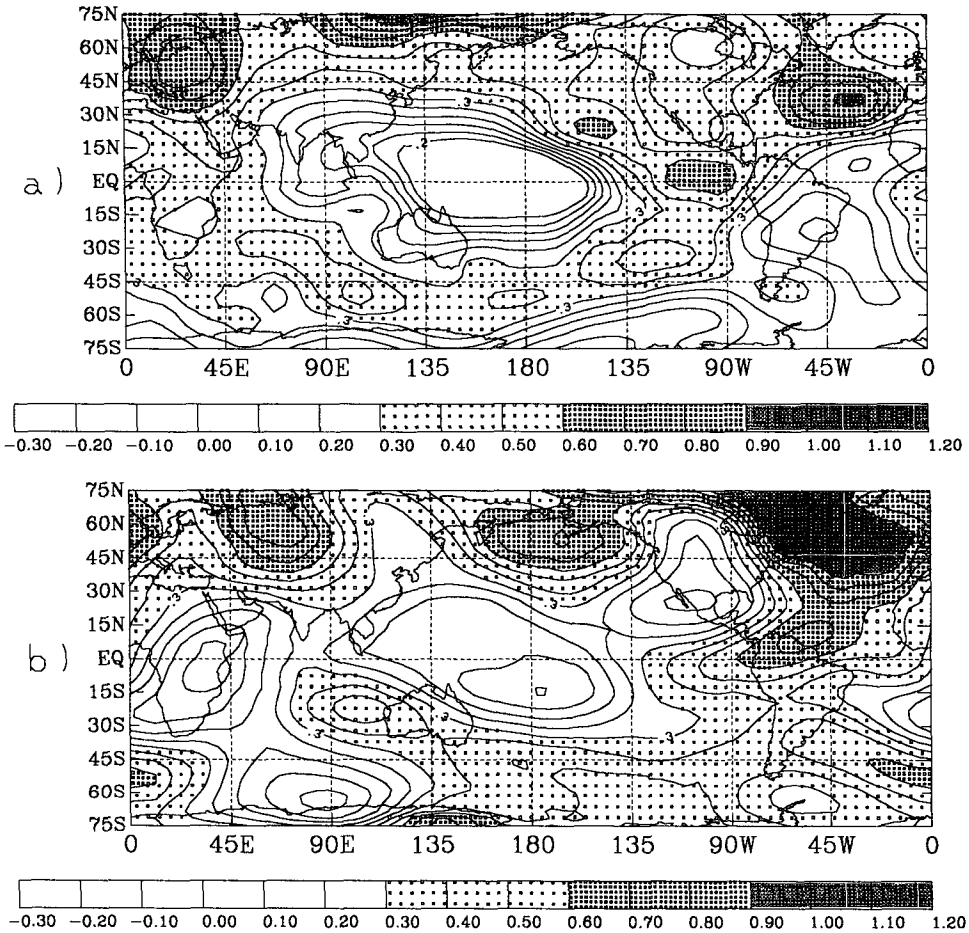

FIG. 18. Internal error parameter a (day^{-1}) for the 200-mb height field for the (a) first half and (b) second half of the JFM87 period.

changes per forecast day in the random mean-square differences between forecasts started one day apart, after Schubert and Suarez (1989), are shown in Figs. 21–23. The random mean-square difference is chosen over the correlation difference to determine if normalization of the correlation errors by the local variance is resulting in lower growth rate estimates in regions of higher variance. Figure 21 is a Hovmöller diagram of the change in the random mean-square difference (RD) per day for the JFM87 200-mb height field along a latitude circle from 45° to 65°N. This diagram shows that the RD growth rate is consistent with the internal error growth estimated using the parameterization (Fig. 15), and is also in basic agreement with the results of Schubert and Suarez (1989), showing the largest growth rates over the Atlantic and Europe, a relative maximum over the eastern Pacific, relative minima over Asia and the western Pacific, and western North America. It also clearly illustrates the eastward advection of large and small error growth rates. In regions where this advection is particularly strong, the error growth pattern may have a double maximum in time, which is inconsistent with the simple increase–decrease pattern expected from exponential error growth and will result in a poorer fit of the parameterization. This problem seems to be small for the first six or so forecast days but becomes more important at longer forecast lead times. This is consistent with a fairly large degradation in the goodness of fit when including more than approximately seven forecast days, as well as the significant change in the estimated values of a in the midlatitudes between the 7- and 10-day cases (Fig. 7).

Figures 22 and 23 show Hovmöller diagrams of the change in RD per day for 200-mb height between 45° and 65°N for the first and second halves of the JFM87 period. As with Fig. 21, these figures qualitatively match the corresponding internal error growth rates well (Fig. 18), with both figures showing similar patterns of high error growth rates over the Atlantic and Europe and lower growth rates over Asia and the western Pacific. These figures also reflect the intraseasonal differences in the eastern Pacific, with higher growth

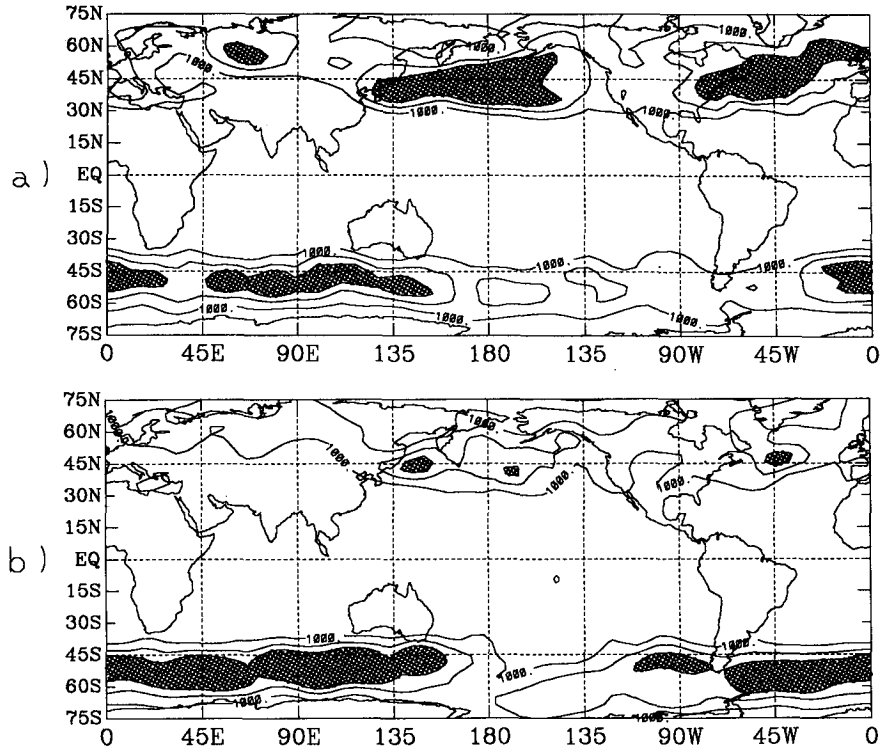

FIG. 19. Variance (m^2) for the bandpass-filtered 500-mb height field for the (a) first half and (b) second half of the JFM87 period. The contour interval is 1000 m^2. Values greater than 3000 m^2 are shaded.

rates during the second part of the season. They also illustrate that as the number of forecasts included in the parameterization is decreased (in this instance, from 90 to 45), the results are subject to more sampling errors and will be more likely dominated by specific events and the corresponding advection of errors related to them.

Both the seasonal and intraseasonal results suggest that blocking regions, rather than storm tracks, have a higher correlation to the local internal error growth rates in the Northern Hemisphere winter. This does not mean that baroclinically unstable synoptic disturbances are not primarily responsible for internal error growth in the midlatitudes but rather that it is the interaction between these transients and the onset of blocking events that dominate error growth on the 5-day time scale. This is not to say that enhanced internal error growth does not occur in regions of cyclogenesis or storm tracks, but rather that this enhanced error growth may occur on the time scales of 1 or 2 days, rather than 5. (A dataset with finer temporal resolution might be used to study this.) This result also illustrates how model deficiencies can sometimes result indirectly in exponential error growth. If the model does a poorer job of simulating the exit region of the Atlantic jet than it does the Pacific jet, model deficiencies are predisposing this area to have larger internal error growth rates.

5) TROPICS

The role of differential advection of errors was also investigated in the tropics; however, the results were inconclusive. Webster and Chang (1988) have theorized that equatorially trapped wave energy will accumulate in regions where dU/dx is negative (to the east of the westerly maxima). If this mechanism of wave energy accumulation is occurring, then it is possible that internal errors may also accumulate in these same regions. These regions of upper-tropospheric westerlies may also be conduits through which errors may propagate from the midlatitudes into the tropics (Webster and Holton 1982; Kiladis and Weickman 1992), or from the tropics into the extratropics (Zhang and Webster 1989). There are regions of the tropics with relatively large values of a. However, we were not able to determine whether these regions were areas of higher internal error growth or differential advection, or if they simply were regions not as seriously affected by the model spindown that takes place during the first few days, which, as noted before, acts to suppress internal

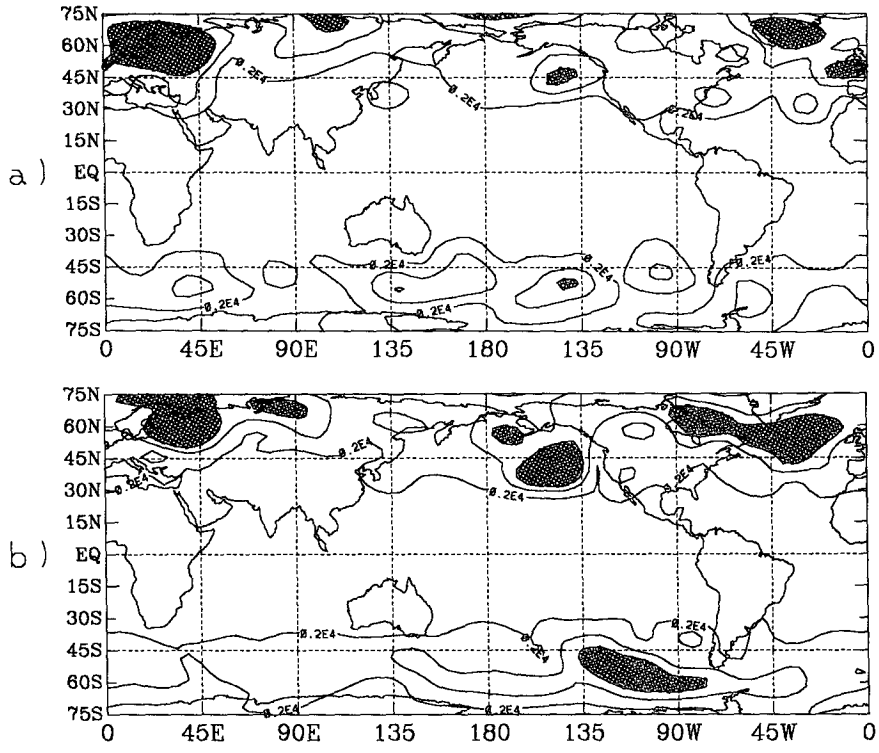

FIG. 20. Variance (m^2) for the low-pass-filtered 500-mb height field for the (a) first half and (b) second half of the JFM87 period. The contour interval is 4000 m^2. Values greater than 10 000 m^2 are shaded.

error growth. This problem illustrates how the severe model adjustment that takes place during the first few forecast days, due to inconsistencies between the analyses and the model climatology, makes this method inadequate for studying the mechanisms of internal error growth in the tropics.

5. Correlation differences

As mentioned earlier, it is assumed that most of the exponential part of the random error growth is due to the self-growth of errors primarily in the initial conditions. For a better understanding of the role of model deficiencies in random error growth, the values of a and z computed for the time correlation between two forecasts started one day apart (CD) are compared with a and z computed for the time correlation between forecasts and analyses (CE). Theoretically, since linear error growth is attributed to the differences between the two systems (the model and the atmosphere), this error growth (z) should be zero for the CD since the two systems being compared are identical.

Figure 24 shows the external and internal differences for the CD fields for the JFM87 200-mb height field. The values of z are indeed everywhere quite close to zero, with only a few regions near Antarctica having values of z greater than 0.06 day^{-1}. The scale for this graph is different than that of the corresponding graph of z calculated using CE fields (Fig. 12) because values of z for this field are much smaller (and some regions are less than zero). This result supports the assumption that the initial linear growth of random errors is caused primarily by model deficiencies.

The values of a for the CD are, in most regions, similar to values of a for the CE for the 200-mb height field (Fig. 15). The locations of the maxima in the Northern Hemisphere are similar for both fields, although the magnitude of a for the CD is smaller than that for the CE. A significant maximum in a for the CE field west of Australia of 0.5 day^{-1} is not seen in the a field calculated for the CD. This may indicate that model deficiencies are also contributing to the exponential error growth. However, a calculated using the CE fields is not everywhere greater than a calculated using the CD fields. The maximum in a centered at 60°S, 110°W is larger for the CD case than for the CE case. This may be owing to the problems caused by the fact that the actual initial conditions are never known, and the initial differences between forecasts started one day apart are not equal to the initial error of a forecast. This will lead to differences in the error growth rates. From these results, it is judged that the assumption that linear error growth is caused by model deficiencies is

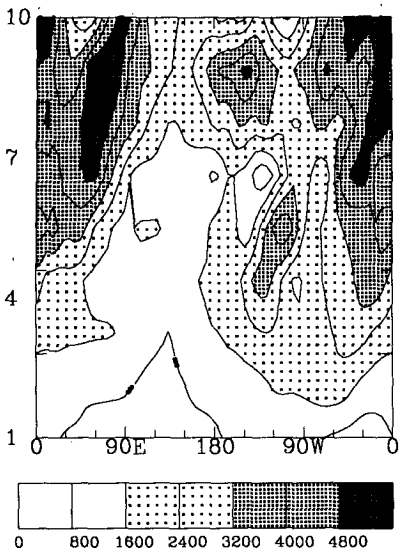

FIG. 21. Hovmöller diagram of the change in RD per forecast day (m^2 day^{-1}) averaged between 45° and 65°N for the JFM87 200-mb height field.

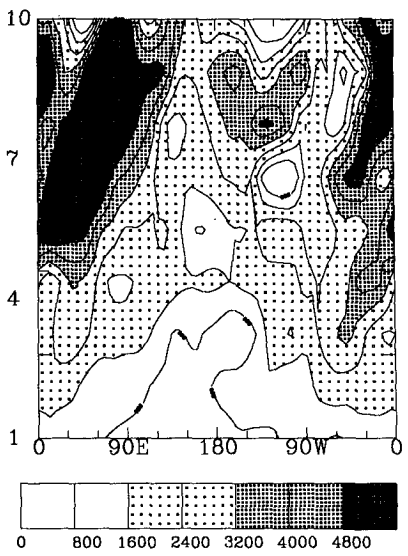

FIG. 22. Hovmöller diagram of the change in RD per forecast day (m^2 day^{-1}) averaged between 45° and 65°N for the JF87 200-mb height field.

good, even on a regional basis, and that exponential error growth is a reasonable, though not perfect, representative of the internal self-growth of errors in the initial conditions.

6. Conclusions

We have presented the results of a parameterization of random error growth designed to distinguish between random error growth due to model deficiencies and the internal growth of errors in the initial conditions. We now discuss conclusions about the impact of these findings on weather forecasting, as well as caveats and shortfalls concerning this study.

a. Evaluation of the performance of the parameterization

Overall, the parameterization proposed in section 2b successfully provides physically meaningful results about the sources of random errors (either internal error growth or model deficiencies) that dominate in different regions. However, there are limitations with this parameterization that should be kept in mind when evaluating the results.

Some simplifying assumptions made in the development of the parameterization may not always be applicable. The parameterization is based on the assumption that the two error coefficients remain constant throughout the forecast time included. However, the internal error results are somewhat sensitive to the number of forecast days included in the parameterization, particularly beyond forecast day 7 or 8, in the midlatitudes (Fig. 7). This may be the result of the advection of errors becoming more important after this time. The effect of the advection of errors means that the local error growth rates do not exhibit the simple increase–decrease pattern expected from exponential error growth (Figs. 21–23), and the fit of the data to the parameterization becomes poorer for long forecast times. This problem becomes more severe as the sample size decreases.

The assumptions that all linear error growth results from differences between the model and the atmo-

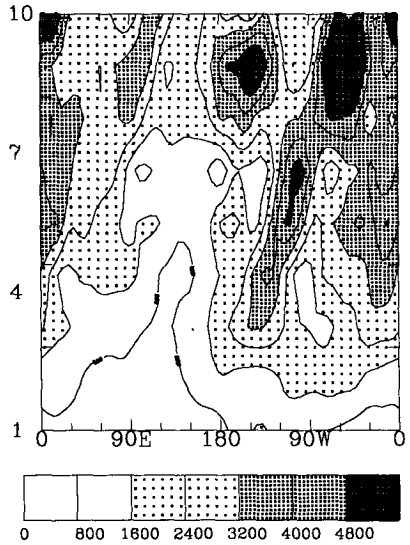

FIG. 23. Hovmöller diagram of the change in RD per forecast day (m^2 day^{-1}) averaged between 45° and 65°N for the FM87 200-mb height field.

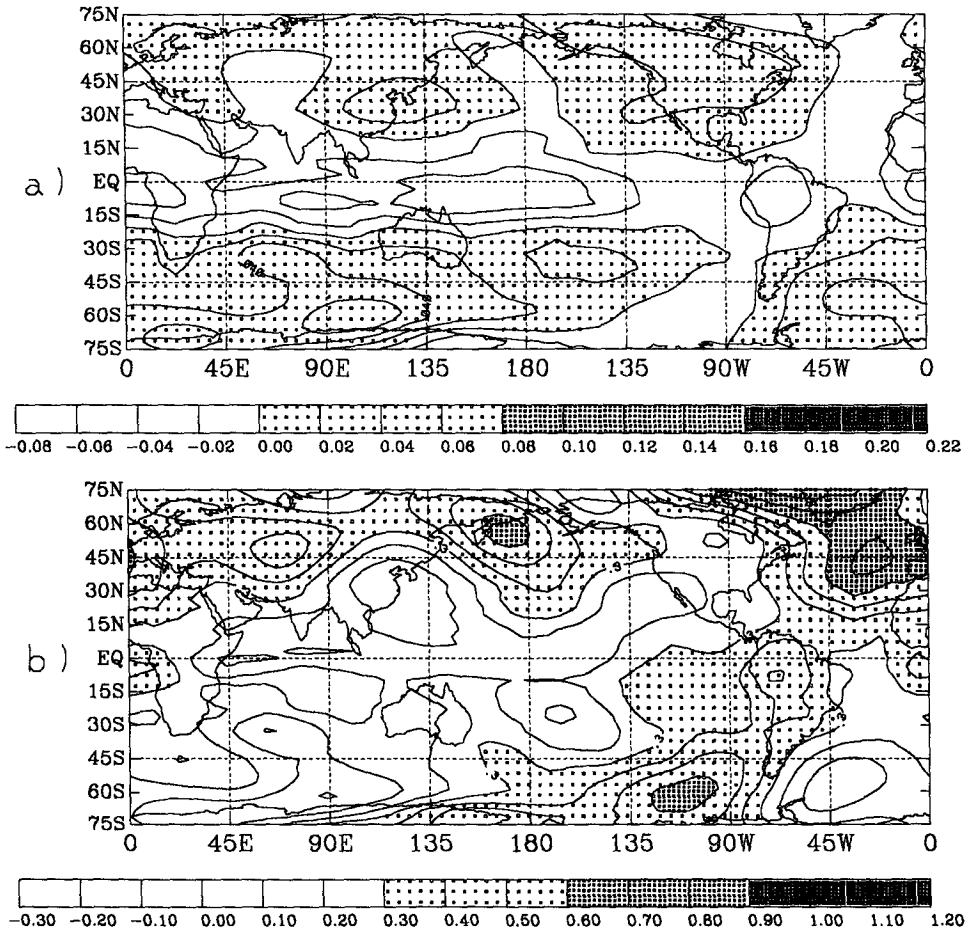

FIG. 24. (a) Linear forecast difference growth rates z (day^{-1}) and (b) internal error growth rates a (day^{-1}) computed for the correlation between forecasts started one day apart for the JFM87 200-mb height field. Note that the shaded values in (a) are different from those in Fig. 12.

sphere and all exponential error growth is due to the self-growth of errors are generally borne out by the results but are not perfect (see section 5 for discussion). External and internal error growth are not independent. An example of this interaction may be that the imperfect formulation of the Rockies suppresses baroclinic wave activity downstream in the storm tracks over the Atlantic (Klinker 1990), which may also have an effect on the difficulties the model has in simulating Atlantic sector blocks. Also, random error growth has been treated independently of systematic errors in the model, which, particularly in the tropics, constitute a significant part of the total forecast error growth (Saha and Alpert 1988). Certainly, the two may be related to each other, as when the systematic error results in a different midlatitude jet structure, which in turn results in random errors in the storm track region, and the systematic decrease in the upper-tropospheric height field in the tropics (White 1988) may be related to random errors in the radiation and convective parameterization schemes.

The parameterization cannot give meaningful high spatial resolution results because of sampling problems and the horizontal smoothing used to overcome this. The local error growth will also be sensitive to the growth and advection of errors associated with particular events. This sensitivity implies that the magnitude of the error growth rates calculated for one region may vary significantly from winter to winter, although the large-scale pattern of highest midlatitude error growth occurring over the Atlantic and Europe is consistent for both winters examined here, as well as the first and second part of the first winter. It would be helpful to examine more wintertime periods to assess the robustness of this general pattern. Also, there is a compensative relationship between a and z. This compensation is not a severe problem if one is just looking at the relative importance of the two error source mechanisms, but it is a hindrance when one wants to examine the spatial structure of one parameter where the other parameter dominates, as in the investigation of internal error growth in the tropics.

The possible existence of systematic analysis errors should be kept in mind when interpreting the results, especially in the tropical and polar regions. The inconsistencies between the analyses and the model climatology result in the severe adjustment process that occurs during the first few forecast days (which has been reduced through use of the new analysis scheme), and the paucity of data in these regions renders an assessment of predictability in the real atmosphere suspect.

Even considering these caveats, the parameterization turns out to be a useful tool for investigating both the internal growth of initial errors as well as locating possible sources of model deficiencies. It is not sensitive enough to pick up minor model improvements above seasonal variations but does reflect major analysis and model improvements, such as in the 850-mb height field between 1987 and 1992. It provides internal error growth rates consistent with previous studies and sheds new light on the nature of error and difference growth rates as a function of location and season.

b. *Implications for modeling*

The results of the correlation parameterization show that model deficiencies dominate random error growth in the tropics. This indicates that significant increases in forecast skill can be made through model improvements. This may be important for global forecasts since tropical forecast fields have been shown to affect midlatitude forecast fields within five days (Baker and Paegle 1983; Paegle and Baker 1983). While the model deficiencies at the 500-mb level and especially the 850-mb level show improvement between 1987 and 1992, there are still very significant errors in the tropics at all three levels, with the external error growth strongly correlated with deep convection at the 200-mb level. The inconsistencies between the analyses and the model's own climatology in the tropics are responsible for producing an initial convergence of forecasts made during consecutive days. However, the slow rate of divergence of the forecasts after this initial adjustment period also indicates that the natural mechanisms of error growth in the tropics on the resolved scales result in considerably slower error growth than the instability mechanisms that dominate in the extratropics. This is encouraging because it indicates that with substantial model improvements the tropics may be predictable for much longer time scales than the midlatitudes (Charney and Shukla 1981).

Unlike the tropics, the extratropical error growth has been shown to be dominated by the intrinsic unstable growth and subsequent advection of errors in the atmosphere, so that future model improvements have a much smaller potential to increase the skill in the extratropics. This suggests that forecast improvements in the midlatitudes will more likely come from improved analyses rather than from improved models.

Regionally, internal error growth in the Atlantic region is larger than internal error growth in the Pacific region during the two winters studied, although this may not necessarily be the case for other winter seasons or for other times of the year. Although in this study no clear relationship was found between internal error growth and storm track activity, it is certainly possible that seasonal variations in the storm track intensity and position, such as the spring and fall maxima in the Pacific storm track (Nakamura 1992), may influence large-scale error growth rates, especially given the close relationship between synoptic-scale eddies and block formation. Even though the model deficiencies are estimated to be less severe in the midlatitudes than in the tropics, improvements in the formulation of orographic forcing in the model should still result in significant improvements in forecast skill, especially for the first few forecast days. External error growth maxima are observed over the Himalayas at 850- and 500-mb levels for all three periods, and over the Rockies at all three levels for the winter periods.

Finally, it should be pointed out that even though the midlatitude error growth is primarily due to the self-growth of errors in the initial conditions, there is still reason to believe that model improvements may result in significant increases in skill in certain regions because model deficiencies appear to be at least indirectly responsible for some of the exponential error growth. Anderson (1993) suggests that it is the climate drift of the mean jets that inhibits the model's ability to form and sustain blocks during the first week of integration, and Klinker (1990) has shown that inadequate representation of the Rockies may lead to a suppression of Atlantic storm track activity. Correction of these types of model deficiencies may also lead to a reduction in the exponential error growth.

Acknowledgments. The authors would like to thank Dr. J. Alpert, Dr. S. Saha, Dr. W. Ebisuzaki, and Dr. Y. Zhou for their help in obtaining and processing the NMC datasets, and Dr. G. White and Dr. W. Chen for helpful discussions concerning characteristics of the MRF's systematic error growth and ability to simulate blocks. The comments of two anonymous reviewers helped to improve and clarify the paper. This research was supported by the Atmospheric Sciences Division of the National Science Foundation under Grant ATM-9003364.

REFERENCES

Alpert, J. C., M. Kanamitsu, P. M. Caplan, J. G. Sela, G. H. White, and E. Kalnay, 1988: Mountain induced gravity wave drag parameterization in the NMC medium range model. Preprints, *Eighth Conf. on Numerical Weather Prediction,* Baltimore, Amer. Meteor. Soc., 726–733.

Anderson, J. L., 1993: The climatology of blocking in a numerical forecast model. *J. Climate,* **6,** 1041–1056.

Arkin, P. A., and P. J. Webster, 1985: Annual and interannual variability of tropical extratropical interaction: An empirical study. *Mon. Wea. Rev.,* **113,** 1510–1523.

Arpe, K., A. Hollingsworth, M. S. Tracton, A. C. Lorenc, S. Uppala, and P. Kallberg, 1985: The response of numerical weather pre-

diction systems to FGGE level IIb data—Part II: Forecast verification and implications for predictability. *Quart. J. Roy. Meteor. Soc.,* **111,** 67–101.

Baker, W. E., and J. Paegle, 1983: The influence of the tropics on the prediction of ultralong waves. Part I: Tropical wind field. *Mon. Wea. Rev.,* **111,** 1341–1355.

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

Caplan, P. M., and G. H. White, 1989: Performance of the National Meteorological Center's Medium-Range Model. *Wea. Forecasting,* **4,** 391–400.

Charney, J. G., and J. Shukla, 1981: Probability of monsoons. *Monsoon Dynamics,* J. Lighthill and R. Pearce, Eds., Cambridge University Press, 99–109.

Chen, W. Y., 1989: Estimates of dynamical predictability from NMC DERF experiments. *Mon. Wea. Rev.,* **117,** 1227–1236.

——, 1990: Interannual variability of skill of NMC medium-range forecasts over the Pacific/North America sector. *Mon. Wea. Rev.,* **118,** 179–188.

Dalcher, A., and E. Kalnay, 1987: Error growth and predictability in operational ECMWF forecasts. *Tellus,* **39A,** 474–491.

Derber, J. C., D. F. Parrish, and S. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. *Wea. Forecasting,* **6,** 538–547.

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

Ferranti, L., T. N. Palmer, F. Molteni, and E. Klinker, 1990: Tropical–extratropical interaction associated with the 30–60 day oscillation, and its impact on medium- and extended-range prediction. *J. Atmos. Sci.,* **47,** 2177–2199.

Horel, J. D., and J. M. Wallace, 1981: Planetary-scale atmospheric phenomena associated with the Southern Oscillation. *Mon. Wea. Rev.,* **109,** 813–829.

Kalnay, E., P. J. Dallavalle, and E. S. Epstein, 1988: Hedging perfect prog forecasts towards climatology. *Research Highlights of the NMC Development Division: 1987–1988,* 197–201. [Available from NMC, Development Division, 5200 Auth Rd., Camp Springs, MD 20746.]

——, M. Kanamitsu, and W. E. Baker, 1990: Global numerical weather prediction at the National Meteorological Center. *Bull. Amer. Meteor. Soc.,* **71,** 1410–1428.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system. *Wea. Forecasting,* **4,** 335–342.

——, J. C. Alpert, K. A. Campana, P. M. Caplan, D. G. Deaven, M. Iredell, B. Katz, H.-L. Pan, J. Sela, and G. H. White, 1991: Recent changes implemented into the global forecast system at NMC. *Wea. Forecasting,* **6,** 425–435.

Kiladis, G. N., and K. M. Weickmann, 1992a: Circulation anomalies associated with tropical convection during northern winter. *Mon. Wea. Rev.,* **120,** 1900–1923.

——, and ——, 1992b: Extratropical forcing of tropical convection during northern winter. *Mon. Wea. Rev.,* **120,** 1924–1938.

Klinker, E., 1990: Investigation of systematic errors by relaxation experiments. *Quart. J. Roy. Meteor. Soc.,* **116,** 573–594.

Leith, C. E., 1971: Atmospheric predictability and two-dimensional flows. *J. Atmos. Sci.,* **28,** 148–161.

——, 1978: Objective methods for weather prediction. *Annu. Rev. Fluid Mech.,* **10,** 107–128.

——, and R. H. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.,* **29,** 1041–1058.

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

——, 1969: The predictability of a flow which possesses many scales of motion. *Tellus,* **21,** 289–307.

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

Miyakoda, K., R. F. Strickler, C. I. Nappo, P. L. Baker, and G. D. Hembree, 1971: The effect of horizontal grid resolution in an atmospheric circulation model. *J. Atmos. Sci.,* **34,** 3–21.

——, D. Hembree, R. F. Strictler, and I. Shulman, 1972: Cumulative results of extended forecast experiments. Part 1: Model performance for winter cases. *Mon. Wea. Rev.,* **100,** 836–855.

Paegle, J., and W. E. Baker, 1983: The influence of the tropics on the prediction of ultralong waves. Part II: Latent heating. *Mon. Wea. Rev.,* **111,** 1356–1371.

Palmer, T. N., 1988: Medium- and extended-range predictability and stability of the Pacific/North American mode. *Quart. J. Roy. Meteor. Soc.,* **114,** 691–714.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.,* **120,** 1747–1763.

Saha, S., and J. C. Alpert, 1988: Systematic errors in NMC medium-range forecasts and their correction. Preprints, *Eighth Conf. on Numerical Weather Prediction,* Baltimore, Amer. Meteor. Soc., 472–477.

——, and H. M. van den Dool, 1988: A measure of the practical limit of predictability. *Mon. Wea. Rev.,* **116,** 2522–2526.

——, E. Kalnay, M. Kanamitsu, and H. M. van den Dool, 1990: DERF90: A long series of extended range (90-day) forecasts. Research Activities in Atmospheric and Oceanic Modeling, Report No. 14, WMO/TD-No. 396, 6.29–6.30.

Schubert, S. D., and M. Suarez, 1989: Dynamical predictability in a simple general circulation model: Average error growth. *J. Atmos. Sci.,* **46,** 353–370.

——, F.-C. Chang, W. Lau, and R. Kistler, 1988: A reduced version of the NMC DERF II data set. Experimental Climate Forecast. NASA/Goddard Laboratory for Atmospheres, 31 pp.

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

——, 1985: Predictability. *Large-Scale Dynamical Processes in the Atmosphere,* B. J. Hoskins and R. P. Pearce, Eds., Academic Press, 87–122.

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

Tracton, M. S., 1990: Predictability and its relationship to scale interaction processes in blocking. *Mon. Wea. Rev.,* **118,** 1666–1695.

——, and R. Kistler, 1988: Activities in dynamic extended range forecasting (DERF) at the National Meteorological Center, Washington, D.C. Preprints, *Eighth Conf. on Numerical Weather Prediction,* Baltimore, Amer. Meteor. Soc., 715–718.

——, K. Mo, W. Chen, E. Kalnay, R. Kistler, and G. White, 1989: Dynamical extended range forecasting (DERF) at the National Meteorological Center. *Mon. Wea. Rev.,* **117,** 1604–1635.

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

Webster, P. J., and J. R. Holton, 1982: Cross-equatorial response to middle-latitude forcing in a zonally varying basic state. *J. Atmos. Sci.,* **39,** 722–733.

——, and H.-R. Chang, 1988: Energy accumulation and emanation regions at low latitudes: Impacts of zonally varying basic states. *J. Atmos. Sci.,* **45,** 803–829.

Wergen, W., 1984: Forced motion in the tropics. *ECMWF Weather Forecasting Workshop on Current Problems in Data Assimilation,* Shinfield Park, United Kingdom, ECMWF, 183–206.

Yang, S., and P. J. Webster, 1990: The effect of summer tropical heating on the location and intensity of the extratropical westerly jet streams. *J. Geophys. Res.,* **95,** 18 705–18 721.

Zhang, C., and P. J. Webster, 1989: Effects of seasonal flows on equatorially trapped waves. *J. Atmos. Sci.,* **46,** 3632–3652.