Simple Extensions of an NWP Model

J. QIN* AND H. M. VAN DEN DOOL

Climate Prediction Center, Washington, D.C.

(Manuscript received 13 February 1995, in final form 5 July 1995)

ABSTRACT

This paper presents a study on simple and inexpensive techniques for extension of NMC's Medium Range Forecasting (MRF) model. Three control forecasts are tested to make 1-day extensions of 500-mb height fields initiated from the MRF at days 0–9. They are persistence (PER), a divergent anomaly vorticity advection model (dAVA), and the empirical wave propagation (EWP) method.

First the traditional 1–10-day global forecasts made by the MRF and the three controls from a common set of 361 initial conditions are discussed. Taking this as a basis, 1-day extension control forecasts starting from MRF prediction over four successive winters are examined next. Experiments show that regardless of the presence or absence of the systematic error in the MRF model output, there exists some point ($T_0 = n$) into the forecast after which the 1-day extension of the day n MRF out to day $n + 1$ by a control forecast is as good as or better than the continued integration of the full blown MRF model. In particular, the EWP provides a 1-day extension that beats the MRF most consistently after about 6 days in the Northern Hemisphere. Decomposition of the forecasts in terms of zonal harmonics further indicates that the skill improvement over the MRF is primarily in the long waves, but contributions from shorter waves are not negligible.

Efforts have been made to understand the mechanisms by which simple methods are superior to complicated models for low-frequency prediction at extended range. It seems that at least two simplifications made in one or all of the control forecasts are crucial in outperforming the MRF beyond day 6. The first one is well known, that is, the contaminating effects of synoptic-scale baroclinic eddies have been filtered out in the simple models considered. More generally, the nonlinear terms (whether barotropic or baroclinic) contribute to skill deterioration beyond day 6. The second reason is the explicit elimination of the divergence process in the control forecasts, as the MRF model may contain significant errors in forecasting the divergence.

1. Introduction

Numerical weather prediction (NWP) models are nowadays used as the primary tool in weather forecasts from 1 to up to 10 days ahead. There is the inevitable question about skill. Is the NWP model of choice skillful out to day 4, 8, or even 15? Since skill is a relative notion, we must always compare a measure of accuracy of NWP forecasts to the same measure for a control forecast. Saha and Van den Dool (1988, hereafter referred to as SD88) showed that at some point in the forecast (around day 6), persistence of the day 6 forecast of the 500-mb height field out to day 7 was as good as or better than the continued integration of NMC's Medium Range Forecasting (MRF) model.

From a rather different point of view, one can argue that errors in the synoptic scales in the current MRF model tend to be large around day 5. Beyond day 5, the skill of the MRF in continued integration depends mainly on the behavior of low-frequency large-scale eddies that carry much of the variance. It is well known that smaller scales can influence the evolution of the larger scales and thus might act to destroy the remaining prediction skill of low-frequency components beyond day 5. It is thus possible that certain models of reduced resolution and physical complexity (or "truncated models") may be used to extend the MRF with more skill than the MRF model itself. Other ways to eliminate the erroneous part of the feedback from synoptic eddies have been suggested in the literature. For example, averaging the members of an ensemble of forecasts (Tracton and Kalnay 1993) might accomplish this goal. Alternatively, Branstator et al. (1993) investigated the possibility to identify highly predictable flow elements for medium- and extended-range forecasts. Their idea is to remove the unpredictable components of the forecast a posteriori by projection onto an EOF basis.

In this paper, we follow the "truncated model" idea. Compared to three inexpensive control forecasts, a systematic study of the skill of MRF 500-mb height forecasts for four winters from 1990 to 1994 is presented.

* Additional affiliation: Scientific Management and Applied Research Technologies, Inc., Silver Spring, Maryland.

Corresponding author address: Dr. Jianchun Qin, Climate Analysis Center, NMC/NWS/NOAA, World Weather Building, 5200 Auth Road, Camp Springs, MD 20746.

The analyses are based on two kinds of experiments. One is for the traditional 1–10-day predictions and the other one is for the 1-day extension forecasts. Efforts are further made to explore the physical reasons as to why simpler forecast methods are or are not better than the MRF at certain forecast leads.

The three control forecasts studied in this paper are persistence (PER), a divergent anomaly vorticity advection model (dAVA), and the so-called empirical wave propagation (EWP) method (Van den Dool and Cai 1994). Short explanations of the control forecasts and the data procedures are given in section 2. Results of the forecast experiments are presented in section 3; followed by conclusions and discussions in section 4.

2. Data and control forecasts

a. Data

The focus of this study is to examine 1-day extensions of the 0–9-day MRF 500-mb geopotential height (Z500) forecasts. The MRF forecasts were made in real time by NMC's operational MRF model (Sela 1980, 1988), running at a horizontal resolution of T80 (1990–92) and T126 (1992–95). The MRF data used consist of daily global Z500 analyses and 1–10-day forecasts on a 2.5° × 2.5° latitude–longitude grid. In total, 361 cases are selected to represent a set of 1–10-day forecasts and verification in four successive winters over the period 15 December 1990–15 March 1994.

The daily control forecasts are made in terms of the Z500 anomaly. To define anomalies, a daily climatology, interpolated to the day in question, is subtracted from the total field. The climatology used is derived from the monthly means of the operational NMC analyses from January 1979 to December 1993 that were archived in the Climate Diagnostics Data Base (CDDB) at NMC.

The control forecasts or 1-day extensions are calculated by the three methods discussed below, starting either from analyses (day 0) or from MRF forecasts at days 1–9. These forecasts are to be compared to the MRF forecast valid at days 1–10, through which we explore the possibility that simple control forecasts would be competitive when the MRF becomes less skillful.

b. Persistence

Among the three control forecasts examined in this study, the most straightforward one is the persistence (PER) forecast. As done in SD88, we persist the daily MRF anomaly for 1 day to get the next day's control forecast. That is, to verify the MRF forecast valid at day $n - 1$ as if it were the new n-day forecast ($n = 1, 2, \cdots, 10$). Results from SD88 indicated that with MRF vintage 1986–87, the 1-day PER will eventually beat the MRF after it reaches the "practical limit of predictability" at day T_0. The concept of "practical limit of predictability," or better perhaps, "practical limit of prediction skill for NWP," can thus be introduced to define the time at which there is no skill in the prediction of the time derivative of the quantity under consideration. Afterall, PER amounts to a forecast with zero tendencies.

c. dAVA model

The anomaly vorticity advection (AVA) model as described in Van den Dool (1991) has been used for several years to aid in the official NMC 6–10-day average forecast. Recently, the model has been modified slightly by the so-called Bolin–Cressman correction (Rinne and Jarvinen 1993), here in anomaly form, to include the dynamical effect of divergence on the phase speed of anomaly waves. In equation form, the new model, dAVA, can be written as

$$\frac{\partial}{\partial t}(\zeta' - \lambda^2\psi') = -\bar{\mathbf{V}}\cdot\nabla\zeta' - \mathbf{V}'\cdot\nabla(\bar{\zeta} + f + \zeta') - D\nabla^4\zeta' - K\zeta', \quad (1)$$

where ζ is the vorticity, $\zeta' = \nabla^2\psi'$, $\mathbf{V}' = \mathbf{k} \times \nabla\psi'$, and the overbar and prime denote climatological mean and the anomaly, respectively. All other symbols are standard. The damping coefficient K in (1) is 2.5E−6 s^{-1}, while the coefficient of the fourth-order diffusion term D is chosen such that at total wavenumber $n = 30$ the linear and diffusive drag on a spherical harmonic wave are the same. The constant in the correction term, λ^2, is taken to be 0.75E−12. For $\lambda^2 = 0$, the dAVA reduces to the original AVA model. The dAVA is applied at 500-mb level and runs at T30 in horizontal resolution. The integration starts from the vorticity anomaly, but at the end, the vorticity is transferred to the geopotential height anomaly based on the linear balance equation (Holton 1979, p. 179). After the anomaly forecast is produced, the climatology valid at the initial time is first put back to form the forecast of the total field, as we keep the climatology constant during the dAVA integration. We then subtract the climatology valid at the verification time for calculation of the forecast scores such as the anomaly correlation (AC).

In (1), the anomaly winds are rotational. The basic-state wind may contain both rotational and irrotational components. Although the accuracy of the climatological divergent wind may be questionable, we decided to include the irrotational components into the basic wind field. Our experiments show that the inclusion or exclusion of the divergent wind in the basic state does not change our conclusion substantially.

One may also notice that a forcing term representing the climatological effects of transient eddy vorticity fluxes has been neglected in (1) (see Johansson and

Saha 1989; Zhao et al. 1994). Numerical experimentation has shown that using prescribed observed transient eddy forcing does not improve the dAVA's performance. We speculate that this is because the one-level model is incapable to produce transient eddy vorticity fluxes that, on average, would balance the prescribed forcing. Based on such reasoning, we opted for not including the eddy forcing in (1). Another way of putting it: the model based on (1) has less climate drift than a model based on (1) including an unbalanced prescribed forcing.

d. Empirical wave propagation

The empirical wave propagation (EWP) method (Van den Dool and Cai 1994) makes forecasts from knowledge of the climatological propagation speeds of zonal harmonic waves in anomaly Z500 fields. It was shown by these authors that a 1-day control forecast made by EWP can easily beat persistence of initial anomalies, particularly in the Southern Hemisphere where phase speeds are high. Technically, the EWP propagates anomaly waves at their climatological phase speeds. The amplitude of each anomaly wave is slightly damped for 1–10-day traditional forecast but is fixed for 1-day extension forecast. The phase speed and the damping function are derived from an observational dataset by applying the "phase-shifting" method (Cai and Van den Dool 1991). For reference, Fig. 1 shows the phase speed by which the zonal harmonic waves are moved for 1-day extension control forecast at 50°N, 50°S and the equator. Speeds at other latitudes (every 2.5°) are available but not shown. The data for Fig. 1 are from an extended time series of Z500 across eight successive winters between 1987 and 1995, for a total of 699 days. The anomaly over 1987–95 is defined in the same way as described in section 2a. In order to refine the calculation, we further exclude nonmoving anomaly waves, if any, by removing pointwise the time mean anomaly of the eight winters.

e. An example of 1-day extensions

Figure 2 gives an example of 1-day extension forecasts made by EWP and dAVA respectively. We start from the analysis on 3 December 1990. While the EWP-based forecast (Fig. 2c) may look similar to the initial condition (Fig. 2a), it does in fact undergo nontrivial changes with time. The EWP method not only moves disturbances zonally but also modulates their amplitudes significantly by the empirical Rossby wave dispersion process. For example, the development of the blocklike high in the Atlantic and the deepening of the low in the northeast Pacific are predicted with some skill.

The 1-day forecast given by the dAVA (Fig. 2d) looks much more like its verification (Fig. 2b) in spatial pattern. The block in the North Atlantic and the circulation pattern in the northeast Pacific are all simulated quite well. Compared to EWP, the 1-day changes produced by dAVA are relatively large.

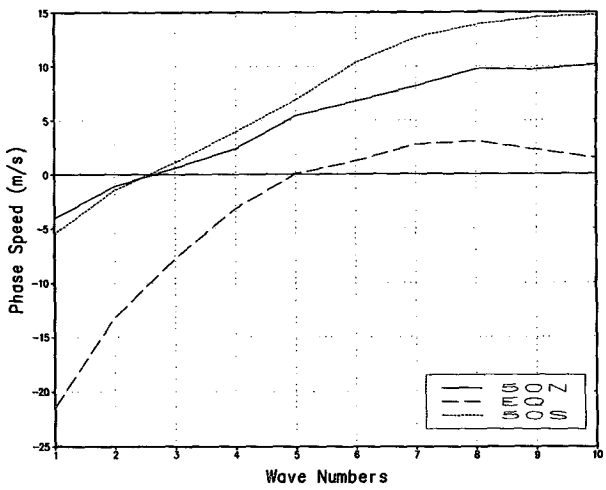

FIG. 1. The mean phase speed of zonal harmonics as a function of zonal wavenumber. The phase speed is calculated from 699 observed height fields at 500-mb level during DJFM 1987–95.

As for EWP, it is noted that while it generates forecasts very similar to the PER, starting from observations (i.e., 1-day extensions from day 0 to day 1), EWP scores better than PER on every single day in the 361-day dataset. Moving the anomaly waves is obviously a safe procedure, conservatively close to PER, but just a bit better, as long as the wave speeds are accurately known.

3. Results

a. Traditional forecasts

For reference we first look at regular 1–10-day forecasts made by the three simple methods described above, as well as for the MRF model. All forecasts start from a common set of initial conditions at day 0, which includes 361 cases, and are verified against a common set of Z500.

The anomaly correlation (AC) for these traditional forecasts made by MRF, PER, dAVA, and EWP are shown in Fig. 3. The ACs are calculated across latitude belts in the Northern Hemisphere (NH) (30°–80°N) (Fig. 3a), the Southern Hemisphere (SH) (30°–80°S) (Fig. 3b), and the Tropics (TROPICS) (20°S–20°N) (Fig. 3c), respectively, and averaged over 361 members. In general, the MRF is obviously by far the best in all areas. The dAVA is the second best until day 7–8 in the extratropics. The scores for EWP and PER are relatively lower in both NH and SH. But EWP beats PER at all times from day 1 to 10. The three controls

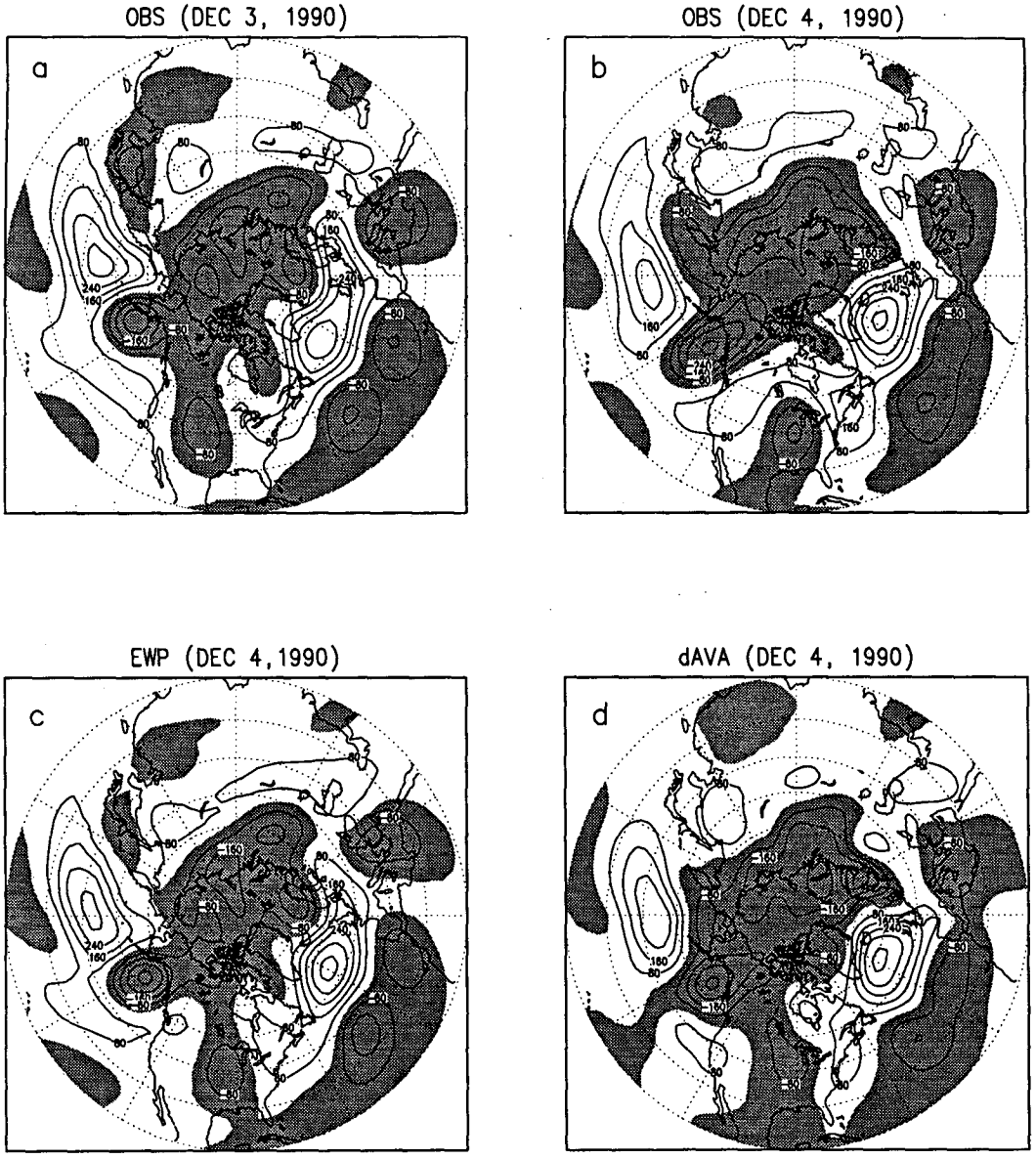

FIG. 2. An example of 1-day EWP and dAVA control forecasts for 500-mb height anomaly. (a) Observed initial condition on 3 December 1990. (b) Verification on 4 December 1990. (c) One-day forecast made by EWP. (d) One-day forecast made by dAVA. The contour interval is 80 gpm, negative values are shaded, and the zero contour is absent.

are close to the MRF in the Tropics because all methods face apparently very severe difficulties right from the start in that part of the world.

The MRF forecasts are better in the NH (Fig. 3a) than in the SH (Fig. 3b). Also, the gain over the nearest competitor (dAVA) is substantially larger in the NH. The MRF reaches AC = 0.6 at day 6.5 (5.5) in the NH (SH). In the NH, at the AC = 0.6 level, the MRF is about 3 (4, 4.5) days ahead of the dAVA (EWP, PER). For the SH the gains of MRF over dAVA (EWP, PER) are about 2 (3.5, 4.5) days at the AC = 0.6 level. Comparing scores of the two hemispheres further, we note that both EWP and dAVA are very close, while MRF and PER are considerably better in the NH. The higher ACs for PER in the NH can be understood by the fact that phase speeds of zonal waves are relatively lower in the NH (see Fig. 1). The higher AC for MRF in the NH cannot be explained readily. If better initial data was the complete explanation, there is no reason as to why dAVA would score virtually equal in the two hemispheres.

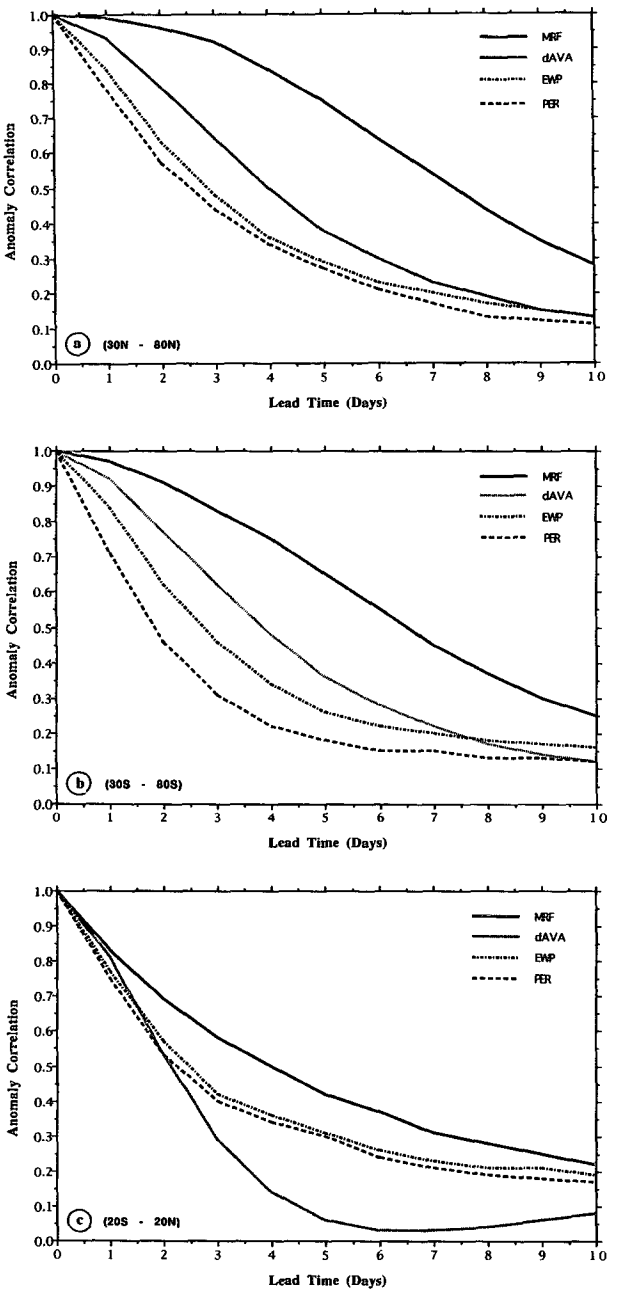

FIG. 3. The mean anomaly correlation (AC) of 361 500-mb height forecasts made by four different methods, as a function of lead time (days). The verification period is 1 December 1990–28 February 1994. The spatial domain is (a) 30°–80°N, (b) 30°–80°S, and (c) 20°S–20°N.

Overall, the scores for dAVA, EWP, and PER drop rather quickly compared to the sophisticated MRF model before day 5. Starting from accurately known initial conditions, the synoptic scale and other complicated physical processes involved in the MRF have plenty of skill. In fact, during that early period, dynamics such as baroclinic instability plays an important role in skillful prediction. Absence of basic dynamics/physics and sufficient model resolution clearly results in large error growth, thus making all control forecasts inferior. Considering Fig. 3a and 3b, it would seem inconceivable that either dAVA or EWP or PER could be serious competition to the MRF. However, beyond day 5 the n-day forecast (the "initial condition" for day $n + 1$ forecast) is no longer accurately known, and the simple methods do become competitive by taking less risks.

b. Control forecasts: 1-day extension

To evaluate 1-day extension forecasts of MRF given by PER, dAVA, and EWP, we first estimate the "practical limit of prediction skill" (T_0). As suggested in SD88, it can be used as a straightforward criterion to determine whether the easily obtainable control forecast at day n actually improves the integration of MRF model from $n - 1$ out to n days. The existence of the T_0 defined by PER in the NH is shown in Fig. 4a. For clarity, only results at days 6–10 are shown because MRF is far superior until day 5. We see that a T_0 exists at day 6, beyond which the MRF has no extra skill over the 1-day persistence extension. Also, the numerical value of T_0 obtained here is almost the same as before (SD88). Similar pictures with similar results can also be constructed for the EWP (Fig. 4b) and the dAVA (Fig. 4c) extensions. The estimates for T_0 based on EWP and dAVA are about day 6 also.

One detail in the calculation needs to be mentioned. The 1-day EWP extension forecasts had to be "cross-validated." Specifically, for each winter from 15 December 1990 to 15 March 1994, the calculation of the climatology/anomaly and the climatological phase speeds are made with that winter withheld. The application of the cross-validation (CV) approach, as far as the EWP extension is concerned, avoids mixing the data from the training period and the data from the validation period. Because of the redefinition of "anomaly," depending on the year withheld, we also felt compelled to express anomalies for dAVA, MRF, and PER in CV mode and verify all the methods in the same way.

It is of interest to examine the impact of the MRF's systematic error on the existence of T_0. In response to SD88, it has often been suggested that a crossover point (T_0) is caused by the growth of the systematic error. Table 1 shows a comparison for the values of T_0 given by the PER and the EWP before and after the systematic error in the MRF is removed a posteriori. The systematic error in the MRF is corrected by using the cross-validation technique, in which the verified winter (between 1990 and 1994) is excluded when the systematic error is calculated over the eight winters from 1987 to 1995. The dAVA was not rerun

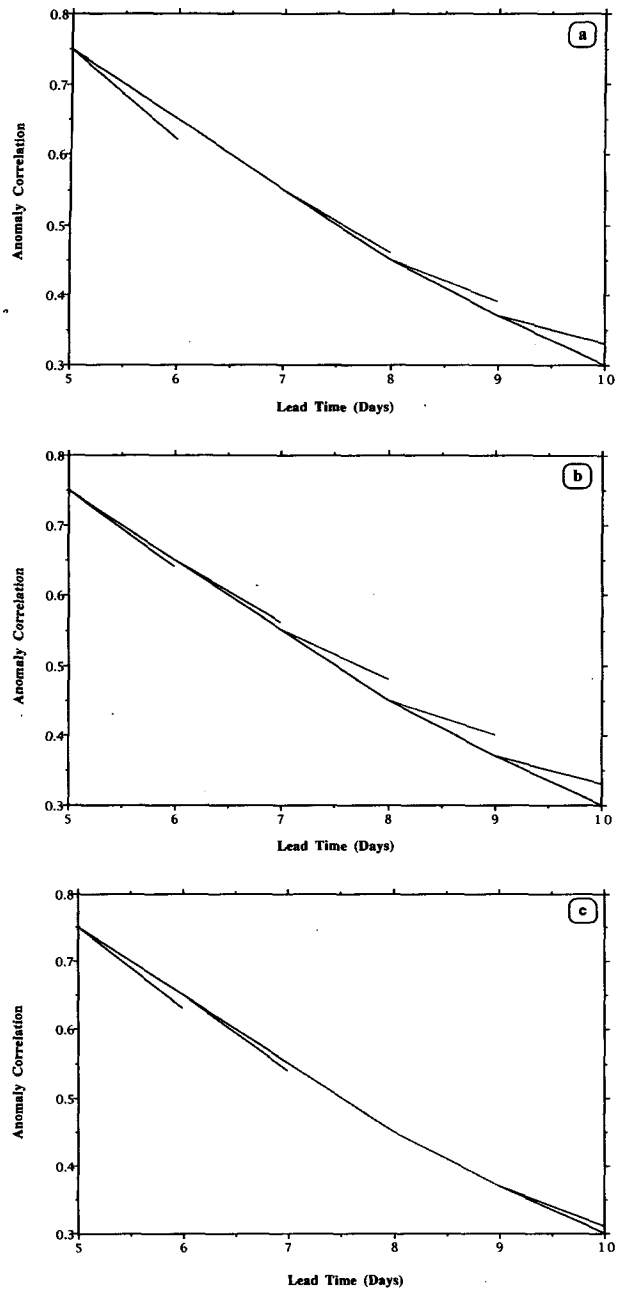

FIG. 4. The mean anomaly correlation (AC) of 361 500-mb height forecasts as a function of forecast lead (days). The continuous curve is the AC from MRF. The five segments represent verification of the day n MRF forecast extended out to $n + 1$ days by (a) PER, (b) EWP, and (c) dAVA controls, $4 < n < 9$. The verification period is 1 December 1990–28 February 1994. The domain is 30°–80°N.

from MRF, corrected for its systematic error. We see that removing the systematic error in the MRF does not address the existence of T_0, nor does it change its value substantially. This agrees with Robinson and Qin (1992), who found a T_0 in a perfect model predictability experiment where the model is free from systematic errors.

Figure 5a further shows a comparison, in histogram form, of scores (AC) of 1-day extensions among PER, dAVA, and EWP, together with that of the MRF in the NH. It is seen that the EWP is most consistently better than the MRF and is the best control after day 6. The dAVA is as good as or better than the MRF but does not beat PER consistently. The 6–10-day-averaged skills for the three control forecasts and the MRF are shown on the right. The improved skill by using simple control forecasts is evident. Further shown are 6–10-day-averaged skills with the systematic error in the MRF removed, which, incidentally, also improves the controls slightly. The contrast between control forecasts and the MRF are roughly the same, no matter whether the systematic error in MRF is corrected or not.

In the SH, there also exists a T_0 for EWP forecast at day 6 (Fig. 5b). But the difference between MRF and EWP after T_0 is small. Moreover, no clear T_0 can be found for either PER or dAVA. Particularly, PER performs rather poorly in this hemisphere, in agreement to what is seen in Fig. 3b. It seems that mechanisms not involved or not well represented in PER, dAVA, and, to some degree, EWP are helpful to forecasts at extended range in the SH, even if the same mechanisms hurt forecasts at the same range in the NH.

Evaluations of the 1-day extension forecast scores in terms of different zonal harmonics bands are shown in Fig. 6 and Table 1. In general, prediction skill beyond day 5 is dominated by planetary-scale waves ($m = 1$–4) (Fig. 6a,c), featuring AC = 0.67 (NH) and 0.53 (SH) at day 6. There is less skill left for shorter waves ($m = 5$–10); the AC is about 0.42 in both the NH and the SH at day 6 (Fig. 6b,d), with almost no skill at all for small waves ($m > 11$) (not shown).

The difference in ACs between the MRF and the 1-day extensions reported above are concentrated mostly in the long-wave spectrum, although contributions from shorter waves are not negligible. The PER performs remarkably well for long waves at almost all days from 5 to 10 in the NH (Fig. 6a), basically because long waves are usually of very low frequency nature and small changes in time are, by definition, hard to predict. In the SH, PER is the worst for $m = 5$–10 since short waves more quickly there. Scores from the 1-day EWP are roughly the same as that for PER for the nearly stationary long anomaly waves but are much better than PER for shorter mobile waves. The reason for this is that shorter waves have large phase speeds, and we know, apparently with some skill, both the phase of these waves at the initial time of the 1-day extension and their speeds when applying EWP.

Similarly, dAVA is on average better than PER for shorter waves. It is, however, worse than PER (and EWP) in the long-wave spectrum because dAVA has

TABLE 1. Practical limit of prediction skill (T_0) in NH derived from PER, dAVA, and EWP extensions, as those shown in Fig. 4.

Control forecast (30°–80°N)	Systematic error included			Systematic error excluded		
	m = 1–20	m = 1–4	m = 5–10	m = 1–20	m = 1–4	m = 5–10
PER	6	6	9	7	6	9
EWP	6	5	6	6	5	6
dAVA	7	6	9	—	—	—

problems of its own with long waves. A detailed comparison (not shown) of the phase speed produced by 1-day integration of the dAVA to those used in the EWP (Fig. 1, based on observation) indicates that, in dAVA, long waves (m = 1 mostly) still retrogress too strong even though this has been corrected to a considerable extent by including the Bolin–Cressman term in (1). This explains why EWP is generally better than dAVA though processes such as wave–wave interaction and barotropic instability are missing in EWP. Apparently, these dynamical processes are known with some accuracy early on in the forecast and should benefit forecast skill of a nonlinear model that starts from an accurately known initial condition. The resulting gain in forecast skill of dAVA over EWP is, however, reduced by the inaccurate wave propagation in the dAVA model.

c. *Possible mechanisms*

Besides the decomposition of the 1-day extension forecasts in terms of different zonal harmonics, other efforts have been made to explore how it is possible for simple methods to be superior to a complicated NWP model for extended range forecasts.

We write the vorticity equation for the anomaly field at the 500-mb level for different forecast methods symbolically as follows:

$$\text{MRF:}\quad \frac{\partial \psi'}{\partial t} = \nabla^{-2}(\text{RWP} + \text{LBI} + \text{NL} + \text{DIS} + \text{DIV} + \text{other}), \quad (2a)$$

$$\text{dAVA-0:}\quad \frac{\partial \psi'}{\partial t} = (\nabla^2 + \lambda^2)^{-1} \times (\text{RWP} + \text{LBI} + \text{NL} + \text{DIS}), \quad (2b)$$

$$\text{dAVA-1:}\quad \frac{\partial \psi'}{\partial t} = (\nabla^2 + \lambda^2)^{-1} \times (\text{RWP} + \text{LBI} + \text{DIS}), \quad (2c)$$

$$\text{dAVA-2:}\quad \frac{\partial \psi'}{\partial t} = (\nabla^2 + \lambda^2)^{-1} \times (\text{RWP} + \text{NL} + \text{DIS}), \quad (2d)$$

$$\text{dAVA-3:}\quad \frac{\partial \psi'}{\partial t} = (\nabla^2 + \lambda^2)^{-1}(\text{RWP} + \text{DIS}), \quad (2e)$$

$$\text{PER:}\quad \frac{\partial \psi'}{\partial t} = 0, \quad (2f)$$

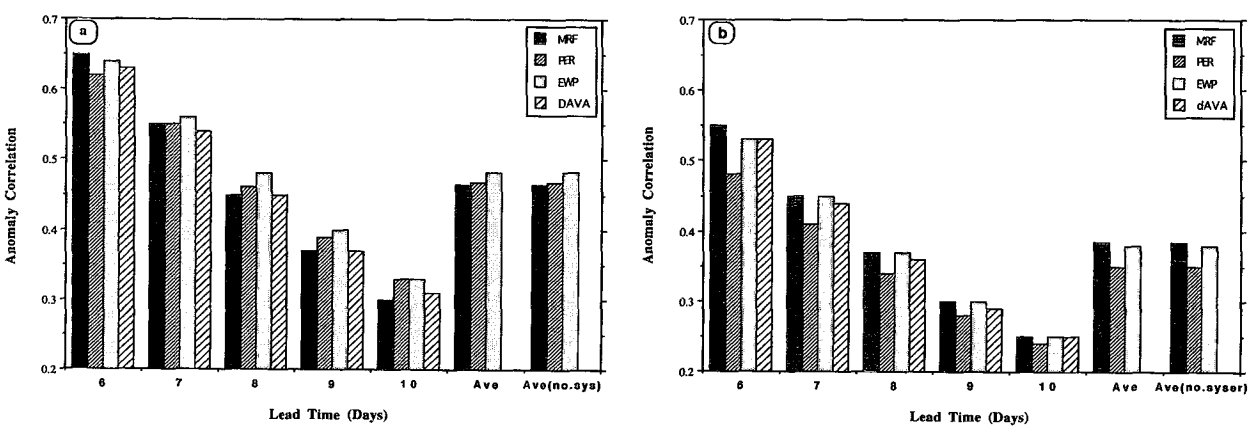

FIG. 5. The mean anomaly correlation (AC) of 361 500-mb height forecasts as a function of lead time (days). For fixed lead, the first bar gives the AC from the MRF model; the second to fourth bars represent verifications of the day n MRF forecast extended out to n + 1 days by PER, dAVA, and EWP methods subsequently. Bars marked with "Ave" denote ACs averaged over day 6–10, and "no sys" stands for ACs calculated with systematic error in MRF removed. The verification period is 1 December 1990–28 February 1994. (a) NH (30°–80°N), (b) SH (30°–80°S).

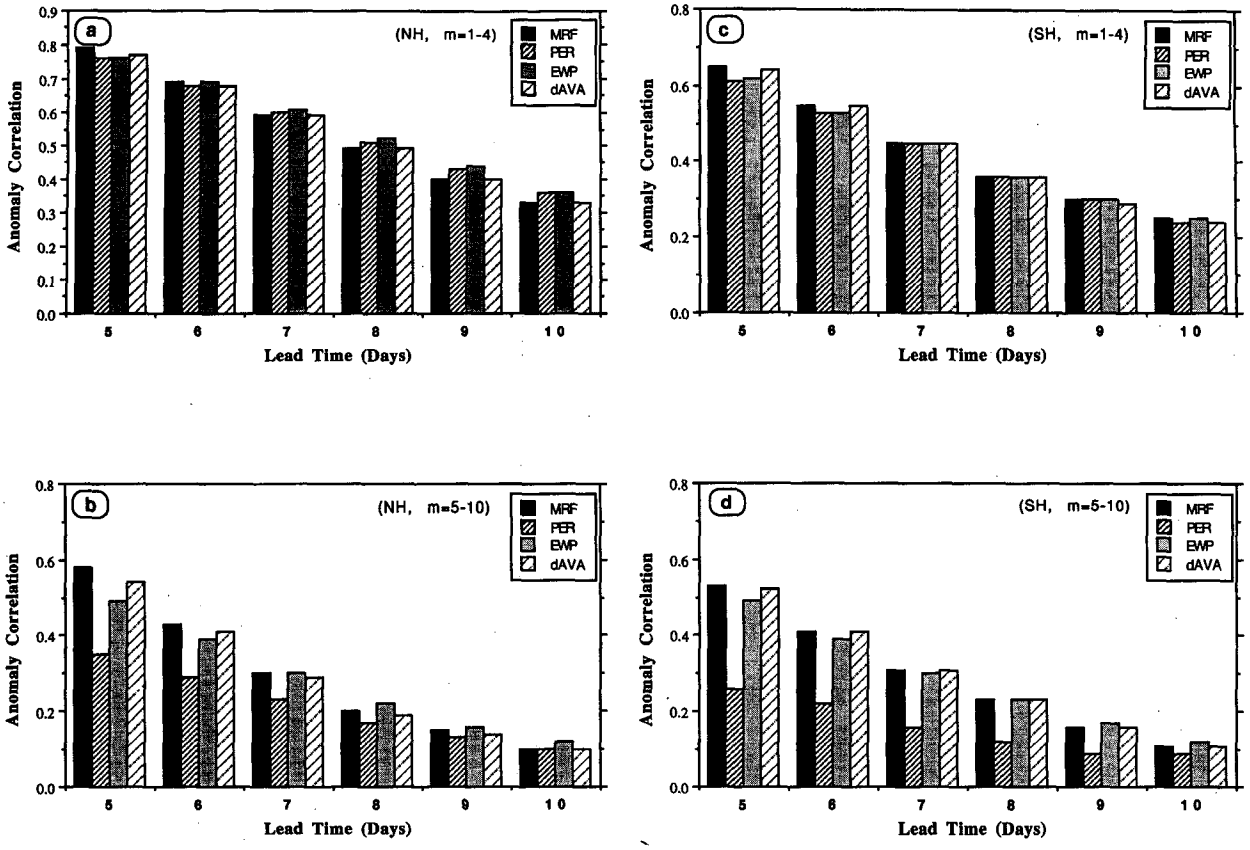

FIG. 6. As Fig. 5. But for height anomalies truncated at zonal harmonics $m = 1$–4 (a) NH, (c) SH and $m = 5$–10 (b) NH, (d) SH only.

where, dAVA-0, -1, -2, and -3 are different versions of the dAVA model. Each term in parentheses denotes one commonly identified physical process in the atmosphere:

RWP Rossby wave propagation: $-[\bar{\mathbf{V}}]\cdot\nabla\zeta' - \bar{\mathbf{V}}'\cdot\nabla([\bar{\zeta}]+f)$,
LBI linear barotropic instability: $-\bar{\mathbf{V}}^*\cdot\nabla\zeta' - \bar{\mathbf{V}}'\cdot\nabla\bar{\zeta}^*$,
NL nonlinear interaction of eddies: $-\mathbf{V}'\cdot\nabla\zeta'$,
DIS diffusion and dissipation processes: $-D\nabla^4\zeta' - K\zeta'$
DIV divergence: $-(f+\bar{\zeta}+\zeta')\nabla\cdot\bar{\mathbf{V}}' - \zeta'\nabla\cdot\bar{\mathbf{V}}$,
other other terms in the original vorticity equation,
ψ' streamfunction anomaly,

where square brackets and asterisks represent zonal mean and departure from zonal mean. The name descriptor chosen to identify each "process" is rather straightforward, except perhaps LBI, linear barotropic instability. Formally, LBI describes advection associated with the standing eddies, but as has been shown empirically in Cai and Van den Dool (1994) this process represents barotropic instability and energy conversion from the climatological eddy flow into the transients. More theoretically, Simmons et al. (1983) have argued the same issue. It also should be pointed out that in (2a)–(2f) EWP is not explicitly mentioned, but we believe that, in this context, EWP is the empirical counterpart of RWP.

We consider first an integration by MRF (2a) and the "full" dAVA (2b) from a common ψ' field (be it observed or an n-day MRF forecast). Difference in skill has to be caused mostly by the DIV term. From the results reported before, it would appear that DIV helps (hurts) forecast skill early (late). As shown by Cai et al. (1994), DIV has large systematic errors in the MRF (and quite possibly even in the analysis). It follows that when errors in DIV are fully built up, the extension forecast by dAVA should be better than continued integration of MRF in which the feedback from DIV errors reduces forecast skill severely. It should be noted that not all of the divergence effect has been removed from the dAVA. The dynamical effect of the divergence has been partly and implicitly included in the dAVA in parameterized form as the Bolin–Cressman correction [see (1)].

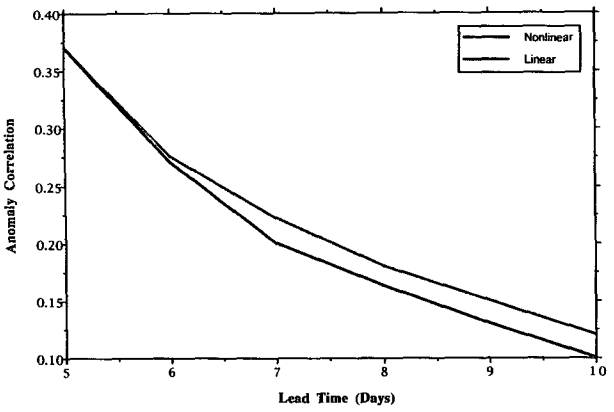

FIG. 7. The mean anomaly correlation (AC) of 90 500-mb height forecasts as a function of lead time (in days) made by nonlinear and linear dAVA models. The linear forecasts are initiated from the nonlinear model output at day 5. The domain is 30°–80°N. The verification period is 1 December 1990–28 February 1991.

We next examine the NL term. Figure 7 gives an example of the impact of that term. Initiating from the dAVA at day 5 and running out to day 10, the linear forecast (2c) beats the nonlinear run (2b) remarkably. Hence, the nonlinear process seems to reduce forecast skill monotonically. The negative effects of the NL term get larger with increasing lead time. An analysis of the NL term applied to the case of 1-day extension forecasts is shown in Fig. 8. The linear extension (2c) becomes superior to the nonlinear forecast (2b) beyond day 5, while earlier on the nonlinearity does help the extension.

Figure 8 also shows ACs at day 1 and day 7 for 1-day extensions of MRF made by all versions of dAVA listed in (2b)–(2e). Removing stationary eddies in the time mean flow (2d), (2e) or turning off LBI in the dAVA enables us to get some understanding about the impact of barotropic instability on forecast skills. A comparison of 1-day extensions made by (2b) versus (2d) and (2c) versus (2e) indicates that presence of barotropic instability never hurts and is usually helpful for gains in forecast skill, especially earlier in the forecast. In fact, no single lead is found from day 1 to day 10 in which the AC is hurt by including LBI.

We did not test any 1-day extension forecast without including the very basic linear RWP process in the dAVA model. It is, however, evident from Fig. 8 that the presence of RWP alone in dAVA (2e) is capable of explaining the bulk of the skill of the full dAVA model. The same results are shown in Fig. 5 in that EWP is the leading control for 1-day extension of MRF beyond day 6. Considering EWP to be the empirical version of RWP, the indication is clear: linear RWP is one of the most fundamental mechanisms that must be included to make skillful forecasts. The positive role of RWP is also lead-independent, helping forecast skill regardless of the advance in the lead time. No matter how poor the forecasts are, EWP's 1-day extensions are always better (or equal) than PER, never worse.

4. Conclusions and discussion

Large scale low-frequency variations in the atmosphere are usually more predictable than the smaller-scale high frequencies (Van den Dool and Saha 1990). For extended-range forecasts, we expect to gain skill mostly from the larger scales. A full forecast model includes both low-frequency oscillations and synoptic-scale high-frequency activities. It is believed that scale interaction is one of the important causes for low-frequency features to lose their predictability (Robinson and Qin 1992). For example, the synoptic eddies lose skill in a short time, after which their forcing of the low-frequency features is simply noise that degrades the forecast of the low-frequency features. Thus, removing the synoptic eddies and simplifying the physics of the model after the time at which they can no longer be skillfully predicted, if done properly, would be helpful for the forecast of the remaining scales. This is our basic motive to study simple forecast methods for low-frequency prediction (LFP).

Three simple and inexpensive control forecasts have been examined in this study. These controls have captured, partly or in full, the characteristics of large-scale low-frequency variability in the atmosphere while reducing the "contamination" from small scales effectively. Numerical experiments based on these control

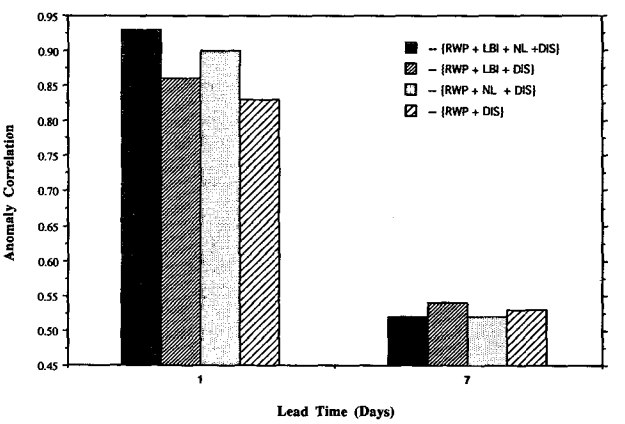

FIG. 8. The mean anomaly correlation (AC) of 90 500-mb height 1-day extension forecasts as a function of lead time (days) made by a different version of the dAVA model. For fixed lead, the first, second, third, and fourth bars show the AC given by the nonlinear dAVA [Eq. (2b)], linear dAVA [Eq. (2c)], nonlinear dAVA but with zonal mean basic flow [Eq. (2d)], and linear dAVA but with zonal mean basic flow [Eq. (2e)], respectively. The domain is 30°–80°N. The verification period is 1 December 1991–28 February 1992.

forecasts have been performed to make 1-day extensions starting from the MRF forecast for day 0–9. Taking the control forecasts as measurements, we have estimated statistically the practical upper limit of prediction skill for the MRF model. Doing so, we have also examined the reasons as to why simple control forecasts can beat the MRF prediction.

Experiments show that, for all three control forecasts in the NH, regardless of the presence or absence of the systematic error in MRF prediction, there exists some point ($T_0 = n$) into the forecast after which 1-day extension of the day n MRF out to day $n + 1$ is as good as or better than the continued integration of the MRF model. The simple control forecasts do improve over the MRF prediction at days 6–10 significantly. In particular, among the three controls, the EWP gives the best 1-day extension after $T_0 = 6$ days in the NH mid- to high latitudes. The gains from the nonlinear dAVA daily extension over the MRF are smaller.

Decomposition of the performance of the control forecasts in terms of zonal harmonics confirms that the improved skill by the control forecasts over the MRF prediction is mainly in the long waves. But contribution from shorter waves are not negligible. This is most clearly demonstrated by the EWP extension. The implication given here is important. Although it is well accepted that synoptic-scale waves become largely unpredictable after one week, our results indicate that not all of the skill in that part of the spectrum is lost by day 10. Extending by EWP is always better than PER. In general, although AC is high for long waves (due to their inertia), we have, as of now, no methods that can forecast changes in the long waves beyond day 6 with remarkable skills over and above PER.

Further efforts have been made to understand why the simple methods can improve over MRF for 1-day extension forecast after the MRF becomes less skillful. The problem is approached by studying the difference between the MRF and several versions of the dAVA model. In doing so, we can identify four processes, namely nonlinearity (NL), divergence (DIV), linear Rossby wave propagation (RWP), and linear barotropic instability (LBI). Results found from intercomparison of different model versions are summarized in Table 2. There are two linear processes (RWP and LBI) that are beneficial to forecast skill all the way from day 0 to day 10. The NL process helps early on (when the anomalies are accurately known) but hurts skill beyond day 5. It is clear that the nonlinear effects have projection not only on the transients but also on the time means that could enlarge the climate drift and intensify the error growth in transients in turn. Similarly, divergence helps out to day 5 but does more harm than good beyond day 5. So even though nature is undoubtedly nonlinear and divergent and models are supposed to be close copies of the reality, there may be circumstances where the inclusion of nonlinearity and divergence in a model is not recommended.

TABLE 2. A comparison of contributions to forecast skill among different physical processes.

Processes	Early	Late
NL	helps	hurts
DIV	helps	hurts
LBI	helps	helps
RWP	helps	helps

We have shown that beyond day 6 the simple controls outperform the MRF. By MRF we mean a single run at the highest affordable resolution. However, since December 1992 the MRF is run many times over, every day, from slightly perturbed initial conditions (Tracton and Kalnay 1993). One might suspect that the simple controls act somewhat like the ensemble average (Van den Dool and Rukhovets 1994), that is, the erroneous impact of high frequency eddies gets explicitly eliminated (EWP, dAVA) or averaged out (ensemble). Indeed the ensemble average improves over the single MRF only after day 5 (see Van den Dool and Rukhovets 1994, Fig. 3) and by a margin comparable to the best of the controls. This is, however, not enough reason to conclude that the control forecasts are a poor man's version of the ensemble average. The direct elimination of divergence in the control forecasts (which has an impact on skills) does not seem to have a counterpart when averaging ensemble members.

Acknowledgments. The authors wish to thank Dr. Hannu Savijarvi for providing valuable advice for this study and Dr. S. Saha for assistance in data acquisition. Thanks are also extended to Drs. W. A. Robinson and M. Cai for useful discussions. Encouraging suggestions and comments from the three anonymous reviewers, which helped strengthen the presentation of this paper, are greatly appreciated.

REFERENCES

Branstator, G., M. Andrew, and B. David, 1993: Identification of highly predictable flow elements for spatial filtering of medium- and extended-range numerical forecasts. *Mon. Wea. Rev.*, **121,** 1786–1802.

Cai, M., and H. M. van den Dool, 1991: Low frequency waves and traveling storm tracks. Part I: Barotropic component. *J. Atmos. Sci.*, **48,** 1420–1436.

——, and ——, 1994: Dynamical decomposition of low-frequency tendencies. *J. Atmos. Sci.*, **51,** 2086–2100.

——, J. S. Whitaker, R. M. Dole, and K. L. Paine, 1994: Dynamics of systematic errors in the NMC Medium Range Forecast Model. *Proc. of the 19th Annual Climate Diagnostics Workshop,* College Park, MD, 389–392.

Holloway, G., 1983: Effects on planetary wave propagation and finite depth on the predictability of atmospheres. *J. Atmos. Sci.*, **40,** 314–327.

Holton, J. R., 1979: *An Introduction to Dynamical Meteorology,* 2d ed. Academic Press, 179 pp.

Johansson, A., and S. Saha, 1989: Simulation of systematic error effects and their reduction in a simple model of the atmosphere. *Mon. Wea. Rev.*, **117,** 1658–1675.

Rinne, J., and H. Jarvinen, 1993: Estimation of the Cressman term for a barotropic model through optimization with use of the adjoint model. *Mon. Wea. Rev.*, **121,** 825–833.

Robinson, W. A., and J. Qin, 1992: Predictability of the zonal index in a global model. *Tellus,* **44A,** 331–338.

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

Sela, J. G., 1980: Spectral modeling at NMC. *Mon. Wea. Rev.*, **108,** 1279–1292.

——, 1988: The new T80 NMC operational spectral model. *Proc. of the Eighth Conf. on Numerical Weather Prediction,* Baltimore, MD, Amer. Meteor. Soc., 312–313.

Simmons, A. J., J. M. Wallace, and G. W. Branstator, 1983: Barotropic wave propagation and instability, and atmospheric teleconnection patters. *J. Atmos. Sci.*, **40,** 1363–1392.

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Wea. Forecasting,* **8,** 379–398.

Van den Dool, H. M., 1991: Mirror images of atmospheric flow. *Mon. Wea. Rev.*, **119,** 2095–2106.

——, and S. Saha, 1990: Frequency dependence in forecast skill. *Mon. Wea. Rev.*, **118,** 128–137.

——, and M. Cai, 1994: An empirical Rossby wave propagation formula. *Proc. of the Int. Symp. on The Life Cycles of Extratropical Cyclones,* Bergen, Norway, Volume II, 96–96.

——, and L. Rukhovets, 1994: On the weights for an ensemble-averaged 6–10-day forecast. *Wea. Forecasting,* **9,** 457–465.

Zhao, Y., M. Cai, and H. M. van den Dool, 1994: Low-frequency variability in an anomaly model and its comparison with observation. *Proc. of the 19th Annual Climate Diagnostics Workshop,* College Park, MD, 285–288.