Sensitivity Analysis of Forecast Errors and the Construction of Optimal Perturbations Using Singular Vectors

R. GELARO,* R. BUIZZA, T. N. PALMER, AND E. KLINKER

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, United Kingdom

(Manuscript received 2 December 1996, in final form 8 September 1997)

ABSTRACT

The sensitivity of forecast errors to initial conditions is used to examine the optimality of perturbations constructed from the singular vectors of the tangent propagator of the European Centre for Medium-Range Weather Forecasts model. Sensitivity and pseudo-inverse perturbations based on the 48-h forecast error are computed as explicit linear combinations of singular vectors optimizing total energy over the Northern Hemisphere. It is assumed that these perturbations are close to the optimal perturbation that can be constructed from a linear combination of these singular vectors. Optimality is measured primarily in terms of the medium-range forecast improvement obtained by adding the perturbations a posteriori to the initial conditions. Several issues are addressed in the context of these experiments, including the ability of singular vectors to describe forecast error growth beyond the optimization interval, the number of singular vectors required, and the implications of nonmodal error growth. Supporting evidence for the use of singular vectors based on a total energy metric for studying atmospheric predictability is also presented.

In general, less than 30 singular vectors capture a large fraction of the variance of the Northern Hemisphere sensitivity pattern obtained from a T63 adjoint model integration, especially in cases of low forecast skill. The sensitivity patterns for these cases tend to be highly localized with structures determined by the dominant singular vectors. Forecast experiments with these perturbations show significant improvements in skill in the medium range, indicating that singular vectors optimized for a short-range forecast continue to provide a useful description of error growth well beyond this time. The results suggest that ensemble perturbations based on 10–30 singular vectors should provide a reasonable description of the medium-range forecast uncertainty, although the inclusion of additional singular vectors is likely to be beneficial.

Nonmodality is a key consideration in the construction of optimal perturbations. There is virtually no projection between the contemporaneous unstable subspaces at the end of one forecast trajectory portion and the beginning of a second, consecutive portion. Sensitivity and ensemble perturbations constructed using the evolved singular vectors from a previous (day−2) forecast are suboptimal for the current (day+0) forecast initial conditions. It is argued that these results have implications for a range of issues in atmospheric predictability including ensemble weather prediction, data assimilation, and the development of adaptive observing techniques.

1. Introduction

It has long been proposed that advances in ensemble weather prediction can provide quantum improvements in assessing forecast skill and utility (Leith 1974). In particular, a primary goal of creating an ensemble of forecasts is not only to produce a mean forecast that is more skillful than the control, but to quantify the uncertainty of the control forecast itself. For realistic models of the atmosphere such as those used at operational forecast centers, ensembles based on a randomly sampled distribution of initial states must be of prohibitively large size to sample adequately the uncertainty in the initial conditions. Thus, a major focus of research has been on developing methods for constructing a limited set of perturbations that grow rapidly, so as to provide the maximum information about the model probability density function from a corresponding limited set of model integrations (Toth and Kalnay 1996; Molteni et al. 1996; Houtekamer and Derome 1995)

Since December 1992, the European Centre for Medium-Range Weather Forecasts (ECMWF) has run as part of its operational suite an ensemble prediction system (EPS) using initial perturbations based on the dominant singular vectors of an approximation to the forward tangent propagator of the ECMWF model linearized about the short-range forecast trajectory (Molteni et al. 1996). These singular vectors describe the principal finite-time linear instabilities of the northern extratropical atmospheric circulation (Buizza and Palmer 1995). The methodology and performance of the ECMWF EPS, as well as recent improvements in the

* Current affiliation: Naval Research Laboratory, Monterey, California.

Corresponding author address: Dr. Ronald Gelaro, Naval Research Laboratory, 7 Grace Hopper Avenue, Monterey, CA 93943-5502.
E-mail: gelaro@nrlmry.navy.mil

system, have been documented in detail by Buizza (1995), Molteni et al. (1996), and Buizza et al. (1997).

The singular vectors of the forward tangent propagator of a nonlinear dynamical system provide a mathematically rigorous, but tractable, approach to quantifying perturbation growth over a finite time interval. The dominant singular vectors describe the most rapidly growing structures with respect to a given metric (norm) over this interval in a tangent linear sense. Although singular vectors have been applied to a variety of problems in meteorology and general geophysical fluid dynamics (e.g., Orr 1907; Lorenz 1965; Farrell 1982, 1988, 1989; Borges and Hartmann 1992; Barkmeijer et al. 1993; Ehrendorfer and Errico 1995; Buizza and Palmer 1995), there remain several important questions regarding their application to predictability problems in general and to ensemble weather prediction in particular.

Several recent studies have begun to shed light on some of the basic foundational issues concerning the application of singular vectors to problems in atmospheric predictability. Palmer et al. (1998), hereafter referred to as P98, examined the selection of appropriate metrics for computing singular vectors for several classes of problems. They showed evidence that perturbation energy is, to first approximation, an appropriate metric for studying atmospheric predictability on timescales of several days. These authors also compared the principle assumptions underlying the use of singular vectors, bred vectors (Toth and Kalnay 1997) and potential vorticity (PV) analysis. Ehrendorfer and Tribbia (1997) discussed the theoretical justification for using singular vectors in ensemble weather prediction. They showed that singular vectors were a key element for making optimal predictions of forecast error covariances in both a three-dimensional Lorenz model and a higher-dimensional barotropic model. It was found that less than 15% of the total number of singular vectors were needed to recover more than 95% of the total forecast error variance in these models. Barkmeijer et al. (1993) used the eigenvectors of a simplified local covariance matrix in a quasigeostrophic model to measure the skill of ECMWF forecasts over Europe. They also found that relatively few eigenvectors were needed to describe a large fraction of the short-range forecast error variance over this limited domain.

In this paper, we examine several questions concerning the use of singular vector perturbations for ensemble weather prediction (and related applications) in the context of an operational forecast model. These include the following.

(i) Do singular vectors optimized for a short-range forecast (in this case, 48 h) provide a useful description of error growth beyond this time, where it is hoped the ensemble will have its greatest potential use?

(ii) How many singular vectors are required to adequately describe the forecast error over a large domain with many degrees of freedom, such as the Northern Hemisphere?

(iii) What are the implications of nonmodal error growth for constructing optimal perturbations; for example, are yesterday's forecast errors useful for constructing optimal perturbations for today's initial conditions?

The general question of whether a given set of perturbations is in some sense optimal is difficult to answer quantitatively. This is due to a combination of factors including the fact that specific measures of optimality are not necessarily well defined, especially in a complex system. The problem is further exacerbated by a general lack of knowledge about the errors in most operational data assimilation schemes and the fact that the largest forecast errors are not necessarily associated with largest errors in the initial conditions.

Nevertheless, recent work by Rabier et al. (1996, hereafter referred to as R96) provides new insight into the dynamic structure and behavior of rapid error growth in a realistic forecast model. Using the adjoint propagator of the ECMWF model, R96 showed that the gradient of a cost function based on the energy of the 48-h forecast error provides a ''sensitivity pattern'' that, neglecting model errors over this interval, can be interpreted as a sum of rapidly growing components of analysis error. Typically, the sensitivity pattern is characterized by small-scale baroclinic structures that have their maximum amplitude in the lower to middle troposphere. These structures evolve rapidly into synoptic-scale features whose structures often closely resemble those of the major forecast errors. Rabier et al. (1996) showed that the a posteriori introduction of an analysis perturbation proportional to this pattern often led to a significant reduction in not only the 48-h forecast error, but the medium-range (days 4–7) forecast error as well. In fact, in many cases, the medium-range forecasts showed the most dramatic improvements in skill. It is also noteworthy that these and subsequent results (F. Rabier 1997, personal communication) appear to indicate that the sensitivity-perturbed analyses fit the observational data as well as, or better, than the unperturbed analyses.

We use an approach similar to that of R96 to investigate questions (i)–(iii) above. However, in contrast to R96, we compute sensitivity and pseudo-inverse estimates of analysis error using a limited number of the dominant singular vectors of a linearized version of the ECMWF forecast model. A primary motivation for the study is to determine whether the forecast error sensitivity pattern obtained from an adjoint model integration spanning a large-dimensional ($\sim 10^6$) space can be adequately described in terms of a much smaller number ($\sim 10^1$) of singular vectors. The sensitivity and pseudo-inverse estimates are used to construct perturbations for a series of forecast experiments with the full nonlinear model. Optimality is measured primarily in terms of the magnitude of the *medium-range* forecast improvement obtained with these perturbations, though the perturbations themselves are computed using a 48-h opti-

mization as in the current EPS. From the point of view of ensemble prediction, these experiments serve as an indicator of the number of singular vectors (which, in turn, determines the number of ensemble members; see Molteni et al. 1996) required to describe the medium-range forecast uncertainty.

This study is, in part, an extension of work reported by Buizza et al. (1997). However, here we focus on the sensitivity estimate of analysis error so as to have a direct qualitative comparison with results presented by R96, and we conduct a much greater number and variety of forecast experiments. We also focus attention on the highly nonmodal nature of the perturbation evolution in these experiments. It is argued that nonmodality has important implications for constructing optimal perturbations for ensemble weather prediction, as well as for data assimilation and the development of so-called adaptive observing strategies recently tested during the Fronts and Atlantic Storm Track Experiment (FASTEX; Snyder 1996). In support of these arguments, we also present results from sensitivity and ensemble prediction experiments using singular vectors optimized for the previous 48-h forecast. Supporting evidence for the use of singular vectors based on a total energy metric for studying atmospheric predictability is also presented.

In section 2, we briefly discuss the selection of cases for this study, highlighting differences in the degree of initial condition sensitivity for two cases selected for closer scrutiny throughout this paper. The mathematical formalism of the singular vector calculations is reviewed in section 3. This section also includes a discussion of the relationship between singular vectors, analysis errors, and sensitivity patterns. The construction of the sensitivity and pseudo-inverse perturbations is described in section 4. Forecast experiments involving these perturbations, including a detailed interpretation of the results for two cases, plus composite results, are presented in section 5. Additional experiments, including ensemble forecasts, which focus on the implications of nonmodal error growth, are presented in section 6. Concluding remarks are presented in section 7.

2. Case selection

The results of this study are based primarily on twelve cases from the winter and spring of 1994. Singular vectors for these cases were computed at T42 horizontal resolution with 19 vertical levels (T42L19) and an optimization time of 48 h. Six of the cases were selected because the control forecasts showed low skill over some part of the Northern Hemisphere in the medium range. We refer to these as ''low-skill'' cases. The remaining six cases, which may be classified as having more or less average skill, were chosen to have starting dates *2* days prior to each of the six low-skill cases. We refer to these as ''average-skill'' cases.

The starting dates for the average-skill cases were selected for two reasons. First, the overall quality of these forecasts provides a marked contrast to the low-skill cases 2 days later (see below). Second, the singular vectors for these cases, which are optimized for the 48-h interval ending at the initial time of a subsequent low-skill forecast, allow us to compare the contemporaneous unstable subspaces corresponding to the end of one forecast trajectory portion and the beginning of a second consecutive trajectory portion. In section 6, we compare these subspaces and examine the impact of perturbations optimized for the forecast interval either proceeding or beginning with the current analysis. It is argued that the results of this comparison are relevant to several areas of research in numerical weather prediction (NWP).

In addition to composite results for the twelve cases, we show more detailed results for two individual cases: the low-skill case 6 April 1994, in which the control forecast over Europe was particularly poor, and the average-skill case 2 days prior to this, 4 April 1994. To illustrate the differences in forecast skill and initial condition sensitivity for these cases, we show in Fig. 1 the anomaly correlation coefficient (ACC) for 500-mb geopotential height over Europe for the ensemble forecasts for each date. The results are based on the EPS configuration used operationally at the time of this writing (see section 6). The control forecast (heavy solid line) for 6 April 1994 loses skill rapidly and completely after day 3, compared with a more gradual and less severe loss of skill after day 5 for 4 April 1994. The difference in sensitivity to initial conditions for these cases is indicated, for example, by the extremely wide range of skill in the medium range for 6 April 1994, compared with a much more narrow range of skill in this time range for 4 April 1994. In the case of 6 April 1994, the ensemble mean is much more skillful than the control forecast after day 5. In the case of 4 April 1994, the ensemble mean and control forecasts have nearly identical skill. In later sections, we postulate that these differences are related to the projection of analysis errors onto the dominant singular vectors for each date.

3. Mathematical formalism

In this section we briefly review the definition of singular vectors and examine their relationship to analysis errors and sensitivity patterns. More detailed discussions on the mathematical development and dynamical interpretation of singular vectors are given by Buizza and Palmer (1995), Farrell (1982), Lacarra and Talagrand (1988), Molteni and Palmer (1993), and Barkmeijer (1995). More detailed discussions of sensitivity analysis based on adjoint methods are given by R96, Langland et al. (1995), Errico and Vukićević (1992), and others.

a. Singular vector definition

Let $\mathbf{L}(\mathbf{x}, t)$ be the forward tangent propagator of a nonlinear dynamical model, $\mathbf{M}(\mathbf{x}, t)$, linearized about a

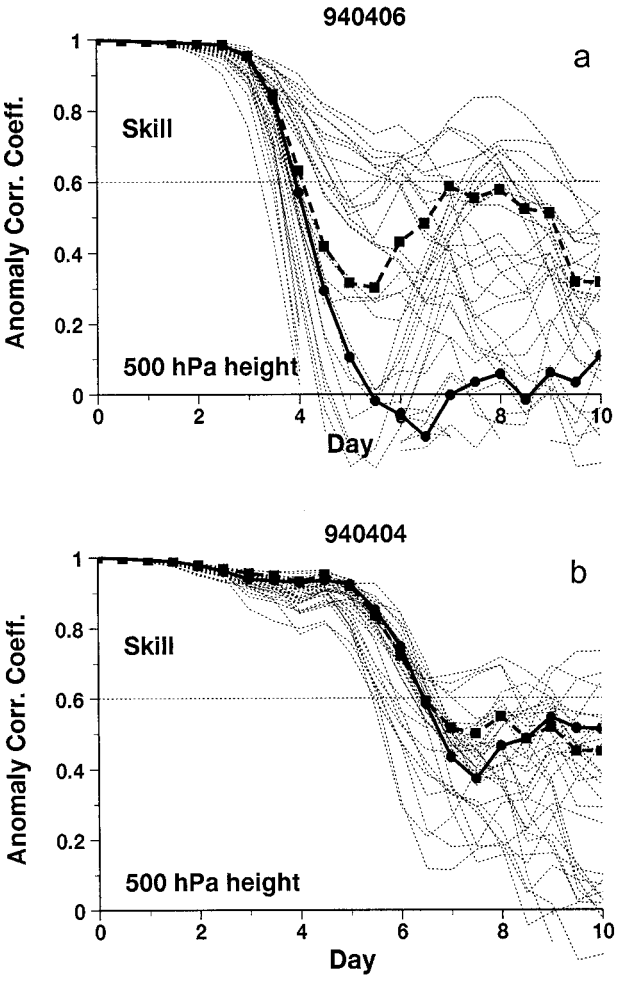

FIG. 1. Ensemble forecast skill over Europe in terms of anomaly correlation coefficient for 500-mb geopotential height for (a) the low-skill case 6 April 1994 and (b) the average-skill case 4 April 1994. Perturbed members indicated by thin dot lines, control forecast indicated by heavy solid line, and ensemble mean indicated by heavy dash line.

portion of a model phase space trajectory $\overline{\mathbf{x}}_t$ over an interval, $0 \rightarrow t$, such that

$$\mathbf{x}(t) = \mathbf{L}\mathbf{x}(0) \tag{1}$$

describes a mapping of the tangent space $\mathcal{T}_{\mathbf{x}(0)}$ to the tangent space $\mathcal{T}_{\mathbf{x}(t)}$. Let $\mathbf{L}^{\mathrm{T}}$ be the adjoint (transpose) of $\mathbf{L}$, which satisfies the definition

$$\langle \mathbf{x};\, \mathbf{L}\mathbf{y} \rangle = \langle \mathbf{L}^{\mathrm{T}}\mathbf{x};\, \mathbf{y} \rangle \tag{2}$$

for any vectors $\mathbf{x}$ and $\mathbf{y}$.

We seek the most unstable dynamical structures over the interval $0 \rightarrow t$ as manifested by some appropriately chosen measure of perturbation growth. Since the state vector $\mathbf{x}$ is multivariate, it is convenient to introduce a transformation of variables,

$$\mathbf{z} = \mathbf{C}^{1/2}\mathbf{x}, \tag{3}$$

where $\mathbf{C}$ is a matrix whose elements are scaling factors applied to the dependent variables. We select $\mathbf{C}$ to be a diagonal matrix such that the total energy $\mathcal{E}$ of a perturbation at time t is given by

$$\mathcal{E}(t) = \langle \mathbf{x}(t);\, \mathbf{C}\mathbf{x}(t) \rangle = \langle \mathbf{z}(t);\, \mathbf{z}(t) \rangle. \tag{4}$$

The inner product in (4) is referred to as the total energy inner product (Buizza and Palmer 1995; R96). Its choice as an appropriate measure of perturbation growth for studying atmospheric predictability is discussed in detail by P98, as well as by Molteni et al. (1996). We give supporting evidence for this choice later.

Applying (1) and (2), we can write (4) in the form

$$\mathcal{E}(t) = \langle \mathbf{L}_{\mathbf{c}}^{\mathrm{T}}\mathbf{L}_{\mathbf{c}}\mathbf{z}(0);\, \mathbf{z}(0) \rangle, \tag{5}$$

where $\mathbf{L}_{\mathbf{c}} = \mathbf{C}^{1/2}\mathbf{L}\mathbf{C}^{-1/2}$. The solutions $\mathbf{u}_i(\mathbf{z})$ that maximize $\mathcal{E}(t)$, subject to the constraint

$$\mathcal{E}(0) = \langle \mathbf{z}(0);\, \mathbf{z}(0) \rangle = 1, \tag{6}$$

satisfy the eigenvalue problem

$$\mathbf{L}_{\mathbf{c}}^{\mathrm{T}}\mathbf{L}_{\mathbf{c}}\mathbf{u}_i = \sigma_i^2\mathbf{u}_i. \tag{7}$$

The eigenvectors $\mathbf{u}_i$ of the compound operator $\mathbf{L}_{\mathbf{c}}^{\mathrm{T}}\mathbf{L}_{\mathbf{c}}$ are referred to as the (right) singular vectors of $\mathbf{L}_{\mathbf{c}}$. The singular values σ_i^2 are real and positive since, as shown for example by P98, $\mathbf{L}_{\mathbf{c}}^{\mathrm{T}}\mathbf{L}_{\mathbf{c}}$ is symmetric, although in general $\mathbf{L}_{\mathbf{c}}$ itself is not (Strang 1976; P98). The singular vectors span a space whose axes represent the most unstable phase space directions of the dynamical system at t_0, with respect to the total energy metric. In practice, we consider only a limited number n of these directions and refer to this as the n-dimensional unstable subspace of the system.

At optimization time t, the singular vectors evolve to

$$\hat{\mathbf{v}}_i = \mathbf{L}\mathbf{u}_i, \tag{8}$$

where $\|\mathbf{u}_i\| = 1$ and $\|\hat{\mathbf{v}}_i\| = \sigma_i$. Thus, it follows that the σ_i are the amplification factors of the singular vectors over the interval $0 \rightarrow t$, as given by

$$\frac{\|\hat{\mathbf{v}}_i\|}{\|\mathbf{u}_i\|} = \sigma_i. \tag{9}$$

The leading singular vectors used in this study have amplification factors $\sigma \sim 10$–20 (see Fig. 2) based on a 48-h optimization that is constrained geographically to the Northern Hemisphere (30°–90°N) by use of a local projection operator (Buizza 1994).

An important characteristic of the singular vectors is that their evolution is not shape preserving; that is, they are strongly nonmodal as indicated by a pronounced upscale energy transformation from initial to final time (Buizza and Palmer 1995). At initial time, the singular vectors have their maximum (predominantly potential) energy in the lower to midtroposphere, while at final time, they have their maximum (predominantly kinetic) energy at the jet level. The dominant mechanism for singular vector growth is consistent with baroclinic instability, although the horizontal structure and tilt of the

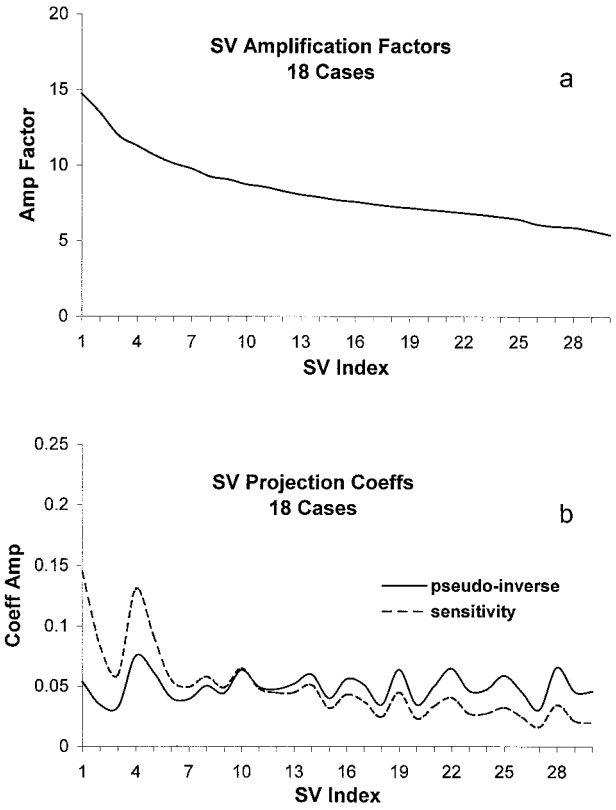

FIG. 2. (a) Average growth rate spectrum of the first $n = 30$ singular vectors for 18 cases from various seasons. (b) Average absolute values of the projection coefficients c_i, based on (22), corresponding to the pseudo-inverse perturbations $\tilde{\mathbf{e}}_0$ (solid), and the scaled sensitivity perturbations $\alpha\tilde{\mathbf{s}}$ (dash). The values of α are based on (25), where $n = 30$.

singular vectors with respect to the basic state indicate that growth also occurs through barotropic instability.

The singular vector structures depend, in part, on the choice of inner product (or metric) used to define the adjoint propagator. Where predictability (and in particular, ensemble prediction) is concerned, the appropriate metric is necessarily determined by both the dynamics of the system and the characteristics of the observing network used to define initial conditions for making forecasts (P98). In sections 4 and 5, we present evidence suggesting that perturbation energy is, to first approximation, an appropriate metric for computing singular vectors for the predictability problem.

b. Forecast error sensitivity

Recent studies have shown that adjoint methods provide an effective means for conducting sensitivity analysis with large NWP models (R96; Langland et al. 1995, Errico et al. 1993). In particular, the adjoint model is an efficient tool for computing the gradient of a forecast measure, J, with respect to the initial conditions. Depending on the choice of J, the gradient (or sensitivity pattern) can be used, for example, to quantify the impact of small perturbations in any model variable on the development of a particular feature, such as an extratropical cyclone, or to identify structures in the initial conditions that might cause large forecast errors.

Following R96, we consider a quadratic forecast measure,

$$J = \frac{1}{2}\langle \mathbf{e}_{48};\, \mathbf{C}\mathbf{e}_{48}\rangle, \tag{10}$$

where $\mathbf{e}_{48} = \mathbf{M}(\mathbf{x}_0) - \mathbf{x}_{48}$ is the Northern Hemisphere (30°–90°N) 48-h forecast error computed as the difference between a 48-h nonlinear model integration from the analyzed initial conditions $\mathbf{x}_0$ and the verifying analysis $\mathbf{x}_{48}$. As before, $\mathbf{C}$ is the diagonal matrix of total energy weights.

A perturbation, $\delta\mathbf{x}_0$, in the initial conditions will lead to a perturbation, $\delta\mathbf{x}_{48}$, that produces a corresponding change in J given, to first approximation, by

$$\delta J = \frac{\partial J}{\partial \mathbf{x}_{48}}\delta\mathbf{x}_{48} = \langle \nabla_{48}J;\, \delta\mathbf{x}_{48}\rangle, \tag{11}$$

where $\nabla_{48}J$ is the gradient of J with respect to the state vector at $t = 48$. Differentiating (10), we obtain

$$\nabla_{48}J = \mathbf{C}\mathbf{e}_{48}. \tag{12}$$

Neglecting model errors over the first 48 h, we have

$$\delta\mathbf{x}_{48} = M(\mathbf{x}_0 + \delta\mathbf{x}_0) - M(\mathbf{x}_0) \approx \mathbf{L}\delta\mathbf{x}_0. \tag{13}$$

Equations (2), (12), and (13) allow us to express the change in J as

$$\delta J = \langle \nabla_0 J;\, \delta\mathbf{x}_0\rangle, \tag{14}$$

where

$$\nabla_0 J = \mathbf{L}^{\mathrm{T}}\mathbf{C}\mathbf{e}_{48} \tag{15}$$

is the gradient of J with respect to the state vector at $t = 0$. Equation (15) represents an integration of the adjoint model, with starting condition $\mathbf{C}\mathbf{e}_{48}$, which maps the forecast error back to initial time. For realistic models of the atmosphere, $\nabla_0 J$ is related, but not equivalent, to the initial (analysis) error $\mathbf{e}_0$, where the latter is obtained only by fully inverting the forward propagator $\mathbf{L}$. Rather, $\nabla_0 J$ is a measure of the sensitivity of $\mathbf{e}_{48}$ to changes in the initial conditions. For convenience, hereafter we denote $\nabla_0 J$ by $\mathbf{s}$. As shown by R96, where $\mathbf{s}$ is large, a change, $\delta\mathbf{x}_0$, in the initial conditions can have a large impact on the forecast error. Conversely, where s is small, a comparable change in the initial conditions has much less impact on the forecast error.

c. Estimates of initial error

The relationship between $\mathbf{s}$ and $\mathbf{e}_0$ can be made explicit in terms of the singular vectors. Using (13) and (15), we can express the sensitivity in the form

$$\mathbf{s} = \mathbf{C}^{1/2}\mathbf{L}_c^{\mathrm{T}}\mathbf{C}^{1/2}\mathbf{e}_{48} = \mathbf{C}^{1/2}\mathbf{L}_c^{\mathrm{T}}\mathbf{L}_c\mathbf{C}^{1/2}\mathbf{e}_0. \tag{16}$$

We can expand $\mathbf{e}_0$ in terms of the singular vectors $\mathbf{u}_i$ at initial time

$$\mathbf{e}_0 = \mathbf{C}^{-1/2} \sum_i c_i \mathbf{u}_i, \tag{17}$$

where

$$c_i = \langle \mathbf{C}^{1/2}\mathbf{u}_i; \mathbf{e}_0 \rangle \tag{18}$$

is a projection coefficient of the initial error onto the *i*th singular vector. Combining (16) and (17), and using the singular vector definition (7), we can express the sensitivity in the form

$$\mathbf{C}^{-1}\mathbf{s} = \mathbf{C}^{-1/2} \sum_i \sigma_i^2 c_i \mathbf{u}_i, \tag{19}$$

where $\mathbf{C}^{-1}\mathbf{s}$ has the same (meteorological) units as $\mathbf{e}_0$. The following can now be readily deduced by comparing (17) and (19). First, it is clear that the sensitivity pattern, as derived for the forecast measure (10), is a weighted measure of the initial error, dominated by the fastest-growing (largest σ) singular vectors. As shown in section 4, for the singular vectors used in this study, the leading components of s typically have weight factors $\sigma^2 \sim 100$ or more. In phase space, the effect of this weighting is to rotate $\mathbf{s}$ away from $\mathbf{e}_0$ toward the most rapidly growing singular vector directions.

Second, it follows that if $\mathbf{s}$ is nonzero, then so is the projection of the analysis error onto the singular vectors at initial time; this is made explicit by (18). This is an important result in view of the general criticism that regions of sensitivity do not necessarily correspond to regions of analysis error. In this case, they correspond by definition.

Finally, since $\mathbf{C}^{-1}\mathbf{s}$ has units of the meteorological fields, it is conceptually straightforward to construct a perturbation, $-\alpha\mathbf{C}^{-1}\mathbf{s}$, where α is an appropriate scale factor, which can be added to the forecast initial conditions. Rabier et al. (1996) have shown that a perturbation of this form can be very useful for diagnosing the effects of initial condition errors on a numerical forecast.

In interpreting certain aspects of the results that follow, it is important to bear in mind that the sensitivity and pseudo-inverse estimates of analysis error are computed within the context of the so-called perfect model assumption (13). A concern then is that in practice these estimates may reflect not only components of the actual analysis error, but also changes to the initial conditions that compensate for model errors. Of course, this will depend on, among other things, the temporal and spatial scales of the problem, as well as the characteristics of the data assimilation and forecast system. For the purposes of this study, we will assume that the forecast error is dominated by growing analysis errors. We make this assumption based on several arguments including our current understanding (e.g., Simmons et al. 1995; Tribbia and Baumhefner 1988) of forecast error growth in models used for NWP (especially for synoptic-scale baroclinic motions on timescales of 2 days or less), evidence from previous studies indicating that the sensitivity-perturbed analyses fit the observational data (section 1), and evidence based on the results of the present study.

4. Sensitivity and pseudo-inverse perturbations

Since the analysis error $\mathbf{e}_0$ is in general not well known, (17) and (19) are of little practical use in their present forms. However, from (2), (7), and (13), it is easily shown that

$$\mathbf{e}_0 = \mathbf{C}^{-1/2} \sum_i \sigma_i^{-1} d_i \mathbf{u}_i, \tag{20}$$

and

$$\mathbf{s} = \mathbf{C}^{-1/2} \sum_i \sigma_i d_i \mathbf{u}_i, \tag{21}$$

where

$$d_i = \langle \mathbf{C}^{1/2}\mathbf{v}_i; \mathbf{e}_{48} \rangle = \sigma_i c_i \tag{22}$$

and $\mathbf{v}_i = \sigma_i^{-1}\hat{\mathbf{v}}_i$ is an orthonormal singular vector at final time. Equation (22) provides a practical way of estimating the analysis error projection onto the singular vectors at initial time [cf. (18)].

For a limited number, *n*, of singular vectors, it is convenient to introduce the matrix forms

$$\tilde{\mathbf{e}}_0 = \mathbf{C}^{-1/2}\mathbf{U}\tilde{\mathbf{\Sigma}}^{-1}\mathbf{V}^{\mathrm{T}}\mathbf{C}^{1/2}\mathbf{e}_{48}, \tag{23}$$

$$\mathbf{C}^{-1}\tilde{\mathbf{s}} = \mathbf{C}^{-1/2}\mathbf{U}\tilde{\mathbf{\Sigma}}\mathbf{V}^{\mathrm{T}}\mathbf{C}^{1/2}\mathbf{e}_{48}, \tag{24}$$

where $\mathbf{U}$ is an $m \times n$ matrix with columns $\mathbf{u}_i$, $\mathbf{V}$ is an $m \times n$ matrix with columns $\mathbf{v}_i$, $\tilde{\mathbf{\Sigma}}$ is an $n \times n$ diagonal matrix with elements σ_i, and the dimension m depends on the model truncation. The tildes ($\tilde{}$) denote that $\tilde{\mathbf{e}}_0$ and $\tilde{\mathbf{s}}$ are projections into an *n*-dimensional subspace of singular vectors. As discussed by Buizza et al. (1997), (23) is the Moore–Penrose pseudo-inverse estimate of the analysis error (Penrose 1955) in the approximation where all the singular values outside the *n*-dimensional subspace are negligible.

Equations (23) and (24) are used to construct the pseudo-inverse and sensitivity perturbations for most of the forecast experiments discussed throughout the remainder of this paper. We focus attention on both types of perturbations since, in the case of $\tilde{\mathbf{s}}$, these perturbations are closely related to the adjoint sensitivity patterns studied by R96, while in the case of $\tilde{\mathbf{e}}_0$, these perturbations are more closely related to those used to construct initial states for the EPS.

Equation (23) can be used directly to construct an initial perturbation, $\mathbf{x}_0 = -\tilde{\mathbf{e}}_0$, that, by definition, will bring about the maximum reduction in the diagnostic function (10) for a given number, *n*, of singular vectors (Buizza et al. 1997). In contrast, any forecast improvement derived from $\tilde{\mathbf{s}}$ will depend on the scaling of a perturbation, $\mathbf{x}_0 = -\alpha\mathbf{C}^{-1}\tilde{\mathbf{s}}$, where α is an appropriate scale factor. Rabier et al. (1996) have discussed various

methods for estimating optimal values of α, including values based on empirical methods, values based on (the inverse of) the leading singular value, and values based on a linear search in the direction of $\nabla_0 J$. Generally, speaking, all of these procedures yield typical values of $\alpha \sim 0.01$ for the Northern Hemisphere domain, although there may be significant case-to-case variations.

Since we consider sensitivity perturbations defined in terms of singular vectors explicitly, it is convenient to derive a scale factor based on the associated singular values. The choice $\alpha = 1/\sigma_1^2$ suggested by R96 seems reasonable provided the leading singular vector dominates the sensitivity projection. Figure 2a shows the average amplification factors σ_i of the first $n = 30$ Northern Hemisphere singular vectors for 18 cases from various seasons, including the 12 cases examined in this study. The spectrum has a moderate slope at leading index values, but there is little evidence of a dominant leading singular vector. In conjunction with this, we show in Fig. 2b the average (absolute) values of the projection coefficients for these cases, based on (22). The solid line corresponds to the coefficients c_i for the pseudo-inverse perturbations $\tilde{\mathbf{e}}_0$. Note the nearly uniform magnitude of these coefficients, indicating that, on average, the analysis error projects equally onto all singular vectors in the unstable subspace. Based on these results, we define a scale factor for the projected sensitivity perturbations,

$$\alpha = \left(\frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 \right)^{-1}, \tag{25}$$

that accounts, in an average sense at least, for error growth associated with all singular vectors in the unstable subspace. The value $n = 30$, although chosen primarily for practical reasons as an upper limit on n for this study, provides robust values of α in the sense that moderately larger or smaller values of n have little impact on the value of α. This is consistent with the fact that the growth rate spectrum in Fig. 2a flattens considerably in this range.

The dashed line in Fig. 2b shows the average values of the projection coefficients $\alpha\sigma_i^2 c_i$ for the scaled sensitivity perturbations, with the values of α based on (25). As expected, the distribution is red with respect to singular vector index (whereas the distribution of c_i is approximately white), owing to the σ_i^2 weighting of the components. On average, the scaled sensitivity perturbations have comparable amplitude to the pseudo-inverse perturbations, as indicated by the similar areas under the curves. The scale factors seem reasonable in this sense and have an average value $\overline{\alpha} = 0.013$. This is in general agreement with the values obtained by R96. The relatively large projection onto the fourth singular vector in this figure is presumably an artifact of the relatively small sample size used to produce these results, which becomes more exaggerated in the case of the sensitivity perturbations.

The roughly uniform values of the projection coefficients c_i in Fig. 2b have implications for assessing the appropriateness of using a metric based on total energy to approximate the analysis error covariance metric (AECM). In principle, the AECM describes the second moment of the analysis error probability distribution. Were such a metric available, then the dominant singular vectors at initial time computed with respect to this metric would be precisely the structures that evolve into the leading EOFs (and thus describe the maximum amount of variance) of the forecast error covariance matrix at optimization time. Palmer et al. (1997) point out that a necessary condition for a trial metric to be a reasonable approximation to the AECM is that the spectra of the dominant singular vectors with respect to this metric be consistent with the spectra of analysis error variance. For example, the spectra of singular vectors should not be dominant in a wavenumber band for which the analysis error variance is relatively small. [Palmer et al. (1997) show that the spectra of singular vectors based on total energy are, in fact, consistent with the spectra of estimates of analysis error variance.] Similarly, if the consistency condition is satisfied, then analysis errors will, on average, project uniformly onto all singular vectors at initial time; that is, the analysis error probability distribution will be isotropic with respect to this metric. The roughly uniform values of c_i shown here clearly support the use of the total energy metric in this regard, although it should be noted that the consistency (or isotropy) condition itself is not sufficient to define an appropriate metric.

5. Perturbation experiments

Control and perturbed 10-day forecast experiments were conducted at T106 horizontal resolution with 31 vertical levels (T106L31) for the cases described in section 2. The control initial conditions were taken from the operational high-resolution (T213L31) optimal interpolation (OI) analysis, truncated to T106. The perturbed initial conditions were obtained by subtracting from the control initial conditions $\tilde{\mathbf{e}}_0$ or $\alpha\mathbf{C}^{-1}\tilde{\mathbf{s}}$ based on the 48-h forecast error for the appropriate starting date, computed using (23), or (24) and (25), respectively. The adjoint and forward tangent models used for the singular vector calculations are not a precise linearization of the T106L31 nonlinear forecast model. For practical purposes, the linearization is based on a T42L19 version of the model with simplified physics, including only surface drag and vertical diffusion as described by Buizza (1994). It should be noted, however, that $\mathbf{L}^{\mathrm{T}}$ is the precise numerical adjoint of $\mathbf{L}$ for this model.

For convenience hereafter, the scaled sensitivity projections will be denoted by simply $\tilde{\mathbf{s}}$, where multiplication by $\alpha\mathbf{C}^{-1}$ is implied unless otherwise stated. The perturbations were interpolated from T42L19 (the resolution of the singular vectors) to T106L31 before being subtracted from the control initial conditions. We begin

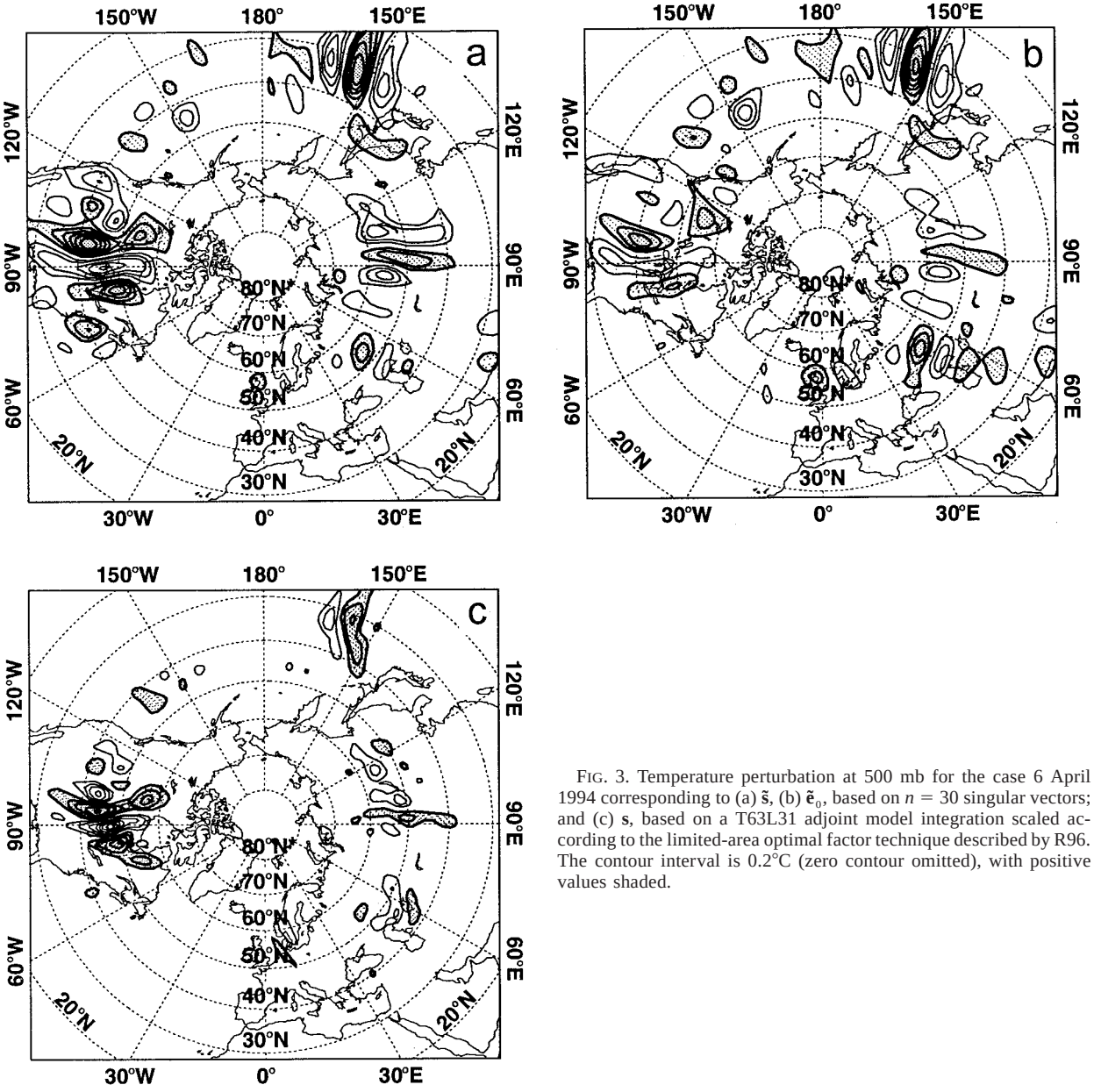

FIG. 3. Temperature perturbation at 500 mb for the case 6 April 1994 corresponding to (a) $\tilde{\mathbf{s}}$, (b) $\tilde{\mathbf{e}}_0$, based on $n = 30$ singular vectors; and (c) $\mathbf{s}$, based on a T63L31 adjoint model integration scaled according to the limited-area optimal factor technique described by R96. The contour interval is 0.2°C (zero contour omitted), with positive values shaded.

this section by presenting results for the low-skill case 6 April 1994 and the average-skill case 4 April 1994 (see Fig. 1). We then present composite results for the twelve cases.

a. Results for 6 April 1994

Figures 3a,b show the 500-mb temperature perturbations for the case 6 April 1994 corresponding to $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$, respectively, using $n = 30$ singular vectors. Equation (25) yields a scale factor of $\alpha = 0.0125$ in this case. For reference, we also show in Fig. 3c the sensitivity perturbation $\mathbf{s}$ based on a T63L31 adjoint model integration, scaled according to the limited-area optimal factor technique described by R96. The following is readily deduced from this figure. First, there is a close similarity between the patterns in all three panels, indicating that the relatively low-dimensional singular vector subspace captures most of the variance produced by the T63 adjoint model integration. Second, there are two primary areas of initial condition sensitivity located over North America and the western Pacific, with a weaker local maximum centered around 90°E. The primary sensitivity maxima occur nearly 90° upstream from the locations of the largest 120-h forecast errors over Europe and North America for this case (cf. Fig. 8c). Third, both the sensitivity and pseudo-inverse perturbations have similar amplitude, with maximum val-

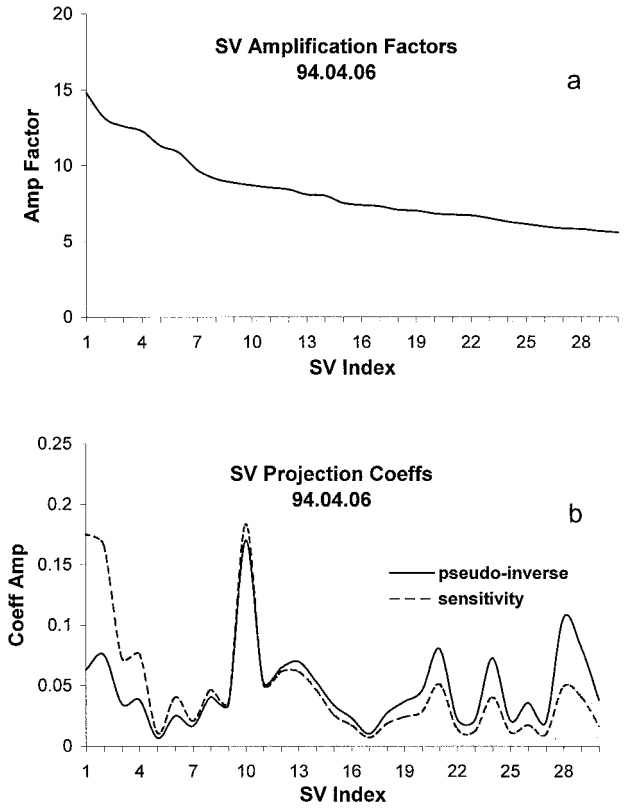

FIG. 4. As in Fig. 2 except for the single case 6 April 1994.

ues on the order of 1°–1.5°C at 500 mb. Note that these values are consistent with estimates of analysis errors in current data assimilation schemes (e.g., Buizza 1994).

Overall, there appear to be only minor qualitative differences between $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$. The similarity (and differences) between these perturbations may be examined in further detail in terms of Figs. 4 and 5. Figure 4 shows the amplification factors and singular vector projection coefficients for $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$ for this date. Figure 4b shows that the analysis error projects strongly onto several singular vectors, including $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_{10}$, $\mathbf{u}_{13}$, $\mathbf{u}_{21}$, $\mathbf{u}_{24}$, and $\mathbf{u}_{28}$. As expected, the projection coefficients for $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$ are similar except for the increased (decreased) weighting of the fastest- (slowest) growing singular vectors in the case of $\tilde{\mathbf{s}}$.

Figure 5 shows the 500-mb temperature structures of singular vectors (a) $\mathbf{u}_1$, (b) $\mathbf{u}_2$, and (c) $\mathbf{u}_{10}$ at $t = 0$. (Note that these are the temperature structures associated with the singular vectors themselves and not their relative contributions to $\tilde{\mathbf{s}}$ or $\tilde{\mathbf{e}}_0$.) Comparing Figs. 3 and 5, it can be seen that these three singular vectors appear to explain a significant fraction of the variance in the primary areas of initial condition sensitivity over North America ($\mathbf{u}_1$ and $\mathbf{u}_2$) and the western Pacific ($\mathbf{u}_{10}$). Again, the similarities between these figures imply that the analysis error has a significant projection onto these singular vectors, assuming the forecast error is dominated by growing analysis errors. Recall that the singular vectors have no explicit knowledge of the analysis errors of the day, only statistical information incorporated through the total energy metric. Finally, note that $\tilde{\mathbf{s}}$ has greater amplitude than $\tilde{\mathbf{e}}_0$ over North America (Figs. 3a,b) consistent with the increased weighting of $\mathbf{u}_1$ and $\mathbf{u}_2$ in $\tilde{\mathbf{s}}$ (dashed line in Fig. 4b).

Figure 6 shows longitudinal cross sections of the temperature and vorticity perturbations associated with $\tilde{\mathbf{s}}$ averaged between 40° and 50°N over North America. The perturbations have their maximum values in the lower to midtroposphere and exhibit a strong westward tilt with height, consistent with the results of R96, as well as other investigators who have examined atmospheric sensitivity patterns based on a wide range of forecast measures (e.g., Langland et al. 1995; Errico and Vukićević 1992; Rabier et al. 1992). Not surprisingly, this structure is also consistent with that of the dominant singular vectors themselves at initial time (not shown).

Figure 7 shows the skill scores of forecast experiments with different initial conditions for this case in terms of the ACC for 500-mb geopotential height over (a) the Northern Hemisphere and (b) Europe. The curves in each panel correspond to the ACC of the control forecast (solid), the forecast perturbed by $\tilde{\mathbf{s}}$ (dash), the forecast perturbed by $\tilde{\mathbf{e}}_0$ (dot), and, for reference, the forecast perturbed by the T63 adjoint model sensitivity perturbation $\mathbf{s}$ (dash–dot). All of the perturbed forecasts show a dramatic improvement in skill over the control forecast, especially over Europe in the medium range. The similarity in skill among the perturbed forecasts in both panels again supports the conclusion that a large fraction of the variance of the full sensitivity pattern is contained within a much lower-dimensional subspace spanned by the dominant singular vectors. The perturbed forecasts show continued improvement over the control well after the 48-h (forecast day+2) optimization time of the singular vectors. That is, the singular vectors continue to grow and provide a useful description of the forecast error evolution well beyond this interval. The fact that the $\tilde{\mathbf{s}}$ forecast performs slightly better than the $\mathbf{s}$ forecast, particularly over Europe in the medium range, may be due, at least in part, to differences in the scaling of the initial perturbations. In contrast, regional differences in skill between the $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$ forecasts may be more intrinsic, resulting from, for example, differences in the respective weighting of the most unstable singular vectors that, in this case, exhibit maximum growth over the North Atlantic–European region (see also Fig. 16).

Figure 8 compares the 120-h forecast error (control forecast minus the verifying analysis) with the evolution of the singular vector perturbations (control forecast minus the $\tilde{\mathbf{s}}$- and $\tilde{\mathbf{e}}_0$-perturbed forecasts) in terms of 500-mb geopotential height. It is clear that the singular vectors continue to describe a significant fraction of the actual forecast error after 120 h. Again, this demon-

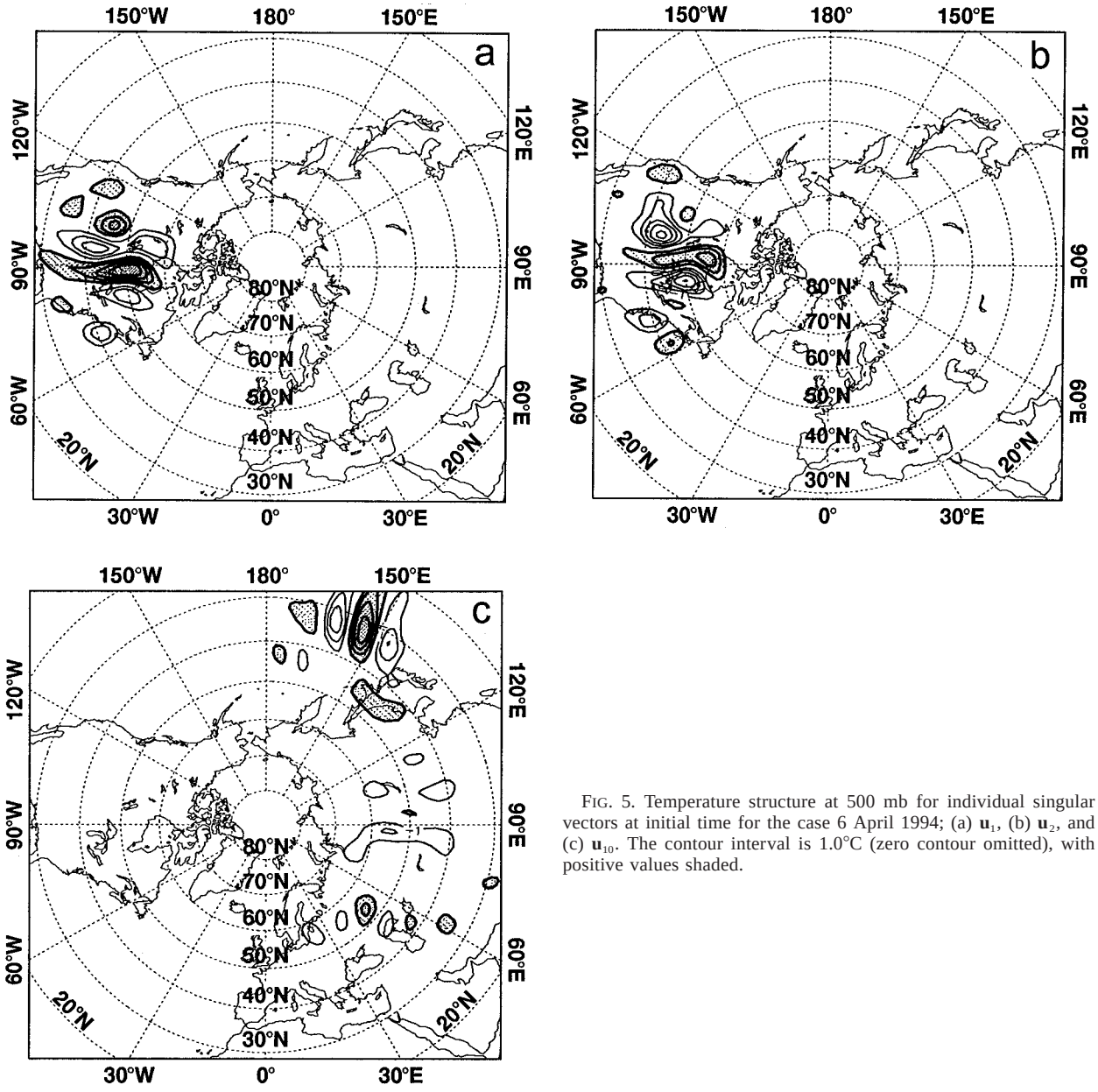

FIG. 5. Temperature structure at 500 mb for individual singular vectors at initial time for the case 6 April 1994; (a) $\mathbf{u}_1$, (b) $\mathbf{u}_2$, and (c) $\mathbf{u}_{10}$. The contour interval is 1.0°C (zero contour omitted), with positive values shaded.

strates their utility well beyond the 48-h optimization interval.

Since the initial perturbations for the EPS are based on a linear combination of singular vectors (without knowledge of the 48-h forecast error), it can be argued that the forecast improvement obtained with $\tilde{\mathbf{e}}_0$ or $\tilde{\mathbf{s}}$ provides a rough estimate of the upper limit on the expected improvement from any ensemble member. Thus, the dependence of the results shown here on the number of singular vectors used to construct $\tilde{\mathbf{e}}_0$ or $\tilde{\mathbf{s}}$ has implications for determining the number of singular vectors required to represent adequately the forecast uncertainty.

Figure 9 shows the spread at $t = 0$ and $t = 120$ h of the perturbations $\tilde{\mathbf{s}}^{30}$, $\tilde{\mathbf{s}}^{16}$, and $\tilde{\mathbf{s}}^{8}$ based on the projected sensitivity pattern, where for clarity, we have introduced an explicit superscript denoting the number of singular vectors used in each projection. The same scale factor [$\alpha = 0.0125$, based on (25) with $n = 30$] has been applied to all of these perturbations, so that the results reflect only differences arising from the number of singular vectors retained in each projection. Also, note the different contour intervals at $t = 0$ (left column panels) and $t = 120$ (right column panels) in this figure.

In all cases, we see dramatic differences in both the amplitude and horizontal scale of the perturbations at $t = 0$ and $t = 120$, indicative of strong nonmodal growth. Reducing the dimension of the sensitivity projection to

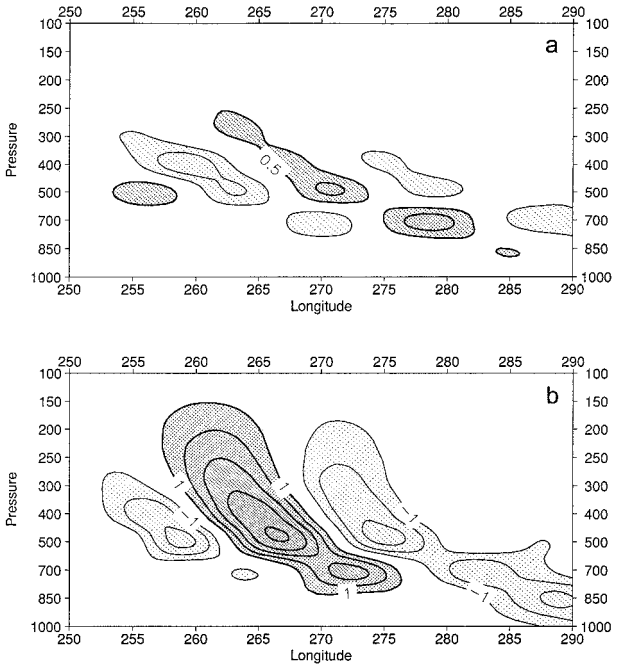

FIG. 6. Longitudinal cross section of perturbation (a) temperature, and (b) vorticity, associated with $\tilde{\mathbf{s}}$, averaged between 40° and 50°N. The contour interval is 0.5°C in (a), and 1.0×10^{-6} s^{-1} in (b), with positive (negative) values heavily (lightly) shaded.

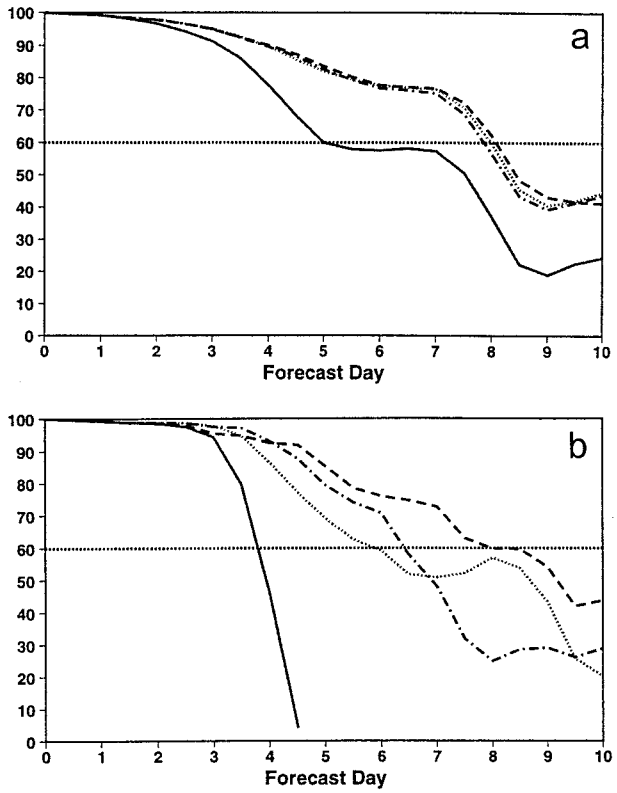

FIG. 7. Anomaly correlation coefficient for 500-mb geopotential height over (a) the Northern Hemisphere, and (b) Europe, for forecasts from 6 April 1994; control forecast (solid), the forecasts perturbed by $\tilde{\mathbf{s}}$ (dash) and $\tilde{\mathbf{e}}_0$ (dot) based on $n = 30$ singular vectors, and the forecast perturbed by **s** (dash–dot) based on the T63L31 adjoint model.

$\tilde{\mathbf{s}}^{16}$ (Fig. 9c) has almost no discernible impact on the evolved perturbation compared with $\tilde{\mathbf{s}}^{30}$ (Figs. 9b,d). Again, this result can be understood in terms of Figs. 4 and 5. Figure 4b shows that the sensitivity projection coefficients for this date (dashed line) are all relatively small beyond $n = 16$. Although the estimated analysis error appears to have a significant projection onto several singular vectors in this range (solid line), the more slowly growing components apparently have little impact compared with the projection onto $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_{10}$ for this date (Fig. 5). This conclusion is supported by qualitatively similar results obtained in experiments using pseudo-inverse perturbations based on different values of n (not shown).

Reducing the dimension of the sensitivity projection still further to $\tilde{\mathbf{s}}^8$ (Fig. 9e) greatly reduces the initial perturbation over the western Pacific associated with $\mathbf{u}_{10}$ (Fig. 5). As expected, however, there is little impact on the initial perturbation over North America associated with $\mathbf{u}_1$ and $\mathbf{u}_2$, or the response at $t = 120$ over Europe (Fig. 9f). A comparison of Fig. 9f with Figs. 9b,d shows the impact of $\mathbf{u}_{10}$ between the date line and 90°W at $t = 120$.

Figure 10 compares the skill of the $\tilde{\mathbf{s}}^{30}$, $\tilde{\mathbf{s}}^{16}$, and $\tilde{\mathbf{s}}^8$ forecasts with that of the control forecast for this date over Europe and North America. Again, only the $\tilde{\mathbf{s}}^8$ forecast over North America differs qualitatively from its counterparts, consistent with the diminution of the initial perturbation over the western Pacific (Fig. 9e).

As might be expected from Figs. 4 and 5, a sensitivity perturbation, $\tilde{\mathbf{s}}^2$, based on only the two leading singular vectors for this date still produces a significant improvement in the medium-range forecast over Europe and the North Atlantic (not shown). While these results do not imply that such a small number of singular vectors will produce such dramatic results in general, they do demonstrate the importance of sampling the leading singular vector directions given that the chance of making an initial error in these directions is as equally likely as any other.

b. Results for 4 April 1994

For this case we present a subset of the results presented for 6 April 1994, in order to demonstrate some notable differences between an average- and low-skill case. Figures 11a,b show the 500-mb temperature perturbations for this case corresponding to $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$, respectively, using $n = 30$ singular vectors and a scale factor $\alpha = 0.0106$. As in Fig. 3, we also show for reference in Fig. 11c the perturbation **s** based on a T63L31 adjoint model integration. In this case, there is considerably less correspondence between the three perturbations compared with the results for 6 April 1994.

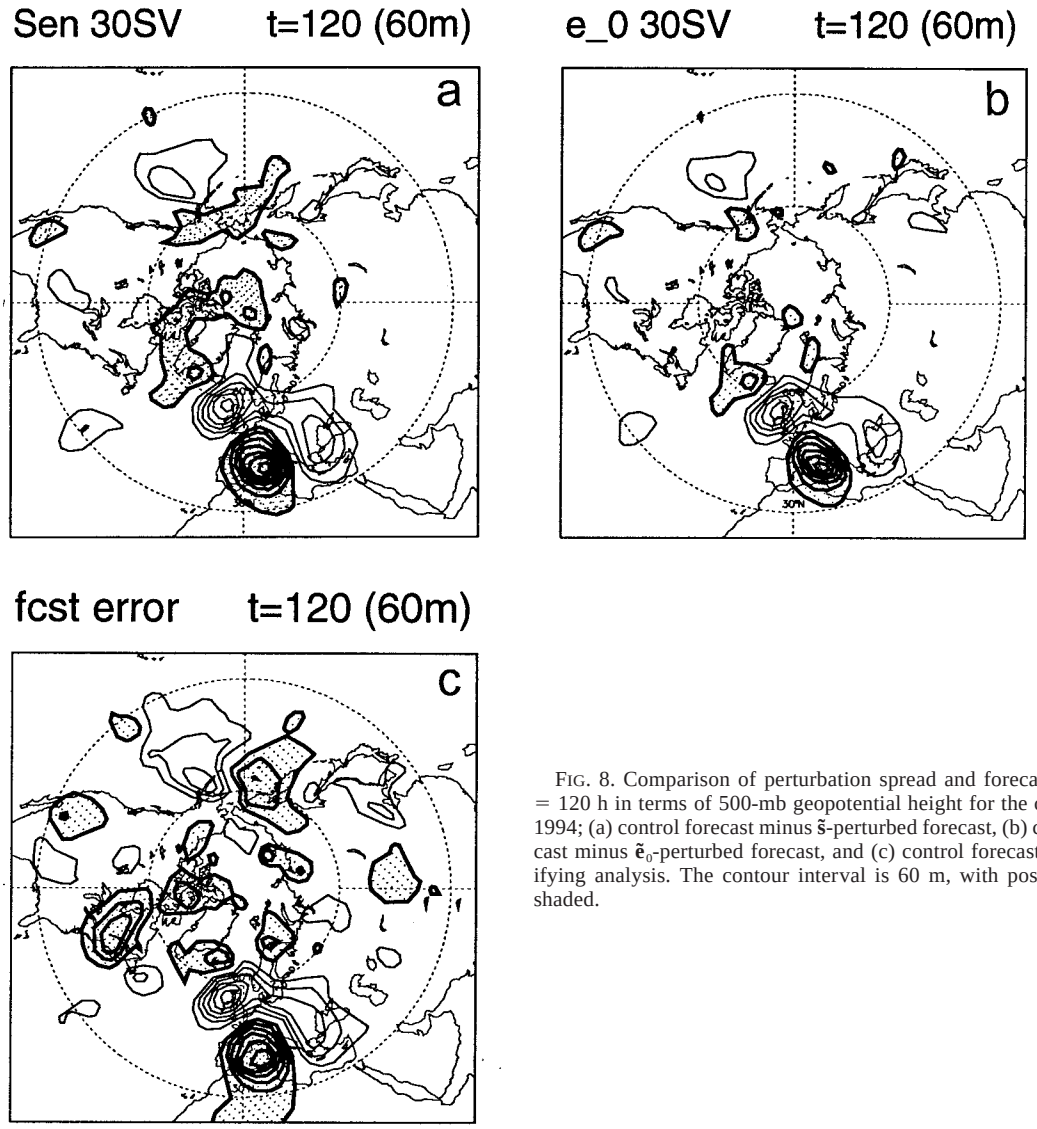

FIG. 8. Comparison of perturbation spread and forecast error at $t = 120$ h in terms of 500-mb geopotential height for the case 6 April 1994; (a) control forecast minus $\tilde{\mathbf{s}}$-perturbed forecast, (b) control forecast minus $\tilde{\mathbf{e}}_0$-perturbed forecast, and (c) control forecast minus verifying analysis. The contour interval is 60 m, with positive values shaded.

In particular, there is far less agreement between $\mathbf{s}$ and $\tilde{\mathbf{s}}$. Although the major sensitivity features over the North Atlantic and Europe are reasonably well represented by the projected field, the full sensitivity pattern has many additional small-scale features with significant amplitude, especially outside the North Atlantic–European region.

The amplification factors and projection coefficients for $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$ for this case are shown in Fig. 12. The growth rates (Fig. 12a) are roughly comparable to those for 6 April 1994 over most of the spectrum, although the leading growth rates are in fact slightly larger in the present case. Generally speaking, the largest projection coefficients are also comparable in magnitude to those for 6 April 1994 (Fig. 4b). However, the characteristics of the error projection onto the singular vectors shows some important qualitative differences. In particular, the error projects predominantly onto singular vectors farther down the spectrum in this case, namely, $\mathbf{u}_{17}$, $\mathbf{u}_{24}$, $\mathbf{u}_{25}$, and $\mathbf{u}_{28}$. Compared with 6 April 1994 (solid line in Fig. 4b), the initial error projection is two to three times smaller for the leading singular vectors in the present case and many times smaller for several singular vectors with index values between 10 and 15. Again, the projection coefficients for $\tilde{\mathbf{s}}$ are more heavily weighted toward the leading singular vectors (dashed line in Fig. 12b).

These differences affect the perturbations in several ways. First, we expect the perturbations to be less localized because the more slowly growing singular vectors have less localized horizontal structures. Figure 13 shows the 500-mb temperature structures associated with (a) $\mathbf{u}_{17}$, (b) $\mathbf{u}_{25}$, and (c) $\mathbf{u}_{28}$ for this case. All three singular vectors have substantial amplitude in several

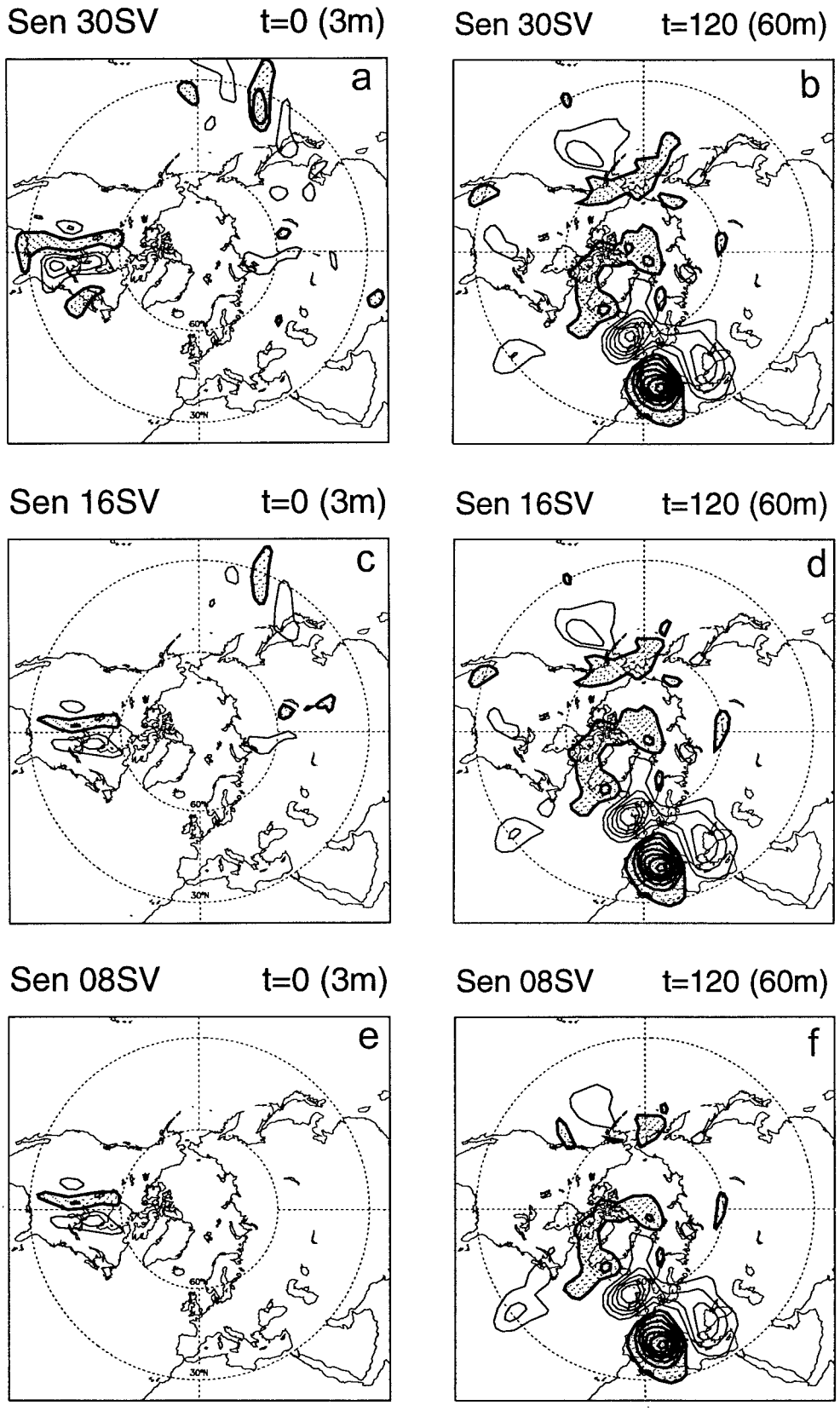

FIG. 9. Perturbation spread at $t = 0$ (left panels), and $t = 120$ h (right panels), in terms of 500-mb geopotential height for the case 6 April 1994; (a)–(b) $\tilde{\mathbf{s}}^{30}$, (c)–(d) $\tilde{\mathbf{s}}^{16}$, and (e)–(f) $\tilde{\mathbf{s}}^{8}$. The same initial scale factor is used in all cases. The contour interval is 3 m at $t = 0$ and 60 m at $t = 120$ h, with positive values shaded.

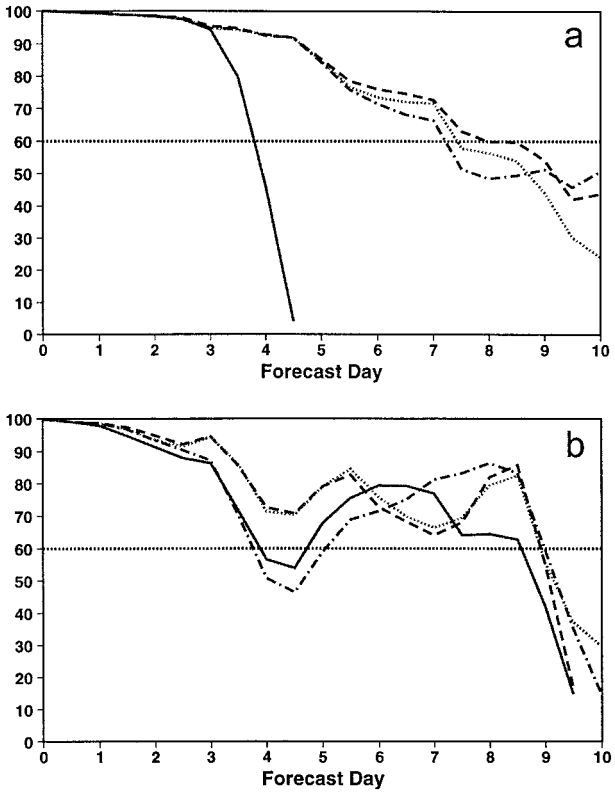

FIG. 10. Anomaly correlation coefficient for 500-mb geopotential height over (a) Europe and (b) North America for forecasts from 6 April 1994; control forecast (solid), and forecasts perturbed by $\tilde{\mathbf{s}}^{30}$ (dash), $\tilde{\mathbf{s}}^{16}$ (dot), and $\tilde{\mathbf{s}}^{8}$ (dash–dot).

locations across the Northern Hemisphere. The structures of these singular vectors are clearly consistent with the noisy structure of the sensitivity and pseudo-inverse perturbations for this case. Second, because of the decreased weighting of the more slowly growing singular vectors in the $\tilde{\mathbf{s}}$ projection, there is less similarity between $\tilde{\mathbf{e}}_0$ and $\tilde{\mathbf{s}}$ when the analysis error projects predominantly onto these singular vectors. Finally, of course, we expect the perturbations to grow more slowly when the initial error projects onto singular vectors that grow more slowly.

Figure 14 shows the skill scores for 4 April 1994 over (a) the Northern Hemisphere and (b) Europe corresponding to the control forecast, the perturbed forecasts using $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$ based on $n = 30$ singular vectors, and the perturbed forecast using $\mathbf{s}$ based on the T63 adjoint model. In this case the control forecast shows roughly average skill. The perturbed forecasts all provide equal (moderate) improvement in medium-range skill over the Northern Hemisphere. Over Europe, the $\mathbf{s}$ forecast provides a more marked improvement than either $\tilde{\mathbf{s}}$ or $\tilde{\mathbf{e}}_0$, although again not as dramatic as for 6 April 1994. The increased performance of $\mathbf{s}$ (which has many more degrees of freedom than either $\tilde{\mathbf{s}}$ or $\tilde{\mathbf{e}}_0$) is consistent with there being a weaker initial error projection onto the leading singular vectors in this case, as shown in Fig. 12b. Thus, initial errors that project onto singular vectors beyond $n = 30$ may not be negligible since their growth rates are not significantly smaller than say $\mathbf{u}_{24}$ and $\mathbf{u}_{25}$. This might explain the additional small-scale structures in $\mathbf{s}$ compared with $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$, although we are unable to examine these projections explicitly in this study.

c. Composite results

In this section we present some composite results based on the 12 cases examined in this study. The growth and structural evolution of the perturbations are summarized in Fig. 15, which shows the composite (a) vertical energy profiles as a function of model level and (b) total wavenumber spectra for $\tilde{\mathbf{e}}_0$ initially (solid) and after 48 h of evolution in the nonlinear model ($\mathbf{M}\tilde{\mathbf{e}}_0$; dash). The composites for the 48-h forecast error $\mathbf{e}_{48}$ over the Northern Hemisphere for these dates are also shown (dot). For visual comparison with $\mathbf{e}_{48}$ in this figure, we have multiplied $\mathbf{M}\tilde{\mathbf{e}}_0$ by a factor of 3, and $\tilde{\mathbf{e}}_0$ by a factor of 300. We show results for $\tilde{\mathbf{e}}_0$ because of its direct relationship to $\tilde{\mathbf{e}}_{48}$ under the perfect model assumption. The results for $\tilde{\mathbf{s}}$ are similar.

The initial perturbations have a clear vertical energy peak at model level 12 (close to 700 mb) and a broad spectrum with maximum energy values between wavenumbers 20 and 25. After 48 h, the evolved perturbations have their vertical peak at model level 9 (approximately 325 mb) near the upper-tropospheric jet and a more pronounced spectral peak between wavenumbers 10 and 15, consistent with synoptic-scale baroclinic disturbances. The spectra and vertical profiles of the evolved perturbations are qualitatively very similar to those of $\mathbf{e}_{48}$. Overall, $\mathbf{M}\tilde{\mathbf{e}}_0$ accounts for approximately 17% of the total energy of $\mathbf{e}_{48}$ over the Northern Hemisphere, as given by the ratio

$$r = \frac{\langle \mathbf{M}\tilde{\mathbf{e}}_0; \mathbf{C}\mathbf{M}\tilde{\mathbf{e}}_0 \rangle}{\langle \mathbf{e}_{48}; \mathbf{C}\mathbf{e}_{48} \rangle}. \tag{26}$$

The dramatic energy growth and highly nonmodal evolution of $\tilde{\mathbf{e}}_0$ shown in these profiles are consistent with those of the sensitivity perturbations described by R96 and with the singular vectors themselves (Buizza and Palmer 1995; Buizza et al. 1997).

Figure 16 shows the composite ACC skill scores over (a) the Northern Hemisphere, (b) Europe, and (c) North America for the control forecasts, and the $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$ forecasts based on $n = 30$ singular vectors. (Perturbed forecasts based on the full sensitivity pattern were not available for all 12 cases.) Over the Northern Hemisphere, $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}_0$ have roughly the same impact on the medium-range forecast. On average, the improvement in skill over the Northern Hemisphere is clearly less than the 48 h lost by waiting for the forecast error information. More dramatic improvements are obtained regionally, especially over Europe, where the $\tilde{\mathbf{s}}$ forecasts show an improvement in medium-range skill approaching 48 h.

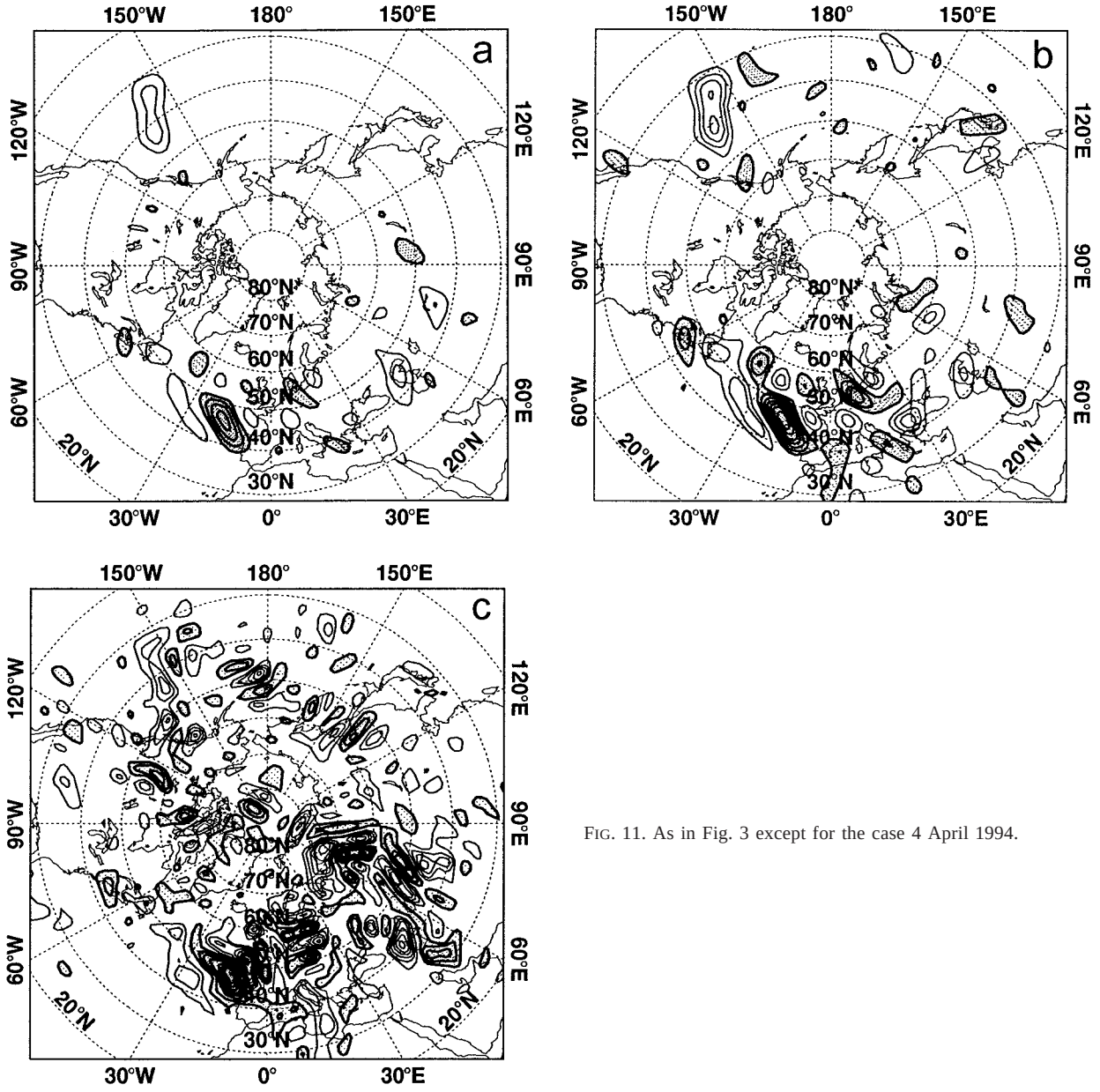

FIG. 11. As in Fig. 3 except for the case 4 April 1994.

The large forecast improvements over Europe and North America are consistent with these regions being located downstream from the primary maxima in Northern Hemisphere baroclinic instability, and their associated maxima in the geographical distribution of the dominant singular vectors (e.g., Buizza et al. 1997).

Figure 17 shows the composite ACC skill scores of the $\tilde{\mathbf{s}}^{30}$, $\tilde{\mathbf{s}}^{16}$, and $\tilde{\mathbf{s}}^{8}$ forecasts compared with the composite skill of the control forecasts. As in the individual cases discussed earlier, there is only moderate dependence on the number of singular vectors retained in the range $n = 8$–30. Similar results were obtained for the pseudo-inverse composites with varying n (not shown). Again, only the medium-range forecasts over Europe show any significant differences between the performance of the sensitivity and pseudo-inverse perturbations. For example, it can be seen that the composite skill of the $\tilde{\mathbf{s}}^{8}$ forecasts is equal to or greater than that of the $\tilde{\mathbf{e}}_0^{30}$ forecasts between days 4.5 and 6.5. This result may be due, in part, to the scale factors α applied to the sensitivity perturbations. However, considering again the geographical distribution of the singular vectors, it is also likely that the increased weighting of the fastest-growing singular vectors in the case of the $\tilde{\mathbf{s}}$ perturbations has a strong forecast impact over the North Atlantic and Europe.

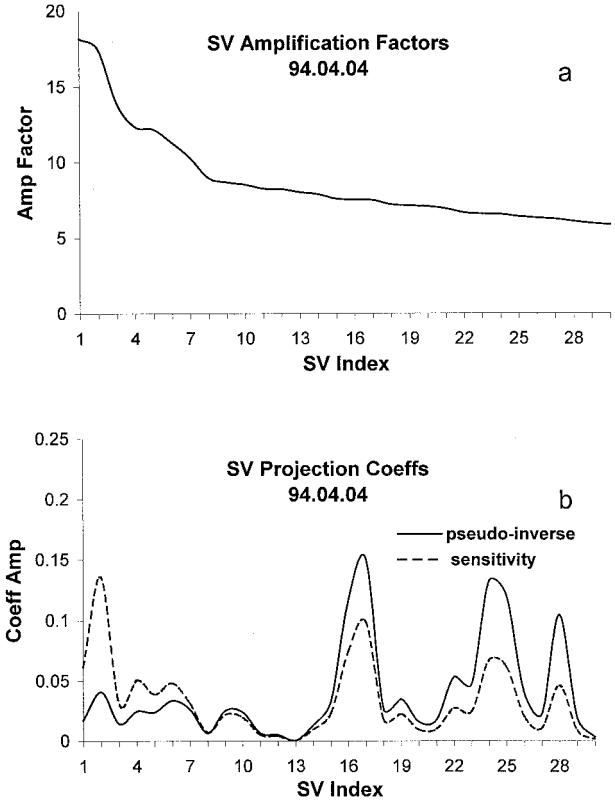

FIG. 12. As in Fig. 2 except for the single case 4 April 1994.

6. Optimality and nonmodal growth

For realistic meteorological problems, the operator $\mathbf{L}_c$ is not self-adjoint. An important consequence of this property is that the singular vectors of $\mathbf{L}_c$ are not shape preserving in their evolution; that is, they are nonmodal (Orr 1907; P98). Thus, for example, the unstable subspaces spanned by the singular vectors at the beginning and end of a given forecast trajectory may be virtually orthogonal to one another. Similarly, if we consider a series of consecutive forecast trajectories (such as that produced by the day-to-day running of an operational forecast model), then there may be little projection between the evolved singular vectors at the end of one forecast trajectory and the initial singular vectors at the beginning of the next trajectory (e.g., Buizza 1998). The weak projectibility of these subspaces can have important implications for constructing optimal perturbations for ensemble weather prediction, as well as for developing future strategies for (adaptive) data observation and assimilation.

To examine some of these implications, let us first consider the six low-skill cases discussed in previous sections. For each case, we denote by $\mathbf{U}_{p,q}$ and $\mathbf{V}_{p,q}$ the singular vectors at initial and final time, respectively, optimized for a forecast trajectory from day$+p$ to day$+q$, where day$+0$ denotes the (initial) dates of these six cases. For the 48-h optimization interval used in this study, we denote the singular vectors at initial and final time for these cases by $\mathbf{U}_{0,2}$ and $\mathbf{V}_{0,2}$, respectively. Now, as described in section 2 and discussed throughout this paper, singular vectors optimized for 48 h were computed for six additional (average-skill) cases, with initial dates 2 days previous to the initial dates of the low-skill cases. Maintaining our reference to day$+0$ as being the initial dates of the low skill cases, then we denote the singular vectors at initial and final time for each of the average-skill cases by $\mathbf{U}_{-2,0}$ and $\mathbf{V}_{-2,0}$, respectively. The singular vector sets $\mathbf{V}_{-2,0}$, from an average-skill case, and $\mathbf{U}_{0,2}$, from the low-skill case 2 days later, thus represent contemporaneous subspaces at the end of one trajectory portion and the beginning of a second consecutive trajectory portion.

We can compare the n leading singular vectors of these subspaces in terms of the projection matrix $\mathbf{W}(\mathbf{X}, \mathbf{Y}; n)$ for two normalized singular vector sets $\mathbf{X}$ and $\mathbf{Y}$ (Buizza 1994, 1995). Each matrix element

$$w_{i,j}(\mathbf{X}, \mathbf{Y}) = \langle \mathbf{x}_i; \mathbf{y}_j \rangle^2 \tag{27}$$

describes the variance of the ith singular vector of $\mathbf{X}$ explained by the jth singular vector of $\mathbf{Y}$. The sum of the matrix elements with one index—for example, i—fixed, denoted by

$$p_i(\mathbf{x}_i, \mathbf{Y}; n) = \sum_{j=1}^{n} w_{i,j}(\mathbf{X}, \mathbf{Y}), \tag{28}$$

describes the degree to which the ith singular vector of $\mathbf{X}$ can be reconstructed from a linear combination of the first n singular vectors of $\mathbf{Y}$. The similarity between the unstable subspaces spanned by the first n singular vectors of $\mathbf{X}$ and $\mathbf{Y}$ can then be defined by the index

$$s(\mathbf{X}, \mathbf{Y}; n) = \frac{1}{n} \sum_{i=1}^{n} p(\mathbf{x}_i, \mathbf{Y}; n). \tag{29}$$

For reference, we note that for parallel subspaces $\mathbf{W}(\mathbf{X}, \mathbf{Y}; n)$ is the identity matrix and $s(\mathbf{X}, \mathbf{Y}; n) = 100\%$. Conversely, for orthogonal subspaces $s(\mathbf{X}, \mathbf{Y}; n) = 0\%$.

Table 1 shows the projection matrix $\mathbf{W}(\mathbf{U}_{0,2}, \mathbf{V}_{-2,0}; 20)$ at day$+0$ for the case 6 April 1994. The index values 1–20 along the left side of the table correspond to singular vectors in $\mathbf{U}_{0,2}$ (the initial singular vectors for 6 April 1994), while those along the top correspond to singular vectors in $\mathbf{V}_{-2,0}$ (the evolved singular vectors for 4 April 1994). The values $p_i(\mathbf{u}_i, \mathbf{V}_{-2,0}; 20)$ and $p_j(\mathbf{v}_j, \mathbf{U}_{0,2}; 20)$ are shown along the right side and bottom of the table, respectively, given by the sum of the elements in the corresponding row or column. The similarity index $s(\mathbf{U}_{0,2}, \mathbf{V}_{-2,0}; 20)$ for these subspaces is shown in the lower right corner. Only nonzero values (%) are shown, rounded to the nearest integer. It is clear from Table 1 that the subspaces spanned by the 20 leading singular vectors of $\mathbf{U}_{0,2}$ and $\mathbf{V}_{-2,0}$ at day$+0$ are nearly orthogonal to one another. For example, only 3% of the variance of the leading singular vector in $\mathbf{U}_{0,2}$ can be explained by a linear combination of the first 20 singular

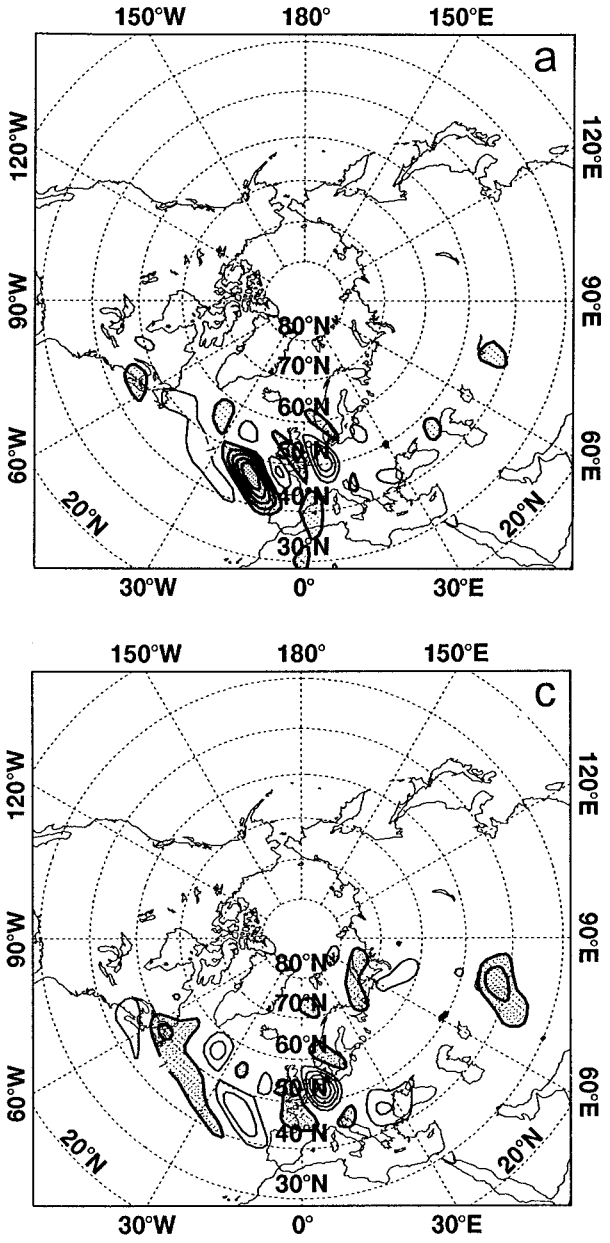

FIG. 13. Temperature structure at 500 mb for individual singular vectors at initial time for the case 4 April 1994; (a) $\mathbf{u}_{17}$, (b) $\mathbf{u}_{25}$, and (c) $\mathbf{u}_{28}$. The contour interval is 1.0°C, with positive values shaded.

vectors in $\mathbf{V}_{-2,0}$. Similarly, only 6% of the variance of the leading singular vector in $\mathbf{V}_{-2,0}$ can be explained by the first 20 singular vectors in $\mathbf{U}_{0,2}$. The similarity index for these subspaces is approximately 2%. Qualitatively similar results were obtained with $n = 30$ singular vectors for this and other cases.

The weak projectibility of $\mathbf{U}_{0,2}$ and $\mathbf{V}_{-2,0}$ at day+0 is due, in part, to differences in the characteristic length scales of the initial and evolved singular vectors. Because the initial singular vectors evolve fairly rapidly toward larger scales, we also examined the projection between $\mathbf{U}_{0,2}$ and $\mathbf{V}_{-2,0}$ at day+1, after integrating each set of singular vectors forward in time an additional 24 h from the day+0 reference date. Table 2 shows the projection matrix $\mathbf{W}(\mathbf{U}_{0,2}, \mathbf{V}_{-2,0}; 20)$ at day+1 for these subspaces, that is, for the contemporaneous subspaces on 7 April 1994 spanned by the singular vectors from 4 April 1994 after 72 h of evolution and the singular vectors from 6 April 1994 after 24 h of evolution. The results show a substantial increase in the subspace projections compared with Table 1, especially for the leading singular vectors in each set, indicating some degree of convergence as the forecast progresses. Anywhere from roughly 10%–60% of the variance of the leading singular vectors in each set can now be explained by a linear combination of the first 20 singular vectors in the other. The variance explained decreases to roughly 10%–25% for singular vectors with index values greater

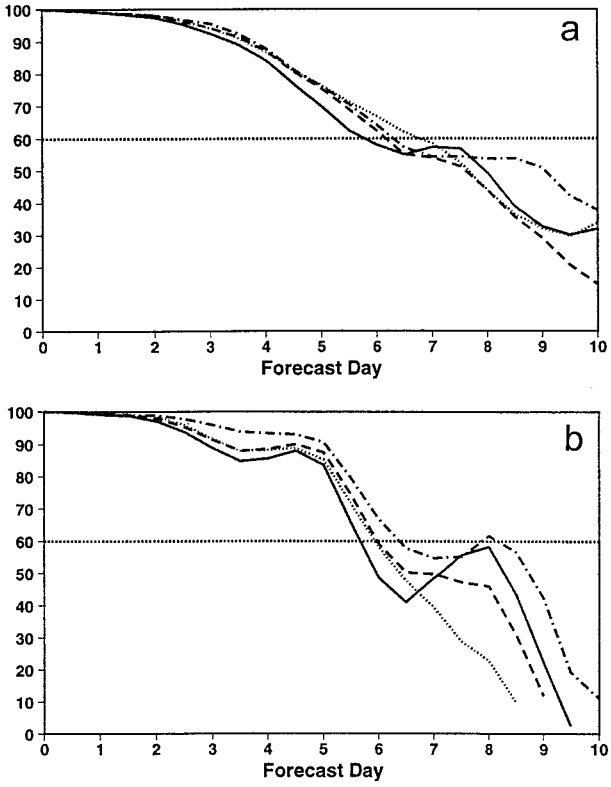

FIG. 14. As in Fig. 7, except for the case 4 April 1994.

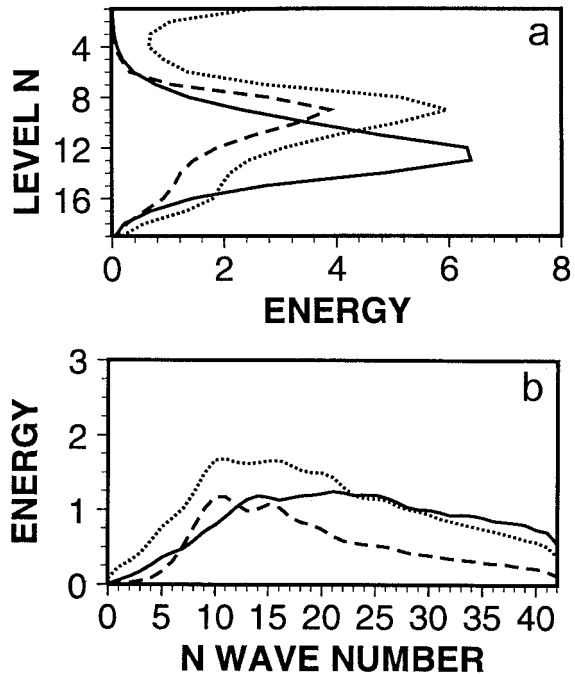

FIG. 15. Composite (a) vertical energy profiles, and (b) total wavenumber spectra, for twelve cases; $\tilde{\mathbf{e}}_0$ at initial time ($\times 300$; solid), $\tilde{\mathbf{e}}_0$ after 48 h of evolution in the nonlinear model ($\times 3$; dash), and $\mathbf{e}_{48}$ (dot). The units are J kg^{-1}.

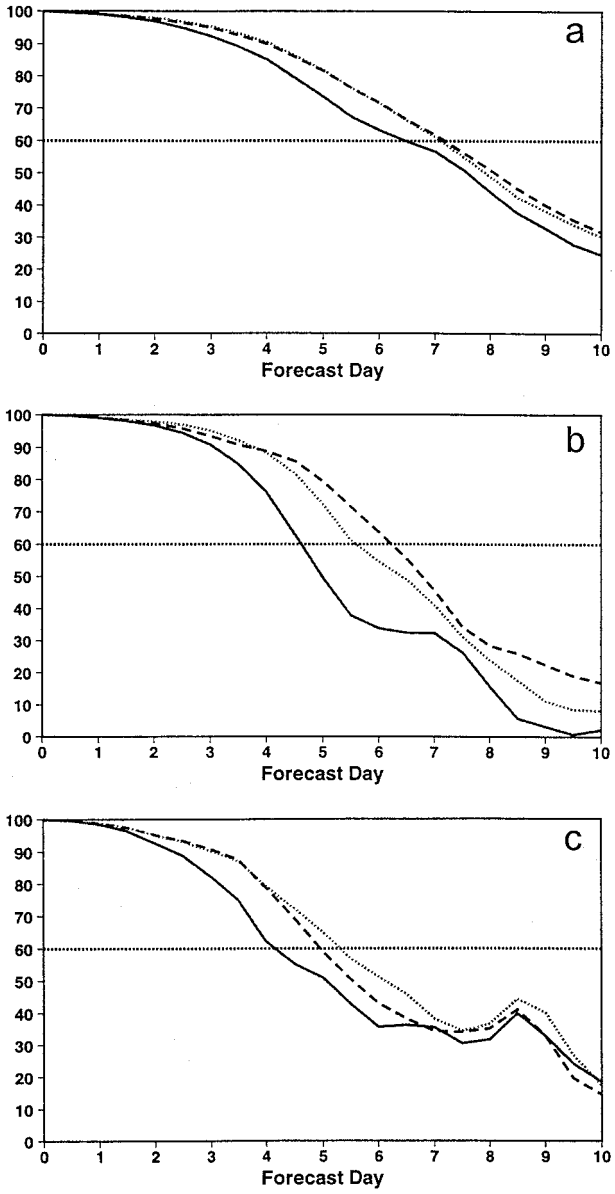

FIG. 16. Composite anomaly correlation coefficient for 500-mb geopotential height over (a) the Northern Hemisphere, (b) Europe, and (c) North America, for 12 cases; control forecasts (solid), and forecasts perturbed by $\tilde{\mathbf{s}}$ (dash) and $\tilde{\mathbf{e}}_0$ (dot) based on $n = 30$ singular vectors.

than 10, and the similarity index for these subspaces is approximately 21%.

Considering the results in Tables 1 and 2, we might ask whether perturbations based on a forecast trajectory either beginning or ending at day+0 evolve similarly beyond this time, despite their initial differences. The forecast experiments described in previous sections involve perturbations based on the singular vectors $\mathbf{U}_{0,2}$ with respect to the day+0 analysis. Below, we examine sensitivity and ensemble forecast experiments in which

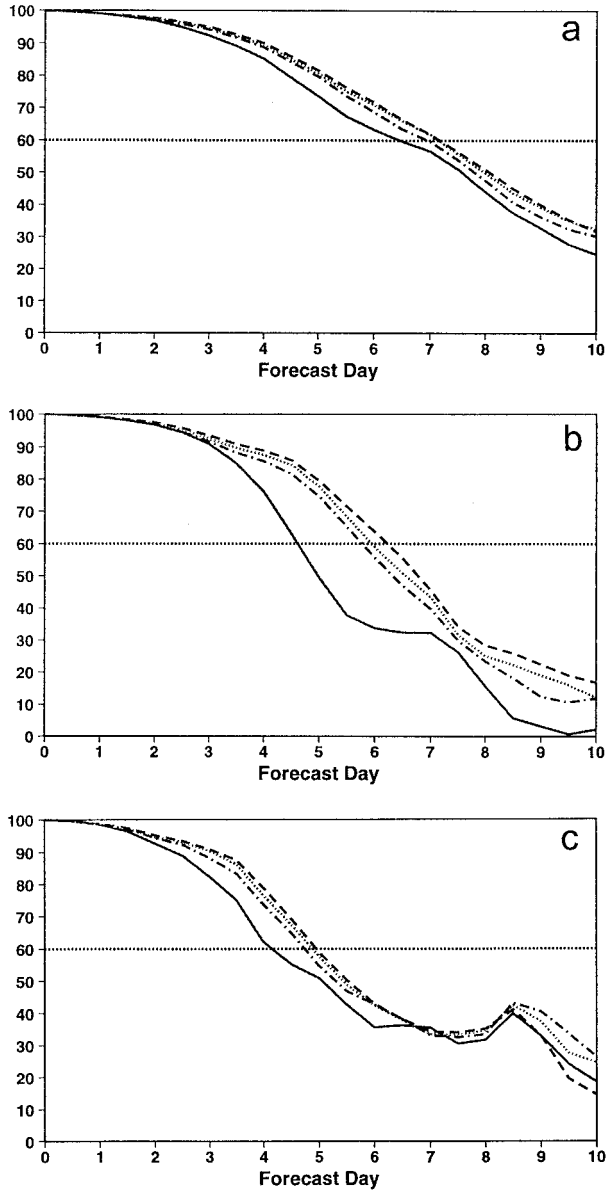

FIG. 17. Composite anomaly correlation coefficient for 500-mb geopotential height over (a) the Northern Hemisphere, (b) Europe, and (c) North America, for 12 cases; control forecasts (solid), and forecasts perturbed by $\tilde{\mathbf{s}}^{30}$ (dash), $\tilde{\mathbf{s}}^{16}$ (dot), and $\tilde{\mathbf{s}}^{8}$ (dash–dot).

the initial perturbations are based on the singular vectors $\mathbf{V}_{-2,0}$ with respect to this analysis.

a. Sensitivity experiments with $\mathbf{V}_{-2,0}$

Figure 18a shows the 500-mb temperature sensitivity perturbation for the case 6 April 1994 projected onto $\mathbf{V}_{-2,0}$ (the evolved singular vectors from 4 April 1994). We denote this perturbation $\mathbf{s}'$, given by

$$\mathbf{s}' = \mathbf{V}_{-2,0}\mathbf{V}^{\mathrm{T}}_{-2,0}\tilde{\mathbf{s}}, \tag{30}$$

where, as before, $\tilde{\mathbf{s}}$ is the scaled sensitivity perturbation based on (24). Here, we use $n = 30$ singular vectors for all projection operations. Note that, since $\tilde{\mathbf{s}}$ is itself a sensitivity projection into $\mathbf{U}_{0,2}$, replacing $\mathbf{V}_{-2,0}$ with $\mathbf{U}_{0,2}$ in (30) gives $\mathbf{s}' = \tilde{\mathbf{s}}$. Also note that in Fig. 18, we have multiplied $\mathbf{s}'$ by the factor 6.59 to set its total energy equal to that of $\tilde{\mathbf{s}}$.

Although $\mathbf{s}'$ has its maximum amplitude in the same general locations as $\tilde{\mathbf{s}}$ (cf. Fig. 3a), there are clear differences in the horizontal scales and amplitudes of these perturbations. For example, over North America the characteristic length scale of $\mathbf{s}'$ has clearly increased compared with $\tilde{\mathbf{s}}$ (Figs. 18a,c), consistent with the scales of evolved singular vectors. Also, in terms of 500-mb temperature, the amplitude of $\mathbf{s}'$ is significantly weaker (the contour interval in Fig. 18a is half that in Fig. 3a) since $\mathbf{V}_{-2,0}$ is dominated by kinetic energy. There is also a significant difference in the vertical structures of these perturbations, as illustrated in Fig. 18b, which shows the vertical energy profiles of $\tilde{\mathbf{s}}$ and $\mathbf{s}'$. The perturbation $\mathbf{s}'$ has its maximum amplitude near the level of the upper tropospheric jet (approximately 300 mb), rather than in the lower to middle troposphere. In general, the wavenumber spectrum and vertical profiles for $\mathbf{s}'$ are more qualitatively similar to those of $\mathbf{e}_{48}$ than $\tilde{\mathbf{s}}$ (cf. Fig. 15).

Forecast experiments analogous to those discussed in section 5 were conducted in which perturbations based on $\mathbf{s}'$ were added to the control initial conditions. Figures 19a,c show the ACC skill scores for the Northern Hemisphere and Europe, respectively, for the case 6 April 1994. The solid and dashed curves in these figures show the control and $\tilde{\mathbf{s}}$ forecasts for this case again for reference. The dotted curve shows the $\mathbf{s}'$ forecast, where the initial perturbation is based on (30) directly. The dash–dot curve shows the forecast based on $\mathbf{s}'$ with total energy at initial time equal to that of $\tilde{\mathbf{s}}$ (as in Fig. 18a). In both cases, the impact of the $\mathbf{s}'$ perturbation is much less influential than that of $\tilde{\mathbf{s}}$; even an increase in the initial perturbation amplitude by the factor 6.59 has little additional effect on the forecast skill. This result strongly implies that structural and dynamic differences between $\tilde{\mathbf{s}}$ and $\mathbf{s}'$ at initial time are the major factors controlling the perturbation growth. Figures 19b,d show the average skill scores for the six (low skill) cases. The mean qualitative behavior is the same as that for 6 April 1994. The results of these sensitivity experiments are clearly consistent with the nonmodality of the singular vectors, and with the weak projectibility of the unstable subspaces at the end of one trajectory portion and the beginning of a second consecutive trajectory portion.

b. Ensemble experiments with $\mathbf{V}_{-2,0}$

Ensemble forecasts experiments were conducted using perturbations at day+0 based on the singular vectors $\mathbf{V}_{-2,0}$. We refer to these experiments as $\mathbf{V}_{-2,0}$ ensembles to distinguish them from the $\mathbf{U}_{0,2}$ ensembles run operationally. Except for using $\mathbf{V}_{-2,0}$ instead of $\mathbf{U}_{0,2}$, we followed the EPS procedure used operationally at the

TABLE 1. The projection matrix **W** ($\mathbf{U}_{0,2}$, $\mathbf{V}_{-2,0}$; 20) at day+0 for the case 6 April 1994. The top row identifies the index number for singular vectors in $\mathbf{V}_{-2,0}$; the leftmost column identifies the index number for singular vectors in $\mathbf{U}_{0,2}$. The boldface values in the bottom row show the percentage of the norm of each singular vector in $\mathbf{V}_{-2,0}$ that can be explained by a linear combination of singular vectors in $\mathbf{U}_{0,2}$, and conversely for the boldface values in the rightmost column. The similarity index for these subspaces is shown in the lower-right corner. Only nonzero values (%) are shown, rounded to nearest integer. See text for details.

$\mathbf{U}_{0,2}$ \ $\mathbf{V}_{-2,0}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	p_i
1					2				1												**3**
2			1																		**2**
3		1																			**2**
4		2																			**2**
5																					**1**
6	2																				**3**
7					1																**2**
8			1						1												**2**
9		1																			**2**
10	1	1					1							1							**3**
11		1																			**2**
12							1								1						**2**
13	1																				**2**
14																					**1**
15	1																				**3**
16																					**1**
17					2																**3**
18			1																		**2**
19													1	1							**2**
20						1									1						**3**
p_j	**6**	**7**	**3**	**1**	**5**	**2**	**2**	**1**	**2**	**1**	**1**	**1**	**2**	**2**	**2**	**1**	**1**	**1**	**1**	**1**	s=**2**

TABLE 2. As in Table 1 except at day+1.

$\mathbf{U}_{0,2}$ \ $\mathbf{V}_{-2,0}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	p_i
1					18				10		1	4				1			1		**36**
2			15		1				2		7	12	1					1	1	2	**42**
3	1	14					5	1							1						**23**
4		19				1	2	1							1		1				**26**
5	1	4		2																	**8**
6	33	2				1	1														**38**
7					20							1				1		1	1	1	**25**
8			20		1				11							1		1	2	1	**38**
9	1	3		1						1		1	2		1			1			**11**
10	4	1				3	3	1		1				3	2		3	2	1	1	**24**
11	1	5						2											1		**11**
12		1		1		2	2	1							5						**14**
13	4	1																3		3	**13**
14	3			1			1													1	**7**
15	10			1				1			1	1	1		2	1		1	1	2	**21**
16	2			1							1						3	2		1	**10**
17			1	1	23				2												**26**
18			6		2				1									1	1	1	**13**
19										1			1				3	2	2		**10**
20						7				1			2	3	3				1		**19**
p_j	**59**	**50**	**42**	**9**	**65**	**15**	**15**	**8**	**26**	**6**	**1**	**20**	**9**	**8**	**15**	**6**	**12**	**15**	**11**	**14**	s=**21**

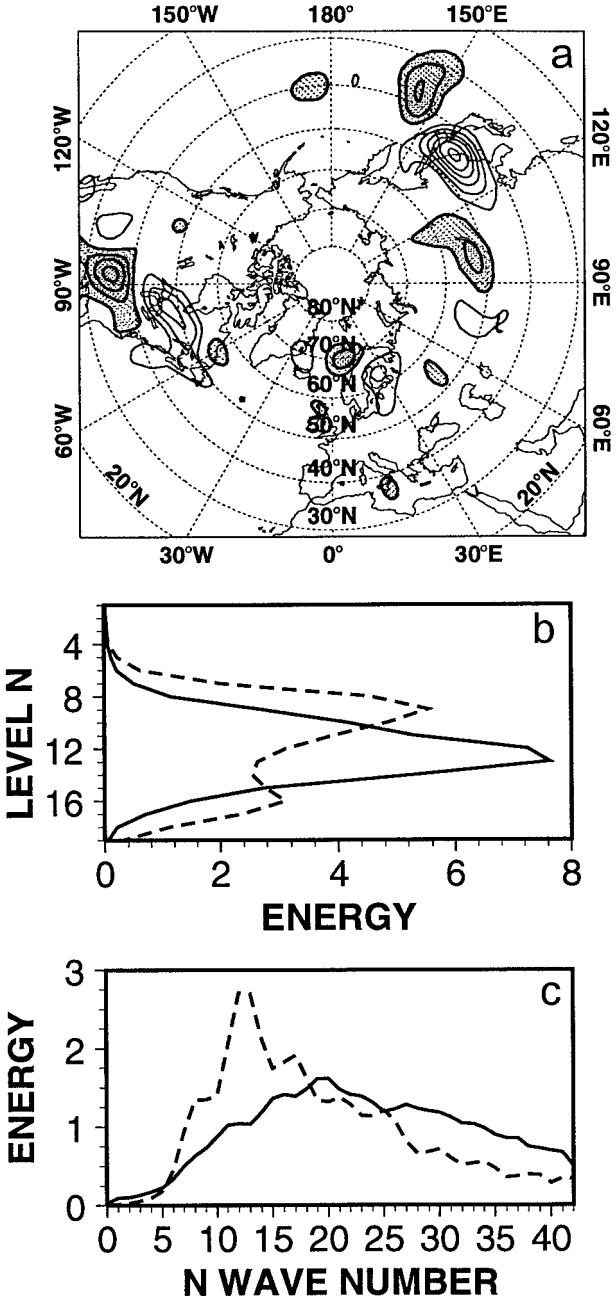

FIG. 18. (a) Scaled temperature perturbation at 500 mb corresponding to $\mathbf{s}'$ for the case 6 April 1994. The contour interval is 0.1°C, with positive values shaded. (See text for details.) (b) Vertical energy profiles and (c) total wavenumber spectra, corresponding to $\tilde{\mathbf{s}}$ (solid), and $\mathbf{s}'$ (dash), for the case 6 April 94. The units are J kg^{-1}.

time of this writing. Having generated the unstable subspace $\mathbf{U}_{0,2}$ or $\mathbf{V}_{-2,0}$, the remaining steps in this procedure included 1) selecting $n = 16$ singular vectors for constructing the initial perturbations; 2) constructing $2n = 32$ perturbations (16 positive–negative pairs) based on scaled, linear combinations of the selected singular vectors; and 3) running 10-day forecasts at T63L19 resolution starting from the 32 perturbed, plus one control, initial states (see Molteni et al. 1996 for details).

Figure 20 shows the spread (top panels) and skill (bottom panels) of the $\mathbf{U}_{0,2}$ (left panels) and $\mathbf{V}_{-2,0}$ (right panels) ensemble forecasts over Europe for the case 6 April 1994 in terms of the ACC for 500-mb geopotential height. In Figs. 20a,b large spread is indicated by small ACC values; that is, a value equal to one indicates no difference between the control forecast and a given perturbed member. In these experiments, the initial perturbations for both ensemble types were given equal amplitude locally, as determined by the operational EPS procedure. Figures 20a,b show that the $\mathbf{V}_{-2,0}$ ensemble has significantly smaller spread than the $\mathbf{U}_{0,2}$ ensemble, especially between days 4 and 8. In fact, most members of the $\mathbf{V}_{-2,0}$ ensemble show relatively little deviation from the control forecast before day+6, with typical ACC values greater than 0.7. This is clearly indicative of the suboptimal growth of the $\mathbf{V}_{-2,0}$ singular vectors over the medium-range forecast trajectory beginning at day+0. Figures 20c,d show for each ensemble type the skill of the individual perturbed members (thin dot), the skill of the ensemble mean (dash), and the skill of the control forecast (solid). The $\mathbf{U}_{0,2}$ ensemble contains many more skillful members than the $\mathbf{V}_{-2,0}$ ensemble and significantly better ensemble mean skill. Note that several members of the $\mathbf{U}_{0,2}$ ensemble show medium-range skill approaching that of the $\mathbf{e}_0$ and $\tilde{\mathbf{s}}$ forecasts for this case (Fig. 7b). In contrast, there is little or no indication that the $\mathbf{V}_{-2,0}$ ensemble contains these alternative solutions.

Figures 21 and 22 show the mean spread and skill, respectively, of the $\mathbf{U}_{0,2}$ (solid) and $\mathbf{V}_{-2,0}$ (dash) ensembles averaged over six cases for the Northern Hemisphere and Europe. The initial perturbations for these experiments were given equal amplitude locally, as described above. It might be argued, however, that poorer performances by the $\mathbf{V}_{-2,0}$ ensembles may result from weak perturbation growth early in the forecast, which can be easily rectified by increasing the perturbation amplitudes initially. To test this, we also show in Figs. 21 and 22 results for an additional set of $\mathbf{V}_{-2,0}$ ensembles in which the initial perturbation amplitudes were increased by a factor of 4, denoted $4\mathbf{V}_{-2,0}$ (dot).

Figure 21 shows a significant difference between the average spread of the $\mathbf{U}_{0,2}$ and $\mathbf{V}_{-2,0}$ ensembles throughout the forecasts. Note that the $\mathbf{V}_{-2,0}$ ensembles actually show larger spread at day+0 since the perturbation amplitudes are constrained only locally and the $\mathbf{V}_{-2,0}$ singular vectors have larger horizontal scale but grow much more slowly up to at least day+4. Increasing the initial perturbation amplitudes has some impact on their growth, as shown by the spread of the $4\mathbf{V}_{-2,0}$ ensembles over both regions. However, note that over Europe the $4\mathbf{V}_{-2,0}$ perturbations have approximately ten times the rms amplitude of the $\mathbf{U}_{0,2}$ perturbations at day+0 but yield only slightly larger spread after day+6.

The results are qualitatively similar in terms of the

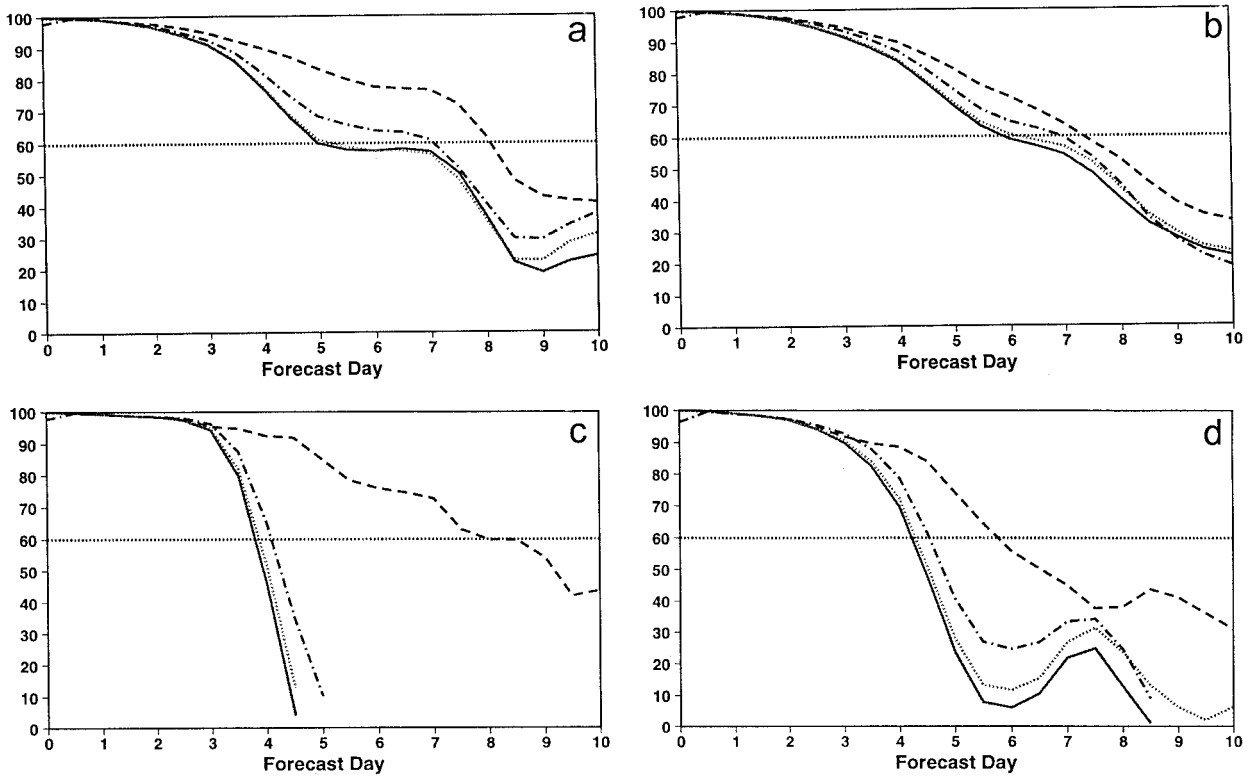

FIG. 19. Anomaly correlation coefficient for 500-mb geopotential height over the Northern Hemisphere (top panels) and Europe (bottom panels), for the single case 6 April 1994 (left panels) and composite for six low-skill cases (right panels); control forecast (solid), and forecasts perturbed by $\tilde{\mathbf{s}}$ (dash), $\mathbf{s}'$ (dot), and $\mathbf{s}'$ with increased initial amplitudes (dash–dot).

skill of the ensemble means, as shown in Fig. 22. The $\mathbf{V}_{-2,0}$ ensembles show lower skill throughout the forecasts, especially beyond about day+4. Again, increasing the initial amplitudes of the $\mathbf{V}_{-2,0}$ perturbations by a factor of 4 improves the skill in the medium range but still has slightly less positive impact than the $\mathbf{U}_{0,2}$ perturbations over most of the forecast range. Of course, it should be kept in mind that the $4\mathbf{V}_{-2,0}$ perturbations have significantly larger amplitude than any estimates (albeit crude) of analysis error obtained from current assimilation schemes (e.g., Buizza 1994).

7. Discussion and conclusions

Using the sensitivity and pseudo-inverse estimates of analysis error, we examined the optimality of perturbations constructed from a limited number of singular vectors in a version of the ECMWF forecast model. Sensitivity and pseudo-inverse perturbations based on the operational 48-h forecast error were computed as explicit linear combinations of singular vectors optimizing total energy over this interval for the Northern Hemisphere. We took the view that these perturbations are close to the optimal perturbation that can be constructed from a linear combination of the selected singular vectors. Optimality was measured primarily in terms of the medium-range forecast improvement in a series of experiments in which these perturbations were added a posteriori to the initial conditions. In light of previous results by R96, a primary objective of this study was to determine whether the sensitivity pattern obtained from an adjoint model integration spanning a large-dimensional (order 10^6) space could be adequately described in terms of a much smaller number (order 10^1) of singular vectors optimized over the same forecast interval.

In principle, we should expect the sensitivity pattern to lie in a limited subspace of the flow since it is a biased estimate of the analysis error weighted toward the most rapidly growing singular vectors. However, an implicit assumption in this is that the analysis error projects with equal probability onto all singular vectors in the unstable subspace. To investigate this, we first examined the projection coefficients for the 30 leading singular vectors in the sensitivity (and pseudo-inverse) expansions. It was shown that these coefficients have roughly equal magnitude on average, suggesting that analysis errors do in fact project with equal probability onto all singular vectors in the unstable subspace. This result supports the use of an energy-based metric as a first-order approximation to the analysis error covariance metric, as suggested by P98, although improved estimates of the true covariance metric are clearly needed. Singular vector calculations using a covariance met-

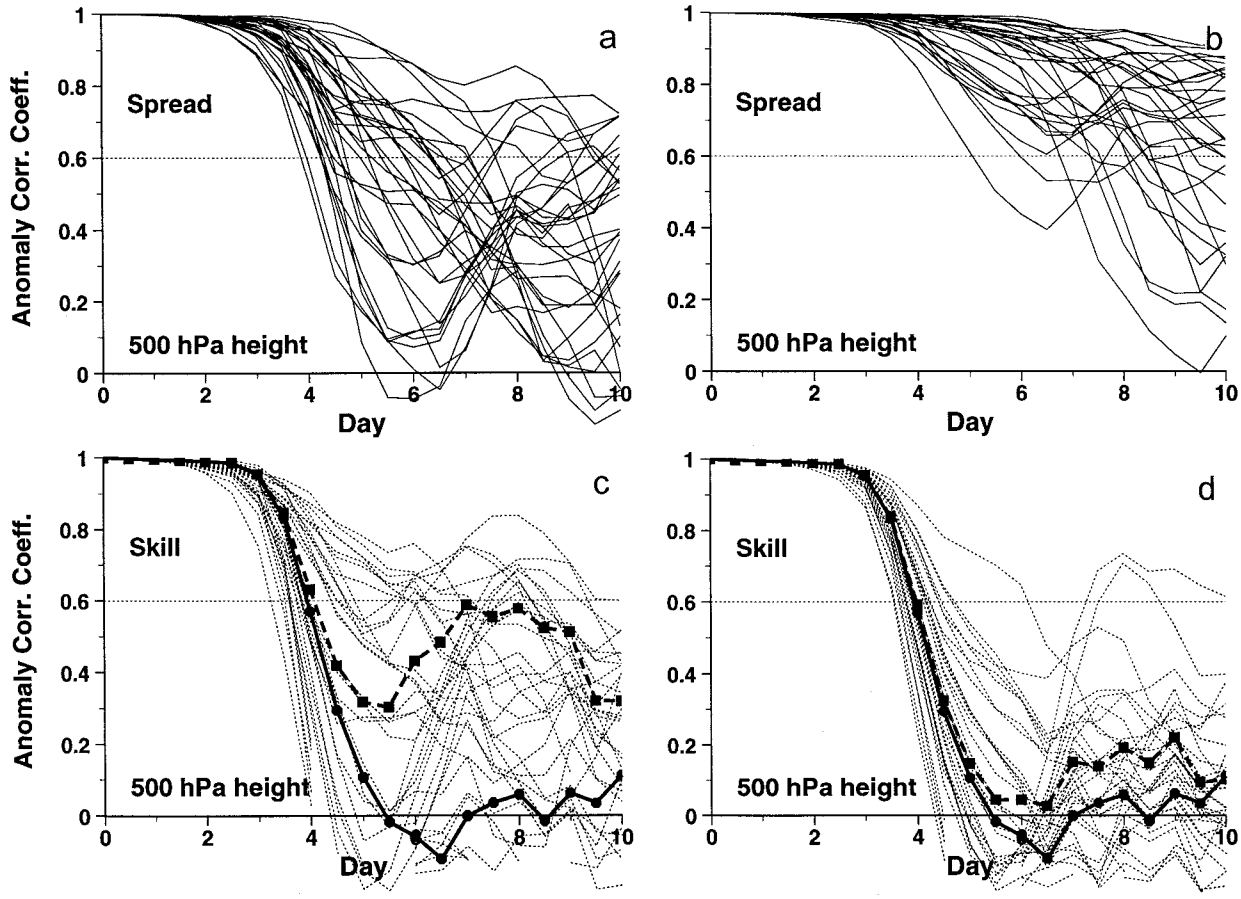

FIG. 20. Spread (top panels) and skill (bottom panels) of the $\mathbf{U}_{0,2}$ (left panels) and $\mathbf{V}_{-2,0}$ (right panels) ensemble forecasts over Europe for the case 6 April 1996 in terms of anomaly correlation coefficient for 500-mb geopotential height. In (a) and (b), perturbed members indicated by thin solid lines. In (c) and (d), perturbed members indicated by thin dot lines, control forecast indicated by heavy solid line, and ensemble mean indicated by heavy dash line.

ric based on the three-dimensional variational data assimilation scheme are being tested at ECMWF.

In most cases examined, it was found that 30 singular vectors at T42 resolution captured a large fraction of the variance of the Northern Hemisphere sensitivity pattern obtained from a T63 adjoint model integration. In cases of low forecast skill (consistent with large initial condition sensitivity), considerably fewer singular vectors were needed, while more singular vectors were needed in less sensitive cases. In low-skill cases, the patterns tended to be quite localized, consistent with a strong error projection onto one or more leading singulars. In less sensitive cases, the patterns were less localized, consistent with a projection onto more slowly growing singular vectors further down the spectrum.

Forecast experiments with the full nonlinear model showed that the introduction of a perturbation based on either the sensitivity or pseudo inverse often produced a significant reduction in forecast errors, especially in low-skill cases. Perturbations based on 8, 16, and 30 singular vectors produced comparable improvements in skill over the Northern Hemisphere, with more varied improvements for specific regions and cases. In many cases, the improvements were comparable with those obtained using the full sensitivity perturbation from the T63 adjoint model, with the most significant improvements obtained well beyond the 48-h optimization interval. For the low-skill case of 6 April 1994, fewer than eight singular vectors accounted for a high percentage of the day+5 forecast error over the Northern Hemisphere. While the results for this case were particularly striking, most cases showed significant improvements in medium-range forecast skill over some part of the Northern Hemisphere using 8–30 singular vectors.

The magnitude of the initial perturbations and their significant positive impact on the forecast support the notion that they primarily reflect the growing components of analysis errors, as opposed to changes in the initial conditions that compensate for model errors. While the latter cannot be ruled out entirely, it was found that the perturbation amplitudes were well within the range of expected analysis errors (a few degrees, or roughly 10 m in geopotential height, at 500 mb). In

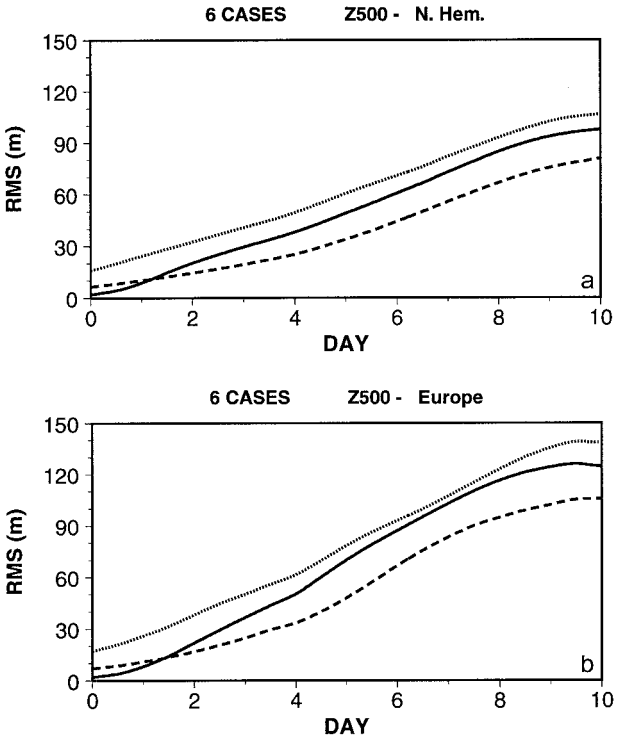

FIG. 21. Ensemble rms spread over (a) the Northern Hemisphere and (b) Europe, in terms of 500-mb geopotential height, averaged over six cases; $\mathbf{U}_{0,2}$ (solid), $\mathbf{V}_{-2,0}$ (dash), and $4\mathbf{V}_{-2,0}$ (dot) ensembles.

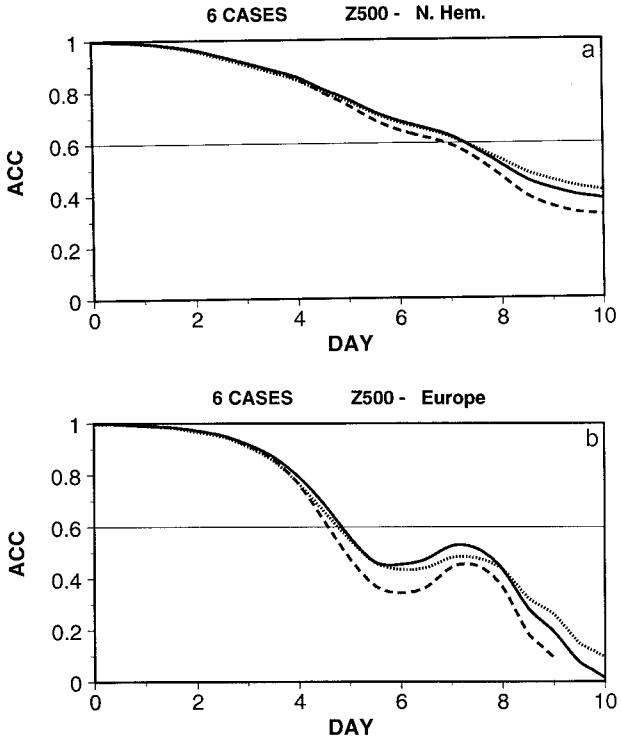

FIG. 22. Ensemble mean skill over (a) the Northern Hemisphere and (b) Europe, in terms of anomaly correlation coefficient for 500-mb geopotential height, averaged over six cases; $\mathbf{U}_{0,2}$ (solid), $\mathbf{V}_{-2,0}$ (dash), and $4\mathbf{V}_{-2,0}$ (dot) ensembles.

addition, had model errors played a significant role, it seems unlikely that these perturbations, which were constructed using relatively coarse-resolution (T42L19), primarily adiabatic linear models, would have performed so well in correcting the high-resolution (T106L31) full nonlinear forecasts.

An important implication of these results is that singular vectors optimized over a short-range forecast trajectory continue to provide a useful description of error growth well beyond this interval, where the current ensemble forecasts are intended to have their maximum value. Moreover, assuming the evolution of the sensitivity pattern provides a reasonable indication of forecast uncertainty arising from errors in the initial conditions, these results suggest that ensemble perturbations based on 10–30 singular vectors (corresponding to ensemble sizes of twice this order to allow for perturbations of both signs) should provide a reasonable description of the Northern Hemisphere forecast PDF in general, although there obviously will be case-to-case variations. Either way, it is likely that increasing the dimension of the unstable subspace will continue to improve the overall performance of the EPS.

Nonmodality is a key feature of singular vector evolution during the early stages of a forecast, during which time there is a significant upscale cascade of energy from subsynoptic to synoptic scales. An important consequence of this nonmodal evolution is that there is little projection between the contemporaneous unstable subspaces at the end of one forecast trajectory portion and the beginning of a second consecutive portion, although the projection increases as the forecast progresses. Perturbations constructed by projecting the sensitivity pattern for day+0 onto the *evolved* singular vectors from 2 days earlier (day−2) had little impact on the forecast initiated at day+0. Similarly, ensemble perturbations constructed from these evolved singular vectors showed significantly less spread, and far fewer skillful members, than those constructed from the initial singular vectors at day+0. Increasing the initial amplitude of the evolved singular vector perturbations improved their performance only moderately, with the largest improvements obtained in the medium range. This indicates that the different perturbations show some degree of convergence as the forecast progresses. However, the initial amplitudes of the perturbations based on the evolved singular vectors significantly exceed current estimates of analysis uncertainty, suggesting that they are likely to overestimate the forecast uncertainty at shorter time ranges as well.

Nonmodality may be an important constraint for newly emerging strategies in atmospheric observing and data assimilation. For example, there is currently great interest in developing so-called adaptive observing strategies, whereby it is proposed that numerical weather

forecasts can be improved by taking additional observations in data-sensitive regions using movable (piloted or preprogrammed) observing platforms. Several strategies have been proposed for identifying these regions, including subjective identification based on the locations of potential vorticity (PV) maxima in the upper troposphere (Snyder 1996), and objective identification based on adjoint sensitivity patterns or singular vectors (Langland and Rohaly 1996; P98), bred modes (Toth and Kalnay 1997), and ensemble transform techniques (Bishop and Toth 1996).

The results shown here indicate that an effective adaptive observing strategy should be based on reducing initial errors that project onto the dominant singular vectors at *initial* time. For example, in the low-skill case of 6 April 1994, the leading singular vector at initial time correctly identified the upstream location where a small change in the initial conditions produced a dramatic improvement in forecast skill both over Europe and the Northern Hemisphere, as confirmed by the forecast impact of the sensitivity perturbation. In principle, this location could have been identified beforehand as a target region for additional observations based on the singular vector calculation alone.

In contrast, a strategy that focuses on identifying maxima in upper-tropospheric PV seems likely to downplay, or even neglect, the importance of small-scale, midtropospheric errors that project onto the initial singular vectors. Instead, this strategy would tend to emphasize structures more consistent with those of the evolved singular vectors, having synoptic length scales and maximum amplitude in the upper troposphere.

Nonmodality may also have important implications for the design of future data assimilation schemes. Returning once more to the case of 6 April 1994, we made additional singular vector calculations to study the projection of the differences between the operational OI and experimental four-dimensional variational (4DVAR) analysis fields for this case (not shown). Preliminary results showed that the impact of 4DVAR on the forecast skill was comparable to that of the sensitivity perturbation projected onto the *evolved* singular vectors of the preceding 24-h forecast (the interval over which the 4DVAR minimization was performed). Again, this impact was significantly less than that of the sensitivity perturbation projected onto the initial singular vectors initiated on 6 April 1994 (or of the full sensitivity perturbation itself). Accordingly, a comparison of these perturbations showed little or no projection of the 4DVAR–OI analysis difference onto the initial singular vectors from 6 April 1994, but a significant projection onto the evolved singular vectors initiated 24 h earlier on 5 April 1994.

These results are not entirely surprising since, as Thépaut et al. (1993) have shown, 4DVAR corrections associated with individual observations can have a substantial projection onto the contemporaneous unstable subspace over which the minimization is performed. One would hope that future analysis systems produce not only an optimal analysis, but an appropriate measure of the uncertainty of that analysis. These results suggest that while 4DVAR analyses may well be a significant improvement over OI in the contemporaneous unstable subspace at the end of the minimization trajectory, the improvement to the analysis in the unstable subspace at the beginning of the forecast trajectory is less certain. Obviously, further research in this area is needed.

Acknowledgments. The authors wish to thank Florence Rabier for her help in preparing the T63 adjoint model integrations. We also thank Jean-Noel Thépaut for providing the 4DVAR analyses discussed in section 7. This work was conducted during the primary author's visit to ECMWF during 1994–95 and was jointly sponsored by ECMWF and the Office of Naval Research (PE 0601153).

REFERENCES

Barkmeijer, J., 1995: Approximating dominant eigenvalues and eigenvectors of the local forecast error covariance matrix. *Tellus,* **47A,** 495–501.

——, P. Houtekamer, and X. Wang, 1993: Validation of a skill prediction method. *Tellus,* **45A,** 424–434.

Bishop, C. H., and Z. Toth, 1996: Using ensembles to identify observations likely to improve forecasts. Preprints, *11th Conf. on Numerical Weather Prediction,* Norfolk, VA, Amer. Meteor. Soc., 72–74.

Borges, M. D., and D. L. Hartmann, 1992: Barotropic instability and optimal perturbations of observed non-zonal flows. *J. Atmos. Sci.,* **49,** 335–354.

Buizza, R., 1994: Localization of optimal perturbations using a projection operator. *Quart. J. Roy. Meteor. Soc.,* **120,** 1647–1681.

——, 1995: Optimal perturbation time evolution and sensitivity of ensemble prediction to perturbation amplitude. *Quart. J. Roy. Meteor. Soc.,* **121,** 1705–1738.

——, 1998: Impact of horizontal diffusion on T21, T42, and T63 singular vectors. *J. Atmos. Sci.,* **55,** 1069–1083.

——, and T. N. Palmer, 1995: The singular vector structure of the atmospheric general circulation. *J. Atmos. Sci.,* **52,** 1434–1456.

——, R. Gelaro, F. Molteni, and T. N. Palmer, 1997: The impact of increased resolution on predictability studies with singular vectors. *Quart. J. Roy. Meteor. Soc.,* **123,** 1007–1033.

Ehrendorfer, M., and R. M. Errico, 1995: Mesoscale predictability and the spectrum of optimal perturbations. *J. Atmos. Sci.,* **52,** 3475–3500.

——, and J. J. Tribbia, 1997: Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.,* **54,** 286–313.

Errico, R. M., and T. Vukićević, 1992: Sensitivity analysis using an adjoint of the PSU–NCAR mesoscale model. *Mon. Wea. Rev.,* **120,** 1644–1660.

——, ——, and K. Raeder, 1993: Examination of the accuracy of a tangent linear model. *Tellus,* **45A,** 462–477.

Farrell, B. F., 1982: The growth of disturbances in a baroclinic flow. *J. Atmos. Sci.,* **39,** 1663–1686.

——, 1988: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.,* **45,** 163–172.

——, 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.,* **46,** 1193–1206.

Houtekamer, P. L., and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.,* **123,** 2181–2196.

Lacarra, J.-F., and O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus,* **40A,** 81–95.

Langland, R. H., and G. D. Rohaly, 1996: Adjoint-based targeting of

observations for FASTEX cyclones. Preprint, *Seventh Conf. Mesoscale Processes*, Reading, United Kingdom, Amer. Meteor. Soc., 369–371.

——, R. L. Elsberry, and R. M. Errico, 1995: Evaluation of physical processes in an idealized extratropical cyclone using adjoint sensitivity. *Quart. J. Roy. Meteor. Soc.,* **121,** 1349–1386.

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

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

Molteni, F., and T. N. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.,* **119,** 269–298.

——, R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.,* **122,** 73–119.

Orr, W. M., 1907: Stability or instability of the steady motions of a perfect liquid. *Proc. Roy. Irish Acad.,* **27,** 9–69.

Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.,* **55,** 633–653.

Penrose, R., 1955: A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.,* **51,** 406–413.

Rabier, F., P. Courtier, and O. Talagrand, 1992: An application of adjoint models to sensitivity analysis. *Beitr. Phys. Atmos.,* **65,** 177–192.

——, E. Klinker, P. Courtier, and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.,* **122,** 121–150.

Simmons, A. J., R. Mureau, and T. Petroliagis, 1995: Error growth and estimates of predictability from the ECMWF forecasting system. *Quart. J. Roy. Meteor. Soc.,* **121,** 1739–1771.

Snyder, C., 1996: Summary of an informal workshop on adaptive observations and FASTEX. *Bull. Amer. Meteor. Soc.,* **77,** 953–961.

Strang, G., 1976: *Linear Algebra and Its Applications.* Academic Press, 374 pp.

Thépaut, J.-N., R. N. Hoffman, and P. Courtier, 1993: Interactions of dynamics and observations in four-dimensional variational assimilation. *Mon. Wea. Rev.,* **121,** 3393–3414.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.,* **125,** 3297–3319.

Tribbia, J. J., and D. P. Baumhefner, 1988: The reliability of improvements in deterministic short-range forecasts in the presence of initial state and modeling deficiencies. *Mon. Wea. Rev.,* **116,** 2276–2288.