Singular Vectors, Metrics, and Adaptive Observations

T. N. PALMER, R. GELARO,* J. BARKMEIJER, AND R. BUIZZA

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

(Manuscript received 3 July 1996, in final form 27 May 1997)

ABSTRACT

Singular vectors of the linearized equations of motion have been used to study the instability properties of the atmosphere–ocean system and its related predictability. A third use of these singular vectors is proposed here: as part of a strategy to target adaptive observations to ''sensitive'' parts of the atmosphere. Such observations could be made using unmanned aircraft, though calculations in this paper are motivated by the upstream component of the Fronts and Atlantic Storm-Track Experiment. Oceanic applications are also discussed. In defining this strategy, it is shown that there is, in principle, no freedom in the choice of inner product or metric for the singular vector calculation. However, the correct metric is dependent on the purpose for making the targeted observations (to study precursor developments or to improve forecast initial conditions). It is argued that for predictability studies, where both the dynamical instability properties of the system and the specification of the operational observing network and associated data assimilation system are important, the appropriate metric will differ from that appropriate to a pure geophysical fluid dynamics (GFD) problem. Based on two different sets of calculations, it is argued that for predictability studies (but not for GFD studies), a first-order approximation to the appropriate metric can be based on perturbation energy. The role of observations in data assimilation procedures (constraining large scales more than small scales) is fundamental in understanding reasons for the requirement for different metrics for the two classes of problems. An index-based tensor approach is used to make explicit the role of the metric.

The strategy for using singular vectors to target adaptive observations is discussed in the context of other possible approaches, specifically, based on breeding vectors, potential vorticity diagnosis, and sensitivity vectors. The basic premises underlying the use of breeding and singular vectors are discussed. A comparison of the growth rates of breeding and singular vectors is made using a T21 quasigeostrophic model.

Singular vectors and subjective potential vorticity (PV) diagnosis are compared for a particular case study. The areas of sensitivity indicated by the two methods only partially agree. Reasons for disagreement hinge around the fact that subjective PV diagnosis emphasizes Lagrangian advection, whereas singular vector analysis emphasizes wave propagation. For the latter, areas of sensitivity may be associated with regions of weak PV gradient, for example, mid to lower troposphere. Amplification of singular vectors propagating from regions of weak PV gradient to regions of strong PV gradient is discussed in terms of pseudomomentum conservation. Evidence is shown that analysis error may be as large in the lower midtroposphere as in the upper troposphere.

1. Introduction

In this paper we consider the dominant singular vectors (or optimal vectors) of the integral tangent propagator, L, for a nonlinear dynamical system A, such as the atmosphere or a coupled ocean–atmosphere system. These vectors provide a precise quantitative means of estimating and characterizing the growth of small perturbations over a finite period of time under very general situations. Singular vectors have been used in two different ways. Following from the work of Orr (1907) in studying the transition to turbulence in Couette flow, these nonmodal entities have been invoked to explain, for example, laboratory fluid behavior (e.g., Trefethen et al. 1993), extratropical cyclogenesis (e.g., Farrell 1982), and even El Niño (Penland and Sardeshmukh 1995). In the geophysical fluid dynamics context, let us refer to these as GFD studies.

However, singular vectors have also been shown to be intimately associated with the notion of predictability, specifically, with the estimation of the evolution of initial errors during the course of a forecast (Lorenz 1965; Farrell 1990). For example, Molteni and Palmer (1993) were able to explain, using singular vectors, the dependence of weather forecast errors on the sign of the Pacific–North American mode. Barkmeijer et al. (1993) have used localized singular vectors to quantify the predictability of short-range weather forecasts over Europe, and singular vectors are used routinely in the construction of the initial perturbations of the European Centre for Medium-Range Weather Forecasts ensemble prediction system (Palmer et al. 1993; Molteni et al. 1996,

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

Corresponding author address: Dr. T. N. Palmer, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, United Kingdom.

hereafter MBPP). A review of the role of singular vectors in predictability of the atmosphere and oceans from days to decades is given in Palmer (1996).

A third potential use of singular vectors is considered in this paper: in defining a strategy for utilizing technology for making adaptive or targeted observations (principally) of the atmosphere. Such observations could be made, for example, from dropsondes released by unmanned aircraft preprogrammed to fly in dynamically sensitive regions. The geographical location of such sensitive regions might vary from day to day, depending on actual and forecast atmospheric circulations. The quantitative location of such regions could be made from suitably defined singular vectors. Such a strategy has been utilized in the FASTEX (Fronts and Atlantic Storm-Track Experiment) campaign (Joly et al. 1997; Snyder 1996). In principle, adaptive observations could revolutionize the methodology for determining the initial conditions for weather prediction by coupling interactively the process of data assimilation and forecast with that of measurement.

Recently, Buizza and Palmer (1995, hereafter BP) have given a detailed account of the calculation of primitive equation atmospheric singular vectors in complete generality, that is, for time-varying basic states optimized for either global or localized energy growth. In BP, the essential nonmodality of singular vector evolution was discussed, with emphasis on the upscale cascade of energy from subcyclone to cyclone scale and the vertical propagation of energy from the baroclinic steering level to the upper-tropospheric jet level. The methods and software used in BP form the basis of calculations in this paper.

In applying the singular vector technique for targeting observations, it is important to know whether the targeted observations are being made to study the precursors of the primary phenomenon (a GFD problem) or to improve the accuracy of initial conditions of a forecast of the primary phenomenon (a predictability problem).

The reason for making this distinction is that the inner product, or metric, with respect to which the singular vectors are computed, will be different depending on whether the calculations are being made for GFD or predictability studies. With respect to the chosen metric, the singular vectors have unit amplitude at initial time and evolve to have some maximal property at optimization time. The dominant singular vector corresponds to the vector with largest amplitude at optimization time, chosen from the set of all unit amplitude vectors at initial time. From a physical point of view, the appropriate metric is one for which each of these unit amplitude vectors is equally likely a priori. An important conclusion of this paper is that a unit vector that is likely from a purely GFD perspective may be very unlikely from a predictability perspective. The reason for this difference is that the predictability problem is constrained both by the observing network and the process of assimilating and analyzing the observations to produce the initial state. The GFD problem is not constrained in this way.

Hence, an important part of this paper is a discussion of the role of the metric in determining the structure of singular vectors. To do this, we use an index-based tensor formalism in which the role of the metric is explicit. This formalism is described in section 2. There is a purpose for not using the more conventional matrix notation. A difficulty of matrix notation is that the same object, the matrix, is used to represent different mathematical concepts, without distinction. For example, a generic matrix $\mathbf{G}$ might represent an operator (such as the tangent model) mapping vectors at initial time to vectors at final time. Alternatively $\mathbf{G}$ might represent products of two vectors (such as the covariance matrix of initial errors). Finally, $\mathbf{G}$ might also represent the inner product or metric mapping a pair of vectors to a scalar number. As discussed in section 4, this lack of distinction in conventional matrix notation can lead to obfuscation concerning metric dependency.

In section 3 examples of the dependence of singular vector structures on four simple choices of metric are discussed explicitly based on calculations made for growth optimized over a limited area relevant to FASTEX.

However, as discussed in section 4, there is, in principle, no arbitrariness in the choice of metric for the application to the adative observation strategy. In section 5 we discuss which of the simple choices of metric defined in section 3 best approximates the predictability and GFD metric.

In addition to the singular vector methodology, two quite different techniques have been proposed for targeting adaptive observations: the breeding method (Toth and Kalnay 1993, 1996) and potential vorticity (PV) diagnosis. An additional purpose of this paper is to compare the singular vector method with these other two techniques. These comparisons are performed in sections 6 and 7, respectively. For completeness, we discuss in section 8 a third proposal for targeting adaptive observations, the gradient sensitivity method. This is, in fact, closely related to the singular vector method. [In fact, at revision time, a further targeting method, based on the inverse tangent operator, was brought to our attention (see Pu et al. 1997, for details).]

In section 9 some specific practical examples for the use of singular vectors as a strategy for targeting potential adaptive observations in the pre-FASTEX period are made. Some remarks on their potential use for tropical cyclone and El Niño prediction are also made. A summary of results is made in section 10.

Throughout most of the paper, it is assumed that the dynamical system A is the midlatitude tropospheric circulation of the atmosphere. However, as mentioned above, in section 9 we consider the straightforward extension of the method to the coupled ocean–atmosphere system.

2. Singular vector formulation

Consider an integration of A that generates a phase-space trajectory from an initial point X_0 to a forecast state X_1. The tangent propagator L^i_j is a linear operator, mapping the tangent space T_{X_0} (i.e., the linear vector space of perturbations at X_0) to the tangent space T_{X_1}. That is to say, if $e^j \in T_{X_0}$, then

$$\hat{e}^i = L^i_j e^j \in T_{X_1}. \tag{1}$$

Now in (1), we have used an index-based tensor notation (e.g., Dodson and Poston 1979) to denote vectors and linear operators. For example, using this notation, we distinguish between a vector e^i with its index ''upstairs,'' and a covector, such as the gradient ∇_i, with its index ''downstairs.'' Note that in (1), and elsewhere, the familiar ''repeated-index'' convention is used, whereby summation over all components is implied when the same index is repeated in a given multiplicative expression.

In practice, for large numerical weather prediction models, the individual components of the tangent propagator are not known explicitly; rather L^i_j is an operator given by a set of FORTRAN statements, which define a linearization of the nonlinear model about a prescribed nonlinear solution (Courtier et al. 1991; BP).

We now define a metric g_{ij}, so that

$$d = g_{ij} e^i f^j \tag{2}$$

denotes the inner product between two arbitrary vectors e^i, f^i.

The inverse (or contravariant) metric g^{ij} can be defined by

$$g^{ij} g_{jk} = \delta^i_k, \tag{3}$$

where δ^i_k is the Kronecker delta. In general, indices will be raised or lowered using these metric forms. Hence, for example, we can define the adjoint of the covariant form L_{ij} by

$$(L^*)_{ij} = L_{ji}. \tag{4}$$

Using (4), the operator form $(L^*)^i_j$ of the adjoint (the adjoint tangent propagator) can be expressed in terms of the metric forms and the forward tangent propagator by

$$(L^*)^i_j = g^{ik} g_{jl} L^l_k. \tag{5}$$

This operator maps the vector $\hat{e}^j \in T_{X1}$ to the vector $(L^*)^i_j \hat{e}^j \in T_{X_0}$. The dependence of the adjoint propagator on the metric is explicit in (5).

Now eigenvectors of the compound operator

$$S^i_j = (L^*)^i_k L^k_j \tag{6}$$

are precisely the singular vectors of L^i_j in T_{X_0}, that is, at initial time. The singular values σ are real and positive, and the singular vectors orthogonal (with respect to g_{ij}) since

$$S_{ij} = S_{ji}, \tag{7}$$

even though for all realistic meteorological problems

$$L_{ij} \neq L_{ji}. \tag{8}$$

The condition (8) has profound consequences for singular vector evolution as many authors have noted (following Orr 1907). One of the most important is that their evolution is not shape preserving, that is, nonmodal. As discussed in BP, spectra of midlatitude singular vectors at initial and final time can show an upscale cascade and quantify what Lorenz (1993) refers to as the ''butterfly effect.''

In the singular vector calculation, the (primary) dynamical propagator can be multiplied with a secondary (filter) operator K^i_k to focus on specific aspects of perturbation growth. For example, Barkmeijer et al. (1993) and BP have studied singular vectors optimized for localized areas. Hartmann et al. (1995) have studied singular vectors optimized for different parts of the wavenumber spectrum. In these examples, the operator K^i_k is a projection operator defined either in geographical or spectral space (equal to unity inside the targeted area, equal to zero outside). In the examples discussed here, we optimize total energy over the geographical domain D (30°–80°N, 30°W–10°E).

[It is also possible to define a secondary projection operator $\tilde{K}^l_j$ at initial time. For example, consider an aircraft based at point p, with total range (out and return) of $2r$. Let $\tilde{K}^l_j$ be a projection operator equal to 1 for points inside the circle centered on p with radius r, and equal to zero for points elsewhere. Then with K^i_k as above, the initial singular vectors of $K^i_k L^k_l \tilde{K}^l_j$ determine the structures within the range of the aircraft that are optimized over D at final time.]

The user is free to choose whatever secondary operator is felt to be most appropriate to the problem at hand. We have chosen to focus on the domain D because of the FASTEX experiment. Moreover, we have chosen to optimize energy (as opposed, e.g., to enstrophy) because winds and temperatures (rather than vorticity) are the primary variables that are used to validate weather forecasts, and hence measure their errors. However, it is important to stress that despite this there is, in principle, no choice in the metric g_{ij}. This is the subject of discussion in sections 4–5.

In relation to its use for gradient sensitivity (e.g., Cacuci 1981; Rabier et al. 1996; Langland and Rohaly 1996), if we choose a scalar J, then the gradient $\nabla_i J \in T^*_{X_1}$ is transformed by the adjoint propagator and the metric to the sensitivity vector

$$s^i = g^{ij} (L^*)^k_j \nabla_k J \in T_{X_0}. \tag{9}$$

Note, from (5) and (9), that the sensitivity vector depends on the metric and on the scalar J. This is entirely analogous to the dependence of the singular vectors on the metric and the secondary operator (respectively). As with the secondary operator in the singular vector calculation, the choice of scalar J is arbitrary (e.g., it can be based on the variance of any chosen field; see section

8). However, we argue in section 4 below that, in principle, there is no freedom in the choice of metric.

3. Dependence of singular vector structure on choice of metric

As mentioned in the introduction, singular vectors have been calculated in a variety of contexts. In addition, a number of metrics have been used in the literature, for example, based on enstrophy,

$$g_{ij}^{(\mathrm{ENS})} e^i f^j = \frac{1}{2} \int \zeta_{(e)}\, \zeta_{(f)}\, dv; \tag{10}$$

total energy,

$$g_{ij}^{(\mathrm{ENE})} \mathbf{e}^i f^j = \frac{1}{2} \int \left(\boldsymbol{\nabla}\Delta^{-1}\zeta_{(e)} \cdot \boldsymbol{\nabla}\Delta^{-1}\zeta_{(f)} + \boldsymbol{\nabla}\Delta^{-1}D_{(e)} \cdot \boldsymbol{\nabla}\Delta^{-1}D_{(f)} + RT_r \ln\pi_{(e)} \ln\pi_{(f)} + \frac{C_p}{T_r} T_{(e)}\, T_{(f)} \right) dv; \tag{11}$$

kinetic energy,

$$g_{ij}^{(\mathrm{KE})} e^i f^j = \frac{1}{2} \int \boldsymbol{\nabla}\Delta^{-1}\zeta_{(e)} \cdot \boldsymbol{\nabla}\Delta^{-1}\zeta_{(f)}\, dv; \tag{12}$$

or streamfunction variance,

$$g_{ij}^{(\mathrm{STR})} e^i f^j = \frac{1}{2} \int \Delta^{-1}\zeta_{(e)}\, \Delta^{-1}\zeta_{(f)}\, dv. \tag{13}$$

In (10)–(13), $\xi_{(e)}$, $D_{(e)}$, $\pi_{(e)}$, and $T_{(e)}$ represent, respectively, the vorticity, divergence, (log) surface pressure, and temperature components of the vector e^i (similarly for f^i), and the integration is taken over the whole volume of the atmosphere; T_r is a reference temperature. The ''dot product'' in (10)–(13) denotes the Euclidean inner product. Of course, the integral denotes a finite sum over all the (horizontal and vertical) components explicitly represented in the model.

To illustrate the impact of the choice of metric on the singular vector structure, singular vector calculations have been made using the ECMWF T42L19 tangent model based on a 48-h total energy optimization within the target area *D*. The initial time is 1200 UTC 5 December 1994. The circulation during this period was fairly zonal across the Atlantic and not atypical of the sort of conditions that were studied during FASTEX.

Figure 1 shows the streamfunction in the midtroposphere (about 500 hPa) of the dominant singular vector at initial time using a metric based on (i) total energy, (ii) kinetic energy, (iii) enstrophy, and (iv) streamfunction variance. Figure 2 shows the same field at optimization time. In all cases, the contour interval at optimization time is 20 times larger than at initial time. For the total energy singular vector (Fig. 1a), the initial amplitude is fairly well localized over the west Atlantic near the North American seaboard. Restricting the metric to kinetic energy (Fig. 1b) has little impact on the structure of the dominant singular vector in terms of the midtroposphere streamfunction, though growth rates are weaker. At initial time, the streamfunction of the kinetic energy metric is larger because the unit kinetic energy norm is defined only in terms of the wind perturbations. The growth of the kinetic energy singular vector is weaker than the growth of the total energy singular vector because the initial temperature perturbations associated with the total energy metric play an important role in growth (see BP). However, the singular values of the dominant kinetic and total energy singular vectors, as measured in terms of their own metric, differ only by a factor of $\sim$1.5; smaller than would be inferred from Figs. 1 and 2 alone.

By contrast to results from either energy metric, the structure of the dominant enstrophy singular vector (Fig. 1c) is fundamentally different, being dominated by planetary-scale perturbations. The structure of the dominant streamfunction singular vector (Fig. 1d), on the other hand, has particularly small-scale structure. Notice that, notwithstanding differences of scale, the maxima of either enstrophy or streamfunction singular vectors do not coincide with those of the energy singular vectors. Hence, if the maxima of the singular vectors were used as a criterion for targeting adaptive observations (cf. section 9), the target regions would be metric dependent.

The final singular vectors show, by construction, amplification over the target domain *D*. The principal maximum of the total energy singular vector (Fig. 2a) is positioned close to the Greenwich meridian. Note that energy has propagated from the initial position at about 25–30° per day, consistent with a typical Rossby wave group speed. For kinetic energy (Fig. 2b) the final singular vector is very similar to that in Fig. 2a. Despite its very different scale at initial time, the predominant scale of the enstrophy singular vector (Fig. 2c) is similar to that of the energy singular vectors. There are differences, however. For example, the center over the eastern Atlantic is more prominent for the enstrophy singular vector, and, compared with the energy singular vectors, there is no high-latitude component. The streamfunction singular vector at optimization time is similar to the energy metric structures.

In Fig. 3 the average energy spectrum of the first 16 singular vectors, computed with each of the energy, enstrophy, and streamfunction metrics, is shown with respect to total wavenumber. The dashed lines show the spectrum at initial time, the solid lines show the spectrum at optimization time. (From here on, kinetic energy singular vectors will not be shown explicitly as they are qualitatively similar, in terms of the diagnostics displayed in this paper, to those of total energy.) The initial spectra have been scaled by a variable factor to make them visible in the figure.

The spectrum of the energy singular vectors (Fig. 3a) is fairly broad at the higher-wavenumber half of the spectrum. The spectrum evolves with an inverse energy

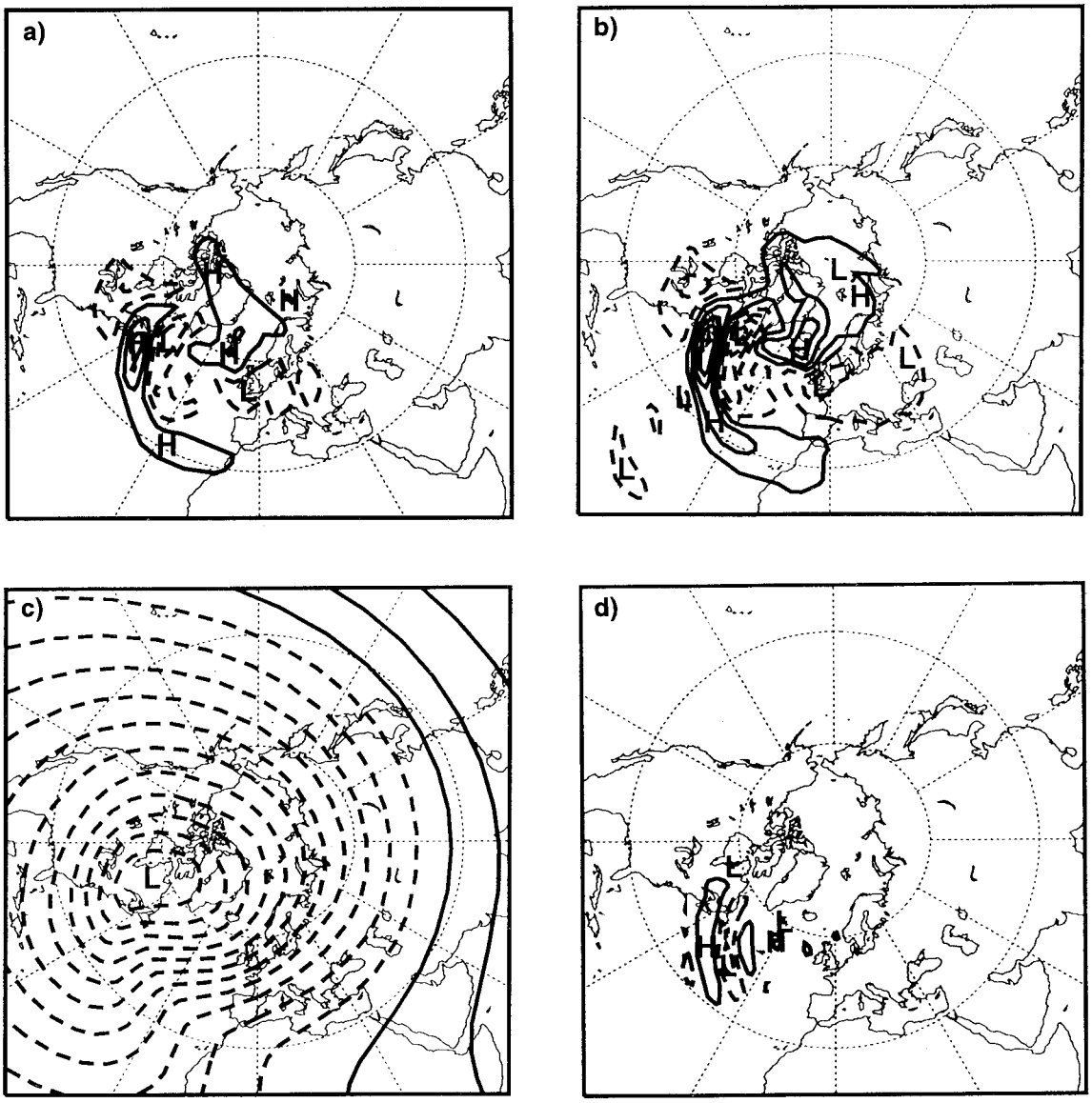

FIG. 1. Streamfunction of dominant 48-h singular vector at initial time on a model level near 500 hPa optimized for energy growth in the target area 30°–80°N, 30°W–10°E. Initial time 1200 UTC 5 December 1994. Metric used: (a) total energy, (b) kinetic energy, (c) enstrophy, (d) streamfunction variance. See Fig. 2 caption for information about contour intervals.

cascade peaking at cyclone scales. By contrast, the spectrum of the enstrophy-metric singular vectors peaks at low wavenumbers and evolves by a energy cascade to cyclone scale and to higher wavenumbers (Fig. 3b). Note that in Fig. 3b there is no scaling between initial and final values; hence, using the enstrophy metric, the singular vectors show virtually no energy growth, even though energy was being optimized at final time. This illustrates what a strong constraint the metric has on growth rates. Finally, the spectrum of the streamfunction-variance singular vectors (Fig. 3c) is highly peaked toward very small scales (indeed it is peaked at the limit of truncation of the tangent propagator). As with the energy metric, there is an inverse cascade to larger scales. However, the scaling factor between initial and final time is an order of magnitude smaller than using the energy metric.

Figure 4 shows the vertical distribution of energy as a function of model level at initial (dashed) and final (solid) time, for the three choices of metric (and with scaling parameters between initial and final time as in Fig. 3). For reference, level 7 corresponds to about 200 hPa, level 13 to about 700 hPa. It can be seen that for all choices, there is a tendency for an energy maximum in the mid- to lower troposphere at initial time, with a maximum in the upper troposphere at final time. As discussed in BP, this is consistent with what would be expected from pseudomomentum conservation of wave

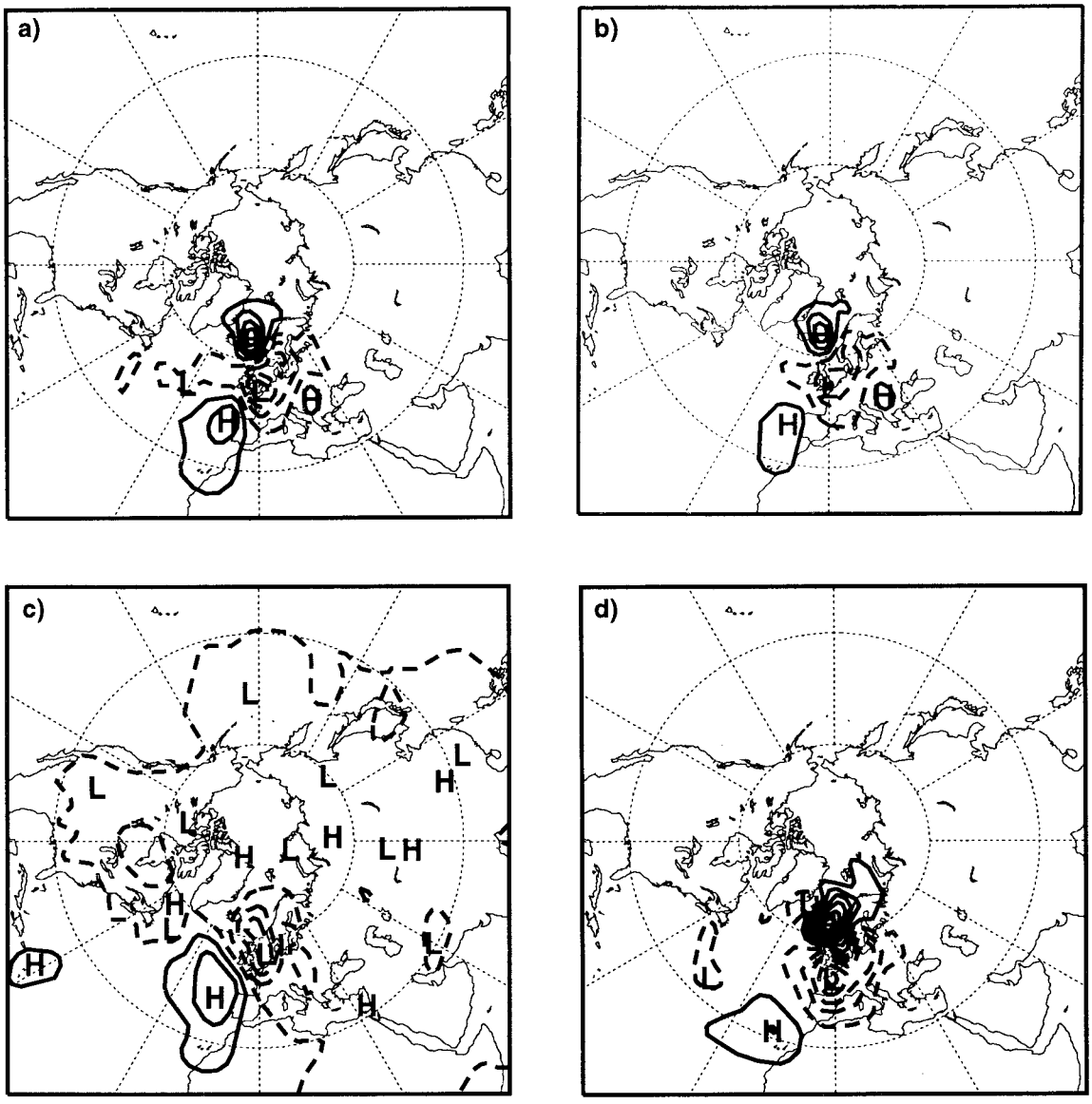

FIG. 2. As in Fig. 1 but at optimization time. For each of (a)–(d), the contour interval is 20 times larger than in the corresponding component of Fig. 1.

activity. We defer further discussion of this issue to section 7.

4. The covariance metric

We now introduce a probability function, or measure, μ. Specifically, $\mu(V)$ gives an a priori probability to a set $V \subset T_{X_0}$ of perturbations. In terms of μ we can address the question: Are the perturbations V likely to occur?

For the predictability problem, it is straightforward to define the appropriate measure μ_{pre}; it gives the probability that the actual initial error (considered as an element of T_{X_0}) is contained within V.

By contrast, given a perturbed set of phase-space points, generated by V and the ''basic state'' point p in phase space, the GFD measure $\mu_{\text{GFD}}(V)$ can be thought of as determining the probability that A's state vector lies in the set of perturbed points. For example, if all the perturbed points lie off A's attractor, then $\mu_{\text{GFD}}(V) = 0$. From a dynamical systems perspective, μ_{GFD} is determined from the invariant measure associated with the system's attractor in phase space (e.g., Frisch 1995).

In general, while μ_{GFD} is only a function of the dynamics of A, μ_{pre} is, in addition, a function of the observing network and the data assimilation system used to create the initial state X_0. For example, from an empirical point of view, a very ''approximate'' attractor A_{dom} of the global atmospheric circulation could be defined from the dominant empirical orthogonal functions

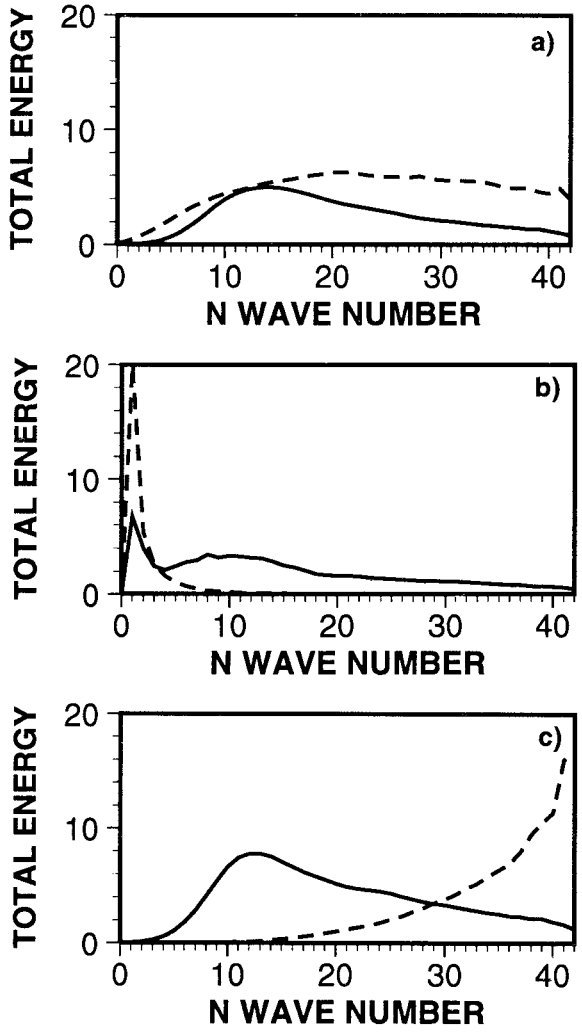

FIG. 3. Energy spectrum of the first 16 singular vectors (as a function of 2D total wavenumber) computed using optimization and initial conditions as in Figs. 1 and 2: (a) total energy, (b) enstrophy, (c) streamfunction variance. Dashed line shows spectrum at initial time, solid line shows spectrum at optimization time. The spectra at final time have been reduced by the following factors (so that initial and final spectra can be plotted together): (a) 100, (b) 1, (c) 10.

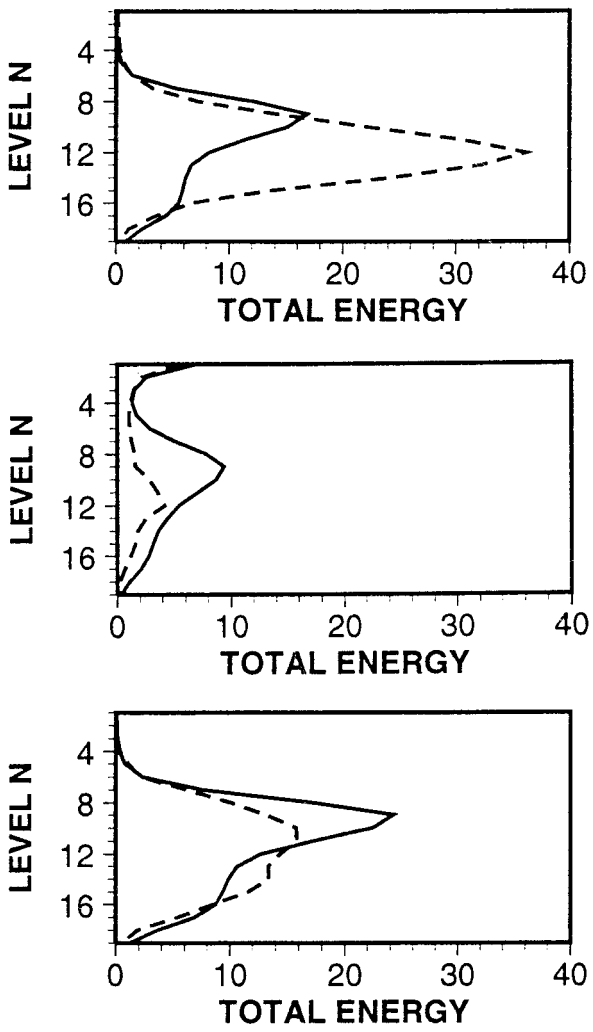

FIG. 4. The vertical distribution of energy averaged over the first 16 singular vectors (as in Fig. 3) at initial (dashed) and optimization (solid) time. The distributions at final time have been reduced by the following factors (so that initial and final distributions can be plotted together): (a) 100, (b) 1, (c) 10.

(EOFs) of the global flow; from this empirical point of view, let $\mu_{\mathrm{GFD}}(V) \approx 0$ if $V \not\subset A_{\mathrm{dom}}$. However, there is no requirement that $\mu_{\mathrm{pre}}(V) \approx 0$ for $V \not\subset A_{\mathrm{dom}}$. To give an extreme example, suppose accurate observations of the atmosphere could be made (e.g., by advanced satellite soundings) on scales of 10 km and greater so that the source of cyclone-scale forecast inaccuracy would be associated with an upscale cascade of mesoscale initial error. We can suppose that atmospheric EOFs with scales of 10 km or less do not contribute to A_{dom}. Hence if V denotes the set of perturbations with scales of 10 km or less, then $\mu_{\mathrm{pre}}(V) > 0$, $\mu_{\mathrm{GFD}}(V) \approx 0$. In general, $\mu_{\mathrm{pre}} \neq \mu_{\mathrm{GFD}}$, as will be confirmed in section 5.

Now, based on a given μ, we can define the first and second moments

$$m^i = \int e^i \, d\mu, \tag{14}$$

$$C^{ij} = \int (e^i - m^i)(e^j - m^j) \, d\mu. \tag{15}$$

The integrand in (15) comprises the tensor product of pairs of vectors. Hence, from (1), C^{ij} is mapped to the forecast error covariance tensor:

$$\hat{C}^{kl} = L^k_i L^l_j C^{ij}. \tag{16}$$

Note from (16) that $\hat{C}^{kl}$ is not itself metric dependent. This is not apparent from the matrix form $\hat{\mathbf{C}} = \mathbf{LCL}^*$ and that the adjoint is metric dependent. As mentioned in the introduction, the matrix form does not distinguish the case when C denotes a covariance tensor, from the

cases when C denotes either a vector operator, or an inner product. In the latter two situations, the equivalent transform is metric dependent.

Let us make the specific choice of contravariant metric

$$g^{ij} = C^{ij} \tag{17}$$

so that, from (5), (6), and (16),

$$\hat{C}^{ij} = S^{ij}. \tag{18}$$

Hence with the specific choice of contravariant metric (17), eigenvectors of $\hat{C}^i_j$ are precisely the evolved singular vectors of L^i_j. The vectors in T_{x_0}, which evolve into these directions are the initial singular vectors of L^i_j.

Note, from (2) and (3), that the (covariant) form of the metric used to define the inner product between pairs of vectors is the inverse of the contravariant form; that is,

$$g_{ij} = C^{-1}_{ij}. \tag{19}$$

In other contexts, the metric (19) is known as the Mahalanobis metric (e.g., Mardia et al. 1979; Stephenson 1997); here we refer to the choice (19) as the covariance metric. For the predictability problem, the metric (19) will be referred to as the analysis error covariance metric (hereafter AECM). For GFD problems, the metric (19) is referred to as the GFD covariance metric (hereafter GFDCM).

The relationship between singular vectors and eigenvectors of the forecast error covariance tensor has also been discussed (in different forms) by Houtekamer (1995) and Ehrendorfer and Tribbia (1997).

It should be noted that, just as the AECM is the appropriate metric with which to define singular vectors for predictability studies, so also is it appropriate to define vectors that give the sensitivity of the forecast error to initial perturbations. From (9) and (17), the AECM sensitivity vector is given by

$$\mathbf{s}^i = C^{ij}(L^*)^k_j \nabla_k J. \tag{20}$$

We can interpret (20) as follows. Let J represent, for example, an rms forecast error, then $\mathbf{s}^i$ determines the sensitivity in forecast error with respect to perturbations drawn with respect to the probability distribution (associated with μ_{pre}) of initial analysis error.

Finally, it can be commented that if the secondary operator K^i_j (see section 2) satisfies $C^i_j = K^i_k K^k_j$, then it is straightforward to show that the singular vectors of $K^i_k L^k_j$ with respect to the AECM will be normalized at initial time with respect to the AECM, but optimized at final time with respect to the Euclidean metric. With respect to the Euclidean metric, these singular vectors satisfy the generalized eigenvector equation $\mathbf{S}^i_j x^j = \lambda C^{-1i}_j x^j$.

5. Relationship between the covariance and other metrics

Overall results in section 3 show that there is a strong dependence of singular vector structure on the choice of metric. We now address the question as to which of the explicitly estimated singular vectors gives the best approximation to the appropriate (AECM or GFDCM) covariance metric.

a. Evidence from analysis fields

For atmospheric predictability, the covariances of the initial or analysis error are not well known and are not readily obtained from operational data assimilation schemes. However MBPP made an estimate of an analysis error field associated with one rather poor ECMWF forecast by taking the difference between the ECMWF analysis and a corresponding Deutsche Wetterdienst (DWD) analysis from which a hindcast (using the same ECMWF model) was much more successful. In this case one can assume that the difference between the ECMWF and DWD analyses was a fair estimate of the ECMWF analysis error.

In this paper, we extend this calculation by estimating the spectrum of a set of 42 difference fields produced from two sets of analyses: the operational ECMWF analysis and an experimental analysis system that was being run in parallel at ECMWF at the time. It can be easily shown that, if the two sets of analyses were each unbiased estimates of truth with uncorrelated error, then the variance of analysis difference fields would have the same spectral shape as the variance of analysis error fields. In practice it is unlikely that this condition is satisfied for these two analysis estimates, though the extent to which the condition is violated is hard to assess. (For example, it is possible that both sets of analyses have correlated deficiencies in their planetary-scale components.) For all cases in this set, the operational analysis system is based on the optimum interpolation (OI) scheme. For half of this set, the experimental analyses were produced using the three-dimensional variational data assimilation system (3DVAR; Courtier et al. 1993). For the other half, the OI scheme was used, though with a significant revision to the model formulation. The 42 cases were all taken from the northern winter.

Figure 5 shows the spectrum of the total energy, enstrophy, and streamfunction squared averaged over all 42 difference fields. It can be seen that the energy spectrum is uniformly weighted over all wavenumbers. By contrast, that the enstrophy spectrum is strongly weighted toward high wavenumbers, whereas the streamfunction spectrum is strongly weighted toward low wavenumbers. These results are consistent with MBPP's single case study.

Now, by construction, the dominant AECM singular vectors must be contained in the subspace of eigenvec-

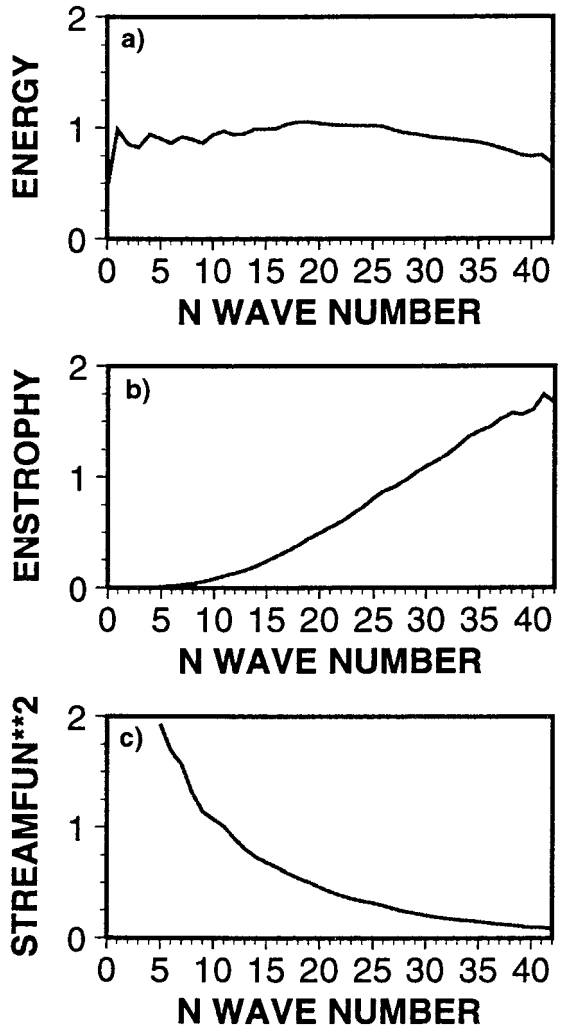

FIG. 5. Spectrum of the variance of a set of 42 analysis difference fields as a function of total wavenumber based on (a) total energy, (b) enstrophy, (c) streamfunction variance. Each field is the difference between two quasi-independent analyses for the same date. The curves have each been scaled by a suitable factor to make them mutually comparable.

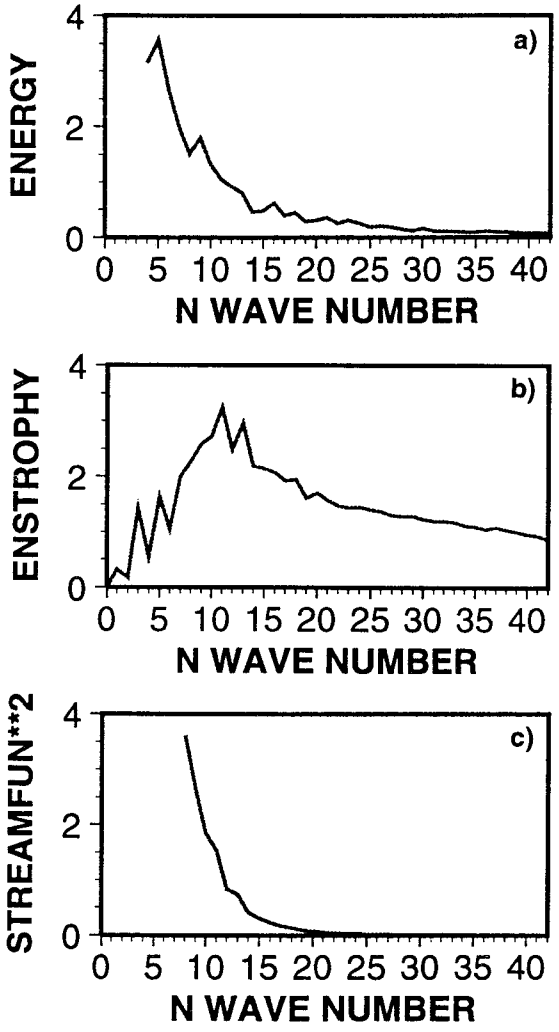

FIG. 6. As in Fig. 5 but for the full analysis fields.

tors of the analysis error covariance matrix. Hence, for a trial metric to be considered an approximation to the AECM, the spectra of the associated dominant singular vectors must not be inconsistent with the spectra of estimates of analysis error variance. Inconsistency is meant in the following sense: spectra of singular vectors are dominant in a wavenumber band for which the analysis error variance is relatively small. This inconsistency condition is clearly not sufficient to define the correct metric, since we do not address here questions relating to the horizontal and vertical correlation of errors. Nevertheless, the inconsistency condition is sufficient to distinguish between enstrophy, energy, and streamfunction as choices of metric.

If energy, enstrophy, and streamfunction variance are considered as candidate AECMs, only the energy metric satisfies our consistency condition. For example, as shown in Fig. 3b, the initial structure of an enstrophy-based singular vector is essentially planetary scale, whereas Fig. 5b shows that the enstrophy of the analysis difference fields is relatively negligible on planetary scales. Similarly, the initial structure of a streamfunction-based singular vector is essentially subcyclone scale (Fig. 3c), whereas Fig. 5c shows that the small-scale streamfunction variance of the analysis difference fields is relatively negligible. For energy, in the range of wavenumbers in which the singular vectors have significant power, the analysis difference fields also have power (cf. Figs. 3a and 5a).

It is important to contrast the spectra of the difference fields with the spectra of the full analysis fields, shown in Fig. 6. The full-field spectra are fundamentally different from those of the difference fields. In particular, both the energy and streamfunction spectra are peaked at low wavenumbers, whereas the enstrophy peaks at cyclone scales with a relatively weak decay to high

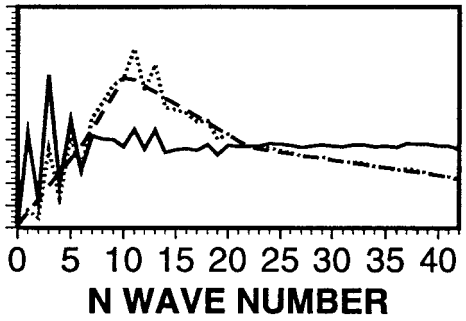

FIG. 7. Enstrophy spectrum from full analysis fields (dotted). Piecewise linear fit to enstrophy spectrum (dashed). Modified-enstrophy spectrum (solid) based on the amplitude of the vorticity fields using a modified-enstrophy metric defined from the dashed line.

wavenumbers. The spectra of the full fields reflect the nonlinear structure of the atmospheric global circulation, and, in particular, the well-known quasigeostrophic enstrophy-cascading subrange.

That the energy spectrum of the analysis difference field is flatter than the energy spectrum of the analysis itself is entirely consistent with the process by which observations are utilized in a data assimilation scheme. As, for example, Daley (1985, 1991) and Hollingsworth (1987) have emphasized, observations necessarily constrain the low-wavenumber components of the analysis more strongly than the high wavenumber components. Hence, from a phase-space perspective, not only will (good) observations ensure that the analysis error has a smaller magnitude than the background error, they will cause the analysis error to rotate away from the background error toward smaller scales.

Whereas for the predictability problem, an energy-based metric appears to be the most suitable of those studied explicitly, it is clear from Fig. 6 that it is not a consistent choice for GFD problems. In fact, none of the simple choices of metric satisfy the consistency condition mentioned above. For example, as mentioned above, the dominant enstrophy singular vectors peak at low wavenumbers, where the analysis enstrophy spectrum shows relatively small amplitude. By contrast the streamfunction and energy singular vectors have most amplitude at subcyclone wavenumbers where the analysis amplitudes are relatively small.

However, one can define empirically a metric for which the singular vector spectra are not inconsistent (in the sense defined above) with the analysis variance spectra. In Fig. 7, the dotted line shows the analysis enstrophy spectrum as in Fig. 6. A piecewise linear function (shown as the dashed line) is fitted to this spectrum and is used to define the modified enstrophy metric $g_{ij}^{(\mathrm{MOD\ ENS})}$. The solid line in Fig. 7 shows the spectrum of enstrophy defined with respect to this metric. To a first approximation, it is reasonably uniform across the wavenumber range (at least for cyclone-scale and higher wavenumbers).

Figure 8 shows the dominant singular vector using $g_{ij}^{(\mathrm{MOD\ ENS})}$ based, as above, for 48-h optimization in domain D at initial (5 December 1994) and final time. Compared with the enstrophy singular vectors, the scale of the modified enstrophy singular vectors is noticeably smaller (though not as small as the initial energy or streamfunction singular vectors). This is confirmed by analyzing wavenumber spectra of the modified enstrophy singular vectors (not shown). At final time the singular vector structure is not dissimilar to those shown in Fig. 2. The metric $g_{ij}^{(\mathrm{MOD\ ENS})}$ is a crude empirical approximation to the GFDCM, having none of the correlation (or teleconnection) structures associated with

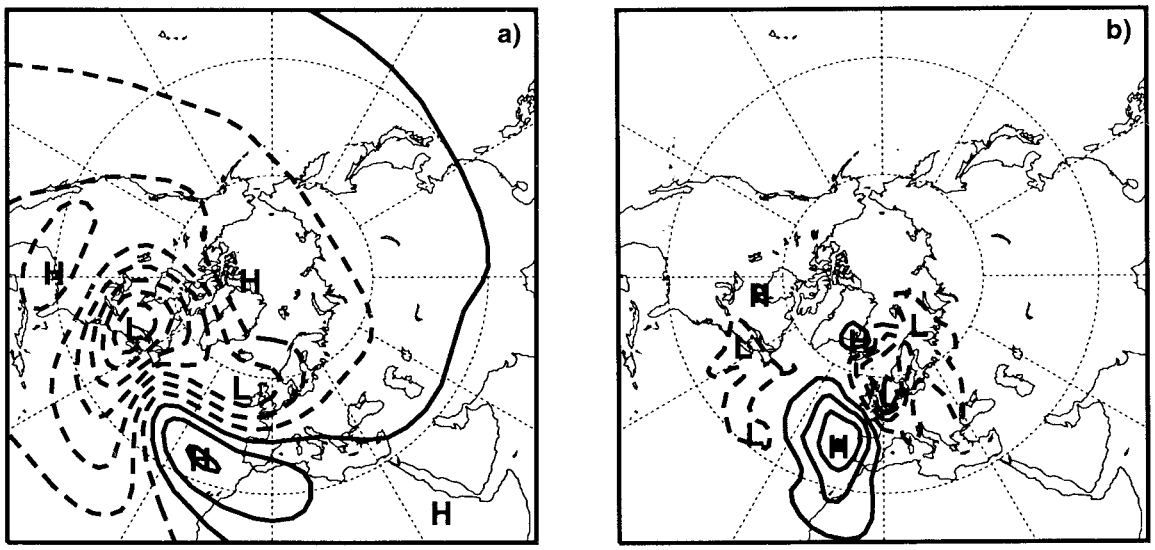

FIG. 8. (a) Initial singular vector, as in Fig. 1c but using the modified enstrophy metric. (b) Singular vector at optimization time, as in Fig. 2c but using the modified enstrophy metric. The contour interval at final time is 20 times larger than at initial time.

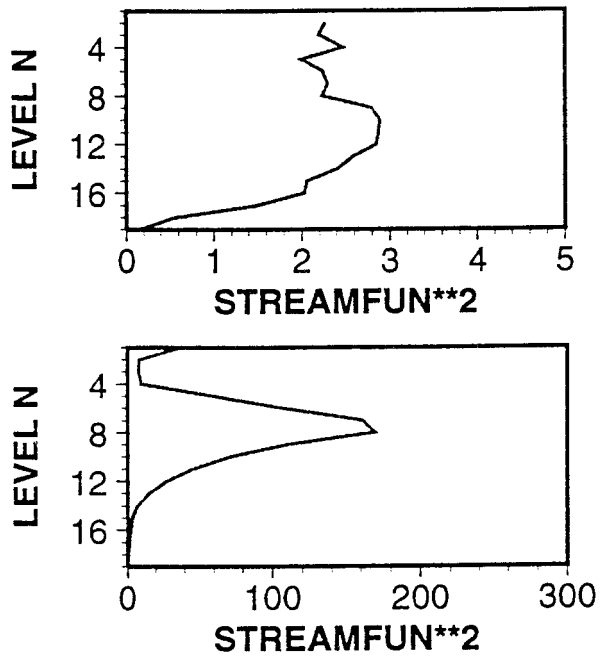

FIG. 9. Streamfunction variance of (a) the analysis difference fields and (b) the full analysis fields, as a function of model level.

the real atmosphere, nor indeed any vertical weighting. Note that the singular vector in Fig. 8 is dependent, through the empirical construct $g_{ij}^{(\mathrm{MOD\ ENS})}$, on aspects of the nonlinear (quasi-turbulent) dynamics of the atmosphere, which determine the spectral characteristics of the atmosphere.

The vertical structure of the analyses difference fields and the full analysis fields are shown in Fig. 9 in terms of streamfunction variance. Once again, note that the structure of the difference fields is quite different from that of the full fields themselves; the full fields peak strongly in the upper troposphere while the difference fields are more uniformly distributed in the vertical. One possible explanation is that analysis differences in those circulation components generated by diabatic processes are comparable with errors associated with adiabatic advection. Since such diabatic processes influence the mid- and lower troposphere more than the upper troposphere, this would tend to give a more uniform vertical distribution to the difference field. In addition, radiance observations from satellites and wind observations from aircraft (which outnumber conventional radiosonde data) have a greater impact on the analyses in the upper troposphere.

b. Evidence from forecast error fields

The conclusion from section 5a was that the only one of our candidate metrics to satisfy the consistency condition (for the analysis error variance) was energy. To give further support to this conclusion, results are shown in this section of the singular vector spectrum of forecast error, based on earlier calculations described in Barkmeijer et al. (1993). The purpose of this is as follows: if we have chosen a good approximation to the AECM, then analysis errors will, on average, project uniformly onto all the initial singular vectors (i.e., the analysis error probability distribution will be isotropic in phase space). If this is true, then the projection of the forecast error onto the evolved singular vectors will be in proportion to the singular value itself.

To demonstrate this, 607 different 48-h operational ECMWF forecasts of 500-hPa streamfunction for a chosen European region were projected onto 23 singular vectors computed from a T21L3 QG model (Marshall and Molteni 1993) using a kinetic energy metric and a secondary projection operator determined by the European region above. Results are shown in Fig. 10. It can be seen that there is excellent agreement between the magnitude of the projection of the forecast error onto the *i*th singular vector and the *i*th singular value itself. In this case all of the variance of the forecast error can be explained by the projection onto these 23 singular vectors, about half of which are growing. This result supports the notion that the probability distribution of analysis error is approximately isotropic with respect to an energy-based metric. Similar results have been obtained for the hemispheric domain (Gelaro et al. 1998; C. Reynolds 1996, personal communication).

c. The Hessian metric

Estimates of the AECM are possible in practice from variational data assimilation schemes (Courtier et al. 1993). Specifically, the Hessian of the variational analysis cost function J (the sum of background and observation cost functions) can be shown to be equal to the inverse of the analysis error covariance matrix (Fisher and Courtier 1995). Hence, the Hessian naturally defines a covariant AECM metric

$$g_{ij}^{(\mathrm{VAR})} = \nabla_i \nabla_j J. \tag{21}$$

With suitable secondary operator, the singular vectors are found by solving a generalized eigenvector equation (see section 4) using a Jacobi–Davidson iterative solver (Sleijpen and van der Vorst 1996).

Calculations of singular vectors using the 3DVAR hessian metric $g_{ij}^{(\mathrm{3DVAR})}$ are currently in progress. In 3DVAR, the background cost function involves the contraction of the first guess fields with (the inverse of) a fixed background covariance metric (Courtier et al. 1993). The formulation of this background metric involves a severe (and probably unrealistic) damping of the baroclinic scales (Thépaut et al. 1996), and at present, it is unclear whether use of the 3DVAR background metric is superior to the energy metric for singular vector calculations. A more satisfactory estimate of the flow-dependent background covariance matrix can be obtained from the Kalman filter technique (Fisher and Courtier 1995), which propagates estimates of analysis

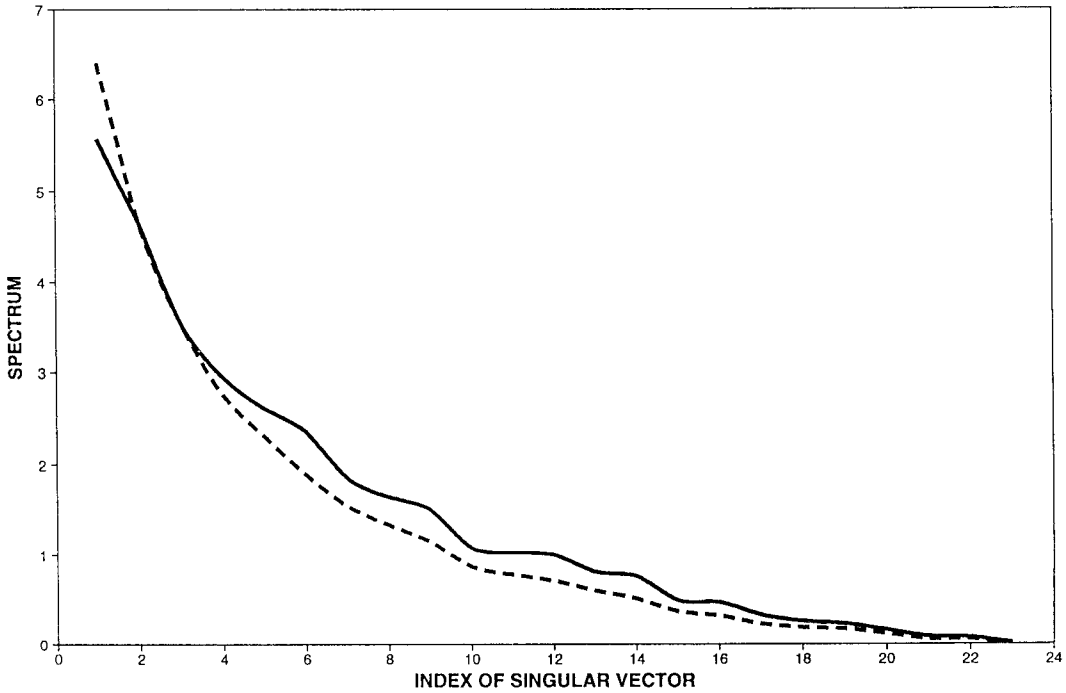

FIG. 10. Solid line: the mean projection of the 2-day ECMWF forecast error for western Europe, projected onto the appropriate 2-day kinetic-energy singular vectors, for 607 separate forecasts. The percentage of explained forecast error variance associated with a particular singular vector is given by the square of the ordinate value of the solid line. Dashed line: the normalized mean singular value spectrum for those 607 days. Further details are given in Barkmeijer et al. (1993).

error covariances using (16). Such a Kalman filter is currently under development at ECMWF (F. Bouttier 1996, personal communication).

6. Singular vectors and breeding vectors

In this and the following two sections, we discuss three alternative strategies for adaptive observations: breeding vectors, potential vorticity diagnosis, and sensitivity vectors.

In the breeding method (Toth and Kalnay 1993, 1996), a random perturbation is repeatedly evolved and rescaled to a specified amplitude over a relatively short cycling time, typically O(1 day), relative to an evolving nonlinear solution (let us call this the control integration trajectory). At any time, the bred perturbation can therefore be obtained as the difference of two nonlinear integrations. In particular, no use is made of the tangent propagators at forecast time. Breeding vectors are used at NCEP as the basis of their ensemble forecasts. In addition, breeding vectors have been suggested as part of a possible strategy for adaptive observations (Bishop and Toth 1996). In the linear limit, the breeding vectors can be thought of as (quasi-) local Lyapunov vectors.

According to the philosophy of breeding, the breeding cycle mimics the analysis cycle. Let us break into this cycle and consider the first-guess error field at some given analysis time. The observations associated with that analysis time reduce (but do not necessarily wipe out) the amplitude of this first-guess error field so that the analysis error also has significant projection onto the growing error carried forward by the first guess.

By contrast, according to the philosophy of singular vectors, the impact of observations in the analysis is to change the structure of the first-guess error, to the extent that the analysis error can have significant projection onto vectors which grow much faster than the local Lyapunov vector. This change in structure is believed to occur because the impact of data in the analysis is scale dependent; on large scales (e.g., cyclone scales), the analysis is strongly constrained by the observations, whereas on small scales (i.e., subcyclone scales) the analysis is relatively poorly constrained by the observations and is determined more by the background field (Daley 1985, 1991; Hollingsworth 1987). This change in structure, equivalent to a phase-space rotation toward small scales, will continually frustrate any tendency of the analysis error to asymptote toward the Lyapunov vector. Of course, in data sparse regions this phase-space rotation would be minimal. It might be argued that such regions occur over the oceans and, in particular, over the North Atlantic where the FASTEX experiment is taking place. However, satellite, ship, buoy, and aircraft data contribute substantially to the operational analyses over the North Atlantic (Kelly et al. 1993), and it would be simplistic to describe such regions as data sparse.

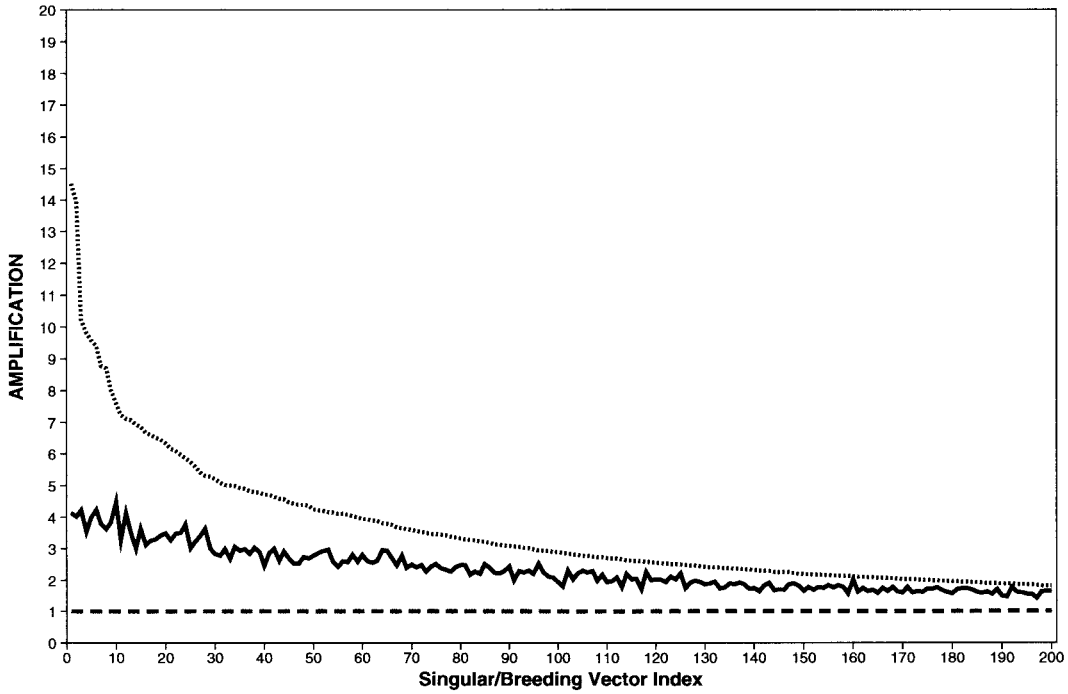

FIG. 11. Solid line: the growth of orthogonal Lyapunov vectors over a typical 48-h period from the QG model. Heavier dashed line: singular values over the same 48-h period. The neutral growth line is also shown for reference.

More generally, in comparing the realism of a given dominant breeding and singular vector, the following question is relevant. How does the growth rate of either vector compare with the projection of the analysis error onto that vector? If the dominant singular vector has an n-fold greater growth than the dominant breeding vector, then it will explain more of the forecast error, provided the projection of the analysis error onto that singular vector is not less than n times the projection of the analysis error onto the breeding vector.

Figure 11 shows an estimate of the growth of the first 200 breeding and (kinetic energy metric) singular vectors in the T21L3 quasigeostrophic model of Marshall and Molteni (1993). Parameter values in the model are the same as in Barkmeijer et al. (1993). Similar calculations of Lyapunov vector growth have been performed using this model by Vannitsem and Nicolis (1997).

To estimate the number of independent growing Lyapunov directions in the QG model, a set of 250 random orthogonal vectors has been constructed with nondimensional kinetic energy of 10^{-10}. These have been used to create initial conditions for a 250-member ensemble. At the end of each 48-h integration, the differences between the perturbed ensemble states and the unperturbed control integration have been orthonormalized with a Gram–Schmidt orthogonalization, rescaling the amplitude to their initial value. This procedure is repeated over a large number (>400) of consecutive 48-h periods. The solid line in Fig. 11 shows the growth rate of the first 200 perturbations over a typical 48-h period where the perturbations are ordered by the Gram–Schmidt procedure. It can be seen that the first orthogonal Lyapunov vector grows at a rate of 4 per 48-h cycle time. The spectrum of growth rates decays rather slowly; the 100th breeding vector grows at a rate of about half the dominant breeding vector.

Singular vectors, optimized over the whole globe, were calculated over the same 48-h period as the breeding vectors were estimated. The spectrum of singular values are shown as the dashed line in Fig. 11. In contrast with singular vectors targeted over Europe, there are at least 200 growing directions, consistent with the breeding vector calculation. The dominant singular values are between 3 and 4 times those of the breeding vectors. Moreover, for the first 30 or so singular vectors, the spectrum of singular values decays much more rapidly than that of the breeding vectors. For example, the growth of the 10th singular vector is about half that of the dominant singular vector.

It should be noted that these results are not consistent with a similar calculation performed by Toth and Kalnay (1996) and Szunyogh et al. (1997) using a T10 primitive equation model. In these calculations, the spectrum of dominant singular values has approximately the same slope as the spectrum of breeding vector growth rates. It is possible that the discrepancy can be explained by noting that the subcyclone to cyclone-scale evolution of a characteristic energy metric extratropical singular vector cannot be simulated in a model with only T10 res-

olution. On the other hand, E. Kalnay (1997, personal communication) has remarked that similar results to those obtained by Toth and Kalnay (1996) have been found in a T62 model at NCEP.

As mentioned above, in assessing the relative importance on atmospheric predictability of the dominant Lyapunov and singular vectors, one must also determine the relative projection of analysis error onto these vectors. If the energy metric were a good representation of the AECM, then these results suggest that more of the forecast error would be explained by the dominant singular vector than the dominant Lyapunov vector. However, this is clearly a contentious issue—and at present it must be concluded that both methods are viable as strategies for targeting adaptive observations.

Returning to Fig. 11, it appears that, on a global basis, there are a substantial number of growing directions. However, when more local target areas are considered (using a secondary spatial projection operator), then the number of growing directions reduces in relation to the number of degrees of freedom in the target area. For example, as Fig. 10 showed, the number of growing directions in the T21L3 quasigeostrophic model for a 48-h optimization over a European region is about 10. This dramatic reduction in the number of growing directions for local domains is clearly relevant for the adaptive observation strategy: the smaller the domain that the observations are targeted for, the fewer the number of sensitive areas that will be revealed by the singular vector calculation. For the FASTEX experiment, singular vector targeting is determined from the first four singular vectors (see section 9).

7. Singular vectors and potential vorticity diagnosis

Potential vorticity has proven to be an extremely powerful diagnostic for analyzing nonlinear advective processes in the atmosphere (e.g., associated with wave breaking: McIntyre and Palmer 1984; Hoskins et al. 1985) and is now widely used to study atmospheric development. It has also been used as a tool for diagnosing the cause of forecast failure (e.g., McIntyre 1988; Appenzeller and Davies 1996). As such, PV diagnosis has also been proposed as a subjective technique for targeting the upstream observation component of FASTEX (K. Browning and A. Thorpe 1996, personal communication).

In diagnosing PV development, it is common to focus on PV maxima on an isentropic surface that crosses the upper troposphere in midlatitudes. The Lagrangian velocity associated with the advection of such PV anomalies is quite unrelated to the group velocity of linear wave activity. For example, in the adiabatic approximation air parcels are constrained to move on isentropic surfaces, whereas wave activity can readily propagate through such surfaces. Insofar as the evolution of singular vectors tends to be readily describable in terms of wave propagation effects, subjective PV targeting and singular vector analysis will not necessarily agree on the location of dynamically sensitive regions.

However, rather than argue from an idealized perspective, it is instructive to consider a particular case study. McIntyre (1988) has described a development occurring during 20–25 September 1982 where a tongue of high PV air breaks off from a stratospheric reservoir near Hudson Bay, advects across the Atlantic, and rolls up to generate a surface cyclone. McIntyre hypothesizes that a poor day-4 ECMWF forecast of this surface cyclone on 24 September may have resulted from small errors in the analysis of upper PV near the Hudson Bay area on 20 September. The 300-K isentropic maps of PV for 20–25 September are shown in Fig 12.

To compare this PV-based estimate of sensitivity, dominant (energy metric) singular vector optimized for the region (50°–60°N, 10°–20°W) for 24 September 1982 (based on reanalysis initial conditions) are shown in Figs. 13–14. Figure 13 shows the initial singular vector for a 48-h optimization, Fig. 14 shows the initial singular vector for a 96-h optimization. For each optimization time, we illustrate the singular vector at two different levels, at model level 9 (about 300 hPa) and model level 15 (about 850 hPa).

Let us first compare the initial singular vector structure for the 48-h optimization, valid for 22 September 1982. In the upper troposphere, the singular vector does not have a simple structure. However, most of the amplitude is concentrated between 40° and 70°N, 70° and 120°W. Insofar as it is possible to define a location with maximum amplitude, it appears to be at about 65°N, 70°W, which would correspond to a point on the western edge of the distorted PV anomaly for 22 September shown in Fig. 12. However, the structure of the singular vector is better defined in the lower troposphere, where it suggests a localized wave train positioned right on the southern tip of the PV anomaly. Hence, in this case, there does appear to be some qualitative correspondence between the horizontal positions of sensitivity as suggested by PV and singular vector analyses, though there is some discrepancy as to what level in the atmosphere the forecast is most sensitive.

The initial singular vector for the 96-h optimization is valid for 20 September. (In practice, initial errors are likely to evolve nonlinearly over 4 days. However, we will ignore such complications here.) In the upper troposphere, the structure of the initial singular vector is rather indistinct, though the largest amplitudes are located in polar latitudes near 80°N, that is, in the region of strong PV anomaly in Fig. 12. However, in the lower troposphere there is again a more clearly defined wave train structure close to 40°N between 60° and 80°W, that is, well south of the southernmost latitude of the PV anomaly for 20 September. Hence, while the 4-day singular vector structure suggests some sensitivity with respect to upper-tropospheric perturbations of PV in polar latitudes, it also highlights an area of sensitivity both

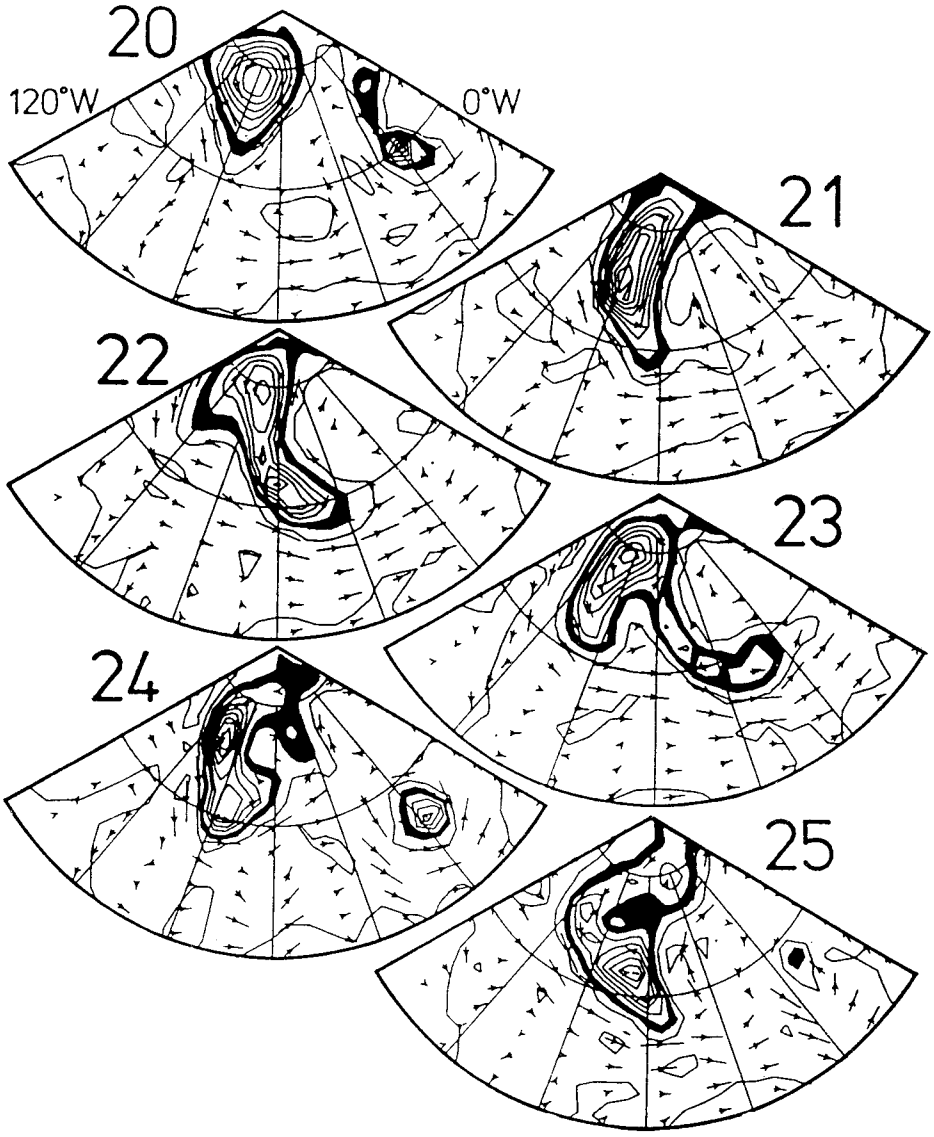

FIG. 12. 300-K isentropic maps of PV for 20–25 September 1982 from ECMWF operational analyses. Latitude circles are 40°N, 60°N, and 80°N; Greenwich meridian on the right, 120°W on the left. See McIntyre (1988) for details.

south of and lower in the atmosphere than would be apparent from the PV analysis.

The reason why the singular vectors have significant structure in the lower troposphere, whereas the PV-based sensitivity estimates do not, is consistent with the difference between the dynamics of Lagrangian advection and wave propagation, as mentioned above. The (potential) enstrophy associated with wave activity propagating from the lower to upper troposphere (or from lower to higher latitudes) will amplify if the wave propagates from a region of small PV gradient into a region of strong PV gradient. This is dictated by pseudomomentum conservation (cf. BP), a simple expression of which is given by the adiabatic frictionless form (e.g., Andrews et al. 1987):

$$\frac{1}{2}\frac{\partial}{\partial t}\left(\frac{\overline{q'^2}}{\overline{q}_y}\right) + \nabla\cdot\mathbf{F} = 0 \qquad (22)$$

for small amplitude Rossby waves propagating on a stationary zonally symmetric basic state. Here, $\overline{q'^2}$ denotes wave potential enstrophy, $\overline{q}_y$ is the meridional gradient of basic-state potential vorticity, and $\mathbf{F}$ is the Eliassen–Palm flux. The westward phase tilt of the initial singular vectors is consistent with a positive vertical component of the Eliassen–Palm flux and, hence, upward propagation into the jet.

The lower-tropospheric 4-day singular vector appears to be located in just such a region of weak PV gradient. Over the 4-day evolution period, the wave propagates both vertically and horizontally into the region of strong

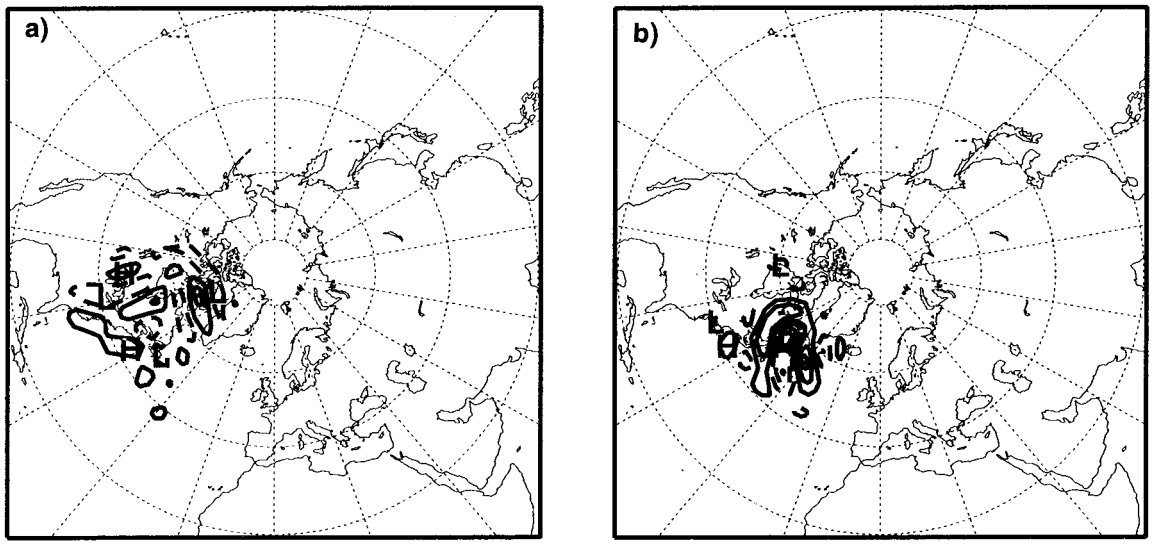

FIG. 13. Streamfunction for initial 48-h singular vector optimized for the region (50°–60°N, 10°–20°W) on 24 September 1982 (and therefore valid for 22 September): (a) near 300 hPa and (b) near 850 hPa.

PV gradient in the upper troposphere over the target region. By contrast, for subjective PV targeting, the role of small PV anomalies in the lower troposphere (e.g., on 20 September) will tend to be downplayed by the eye in comparison with the much larger PV anomalies in the region of strong gradient in the upper troposphere.

The crucial issue, as far as the relationship between these lower-tropospheric singular vector structures and forecast sensitivity is concerned, is whether the component of analysis error in lower-tropospheric (and lower latitude) PV anomalies projecting onto vertically propagating waves is large enough to be responsible for forecast errors in upper-tropospheric PV. At present it is not possible to give a definitive answer to this. However, the answer does depend on the vertical distribution of variance in the AECM. As shown in Fig. 9, the vertical distribution of variance associated with the (surrogate) analysis error fields is much broader than the distribution associated with the analyses themselves. The results in Fig. 9 suggest that the a priori probability of an analysis error of given magnitude in the lower troposphere may not be so much different from the probability of an analysis error with the same magnitude in the upper troposphere. (Of course, these variance distributions do not prove that the initial error has sufficient projection onto vertically propagating modes to explain

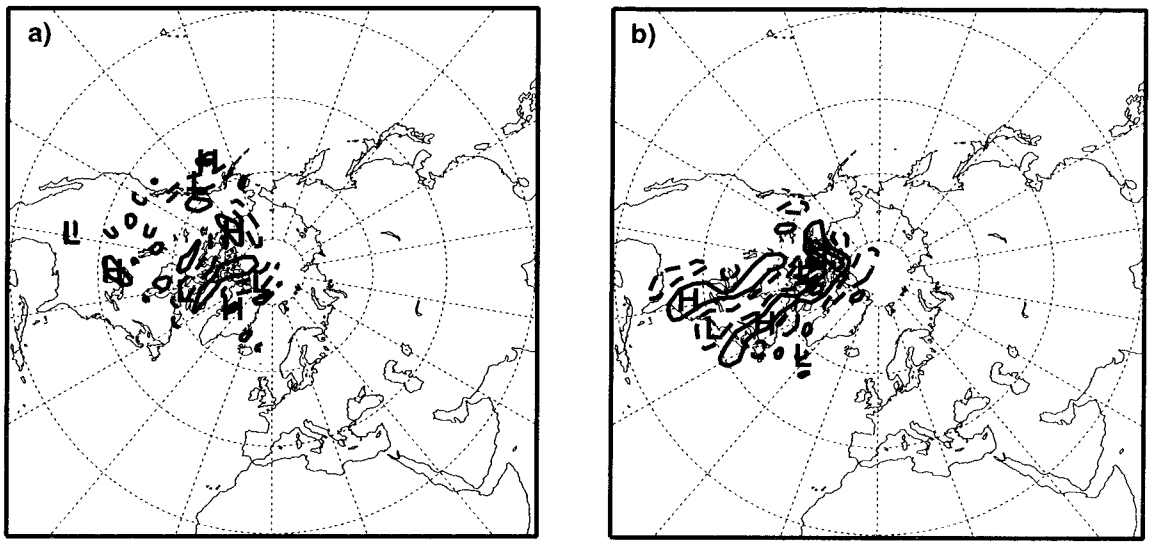

FIG. 14. Streamfunction for initial 96-h singular vector optimized for the region (50°–60°N, 10°–20°W) on 24 September 1982 (and therefore valid for 20 September): (a) near 300 hPa and (b) near 850 hPa.

upper-tropospheric forecast error.) As remarked above, possible explanations for these results could be the role of diabatic heating errors in the analysis, the relative accuracy of satellite sounding data in the upper troposphere, and the relative abundance of aircraft wind data in the upper troposphere.

It can be asked whether perturbations in lower-tropospheric PV act as a trigger for upper-tropospheric PV development, through pseudomomentum propagation. This is essentially a GFD question rather than a predictability question. Compared with the analysis differences, the vertical distribution of streamfunction variance in the analysis is strongly weighted toward the upper troposphere (Fig. 9). Conceivably this weighting (which would be implicit in the GFDCM) might be a sufficiently strong constraint as to eliminate these vertically propagating waves in GFD singular vector computations in comparison with amplification by horizontal wave propagation in the upper troposphere. Such a calculation has not been performed.

8. Singular vectors and the sensitivity vector

The third proposal (Langland and Rohaly 1996) for a strategy for targeting adaptive observations is based on the sensitivity vector (9). In this proposal J is the lower-tropospheric enstrophy of the forecast state. From a predictability perspective, this strategy can be thought of as targeting observations on the sensitivity of the component of forecast error that is correlated with the forecast state itself.

The sensitivity and singular vector methods are closely related. For example, if forecast error is expanded in terms of the (final time) singular vectors, then the sensitivity vector can be written in terms of the initial singular vectors with coefficients of projection weighted by the singular values (Rabier et al. 1996).

In the limit where the secondary operator is a projection operator for a single model variable on a single grid point at a single level (i.e., there is only one degree of freedom left in the optimization), then the initial singular vector and the sensitivity vector will agree precisely. One can anticipate that the sensitivity and singular vector methods will disagree more significantly when there are a number of singular vectors with comparable singular values. In this situation, the targeting domain based on the singular vector strategy may be more extensive than that based on the sensitivity analysis.

9. Targeting products based on singular vectors

Some examples of the use of the potential use of singular vectors to target observations have been given in Montani et al. (1996). In this section we show two examples of the use of singular vectors for targeting observations for the upstream component of FASTEX in Fig. 15, based on a the period immediately preceding the start of the FASTEX experiment (1200 UTC 23 December 1996 and 1200 UTC 30 December 1996). This period is of interest as the flow over the Atlantic region was rapidly developing (toward the establishment of a blocking anticyclone). The area shown covered with symbols denotes the region(s) over which aircraft dropsonde observations should be made, based on a set of the first four singular vectors optimized at 48 h over the FASTEX intensive observing area. The calculations have been performed on a 48-h forecast trajectory from $T + 24$ to $T + 72$ h. In this way, the singular vector results can be disseminated in time for adequate flight preparation. For the FASTEX experi-

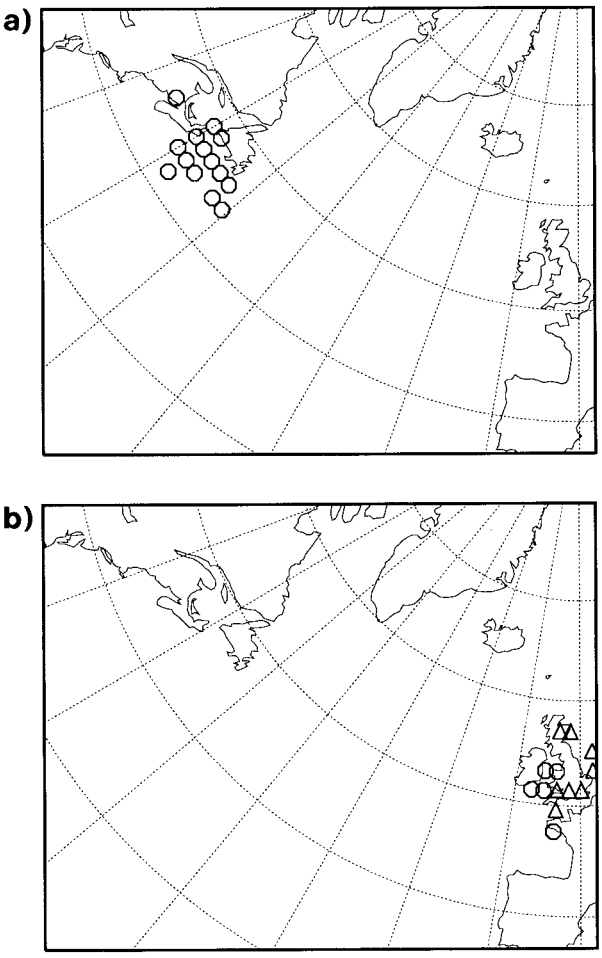

FIG. 15. Targets for adaptive observations based on 48-h energymetric singular vectors optimized for the FASTEX intensive observation area over the eastern Atlantic. The symbols show grid points where the vorticity of any of the first four singular vectors (weighted by the corresponding singular value) exceeds 0.7 times the maximum of the dominant singular vector. In (a) (1200 UTC 23 December 1996) the target region is given only by the first singular vector (circles). In (b) (1200 UTC 30 December 1996) the target region is defined from both the first (circles) and second (triangles) singular vectors. The difference in targeting regions between (a) and (b) reflects the changing large-scale atmospheric circulation between 23 and 30 December. (Figure produced by A. Montani, Reading University, United Kingdom).

ment, the choice of a 48-h optimization was considered desirable since, on the one hand, the most likely upstream targeting area would be easily within range of aircraft based on the east coast of the North American continent and, on the other hand, the optimization period would not be so long as to jeopardize the assumption of linear error growth.

Specifically, the symbols denote regions where the amplitude of any of the first four singular vectors weighted by their respective singular value exceeds 0.7 of the maximum amplitude of the dominant singular vector. This amplitude is calculated separately in terms of vorticity and temperature, though only the vorticity targets are shown in this paper.

For the first example shown (Fig. 15a), the only targeted regions shown are associated with the dominant singular vector only. Moreover (not shown), the regions associated with maximum temperature and maximum vorticity roughly agree—indicating that the extra dropsonde observations should be made to the south of Newfoundland.

For the second example (Fig. 15b), the targeted region is associated with an area over part of the United Kingdom and North Sea, well to the east of the FASTEX intensive observing area. The targeted region involves both the first and second singular vectors, but these roughly agree as far as the general geographical location of targeting is concerned.

These two examples show, for the same intensive observation area, how flow dependent the targeted region can be. The reason the targeted region changes so much between the first and second dates can be explained in terms of the development of the strong blocking anticyclone over northern Europe and the western Atlantic during the last two weeks of December 1996.

In these examples, the singular vector calculations each indicate a relatively small targeting area. For other flow configurations there may be no single region. In fact for the date of the first example, 24-h singular vectors did not indicate a single well-defined region: singular vectors 1 and 2 defining an area in the mid-Atlantic, singular vectors 3 and 4 indicating a region over the United Kingdom itself (consistent with the developing blocking anticyclone, as mentioned above). In this situation it would not be possible to define a unique target region.

Although much of this paper has been motivated by the upstream component of FASTEX, there is a developing technology, described by Holland et al. (1992) and Langford and Emanuel (1993) for which the singular vector targeting strategy could in principle be utilized on a more routine basis. The technology is based on the use of unmanned aircraft, which have been shown to be capable of flight to 100 hPa with a range of many thousands of kilometers. Such aircraft are capable of taking in situ measurements or deploying a number of dropsondes during flight. Autonomous navigation would be possible through the Global Positioning System, and data would be retrieved by satellite relay.

It should be noted that much of the original interest in unmanned aircraft was motivated by the tropical cyclone prediction problem (Holland et al. 1992; Langford and Emanuel 1993). Singular vectors with moist physics included (e.g., Vukicevic and Errico 1993; Buizza et al. 1996; J. F. Mahfouf 1996, personal communication), can, in principle, be used to target observations for tropical cyclones. Obviously, in this case, the secondary operator would not be fixed to a specific geographical area, but would be centered on the specific forecast tropical cyclone of interest.

While all applications so far have been atmospheric, there is no reason in principle why singular vectors should not be used as a strategy for adaptive observations in the ocean. For example, Chen et al. (1997) have calculated 6-month optimized singular vectors in an intermediate coupled ocean–atmosphere model of the tropical Pacific. In all situations studied, at initial time, the sea surface temperature of the coupled singular vector has a dipole structure oriented across the tropical Pacific basin in a northwest–southeast direction. The principal center of action appears to be located close to, or just to the west of the date line, varying in location by about 20° long depending on the season and state of El Niño in the basic state. An adaptive observation strategy making use of ship-based XBT measurements, or maybe autonomous submarines, targeting the singular vector maximum in the west Pacific could be envisaged based on such analysis.

The next phase of research is to demonstrate the feasibility and impact of this strategy in data assimilation and forecast mode. Comparison of forecasts made from initial conditions with and without targeted observations are required, though the impact of these forecasts should also be compared with forecasts made using randomly made extra observations. To some extent this could be done through (targeted) observation simulation experiments.

10. Conclusions and summary

Singular vectors (of the linearized equations of motion) have been discussed in the scientific literature as a means of quantifying the growth of perturbations in the atmosphere and oceans in order to explain the geophysical fluid dynamics of a particular climatic phenomenon (such as cyclogenesis or El Niño). Singular vectors have also been used to quantify how dynamically active errors in forecast initial conditions evolve through the forecast. For this class of problem the instability of the system, the specification of the network used to observe the system, and the process of assimilating the observations are all relevant. In this paper, we discuss a third use of singular vectors: to define a strategy for targeting adaptive observations of the atmosphere (or the oceans). In such a strategy, observa-

tions would be made in particularly flow-dependent ''sensitive'' regions determined by the location of the singular vectors at initial time. The developing unmanned aircraft technology could provide the practical means of realizing such a strategy, though many of the calculations in this paper were motivated by the FASTEX experiment.

In discussing their potential application as a strategy for targeting adaptive observations, it is important to decide whether the singular vector method is being applied in the GFD or predictability sense of the paragraph above. Such a decision determines the most appropriate metric, or inner product, with which to compute the singular vectors. To make this metric dependence explicit and to distinguish the mathematical objects used to define metrics, covariance matrices and tangent operators, an index-based tensor formalism is used in this paper to define the singular vectors and related quantities.

If the purpose of the targeted measurements is to improve the skill of forecasts of the primary phenomenon, then the singular vector metric should be consistent with the predictability measure μ_{pre} that gives the probability that an atmospheric state determined by a data assimilation system corresponds to the true state of the atmosphere at the analysis time. By choosing a metric based on the second moment of the analysis error probability distribution, then the singular vectors at optimization time are the eigenvectors of the forecast error covariance tensor. With respect to this metric (which we refer to as the analysis error covariance metric, AECM), the singular vector calculation is an optimization taking into account the flow-dependent dynamical growth of any particular perturbation, constrained by the a priori probability that the analysis error might be described by that perturbation.

By contrast, if the purpose of the targeted observations is to study the precursors of the primary phenomenon, then the singular vector metric should be determined purely by the dynamical equations of motion alone. The appropriate probability measure μ_{GFD} and associated GFD covariance metric (GFDCM) will be given (in principle) by the invariant measure of the attractor determined solely by the dyncamical equations of motion. In practice, the GFDCM could be obtained by a truncated EOF basis.

In this paper, energy, enstrophy, and streamfunction variance were all considered as candidate metrics for singular vector calculations. From two independent sets of calculations based on analyses, and short-range forecast data, it was shown that of these three choices, energy is the most appropriate metric for the predictability problem. More accurate estimates of the AECM can in principle be obtained from variational and Kalman filter data assimilation techniques.

While the energy metric may be a reasonable first-order estimate of the AECM, it does not appear a suitable inner product for the GFD problem. A fundamental reason for this is that spectrum of analysis error appears to be relatively white (in terms of horizontal total wavenumber) compared with the equivalent spectrum of atmospheric variability (obtained from the analysis fields themselves). With an energy inner product, initial singular vector structures are predominantly subcyclone scale, consistent with the spectra of analysis error but not with the spectra of the analyses themselves.

The reason why the analysis error spectra appear to be whiter than the spectra of the analyses themselves is basic to the methodology of data assimilation. In data assimilation, the observations are added to a background field, itself a forecast field from an earlier initial analysis. However, the impact of such observations is scale dependent, constraining the larger scales (well resolved by the observing network) much more strongly than the smaller scales. Hence the role of observations is to damp preferentially the amplitude of background forecast error at low wavenumbers, and hence to generate a whiter analysis error spectrum than the background error spectrum.

For FASTEX, the main purpose of the targeted observations is to improve the predictive skill of forecasts for the intensive observation area over the eastern Atlantic. Examples of products produced in the pre-FASTEX period are shown based on the first four energy metric singular vectors, optimized for the intensive observation area.

In addition to singular vectors, three other techniques have been proposed for adaptive observation strategies: the breeding method, subjective potential vorticity analysis, and sensitivity analysis. These additional methods were compared with the singular vector strategy. With the exception of sensitivity analysis, which is also a metric-dependent adjoint method, these methods can generate quite different regions for observation targeting from that associated with the singular vector method.

Reasons for these differences are discussed in the body of the paper. Differences between breeding and singular vectors reflect the differences between Lyapunov and optimal growth. Their relevance to targeting for predictability depends on the growth rates of these vectors and on their projection onto analysis error. Discrepancies in targeting based on singular vectors and PV analysis can be explained in terms of the different perspectives associated with Lagrangian advection and wave propagation. For the latter, pseudomomentum conservation can lead to regions of sensitivity being located in areas of weak PV gradient.

Ultimately an assessment of the optimal strategy for observation targeting can only be made by comparing the impact of extra observations made using each of the candidate strategies on forecast skill.

Acknowledgments. Our thanks to P. Courtier, R. Daley, K. Emanuel, A. Hollingsworth, B. Hoskins, P. Ioannou, A. Joly, E. Kalnay, M. McIntyre, A. Persson, A. Simmons, C. Snyder, D. Stephenson, A. Thorpe, and Z.

Toth for many helpful comments and discussions. We thank A. Montani of Reading University for Fig. 15.

REFERENCES

Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics.* Academic Press, 489 pp.

Appenzeller, C., H. C. Davies, J. M. Popovic, S. Nickovic, and M. B. Gavrilov, 1996: PV morphology of a frontal-wave development. *Meteor. Atmos. Phys.*, **58,** 21–40.

Barkmeijer, J., 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.

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

——, ——, J. Barkmeijer, R. Gelaro, and J. F. Mahfouf, 1996: Singular vectors, norms and large-scale condensation. Preprints, *11th Conf. on Numerical Weather Prediction*, Norfolk, VA, Amer. Meteor. Soc., 50–51.

Cacuci, D. G., 1981: Sensitivity theory for nonlinear systems. I: Nonlinear functional analysis approach. *J. Math. Phys.*, **22,** 2794–2802.

Chen, Y.-Q., D. S. Battisti, T. N. Palmer, J. Barsugli, and E. S. Sarachik, 1997: A study of the predictability of tropical Pacific SST in a coupled atmosphere–ocean model using singular vector analysis: The role of the annual cycle and the ENSO cycle. *Mon. Wea. Rev.,* **125,** 224–238.

Courtier, P., C. Freydier, J.-F. Geleyn, F. Rabier, and M. Rochas, 1991: The Arpege project at Meteo-France. *Proc. ECMWF Seminar on Numerical Methods in Atmospheric Models*, Vol. 2, Reading, United Kingdom, ECMWF. 193–231.

——, E. Andersson, W. Heekley, G. Kelly, J. Pailleux, F. Rabier, J. N. Thepaut, P. Unden, D. V. Vasiljevic, C. Cardinali, J. Eyre, M. Hamrud, J. Haseler, A. Hollingsworth, A. McNally, and A. Stofflen, 1993: Variational assimilation at ECMWF. ECMWF Research Department Tech. Memo 194, 84 pp. [Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom.]

Daley, R., 1985: The analysis of synoptic scale divergence by a statistical interpolation scheme. *Mon. Wea. Rev.*, **113,** 1066–1079.

——, 1991: *Atmospheric Data Analysis.* Cambridge University Press. 457 pp.

Dodson, C. T. J., and T. Poston, 1979: *Tensor Geometry.* Pitman, 598 pp.

Ehrendorfer, M., and J. J. Tribbia, 1997: Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.*, **53,** 286–313.

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

——, 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.,* **47,** 1193–1206.

Fisher, M., and P. Courtier, 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. ECMWF Tech. Memo. 220, 28 pp. [Available from Research Department, ECMWF Shinfield Park, Reading RG2 9AX, United Kingdom.]

Frisch, U., 1995: *Turbulence.* Cambridge University Press. 296 pp.

Gelaro, R., R. Buizza, T. N. Palmer, and E. Klinker, 1998: Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. *J. Atmos. Sci.,* **55,** 1012–1037.

Hartmann, D. L., R. Buizza, and T. N. Palmer, 1995: Singular vectors: The effect of spatial scale on linear growth of disturbances. *J. Atmos. Sci.,* **52,** 3885–3894.

Holland, G. J., T. McGeer, and H. Youngren, 1992: Autonomous aerosondes for economical atmospheric soundings anywhere on the globe. *Bull. Amer. Meteor. Soc.*, **73,** 1987–1998.

Hollingsworth, A., 1987: Objective analysis for numerical weather prediction. *Short and Medium Range Numerical Weather Prediction. Collected Papers WMO/IUGG NWP Symp.,* Tokyo, Japan, Meteorological Society of Japan, 11–59.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111,** 877–946.

Houtekamer, P. L., 1995: The construction of optimal perturbations. *Mon. Wea. Rev.*, **123,** 2888–2898.

Joly, A., and Coauthors, 1997: R. M. Wakimoto, 1996: The Fronts and Atlantic Storm-Track Experiment (FASTEX): Scientific objectives and experimental design. *Bull. Amer. Meteor. Soc.*, **78,** 1917–1940.

Kelly, G., J. Pailleux, F. Rabier, and J.-N. Thépaut, 1993: Observing system experiments made with the ECMWF system: World Weather Watch Tech. Rep., 16, WMO/TD 594, 91 pp. [Available from World Meteor. Org., Case Postale 2300, CH–1211 Geneva, Switzerland.]

Langford, J. S., and K. A. Emanuel, 1993: An unmanned aircraft for dropwindsonde deployment and hurricane research. *Bull. Amer. Meteor. Soc.*, **74,** 367–375.

Langland, R. H., and G. D. Rohaly, 1996: Adjoint-based targeting of observations for FASTEX cyclones. *Proc. Seventh Conf. on Mesoscale Processes*, Reading, United Kingdom, Amer. Meteor. Soc., 369–371

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

——, 1993: *The Essence of Chaos.* University of Washington Press, 227 pp.

Mardia, K. V., J. T. Kent, and J. M. Bibby, 1979: *Multivariate Analysis.* Academic Press, 518 pp.

Marshall, J., and F. Molteni, 1993: Toward a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.,* **50,** 1792–1818.

McIntyre, M. E., 1988: The use of potential vorticity and low-level temperature/moisture to understand extratropical cyclogenesis. *Proc. ECMWF Seminar*, Vol. 1, Shinfield Park, Reading, United Kingdom, ECMWF, 261–280.

——, and T. N. Palmer, 1984: The ''surf zone'' in the stratosphere. *J. Atmos. Terr. Phys.,* **46,** 825–849.

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,** 73–120.

Montani, A., R. Buizza, and A. J. Thorpe, 1996: Singular vector calculations for cases of cyclogenesis in the North Atlantic storm-track: Preprints, *Seventh Conf. on Mesoscale Processes*, Reading, United Kingdom, Amer. Meteor. Soc., 391–392.

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

Palmer, T. N., 1996: Predictability of the atmosphere and oceans: From days to decades. *Decadal Climate Variability—Dynamics and Predictability*, NATO ASI Series, Subseries I, *Global Environmental Change*, Vol. 44, D. Anderson and J. Willebrand, Eds., Springer, 83–155.

——, F. Molteni, R. Mureau, R. Buizza P. Chapelet, and J. Tribbia, 1993: Ensemble prediction. *Proc. ECMWF Seminar Proc. on Validation of Models over Europe*, Vol. 1, Shinfield Park, Reading, United Kingdom, ECMWF, 21–66.

Penland, C., and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperatures. *J. Climate*, **8,** 1999–2024.

Pu, Z., E. Kalnay, J. Sela, and I. Szunyogh, 1997: Sensitivity of forecast errors to initial conditions with a quasi-inverse linear model. *Mon. Wea. Rev.,* **125,** 2479–2503.

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

Sleijpen, G. L. G., and H. A. van der Vorst, 1996: A Jacobi–Davidson

iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.,* **17,** 401–425.
Snyder, C., 1996: Summary of an informal workshop on adaptive observations and FASTEX. *Bull. Amer. Meteor. Soc.,* **77,** 953–961.
Stephenson, D. B., 1997: Correlation of spatial climate/weather maps and the advantages of using the Mahalanobis metric in predictions. *Tellus,* **49A,** 513–527.
Szunyogh, I., E. Kalnay, and Z. Toth, 1997: A comparison of Lyapunov and optimal vectors in a low resolution GCM. *Tellus,* **49A,** 200–227.
Thépaut, J. N., P. Courtier, G. Belaud, and G. Lemaître, 1996: Dynamical structure functions in a four-dimensional variational assimilation: A case study. *Quart. J. Roy. Meteor. Soc.,* **122,** 535–562.
Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.,* **74,** 2317–2330.
——, and ——, 1996: Ensemble forecasting at NCEP. *ECMWF Seminar on Predictability,* Vol. 2, Shinfield Park, Reading, United Kingdom, ECMWF, 259 pp.
Trefethen, L. N., A. E. Trefethen, S. C. Reddy, and T. A. Driscoll, 1993: Hydrodynamic stability without eigenvalues. *Science,* **261,** 578–584.
Vannitsem, S., and C. Nicolis, 1997: Lyapunov vectors and error growth patterns in a T21L3 quasi-geostrophic model. *J. Atmos. Sci.,* **54,** 347–361.
Vukicevic, T., and R. M. Errico, 1993: Linearization and adjoint of parametrized moist diabatic processes. *Tellus,* **45A,** 493–510.