Decaying Singular Vectors and Their Impact on Analysis and Forecast Correction

C. A. REYNOLDS

Naval Research Laboratory, Monterey, California

T. N. PALMER

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

(Manuscript received 21 March 1997, in final form 22 December 1997)

ABSTRACT

The full set of kinetic energy singular values and singular vectors for the forward tangent propagator of a quasigeostrophic potential vorticity model is examined. In contrast to the fastest growing singular vectors, the fastest decaying vectors exhibit a downward and downscale transfer of energy and an eastward tilt with height. The near-neutral singular vectors resemble small-scale noise with no localized structure or coherence between levels.

Post-time forecast and analysis correction techniques are examined as a function of the number of singular vectors included in the representation of the inverse of the forward tangent propagator. It is found that for the case when the forecast error is known exactly, the best corrections are obtained when using the full inverse, which includes all of the singular vectors. It is also found that the erroneous projection of the analysis uncertainty onto the fastest decaying singular vectors has a significant detrimental effect on the estimation of analysis error. Therefore, for the more realistic case where the forecast error is known imperfectly, use of the full inverse will result in an inaccurate estimate of analysis errors, and the best corrections are obtained when using an inverse composed only of the growing singular vectors. Running the tangent equations with a negative time step is a very good approximation to using the full inverse of the forward tangent propagator.

1. Introduction

Medium-range forecast models have improved to the point where the systematic component of the forecast error has been reduced considerably. Systematic error now accounts for only a small part of the total extratropical forecast error (Palmer et al. 1990; Saha 1992; Reynolds et al. 1994; Simmons et al. 1995). However, there still appears to be considerable potential for improvement of deterministic forecasts through a reduction in both the analysis error and the random component of model error. Simmons et al. (1995) suggest that five-day forecasts could be as good as current three-day forecasts if the one-day forecast error could be reduced by the same factor as it was over the 1981–94 time period. This one-day forecast error improvement could come from both improvements in the initial analysis as well as improvements in the forecast model. Intensive efforts are currently on going to improve analyses, such as through the inclusion of satellite radiance data (Andersson et al. 1994; Baker et al. 1995) and the development of advanced analysis techniques such as four-dimensional variational data assimilation (4DVAR, Derber 1987; Rabier et al. 1993; Zupanski 1993; Rabier et al. 1996a) and Kalman filtering (Fisher and Courtier 1995; Ménard and Daley 1996). Several operational centers have also introduced ensemble forecasts as an attempt to deal quantitatively with the inevitable uncertainty in the initial analysis (Toth and Kalnay 1993; Molteni et al. 1996).

Some recent efforts to understand the nature of analysis error and its impact on forecast error have utilized linear techniques. Rabier et al. (1996b) have used the adjoint of a simplified version of the operational European Centre for Medium-Range Weather Forecasts (ECMWF) forecast model to find the sensitivity of forecast errors to initial condition errors. By appropriately scaling this sensitivity and subtracting this perturbation from the initial conditions, subsequent nonlinear forecasts made from these corrected analyses produce significantly improved forecasts. Buizza et al. (1997) have applied similar techniques involving estimates of sensitivity and analysis error obtained by using the first 30 singular vectors. These techniques improve forecasts in a post-time setting, but they may also show some promise for real-time correction (Gelaro et al. 1996).

Both of the above-mentioned techniques produce corrections to the analyses that are smaller than the esti-

Corresponding author address: Carolyn Reynolds, Naval Research Laboratory, 7 Grace Hopper Ave., Monterey, CA 93943-5502.
E-mail: reynolds@nrlmry.navy.mil

mated size of the analysis error, yet result in very significant improvements to the nonlinear forecasts. While these techniques have proved successful in capturing the fastest growing components of the analysis error, there is still incentive to learn about the full analysis error. Improved analyses are of interest in themselves for studies of atmospheric processes. Also, while the above studies find the forecast error dominated by the fastest growing perturbations, they do not capture the entire growing subspace and therefore do not account for the full linear growth of analysis errors.

Pu et al. (1997) have proposed an alternative technique that theoretically will produce approximations of the full analysis error, not just the fastest growing part of that error. An approximation of the full inverse of the forward tangent propagator is obtained by running the tangent equations with a negative time step and negative dissipative terms. The tangent equations are the equations of the model that have been linearized about a nonlinear trajectory. The tangent equations for most global forecast models are currently based on simplified versions of the full nonlinear forecast model that do not include many of the diabatic parameterizations, such as moist physics. Therefore, this technique, and any other technique using a primarily adiabatic tangent propagator, cannot properly describe the components of analysis error that amplify through nonlinear or diabatic processes. Despite these limitations, Pu et al. (1997) are able to obtain corrected analyses that produce significantly improved forecasts. They also find that in ''perfect-model'' experiments, where both the analysis and forecast differences are known exactly, this technique is able to produce a very close approximation to the prescribed analysis difference. This technique also has the advantage of being computationally very economical.

The above experiments have been conducted using versions of operational forecast models that, though considerably simplified, are still too large for matrix forms of the forward tangent propagator to be represented and inverted explicitly. In this study, a simple model is used to study the properties of the forward tangent propagator and its inverse, and assess how the above approximations compare to the use of the exact inverse. A brief discussion of the pertinent linear theory is given in the following section. In section 3 the model used in this study is briefly described. In section 4, the method employed is discussed in detail. In section 5, characteristics of the full spectrum of singular values and the structures of growing, decaying, and neutral singular vectors are described. In section 6, results using different methods of forecast and analysis difference correction are presented. Section 7 contains the summary and conclusions.

2. Linear theory

This study is done in the context of a ''perfect model'' assumption. That is, one nonlinear trajectory is taken as ''truth,'' and another nonlinear trajectory, started from slightly different initial conditions, is taken as the ''forecast.'' In this situation, the model and the ''atmosphere'' behave in the same way, and the forecast ''error'' is due entirely to differences in the initial conditions. Therefore, for this study, only forecast differences, not true forecast errors (which are also caused by model deficiencies), are examined. Likewise, the initial difference between the two trajectories are actually differences between analyses, not true analysis errors.

For a perfect model:

$$\mathbf{L}\mathbf{e}_o = \mathbf{e}_{fl} = \mathbf{e}_f - \mathbf{e}_{fn}, \tag{1}$$

where $\mathbf{L}$ is the forward tangent propagator of the model, $\mathbf{e}_o$ is the error at initial time and $\mathbf{e}_f$ is the full nonlinear forecast error. The linear component of the forecast error is $\mathbf{e}_{fl}$, while $\mathbf{e}_{fn}$ is the difference between $\mathbf{e}_f$ and $\mathbf{e}_{fl}$ due to nonlinearities.

From (1), following Buizza et al. (1997) and Pu et al. (1997),

$$\mathbf{e}_o \cong \mathbf{L}^{-1}\mathbf{e}_f, \tag{2}$$

which indicates that if the inverse of the forward tangent propagator, or an approximation of it, is known, and the nonlinearities ($\mathbf{e}_{fn}$) are very small, then operating on the forecast error with $\mathbf{L}^{-1}$ should approximate the analysis error.

Here $\mathbf{L}$ can be expressed in terms of singular vectors and singular values (see Molteni and Palmer 1993), such that

$$\mathbf{K}^{1/2}\mathbf{L}\mathbf{K}^{-1/2} = \mathbf{V}_K\mathbf{\Sigma}_K\mathbf{U}_K^T, \tag{3}$$

where $\mathbf{K}$ is a matrix to be defined below, $\mathbf{V}_K$ is a matrix with columns that are the (left) singular vectors of $\mathbf{K}^{1/2}\mathbf{L}\mathbf{K}^{-1/2}$ at final time, $\mathbf{U}_K$ is a matrix with columns that are the (right) singular vectors of $\mathbf{K}^{1/2}\mathbf{L}\mathbf{K}^{-1/2}$ at initial time, and $\mathbf{\Sigma}_K$ is a diagonal matrix whose elements are the singular values of $\mathbf{K}^{1/2}\mathbf{L}\mathbf{K}^{-1/2}$, ranked in order of descending magnitude. Note that all matrices that are dependent upon the metric have the subscript $\mathbf{K}$. The first column of $\mathbf{U}_K$ represents the fastest growing perturbation for the specific trajectory considered. This perturbation evolves into the vector of the first column of $\mathbf{V}_K$ and has an amplification factor given by the first singular value. The last column of $\mathbf{U}_K$ represents the fastest decaying initial perturbation that evolves into the last column of $\mathbf{V}_K$ with an amplification factor given by the last singular value.

The singular vectors and singular values depend on the metric $\mathbf{K}$. For studying predictability and forecast error growth, an appropriate metric should be based on analysis error covariances (Ehrendorfer and Tribbia 1997; Barkmeijer et al. 1998). However, given the uncertainties in current estimates of analysis error covariances, a kinetic energy metric is used in this study. With this choice, $\mathbf{K}$ is a diagonal matrix such that for a state vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{K}\mathbf{x}$ gives the kinetic energy of the system. The singular values (vectors) of $\mathbf{K}^{1/2}\mathbf{L}\mathbf{K}^{-1/2}$ will be

referred to as the KE singular values (vectors) of $\mathbf{L}$. The appropriateness of an energy-based metric for these applications is discussed in detail in Palmer et al. (1998).

The inverse of the forward tangent propagator may be expressed in terms of the same singular values and singular vectors as above:

$$\mathbf{K}^{1/2}\mathbf{L}^{-1}\mathbf{K}^{-1/2} = \mathbf{U}_{\mathrm{K}}\mathbf{\Sigma}_{\mathrm{K}}^{-1}\mathbf{V}_{\mathrm{K}}^{\mathrm{T}}. \quad (4)$$

A related quantity is the adjoint (transpose) tangent propagator, which can be expressed as

$$\mathbf{K}^{-1/2}\mathbf{L}^{\mathrm{T}}\mathbf{K}^{1/2} = \mathbf{U}_{\mathrm{K}}\mathbf{\Sigma}_{\mathrm{K}}\mathbf{V}_{\mathrm{K}}^{\mathrm{T}}. \quad (5)$$

Note from (4) and (5) that the inverse and the adjoint differ by the weighting of the singular values. For the Euclidean metric (i.e., $\mathbf{K}$ is the identity matrix), $\mathbf{L}^{-1}$ and $\mathbf{L}^{\mathrm{T}}$ are identical if and only if $\mathbf{L}$ is unitary.

Because $\mathbf{L}$ is an $m \times m$ matrix, where m is the number of nonzero components of the state vector of the model, the full matrix for even a relatively low-resolution adiabatic version of a forecast model is too large to represent explicitly and invert given current computing capabilities. However, Rabier et al. (1996b) have used the adjoint equations operating on the forecast error to find the sensitivity of the forecast error to changes in the initial conditions, given by

$$\mathbf{s} = \mathbf{L}^{\mathrm{T}}\mathbf{K}\mathbf{e}_{\mathrm{f}}. \quad (6)$$

It is clear from (6) that the sensitivity is metric dependent. Rabier et al. (1996b) showed that an analysis perturbation $\mathbf{e}_{\mathrm{o}}^{s}$, found by appropriately scaling $\mathbf{s}$, can significantly reduce errors in a nonlinear forecast. The magnitude of $\mathbf{e}_{\mathrm{o}}^{s}$ is well within the bounds of current analysis uncertainty estimates. Buizza et al. (1997) have similarly been able to reduce the forecast error by using the ''pseudoinverse'' method. The pseudoinverse method uses $\mathbf{L}^{-1}$ as described in (4) but composed of only a limited subset of the fastest growing singular vectors in order to find the fastest growing components of $\mathbf{e}_{\mathrm{o}}$. The pseudoinverse $\mathbf{L}_{\mathrm{K},n}^{-1}$ is defined as

$$\mathbf{L}_{\mathrm{K},n}^{-1} = \mathbf{U}_{\mathrm{K},n}\mathbf{\Sigma}_{\mathrm{K},n}^{-1}\mathbf{V}_{\mathrm{K},n}^{\mathrm{T}}, \quad (7)$$

where $\mathbf{U}_{\mathrm{K},n}$ is a rectangular matrix with columns composed of the first n initial-time singular vectors of $\mathbf{K}^{1/2}\mathbf{L}\mathbf{K}^{-1/2}$, $\mathbf{V}_{\mathrm{K},n}$ is a rectangular matrix with columns composed of the first n final-time singular vectors of $\mathbf{K}^{1/2}\mathbf{L}\mathbf{K}^{-1/2}$, and $\mathbf{\Sigma}_{\mathrm{K},n}$ is a diagonal matrix composed of the first n singular values of $\mathbf{K}^{1/2}\mathbf{L}\mathbf{K}^{-1/2}$. The fastest growing part of the initial error can then be defined as

$$\mathbf{e}_{\mathrm{o}}^{n} = \mathbf{K}^{-1/2}\mathbf{L}_{\mathrm{K},n}^{-1}\mathbf{K}^{1/2}\mathbf{e}_{\mathrm{f}}. \quad (8)$$

Note from (4) and (7) that when $n = 1449$, $\mathbf{L}_{\mathrm{K},n}^{-1} = \mathbf{K}^{1/2}\mathbf{L}^{-1}\mathbf{K}^{-1/2}$, which reduces (8) to $\mathbf{e}_{\mathrm{o}} = \mathbf{L}^{-1}\mathbf{e}_{\mathrm{f}}$. Thus, the calculation in (8) is not a function of the metric $\mathbf{K}$ when all of the singular vectors are included. For larger models, the first n singular vectors are found through an iterative technique based on the Lanczos algorithm and the tangent equations of the model. Gelaro et al. (1998) have shown $\mathbf{e}_{\mathrm{o}}^{s}$ and $\mathbf{e}_{\mathrm{o}}^{n}$ to have similar spatial patterns. The similarity of the spatial patterns of these fields is not surprising given that both $\mathbf{e}_{\mathrm{o}}^{s}$ and $\mathbf{e}_{\mathrm{o}}^{n}$ are dominated by the fastest growing singular vectors; $\mathbf{e}_{\mathrm{o}}^{s}$ because of the weighting of the singular values in the sensitivity in (6) and $\mathbf{e}_{\mathrm{o}}^{n}$ because only the fastest growing singular vectors are included in the calculation. Thus, both techniques give information about the part of the initial error that grows the fastest. No knowledge is gained about the other parts of the initial error that either grow slowly, remain neutral, or actually decay.

Pu et al. (1997) have proposed a technique that may find an approximation to the full analysis error, not just the fastest growing components. They suggest that running the tangent equations with a negative time step starting with the forecast error is a good approximation to the full matrix $\mathbf{L}^{-1}$. In their experiments, they also reverse the sign of the dissipation terms to ensure numerical stability. They find that this technique yields forecast corrections as good as those using five iterations of the adjoint (Rabier et al. 1996b). They also find that this technique yields good estimates of the analysis difference in perfect-model experiments where the forecast difference is known exactly.

The characteristics of a full matrix $\mathbf{L}$ for a simple model are examined in order to answer some basic questions about the nature of $\mathbf{L}$ in the context of the abovementioned work. Are the magnitudes of the trailing singular values so close to zero that $\mathbf{L}$ is effectively not invertible? If the trailing singular values are small, would the projection of any noise or uncertainty onto these fastest decaying singular vectors cause the field to become unrealistically large when integrated backward in time? If this is true, then what is the optimal size of the pseudoinverse? The appropriateness of running the tangent equations with a negative time step as an alternative to using the full $\mathbf{L}^{-1}$ is also assessed.

3. Model description

The model used in this study is the quasigeostrophic potential vorticity model described in Marshall and Molteni (1993). The model is run with a T21 truncation and has three levels corresponding to 800, 500, and 200 hPa. The state vector has 1449 nonzero components, and therefore is small enough so that $\mathbf{L}$ can be calculated and inverted explicitly. The condition number of $\mathbf{L}$ is relatively small (on the order of 200), so that the inversion is accurate. The model forcing is composed of specified source terms of PV that are spatially varying but temporally constant and correspond to a Northern winter climatology. The model has three types of dissipative forcing; Newtonian relaxation of temperature between levels of 25 d^{-1}, an 800-hPa linear drag term that varies with topography from 3 to 1.5 d^{-1}, and horizontal scale-selective (∇^8) dissipation such that spherical harmonics with a total wavenumber of 21 are damped on a two-day timescale.

The model has been used in several previous studies

involving singular vectors. Molteni and Palmer (1993) used it to look at finite time instability. Mureau et al. (1993) used it to look at singular vectors as the basis for the initial perturbations in ensemble prediction. Barkmeijer et al. (1993) have used it to evaluate different skill prediction methods. The model is well suited for this study, since it is complex enough to capture baroclinic synoptic-scale processes important in forecast error growth, as well as being small enough so that $\mathbf{L}$ can be calculated and inverted explicitly.

4. Method

The matrix $\mathbf{L}$ is constructed by running the tangent equations 1449 times. The tangent equations are the equations of the model linearized about a nonlinear trajectory. Each time, the tangent equations operate on a state vector composed of 1448 zeros and a one, with the position of the nonzero element varying from 1 to 1449. Once $\mathbf{L}$ is computed, numerical algorithm routines are used to calculate both the inverse and the KE singular values and vectors of $\mathbf{L}$. All results shown in this paper correspond to an integration time of 48 h, and a kinetic energy metric. In the first part of this study, the full spectrum of KE singular values and the structure of growing, decaying, and neutral KE singular vectors are examined.

The second part of this study involves post-time forecast and analysis correction in a perfect-model setting. Two 48-h nonlinear integrations of the quasigeostrophic model are performed using different ECMWF analyses for the same date. As mentioned previously, one of the integrations is taken as truth and the other as the forecast. For example, a nonlinear integration starting from an ECMWF experimental 4DVAR analysis for 1200 UTC 21 January 1996, referred to as $\mathbf{f}_{4\mathrm{DVAR}}$, is taken as the truth, and a 48-h nonlinear integration starting from the operational 3DVAR analysis (Andersson et al. 1995) for the same time, $\mathbf{f}_{3\mathrm{DVAR}}$, is taken as the forecast. The terms forecast and analysis differences are used to reflect that true forecast and analysis errors are not being used here. In this setting, the analysis difference $\mathbf{e}_\mathrm{o}$ and the 48-h forecast difference $\mathbf{e}_\mathrm{f}$ are known exactly. The forecast difference is defined as

$$\mathbf{e}_\mathrm{f} = \mathbf{f}_{3\mathrm{DVAR}} - \mathbf{f}_{4\mathrm{DVAR}}. \tag{9}$$

The forecast difference is operated on with $\mathbf{L}^{-1}$ based on the 3DVAR 48-h trajectory to get an estimate of the analysis difference $\mathbf{e}_{\mathrm{o}1}$; that is,

$$\mathbf{e}_{\mathrm{o}1} = \mathbf{L}^{-1}\mathbf{e}_\mathrm{f}. \tag{10}$$

Then, $\mathbf{e}_{\mathrm{o}1}$ is subtracted from the 3DVAR analysis and a third nonlinear integration, referred to as $\mathbf{f}_1$, is run from these initial conditions. The corrected forecast difference is then given by

$$\mathbf{e}_{\mathrm{f}1} = \mathbf{f}_1 - \mathbf{f}_{4\mathrm{DVAR}}. \tag{11}$$

The above experiments are done for 12 sets of cases. Seven cases are based on differences between 3DVAR and 4DVAR analyses for 21–27 January 1996. Four cases are based on differences between optimal interpolation (OI) analysis and 3DVAR analyses for 6–9 January 1996. One case is based on the ECMWF and German operational analyses for 3 January 1994. In the last case the ECMWF operational forecast performed very poorly in comparison to the German operational forecast (Molteni et al. 1996). Unless stated otherwise, the results presented here are based on the 12-case average.

The above sets of experiments represent a situation where the forecast difference is known exactly. To simulate a more realistic situation, uncertainty, δ, is added to the forecast difference, $\mathbf{e}_\mathrm{f}$, before being operated on with $\mathbf{L}^{-1}$ such that

$$\mathbf{e}_{\mathrm{o}2} = \mathbf{L}^{-1}(\mathbf{e}_\mathrm{f} + \delta), \tag{12}$$

where δ is equal to the difference between the 4DVAR analysis and the 3DVAR analysis valid 48 h after the start of the integration time. Thus, δ represents the uncertainty in the forecast difference due to uncertainty in the final-time analysis. It is approximately the same magnitude as the analysis difference 48 h earlier. The final step is to perform a fourth nonlinear integration with the initial conditions based on the 3DVAR analysis minus $\mathbf{e}_{\mathrm{o}2}$ yielding the second corrected forecast $\mathbf{f}_2$, such that

$$\mathbf{e}_{\mathrm{f}2} = \mathbf{f}_2 - \mathbf{f}_{4\mathrm{DVAR}}. \tag{13}$$

Thus, the estimates of initial analysis difference and forecast correction with the subscript 1 refer to the unrealistic case where the forecast difference is known exactly. The estimates of initial analysis difference and forecast correction with the subscript 2 refer to the more realistic case where the forecast difference is not known exactly. It is the interaction between this uncertainty in the forecast difference and the fastest decaying vectors that is of primary interest in this study.

Both sets of experiments described above are also performed using the tangent equations with a negative time step in place of using the actual $\mathbf{L}^{-1}$. This approximation to $\mathbf{L}^{-1}$ is used by Pu et al. (1997) and is necessary for more realistic cases where the model is too large for the full matrix $\mathbf{L}$ to be calculated and inverted explicitly. Performing these experiments allows for an assessment of how close this approximate method is to using the full inverse. In addition to using a negative time step, Pu et al. (1997) also reversed the sign on the dissipative terms to ensure numerical stability. This is not necessary in our case, perhaps because of the stability of the Runge–Kutta time scheme used here. A simplified example of this approximation, hereafter referred to as the backward integration technique, for a case with only one time step, is presented below. In its simplest form, the forward time-integration of the tangent linear equations can be represented as

$$\mathbf{e}_{t+1} = \mathbf{e}_t + \Delta t\ \mathrm{TEQ}_t\mathbf{e}_t, \tag{14}$$

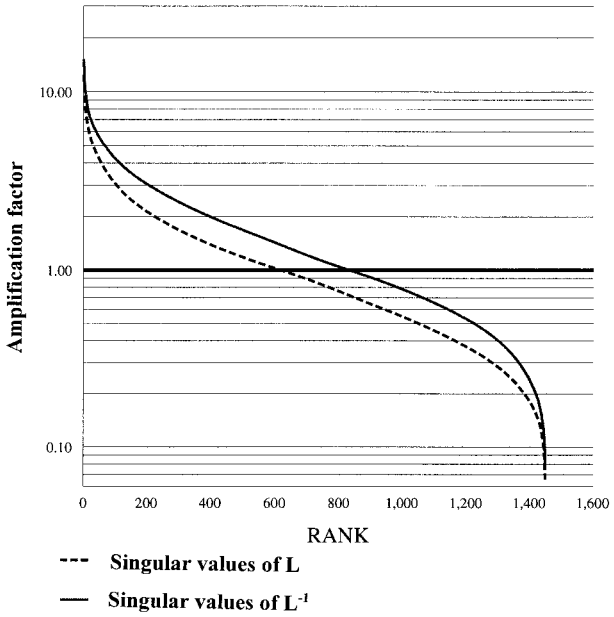


FIG. 1. The 48-h KE singular values (amplification factors) of $\mathbf{L}$ (dashed line) and $\mathbf{L}^{-1}$ (solid line) as a function of rank. Results shown are for the 12-case average.

where Δt is the time step and TEQ_t is the tangent equations linearized about a nonlinear trajectory at time t. [For most time-differencing schemes, the last term in (14) is actually more complex.] From (14) we see that

$$\mathbf{e}_t = \mathbf{e}_{t+1} - \Delta t\ \mathrm{TEQ}_t \mathbf{e}_t. \tag{15}$$

Starting from a given error at time $t + 1$, we want to find the error at time t. Instead of solving for $\mathbf{e}_t$ implicitly, we approximate (15) with

$$\mathbf{e}_{t\mathrm{approx}} = \mathbf{e}_{t+1} - \Delta t\ \mathrm{TEQ}_t \mathbf{e}_{t+1}. \tag{16}$$

The analysis difference estimates and the corrected forecast differences for these backward experiments will be referred to as $\mathbf{e}^{\mathrm{B}}_{\mathrm{o}1}$, $\mathbf{e}^{\mathrm{B}}_{\mathrm{o}2}$, $\mathbf{e}^{\mathrm{B}}_{\mathrm{f}1}$, and $\mathbf{e}^{\mathrm{B}}_{\mathrm{f}2}$.

Both sets of perfect-model experiments are also repeated using the pseudoinverse method described in the introduction and used by Buizza et al. (1997). That is, the analysis difference estimates are defined as $\mathbf{e}^n_{\mathrm{o}1} = \mathbf{K}^{-1/2}\mathbf{L}^{-1}_{\mathrm{K},n}\mathbf{K}^{1/2}(\mathbf{e}_{\mathrm{f}})$ and $\mathbf{e}^n_{\mathrm{o}2} = \mathbf{K}^{-1/2}\mathbf{L}^{-1}_{\mathrm{K},n}\mathbf{K}^{1/2}(\mathbf{e}_{\mathrm{f}} + \delta)$. The subsequent corrected forecast differences for these experiments will be referred to as $\mathbf{e}^n_{\mathrm{f}1}$ and $\mathbf{e}^n_{\mathrm{f}2}$. As before, n refers to the number of KE singular vectors included in $\mathbf{L}^{-1}_{\mathrm{K},n}$ and ranges from 15 to 1449.

5. Growing, decaying, and neutral singular vectors

a. Structure of singular vectors

The full set of KE singular values (amplification factors) for $\mathbf{L}$ and $\mathbf{L}^{-1}$ averaged over the 12 cases is shown in Fig. 1. The average ratio of the largest to smallest singular value is 197, ranging from 125 to 279 for individual cases. These relatively small condition numbers indicate that $\mathbf{L}$ can be inverted accurately. On average, 618 out of the 1449 singular values are greater than 1. This means that, for a 48-h integration time, about 43% of the singular vectors grow while 57% of the singular vectors decay with time. The exact percentage of singular values that are greater than one will vary with the trajectory and the metric. These percentages are also dependent on the values of the dissipative parameters in the quasigeostrophic model. Also note that more than 850 singular values are between 2.0 and 0.5, indicating that the magnitude of 60% of all singular vectors changes by less than a factor of 2. Lorenz (1965) and Farrell (1990) point out that the mean square of the singular values represents the average variance perturbation growth, given that the initial error projects equally onto all the initial singular values. The mean square of the KE singular values for $\mathbf{L}$ averaged for the 12 cases is 3.01, with the individual cases varying from 2.82 to 3.16. This implies that the rms of the forecast difference due to the linear growth of the analysis difference would increase by a factor of 1.73 over two days.

From (4) we can see that the nth KE singular value of $\mathbf{L}^{-1}$ is the inverse of the KE singular value of rank $1450 - n$ of $\mathbf{L}$. (Hereafter, rank refers to the position of the singular value when ranked in descending order). The KE singular values of $\mathbf{L}^{-1}$, also shown in Fig. 1, are larger than those of $\mathbf{L}$. The average mean square of the singular values of $\mathbf{L}^{-1}$ is 5.17, implying that the rms of the analysis difference when operated on by $\mathbf{L}^{-1}$ would increase by a factor of 2.27. Thus, a field that projects onto all the initial and final-time singular vectors equally will grow faster going backward in time than when going forward in time. As dissipation in the model is increased, all KE singular values of $\mathbf{L}$ decrease (therefore all values of $\mathbf{L}^{-1}$ increase) and the gap between these two curves increases. Likewise, as dissipation is decreased, these two curves approach each other, and the number of growing vectors approaches one-half of the total [see Pu et al. (1997) for a discussion of singular values for Hamiltonian systems]. The percentage of singular values that are greater than one is similar to that found by Farrell (1990) for the case of an undamped baroclinic shear flow. However, other studies have found the percentage of singular vectors that are growing to be far smaller (Borges and Hartmann 1992; Buizza et al. 1993; Ehrendorfer and Errico 1995; Sardeshmukh et al. 1997). The relatively larger percentage of growing singular vectors found in this study may be due to the small amount of dissipation present in the model version used here.

Figure 2 shows the fastest growing KE singular vector converted to geopotential height perturbations for one case. The lengths of both the initial-time and final-time singular vector have been normalized so that the total kinetic energy is 1. The singular value (amplification factor) for this vector is 11.2. At initial time, the singular vector has very little energy at 200 hPa, and the anom-

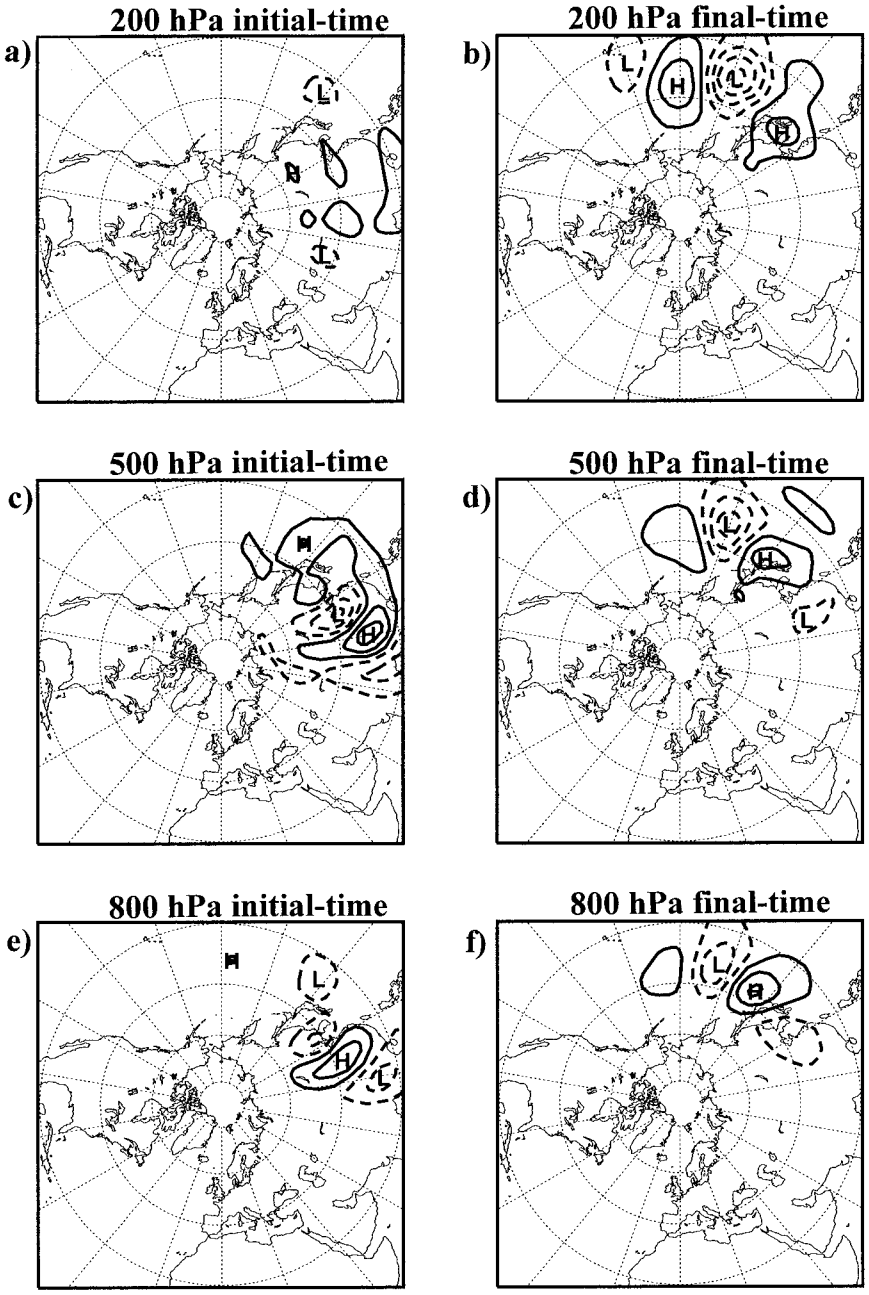


FIG. 2. The geopotential height perturbations associated with the first KE singular vector for (a) initial time at 200 hPa; (b) final time at 200 hPa; (c) initial time at 500 hPa; (d) final time at 500 hPa; (e) initial time at 800 hPa; (f) final time at 800 hPa. Both final and initial KE singular vectors have normalized length and the contour interval is constant for all six panels.

alies at the lower levels exhibit a westward tilt with height (vertical cross sections confirm this phase tilt). At final time, the singular vector has much more energy at 200 hPa, and the westward tilt with height has been reduced. The perturbations at initial time also appear to be of smaller spatial scale than those at final time. These characteristics are very similar to those found in previous studies examining the nature of the fastest growing singular vectors (Molteni and Palmer 1993; Buizza and Palmer 1995).

The geopotential height perturbations corresponding to the fastest decaying KE singular vector for the same case are shown in Fig. 3. This singular vector decays by a factor of 0.07. In contrast to the fastest growing singular vector shown in Fig. 2, this vector shows a significant decrease in energy at 200 hPa between initial

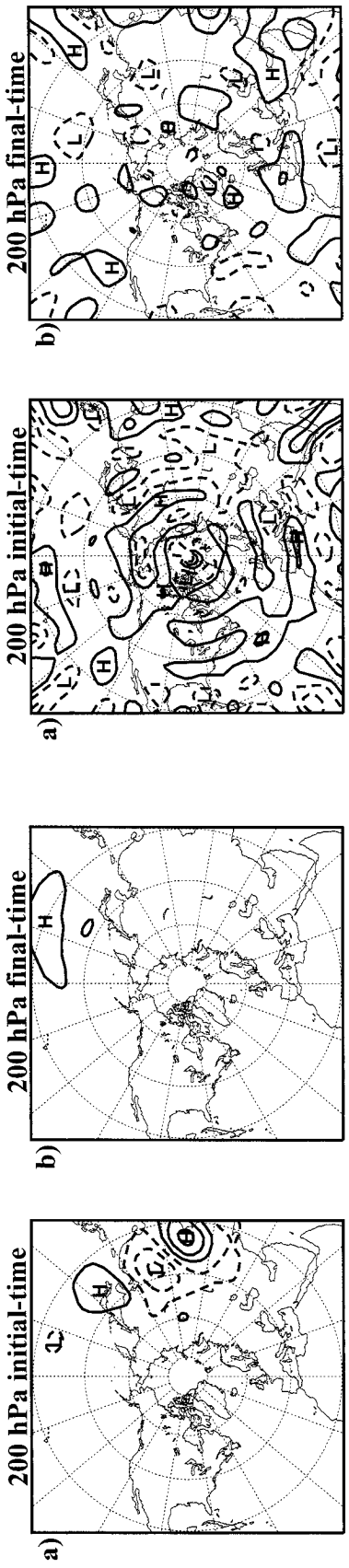


FIG. 3. Same as Fig. 2 but for the (fastest decaying) KE singular vector number 1449.

FIG. 4. Same as Fig. 2 but for the near-neutral KE singular vector number 600.

and final times. The vertical structure exhibits an eastward tilt with height that is more pronounced at final time than at initial time. The decay associated with an eastward-tilted midlatitude perturbation may be understood through the pseudomomentum-conserving Rossby wave theory used to explain the growth of westward-tilted perturbations in Buizza and Palmer (1995) under conditions where the WKBJ approximation is valid. In the case of an energetically amplifying Rossby wave packet, a westward phase tilt with height below the jet maximum leads to the group velocity being focused toward the jet core. Buizza and Palmer (1995) explain that this focusing of the pseudomomentum, or wave activity, into the jet leads to an increase in the intrinsic frequency and therefore energy growth, while the propagation and refraction of the wave packet into the jet leads to a decrease in the vertical tilt with time. In the case of these decaying singular vectors, the eastward phase tilt with height leads to the group velocity being focused away from the jet core (downward), which leads to energy decay, whereas the propagation and refraction of the wave packet away from the jet leads to an enhancement of the eastward tilt with time.

Alternatively, the rapid decay associated with an eastward-tilted system may be explained using the same quasigeostrophic reasoning used to explain the growth of westward-tilted disturbances (Holton 1979). In the case of a westward-tilted disturbance, the cold air advection that decreases with height to the east (west) of the surface high (low) would act to deepen the upper-level trough. In an eastward-tilted disturbance, the cold air advection that decreases with height in front of the surface high would now act to weaken the upper-level ridge. Also, the differential relative vorticity advection acts to propagate the upper-level disturbance eastward at a faster rate than the lower-level disturbance. In a westward-tilted disturbance, this is countered by the vorticity stretching term, which acts to slow the eastward propagation at upper levels and enhance it at lower levels, thus acting to maintain the westward tilt. In an eastward-tilted disturbance, the vorticity stretching term is no longer positioned correctly to provide this counterbalance, thus allowing the eastward tilt to become more pronounced with time.

The height perturbations corresponding to KE singular vector number 600, which is almost neutral (amplification factor of 1.03), are shown in Fig. 4. This vector, as with all other near-neutral vectors examined, exhibits a noisy small-scale pattern that does not have any spatially localized structure or any coherence between levels. Ehrendorfer and Errico (1995) also find that the singular vectors for a mesoscale model lose their localized structures as the amplification factors approach neutrality.

Figure 5 shows the kinetic energy as a function of total wavenumber averaged for the 10 fastest growing, fastest decaying, and near-neutral normalized KE singular vectors. The growing vectors (Figs. 5a,b) indicate an upward and upscale transfer of energy between initial and final times, again similar to previous results (Buizza and Palmer 1995). The decaying vectors (Figs. 5c,d) indicate a downward transfer of energy along with a slight downscale transfer of energy, although the scale separation between final and initial times is less dramatic for the decaying vectors than it is for the growing vectors. Figures 5e,f indicate that the neutral vectors contain the most energy at the smaller scales. The kinetic energy distributions described above do not change qualitatively as the dissipation is decreased. However, if the dissipation is doubled, then the trailing singular vectors change dramatically. At both initial and final times, the trailing singular vectors no longer exhibit any spatially localized structure. Instead they display globalwide, small-scale patterns and almost all their energy is contained in total wavenumbers 18–21.

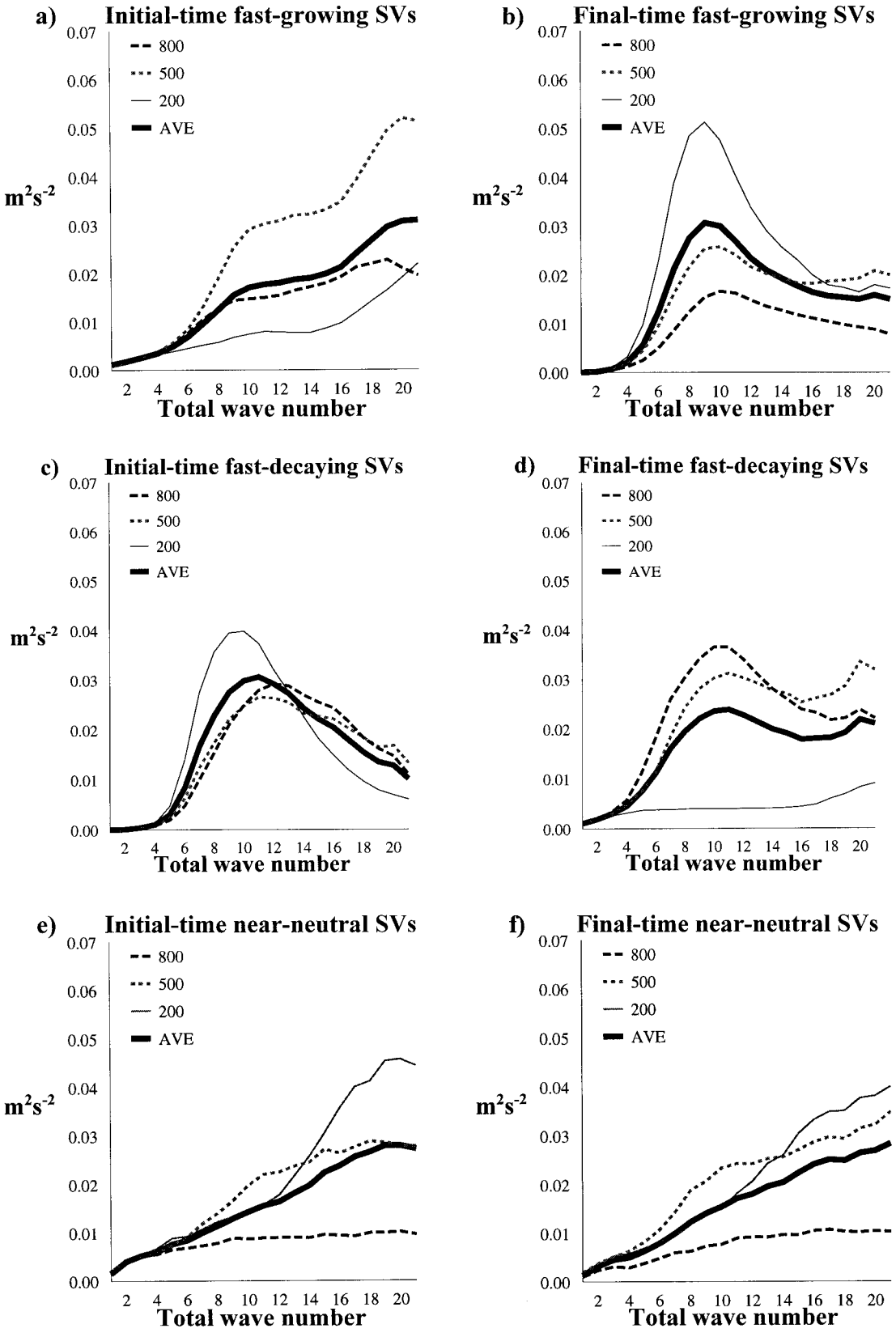
b. Projection of difference fields onto singular vectors

One of the premises for using singular vectors as a method for determining the initial perturbations for ensemble forecasts is that the analysis difference is approximately isotropic with respect to the initial-time singular vectors. Even though this will not be strictly true for individual cases, what is important is that the analysis errors do not have an anomolously very small projection onto the leading singular vectors. Gelaro et al. (1998) show that the projection of the forecast errors onto the first 30 final-time singular vectors used in the operational ensemble prediction system at ECMWF are approximately proportional to their singular values. They point out that this implies that the initial analysis errors are approximately isotropic with respect to these leading initial-time singular vectors, assuming short-term error growth is well approximated by the tangent forward propagator.

In this study, since the initial analysis ''error'' is really a prescribed analysis difference, the projection of these initial differences onto the singular vectors can be calculated explicitly. Figure 6 shows the magnitude of the projection of the analysis differences onto the corresponding initial-time KE singular vectors. The thin

→

FIG. 5. The kinetic energy as a function of total wavenumber averaged for (a) KE singular vectors 1–10 at initial time; (b) KE singular vectors 1–10 at final time; (c) KE singular vectors 1440–1449 at initial time; (d) KE singular vectors 1440–1449 at final time; (e) KE singular vectors 621–630 at initial time; (f) KE singular vectors 621–630 at final time. Kinetic energy is shown at 800 hPa (dashed curves), 500 hPa (dotted curves), 200 hPa (thin solid curve), and three-level average (thick solid curve). Results are shown for the 12-case averages.



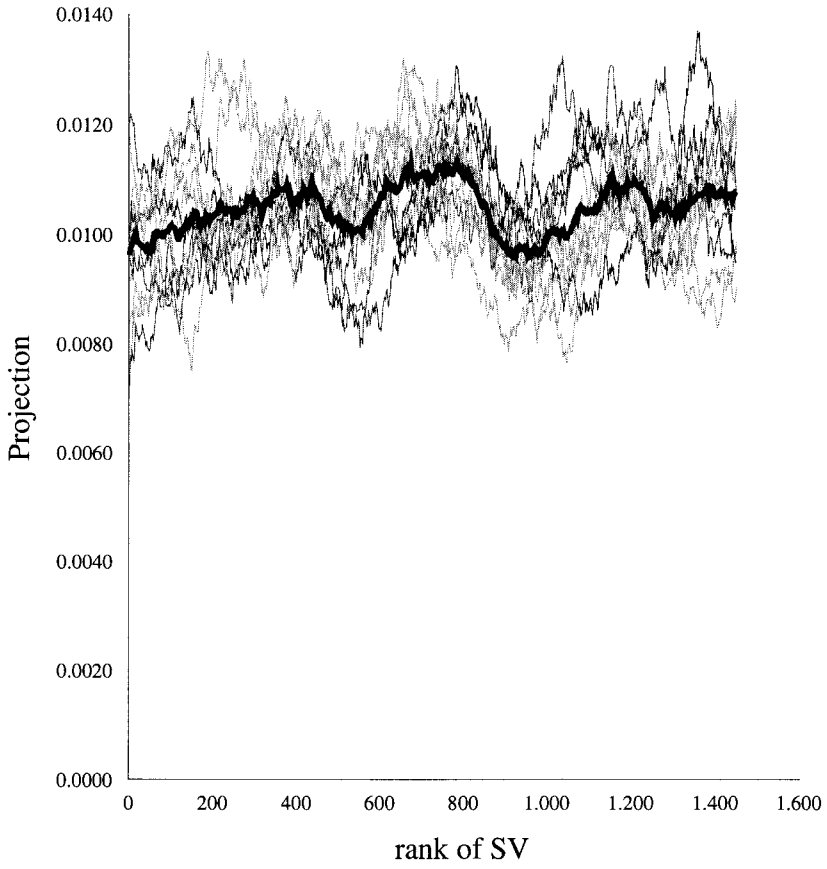


FIG. 6. The projection of the analysis differences onto the initial-time KE singular vectors as a function of rank. Thin curves represent individual cases; the thick black curve shows 12-case average. All curves have been smoothed using a 100-member running mean.

curves represent the individual cases and the thick curve indicates the 12-case mean. All curves have also been smoothed using a 100-member running mean. One can see that there is much case-to-case variability. Examination of the unsmoothed curves (not shown) indicates that the projection onto singular vectors varies widely from one vector to another. There is a slight suggestion of a positive trend with rank. However, the large degree of case-to-case variability implies that any structure in the 12-case mean may be due to sampling error. In any case, it appears that the projection of the analysis difference onto the initial-time singular vectors is not a strong function of rank. [There is some indication that the 3DVAR-4DVAR analysis differences do have a larger projection onto the fastest growing singular vectors than the OI-3DVAR analysis differences on average. This is consistent with the results of Thépaut et al. (1996), who find that the analysis increments with 4DVAR have more shortwave energy than the analysis increments with 3DVAR.] There is no a priori expectation that the projection of analysis error onto the singular vectors, if known, would be a strong function of the rank of the singular values either, even though the structures of the analysis differences and the true analysis errors are probably significantly different.

The analysis differences are approximately isotropic with respect to the growing and decaying initial-time KE singular vectors. Therefore, one would expect that the forecast differences would have a strong projection onto the growing vectors and a weak projection onto the decaying vectors at final time. Figure 7 shows the magnitude of the projection onto the final-time KE singular vectors of the exact forecast difference, the forecast difference plus analysis uncertainty, and the analysis uncertainty itself, averaged over the 12 cases. All three fields have been normalized so that the mean square of the projection of all three fields is the same. As expected, Fig. 7 indicates that the strongest projection of the forecast difference is onto the fastest growing vectors, and the weakest projection is onto the fastest decaying vectors. Note that for perfectly linear error growth and an initial error field that is isotropic with respect to the initial-time singular vectors, this curve would look like the curve of singular values. However, when uncertainty is added to the forecast difference, the projection onto the fastest growing vectors is decreased,

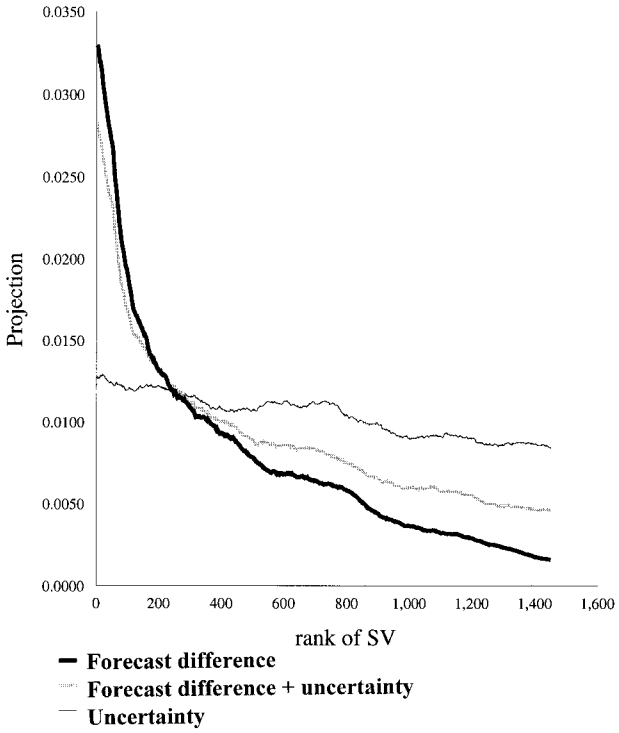


FIG. 7. The projection onto the final-time KE singular vectors as a function of rank of exact forecast difference (thick solid curve), forecast difference with analysis uncertainty (thick gray curve), and analysis uncertainty (thin black curve). All three fields have normalized length. All curves are for 12-case averages.

while the projection onto the fastest decaying vectors is increased. Note that the projection onto the fastest growing vectors is decreased by approximately 15%, while the projection onto the fastest decaying vectors is increased by almost 150%. The projection of the analysis differences onto the KE singular vectors at final time exhibits a small negative trend with rank. This may be due to the extremely small amount of energy at 200 hPa in the trailing singular vectors at final time. However, the curve is still relatively flat compared to the forecast difference curves.

Given that the forecast differences have the strongest projections onto the fastest growing singular vectors, it is expected that the spectral energy distribution of the forecast differences would look more like that of the leading singular vectors at final time than would the spectral energy distribution of the analysis differences. The kinetic energy as a function of total wavenumber for the analysis differences and the forecast differences is shown in Fig. 8. One can see that the forecast differences show an increase of energy at 200 hPa, as well as an increase in energy at wavenumbers 10–12, over the energy profile of the analysis differences. These are similar to the changes that occur for the fastest growing singular vectors between initial and final time (Fig. 5).

6. Forecast and analysis correction techniques

a. Full inverse and backward integration techniques

Results are now presented for the analysis and forecast correction techniques using both the full $\mathbf{L}^{-1}$ and the method of running the tangent equations with a negative time step (the backward integration technique).

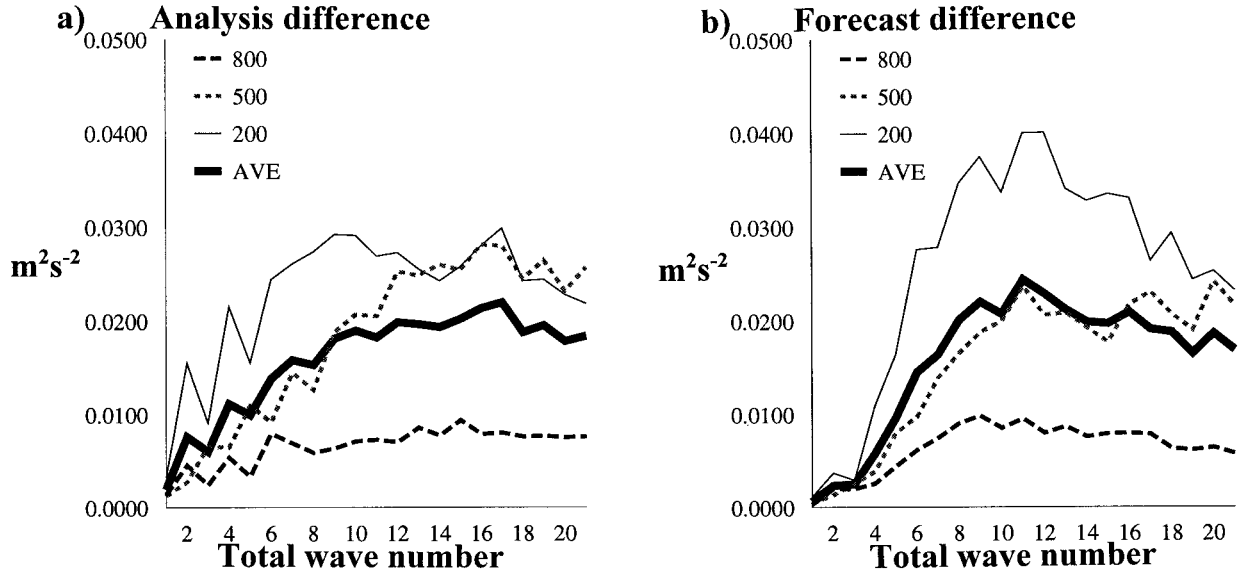


FIG. 8. The kinetic energy as a function of total wavenumber for (a) analysis differences and (b) exact forecast differences for 800 hPa (dashed), 500 hPa (dotted), 200 hPa (thin solid), and three-level average (thick solid). Both fields have normalized length and are for 12-case averages.

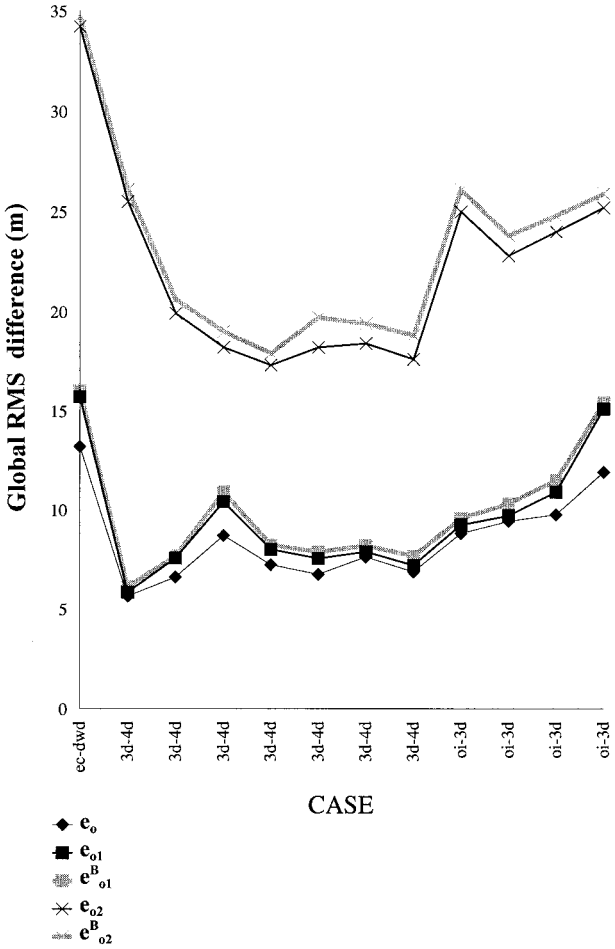


FIG. 9. The 500-hPa global rms (m) for each of the 12 cases for the analysis difference $\mathbf{e}_o$ (black diamonds); the full inverse estimate $\mathbf{e}_{o1}$ based on the exact forecast difference (black squares); the backward integration estimate $\mathbf{e}^B_{o1}$ based on the exact forecast difference (gray squares); the full inverse estimate $\mathbf{e}_{o2}$ based on the uncertain forecast difference (black crosses); the backward integration estimate $\mathbf{e}^B_{o2}$ based on the uncertain forecast difference (gray crosses).

Figure 9 shows the prescribed analysis difference ($\mathbf{e}_o$), as well as the estimates of the analysis difference, for each of the 12 cases. The estimates of analysis difference as derived from the full inverse technique [$\mathbf{e}_{o1}$ and $\mathbf{e}_{o2}$ as given by (10) and (12)], and the corresponding estimates derived from the backward integration technique ($\mathbf{e}^B_{o1}$ and $\mathbf{e}^B_{o2}$) are both shown. Because of its common usage in forecast error statistics, these differences are shown in terms of the corresponding 500-hPa rms. Figure 9 shows that both the full inverse and backward integration techniques result in estimates of analysis difference that are very close to the observed magnitude of the prescribed analysis difference, although slightly too large. The rms of $\mathbf{e}_o$ and $\mathbf{e}_{o1}$ have an average percent difference of 45%. These differences are due solely to nonlinearities in the forward 48-h integrations. The percent differences between $\mathbf{e}_{o1}$ and $\mathbf{e}^B_{o1}$ are less than 10%, indicating that the backward integration technique is a very good approximation to the full inverse technique.

TABLE 1. Forward and backward growth of the global, three-level average rms of analysis uncertainty over 48 h as a function of dissipation. Results are also shown for the cases where the dissipation terms are negative in the backward integration. Results are for the 12-case average.

	Forward growth	Backward growth	Backward growth negative dissipation
Strong dissipation	1.37	2.43	1.33
Control dissipation	1.56	2.00	1.49
Weak dissipation	1.82	1.77	1.72

When analysis uncertainty is added to the forecast difference, then the estimates of analysis difference degrade considerably. Figure 9 shows that $\mathbf{e}_{o2}$ and $\mathbf{e}^B_{o2}$ are on average more than twice as big as $\mathbf{e}_o$, although $\mathbf{e}^B_{o2}$ is still a very good approximation to $\mathbf{e}_{o2}$. These results show that both the full inverse and the backward integration technique give good approximations of the analysis difference when the forecast difference is known exactly. However, these techniques do not provide a useful estimate of analysis difference in the more realistic case where analysis uncertainties result in the forecast difference being known only approximately. As shown in the next subsection, this is a result of the spurious projection of the analysis uncertainty onto the trailing singular vectors, which grow by more than an order of magnitude when operated on by $\mathbf{L}^{-1}$.

Table 1 illustrates the fact that the error growth due to analysis uncertainty is larger when going backward in time than when going forward in time. The first column of Table 1 is the rms of the nonlinear forecast differences divided by the rms of the analysis differences (here the rms is computed based on the three-level global geopotential height fields). This represents the nonlinear growth of the analysis uncertainty in the forward direction, which, at this 48-h time interval, is very similar to the linear error growth. For the experiments with control dissipation, the average rms growth over the two-day interval is 1.56 for the 12 cases shown here. This growth is comparable to the 1.73 amplification factor predicted from the rms of the singular values (Lorenz 1965; Farrell 1990) that would occur for linear error growth given that the initial perturbation is perfectly isotropic with respect to the initial-time singular vectors. It is noteworthy that the average amplification factor for the 3DVAR−4DVAR differences is 1.61, whereas the average amplification factor for the OI−3DVAR differences is 1.46. This may indicate that the 4DVAR analyses are correcting some of the fastest growing components of the error in the 3DVAR analyses. The limited sample size of this study allows only for speculation on this point as yet. However, these results are consistent with those of Rabier et al. (1996a),

TABLE 2. The global, three-level average of the geopotential height rms of (estimated $\mathbf{e}_0$ − true $\mathbf{e}_0$) as a function of dissipation. Results are also shown for the cases where the dissipative terms are negative in the backward integration. Units are m. Results are for the 12-case average.

	$\mathbf{e}_0$	$(\mathbf{e}^B_{01} - \mathbf{e}_0)$ no uncert.	$(\mathbf{e}^B_{01} - \mathbf{e}_0)$ no uncert. neg. diss.	$(\mathbf{e}^B_{02} - \mathbf{e}_0)$ with uncert.	$(\mathbf{e}^B_{02} - \mathbf{e}_0)$ with uncert. neg. diss.
Strong dissipation	10.8	4.95	7.57	26.08	16.51
Control dissipation	10.8	5.22	6.39	21.88	17.41
Weak dissipation	10.8	5.73	5.62	19.68	19.15

who find similarities between the 4DVAR−3DVAR analysis difference and the related sensitivity pattern.

The second column in Table 1 is the rms of the backward integration of the analysis uncertainty divided by the rms of the analysis uncertainty. This represents the linear growth of the analysis uncertainty in the backward direction. The average amplification factor for the control dissipation cases is 2.0. This is also in fairly close agreement with the value of 2.27 predicted by the rms of the singular values. It is this backward growth of the analysis uncertainty that causes the estimate of the analysis difference $\mathbf{e}_{o2}$ to become so unrealistically large. In order to approximate more closely the method described by Pu et al. (1997), the backward integration is also performed with the sign of the dissipative terms reversed. The third column of Table 1 shows that reversing the sign of the dissipative terms decreases the backward error growth from 2.0 to 1.49 for control dissipation.

When the strength of the dissipative forcing is increased by a factor of 2, then the forward amplification factor over 48 h decreases to 1.37, whereas the backward amplification factor increases to 2.43. When the dissipative forcing in the model is weakened by an order of magnitude, the forward and backward amplification factors over the 48-h time interval are 1.82 and 1.77, respectively. The trend of the forward (backward) amplification factor decreasing (increasing) as dissipation is strengthened is consistent with the fact that all the singular values of $\mathbf{L}$ decrease (and all singular values of $\mathbf{L}^{-1}$ increase) as dissipation is increased. Note that for the negative dissipation cases, the backward amplification factor decreases as the magnitude of the dissipation is increased.

Reversing the sign of the dissipative terms in the backward integration will result in a less accurate approximation of $\mathbf{L}^{-1}$. This is because the forward and backward tangent equations [TEQ in (14) and (15)] are no longer identical. However, negative dissipative terms also result in a smaller backward growth of uncertainties, and therefore may have a beneficial effect. Table 2, showing the rms of $(\mathbf{e}^B_{o1} - \mathbf{e}_o)$ and the rms of $(\mathbf{e}^B_{o2} - \mathbf{e}_o)$ for the positive and negative dissipation cases, confirms this for the control and strong dissipation cases. For the case where the forecast difference is known exactly, the best estimate of the initial analysis difference, $\mathbf{e}_o$, is obtained by using the backward integration technique with positive dissipation, which is a closer approximation to $\mathbf{L}^{-1}$. However, for the case where uncertainty is added to the forecast difference, then the backward integration technique with negative dissipation provides a less poor estimate of $\mathbf{e}_o$, even though the error in the estimate is still prohibitively large. For weak dissipation, a change in the sign of the dissipation terms does not have a significant effect on the analysis difference estimate.

The estimates of analysis differences shown in Fig. 9 are used to correct the initial conditions from which another nonlinear forecast is run. Figure 10 shows the uncorrected forecast difference ($\mathbf{e}_f$) as well as the forecast difference corrected using the inverse technique ($\mathbf{e}_{f1}$ and $\mathbf{e}_{f2}$) and the backward integration technique ($\mathbf{e}^B_{f1}$ and $\mathbf{e}^B_{f2}$) for the 12 cases. When the forecast difference is known exactly, the rms of both $\mathbf{e}_{f1}$ and $\mathbf{e}^B_{f1}$ are approximately 79% less than the uncorrected forecast difference on average, with $\mathbf{e}^B_{f1}$ only slightly larger than $\mathbf{e}_{f1}$ in almost all cases. When uncertainty is added to the forecast difference, both full inverse and backward integration techniques still result in a significant reduction in forecast difference in most cases ($\mathbf{e}_{f2}$ and $\mathbf{e}^B_{f2}$ are both about 21% less than $\mathbf{e}_f$ on average). However, the improvement is significantly degraded from the results where the forecast difference is known exactly. Significant forecast improvement is still possible, given the very large amplitude of $\mathbf{e}_{o2}$ and $\mathbf{e}^B_{o2}$, because the singular vectors that grow when going backward in time, now decay when going forward in time.

Table 3 summarizes the result for the forecast difference corrections obtained using both positive and negative dissipative terms in the backward integration. As in Table 2, the average global, three-level geopotential height rms of the corrected forecast differences is shown as a function of the magnitude of the dissipation. For the case where the forecast differences are known exactly, the best correction is obtained by using the best approximation to $\mathbf{L}^{-1}$—that is, the backward integration with the positive dissipative terms. However, for the case where the forecast difference is known inexactly, the damping effect of the negative dissipation terms on the backward growth of uncertainties results in a better forecast correction. Again, the differences between the two backward integration techniques are minimal when the magnitude of the dissipation is small.

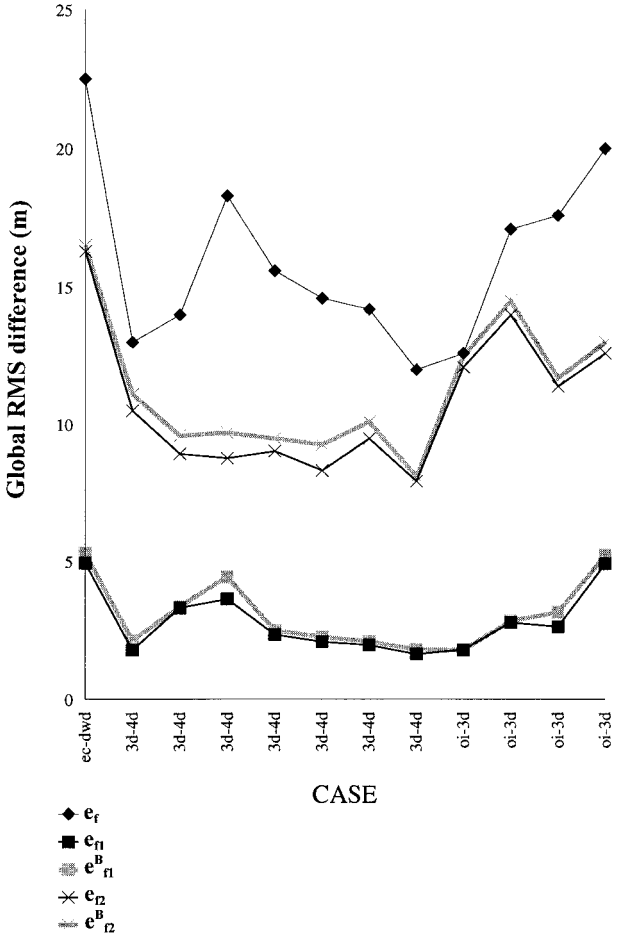


FIG. 10. The 500-hPa Z global rms (m) for the 12 cases of the uncorrected forecast difference $\mathbf{e}_f$ (black diamonds); the inverse-corrected forecast difference $\mathbf{e}_{f1}$ based on the exact forecast difference (black squares); the backward integration–corrected forecast difference $\mathbf{e}^B_{f1}$ based on the exact forecast difference (gray squares); the inverse-corrected forecast difference $\mathbf{e}_{f2}$ based on the uncertain forecast difference (black crosses); the backward integration–corrected forecast difference $\mathbf{e}^B_{f2}$ based on the uncertain forecast difference (gray crosses).

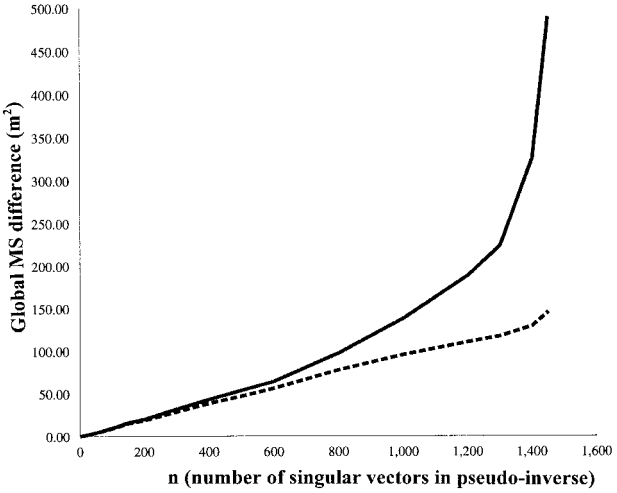


FIG. 11. The global geopotential height variance (m^2) of the estimated analysis difference as a function of the number of KE singular vectors included in the pseudoinverse. Dashed curve indicates $\mathbf{e}^n_{o1}$ where the forecast difference is known exactly. Solid curve represents $\mathbf{e}^n_{o2}$ where uncertainty is included in the forecast error. Results are for the 12-case averages.

b. Pseudoinverse technique

Results are now presented for the estimates of analysis differences $\mathbf{e}^n_{o1}$ and $\mathbf{e}^n_{o2}$ based on the pseudoinverse technique (Buizza et al. 1997) averaged over the 12 cases. The reader is reminded that $\mathbf{e}^n_{o1}$ is found by operating on $\mathbf{e}_f$ using the pseudoinverse that contains the first *n* KE singular vectors only. Figure 11 shows the variance of $\mathbf{e}^n_{o1}$ and $\mathbf{e}^n_{o2}$ as a function of *n.* Results are now presented in terms of variance rather than rms. This is done in order to illustrate how the variance explained by including more and more singular vectors in the pseudoinverse increases linearly as a function of *n* when the forecast difference is known exactly. This is expected from the relatively flat projection of the analysis differences onto the KE singular vectors at initial time as shown in Fig. 6. When uncertainty is added to the forecast difference, then the analysis uncertainty estimates become exponentially larger as the decaying vectors are added to the pseudoinverse. One can see that the higher the rank of the decaying singular vectors included in the pseudoinverse, the larger the detrimental effect is on the estimate of $\mathbf{e}_o$.

Figure 12 shows the variance of $\mathbf{e}_o - \mathbf{e}^n_{o1}$ and $\mathbf{e}_o - \mathbf{e}^n_{o2}$. The fact that $\mathbf{e}_o - \mathbf{e}^n_{o1}$ reaches a minimum when slightly less than the full 1449 vectors are used is due to the effects of nonlinearities projecting onto the trailing singular vectors. For perfectly linear error growth, $\mathbf{e}_o - \mathbf{e}^n_{o1}$ would go to zero as *n* goes to 1449. For the

TABLE 3. The global, three-level average of the geopotential height rms of the uncorrected and corrected forecast difference as a function of dissipation. Results are also shown for the cases where the dissipative terms are negative in the backward integration. Units are m. Results are for the 12-case average.

	Fcst diff $\mathbf{e}_f$	Corrected fcst diff $\mathbf{e}^B_{f1}$ no uncert.	Corrected fcst diff $\mathbf{e}^B_{f1}$ no uncert. neg. diss.	Corrected fcst diff $\mathbf{e}^B_{f2}$ with uncert.	Corrected fcst diff $\mathbf{e}^B_{f2}$ with uncert. neg. diss.
Strong dissipation	14.76	2.90	4.98	13.48	11.64
Control dissipation	16.84	3.45	4.67	13.45	12.27
Weak dissipation	19.64	4.29	4.20	13.75	13.50

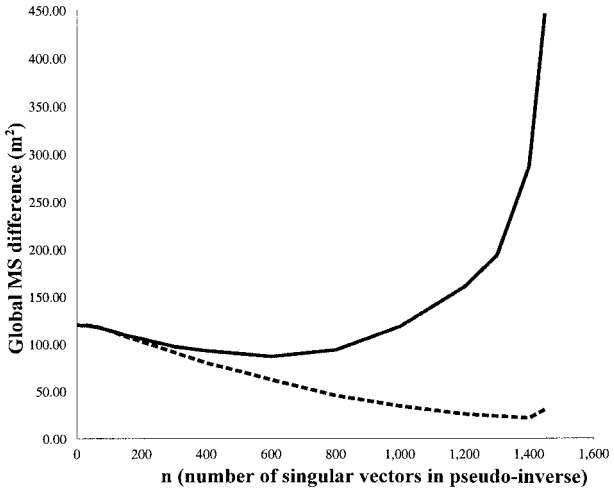


FIG. 12. The global geopotential height variance (m^2) of the true analysis difference minus the estimated analysis difference as a function of the number of KE singular vectors included in the pseudoinverse. The dashed curve indicates ($\mathbf{e}_o - \mathbf{e}^n_{o1}$) where the forecast difference is known exactly. The solid curve represents ($\mathbf{e}_o - \mathbf{e}^n_{o2}$) where uncertainty is included in the forecast difference. Results are for the 12-case averages.

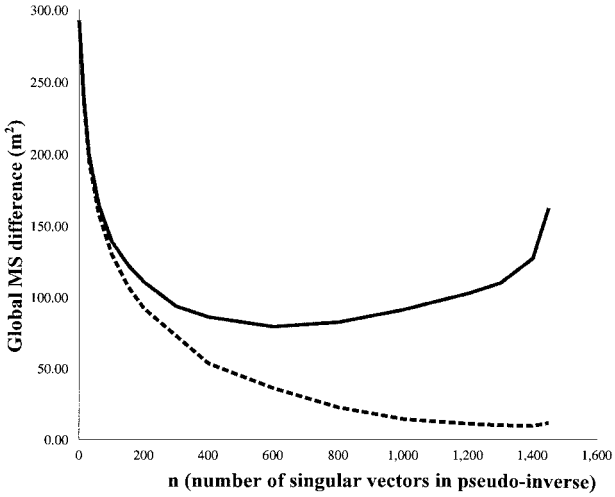


FIG. 13. The global geopotential height variance (m^2) of the corrected forecast difference as a function of the number of KE singular vectors included in the pseudoinverse. The dashed curve indicates $\mathbf{e}^n_{f1}$ where the forecast difference is known exactly. The solid curve indicates $\mathbf{e}^n_{f2}$ where uncertainty has been added to the forecast difference. Results are for the 12-case averages.

case where the forecast difference is known only inexactly, the best analysis estimate comes from using the first 600 KE singular vectors only. As more decaying singular vectors are included in the pseudoinverse, estimates of analysis difference become worse. The largest errors are introduced by the inclusion of the fastest decaying singular vectors. These results are consistent with those of Klinker et al. (1998), who perform analysis correction experiments using an iterative minimization technique involving the adjoint. They find that for the first few iterations of their minimization of the forecast error, the fit of the analysis to the observations improves. However, for 10 iterations, the fit of the analysis to the observations becomes worse. They hypothesize that this is due to the inversion of inaccuracies in the forecast error field. The point at which further iterations no longer result in an improved analysis may be analogous to the inclusion of the decaying singular vectors in the pseudoinverse technique.

Figure 13 shows the variance of the corrected forecast difference $\mathbf{e}^n_{f1}$ and $\mathbf{e}^n_{f2}$, again as a function of *n.* For the case where the forecast difference is known exactly, the best correction comes from using the full inverse (or at least something close to it). However, unlike the estimates of analysis difference, the decrease in forecast difference variance is not linear. The correction per vector decreases as the rank of the singular vector increases. In fact, the variance of the forecast difference can be cut in half by using only the first 100 singular vectors (less than 7% of the total). When analysis uncertainty is added to the forecast difference, the best correction no longer comes from using the full inverse, but rather from using only the first 600 (i.e., only the growing) KE singular vectors in the pseudoinverse. As with estimates of the analysis difference, the fastest decaying singular vectors have the most detrimental effect on the forecast correction. However, it is encouraging to note that the inclusion of this uncertainty does not seem to have a significant adverse effect on the forecast correction until more than 200 or so singular vectors are included in the pseudoinverse. For the more realistic case where uncertainty is included in the forecast difference, the full inverse (pseudoinverse with $n = 1449$) and backward integration techniques are not the most optimal types of correction. However, these forecast corrections are as good as those obtained by using the pseudoinverse with the first 60 to 100 singular vectors, at considerably less expense.

c. Local projection operators

All of the results presented so far have been for global error growth. Oftentimes the region of interest is considerably smaller, such as the Northern Hemisphere extratropics, or the region over Europe. Following Barkmeijer (1992) and Buizza (1994) a local projection operator is applied to the calculation of the singular vectors. This is done by finding the KE singular values not of $\mathbf{L}$, but of $\mathbf{PL}$, where the local projection operator $\mathbf{P} = \mathbf{S}^{-1}\mathbf{GS}$. Here $\mathbf{S}$ is the matrix representation of the spectral to grid point transform, and $\mathbf{G}$ is a diagonal matrix with diagonal elements that are 1 within the region of interest and 0 outside this region. Note that the full inverse of $\mathbf{PL}$ does not exist, and that a number of the trailing singular values will be effectively zero, this number being a function of the size of the projection area. Applying this projection operator constrains the

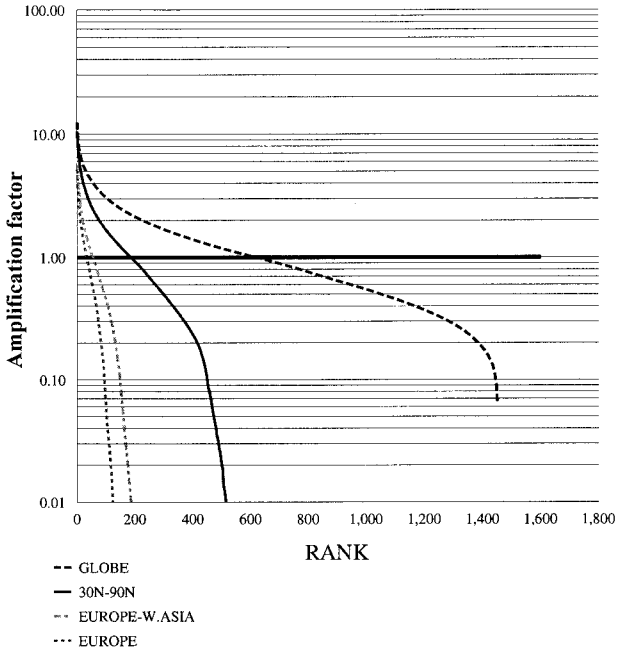


FIG. 14. The KE singular values for the global domain (dashed line) and for the northern extratropical (solid), Europe–western Asian (dot–dashed), and European (dotted) local projection operators. Results are for the 12-case averages.

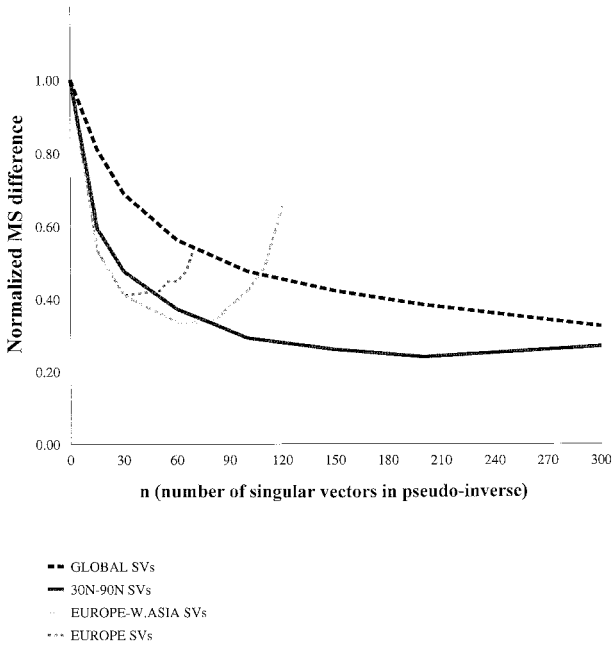


FIG. 15. The variance of the corrected forecast difference normalized by the variance of the uncorrected forecast difference—i.e., $\mathbf{e}_{f2}^{n}/\mathbf{e}_{f}$ as a function of n, the number of KE singular vectors included in the pseudoinverse. The normalized corrected forecast difference is shown for the global region (dashed), northern extratropical region (solid), Europe–western Asian region (dot–dashed), and European region (dotted). The forecast corrections have been made using the KE singular vectors for that region. Results are for the 12-case averages.

final-time singular vectors to optimize perturbation growth within the region of interest. The initial-time singular vectors may occur outside that region though. In this study three local projection operators are applied. The Northern Hemisphere extratropical region includes everything north of 30°N. The Europe–western Asian region is bounded by 30°W, 80°E, 30°N, and 80°N. The European region is bounded by 30°W, 40°E, 30°N, and 80°N.

The KE singular values averaged for the 12 cases for these three regions of interest, along with the global KE singular values, are shown in Fig. 14. The lead singular value decreases as the area of the target region decreases. The lead singular value decreases from an average of 12.4 for the globe, to 10.5 for the northern extratropics, to 5.7 and 5.1 for the Europe–western Asian and European regions, respectively. This is not surprising given that the application of the local projection operator acts as a constraint on where error growth is maximized. (This would not always necessarily be the case. If the fastest growing global singular vector at final time is completely contained within the region of interest, the leading global and regional singular values would be the same.) The number of growing vectors also decreases as the area of the target region decreases. Note that the ratio between the number of growing vectors for the local projection and the number of growing vectors for the globe is roughly equal to (somewhat less than) the ratio between the area of the local projection and the area of the globe. For example, the northern extratropical region is one-third the area of the globe in gridpoint space, whereas the number of northern extratropical growing vectors (188) is roughly one-third of the number of growing vectors over the globe (618).

Figure 15 shows the normalized corrected forecast difference variance as a function of the number of KE singular vectors included in the pseudoinverse. Analysis uncertainty has been included in the forecast difference fields. Each curve shows the corrected forecast difference over the local projection area using the pseudoinverse based on the singular vectors for that area. Figure 15 shows that the largest correction per singular vector occurs for the smallest target areas. Note that for each area, a degradation in forecast correction starts to occur when the pseudoinverse starts to include decaying vectors. The optimal correction obtained for a particular target area for a given number of singular vectors is obtained by using the singular vectors calculated for that target area (not shown). For the cases shown here, each singular vector accounts for a larger percentage of forecast difference variance as the target area decreases in size, allowing for more computational economy. However, the percentage of forecast difference variance accounted for by each singular vector will also be a function of the stability of the region chosen. The results

might be quite different, for example, for a tropical region.

7. Summary and conclusions

A simple quasigeostrophic model has proven to be a useful tool for examining aspects of linear error growth through the use of the full set of singular vectors. The size of the model allows for the explicit representation and inversion of matrices that can only be approximated for more complex models, and therefore allows for an evaluation of the validity of approximations used with these larger systems.

It is found that $\mathbf{L}$ is not strictly ill-conditioned (i.e., it is invertible). The trailing KE singular vectors have coherent localized structure, and some time-dependent properties (downscale and downward energy transfer, eastward tilt with height) that contrast with those of the leading KE singular vectors. The near-neutral KE singular vectors represent small-scale noise. For dissipative systems, the nth singular value of $\mathbf{L}^{-1}$ is greater than the nth singular value of $\mathbf{L}$ (Fig. 1). Because the projection of the analysis differences onto the singular vectors is fairly isotropic, $\mathbf{L}^{-1}\mathbf{e}_o > \mathbf{L}\mathbf{e}_o$. In other words, the growth of analysis uncertainty as one goes backward in time will be greater than the growth of analysis uncertainty as one goes forward in time (see Table 1). Thus, while $\mathbf{L}^{-1}\mathbf{e}_f$ may yield a good estimate of $\mathbf{e}_o$ when $\mathbf{e}_f$ is known exactly, it cannot yield a good estimate of $\mathbf{e}_o$ when $\mathbf{e}_f$ includes errors of the same magnitude as the analysis errors. Thus, the full inverse and backward integration techniques are not suitable for analysis error estimates in more realistic situations where the forecast difference or error is known only approximately, such as in an operational setting. For these cases, the best forecast corrections are obtained by using the pseudoinverse technique (Buizza et al. 1997) including all the growing singular vectors only. However, increasing the size of the pseudoinverse results in diminishing returns as the number of singular vectors included increases. Half of the global forecast difference variance can be explained by using just the first 7% of all the singular vectors. It is also shown that local projection operators can yield more effective singular vectors. That is, the smaller the region of interest, the larger the fraction of forecast difference explained by a given number of singular vectors computed with the appropriate local projection operator, resulting in less computational expense.

The fact that only growing singular vectors can be useful in correcting initial-time analysis errors when errors are also present at final time can be interpreted intuitively. Initial analysis errors that remain neutral or decay with time will evolve into perturbations that are comparable to, or smaller than, analysis uncertainty at final time. The signal-to-noise ratio associated with the projection of the forecast error onto these neutral and decaying final-time singular vectors will be approximately equal to or less than one, and therefore no confidence should be placed in their backward evolution in time. Also, the stronger the dissipation in the system, the larger the decaying subspace becomes, resulting in even poorer analysis error estimates (Table 2). These results have been anticipated by several variational analysis studies. Cohn et al. (1994) point out that ''as unstable trajectories diverge in phase space, one does expect knowledge of future states to pin down the current state well, especially for a 'perfect' model. On the other hand, knowledge of future converging (diffusive) trajectories does little to define the current state.'' Their fixed-lag Kalman smoother assimilation experiments also show a strong sensitivity to model diffusivity or instability. Pires et al. (1996) point out that assimilation of future observations would reduce the error along unstable components, while assimilation of past observations essentially reduces the error along the stable components of the flow.

The full inverse and backward integration techniques still yield forecast corrections that, while not optimal, are as good as those obtained using the pseudoinverse with the first 60–100 singular vectors. The backward integration technique employed here, based on and similar to the one used by Pu et al. (1997), does indeed prove to be a very close approximation to the use of the full inverse, while being extremely computationally economical. The magnitude and treatment of the dissipation terms have a significant impact on the results. The percent reduction in forecast difference decreases as the magnitude of the dissipation is increased (see Table 3). Reversing the sign of the dissipation during the backward integration results in a less accurate representation of $\mathbf{L}^{-1}$, and worse analysis and forecast corrections, when the forecast difference is known exactly. However, when the forecast difference is not known exactly, the negative dissipation produces a slight beneficial effect by damping the backward growth of analysis uncertainty.

Several simplifications have been made in this study that should be considered when trying to make quantitative inferences about operational forecast and analysis corrections from these results. These experiments were conducted in a ''perfect model'' setting. That is, only forecast differences, not real forecast errors, were considered. The effects of forecast model deficiencies were not addressed. Ménard and Daley (1996) find that applying a 4DVAR algorithm (using a perfect-model assumption) when the model is imperfect, may result in larger errors, especially at the beginning and end of the assimilation period. The δ that represents analysis uncertainty in this study may alternatively be interpreted as cumulative model errors. As with analysis uncertainty, these model errors could also result in a spurious projection onto the decaying singular vectors at final time. With an imperfect model, it should be pointed out that the initial analysis that results in the maximum re-

duction in forecast error is not necessarily the initial analysis that is closest to the truth.

There are other characteristics of this experimental system that differ significantly from operational forecast systems. The forward tangent propagator used in this study is the full linearized equations of the quasigeostrophic model. In contrast to the current forward tangent and adjoint equations for operational forecast models, no approximations concerning highly nonlinear diabatic terms are necessary. In this system, on average, the growing subspace was over 40% of the entire singular vector space. This percentage will vary with the metric and the trajectory, and is a strong function of the dissipation. As dissipation increases, this growing subspace decreases in size. The ratio of growing singular vectors to total singular vectors may be much smaller for the operational models if the dissipation in those models is considerably higher. Also, one might assume that as resolution is increased, a large number of the extra singular vectors will be near-neutral vectors. If both of these assumptions are true, then a significantly larger portion of the growing subspace can be captured by a significantly smaller portion of singular vectors in a more complex system. Finally, analysis differences have been used as a proxy for analysis errors in this study. It is expected that the characteristics of true analysis errors might differ substantially from analysis differences. The exact nature of analysis error remains unknown. Therefore, it is impossible at this time to determine whether analysis errors are approximately isotropic with respect to the initial-time singular vectors, although there is indirect evidence that this assumption is true for the first 30 total energy singular vectors (Gelaro et al. 1998). As long as the projection of the analysis errors onto the initial-time singular vectors is relatively flat compared to the slope of the singular values, these results should remain qualitatively accurate

Acknowledgments. The authors would like to thank Jan Barkmeijer, Ron Errico, Ron Gelaro, Tony Hollingsworth, Ernst Klinker, and Florence Rabier for helpful discussions. The three-level quasigeostrophic model was developed by Franco Molteni. Jan Barkmeijer provided much of the software. Support of the sponsor, Office of Naval Research, and the program manager, Naval Research Laboratory, Project Element 0601153N, is gratefully acknowledged. Much of this work was done while the first author was visiting the European Centre for Medium-Range Weather Forecasts.

REFERENCES

Andersson, E., J. Pailleux, J. N. Thépaut, J. R. Eyre, A. P. McNally, G. A. Kelly, and P. Courtier, 1994: Use of cloud-cleared radiances in three/four-dimensional variational data assimilation, *Quart. J. Roy. Meteor. Soc.,* **120,** 627–653.

——, P. Courtier, W. Heckley, F. Rabier, J. N. Thépau, P. Undén, and D. Vasiljevic, 1995: Pre-operational tests with a three-dimensional variational analysis scheme. *Proc. Second WMO Int. Symp. on Assimilation of Observations in Meteorology and Oceanography,* Tokyo, Japan, WMO, 43–48.

Baker, W., H. Fleming, M. Goldberg, B. Katz, and J. Derber, 1995: Development and implementation of satellite temperature soundings produced interactively. NWS Tech. Procedures Bull. No. 422, 8 pp. [Available from Office of Meteorology, National Weather Service, Silver Spring, MD 20910.]

Barkmeijer, J., 1992: Local error growth in a barotropic model. *Tellus,* **44A,** 314–323.

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

——, M. van Gijzen, and F. Bouttier, 1998: Singular vectors and the analysis error covariance metric. *Quart. J. Roy. Meteor. Soc.,* **124,** 1695–1713

Borges, M. D., and D. L. Hartmann, 1992: Barotropic instability and optimal perturbations of observed nonzonal 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.

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

——, J. Tribbia, F. Molteni, and T. Palmer, 1993: Computation of optimal unstable structures for a numerical weather prediction model. *Tellus,* **45A,** 388–407.

——, 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.

Cohn, S. E., N. S. Sivakumaran, and R. Todling, 1994: A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.,* **122,** 2838–2867.

Derber. J. C., 1987: Variational four-dimensional analysis using quasi-geostrophic constraints. *Mon. Wea. Rev.,* **115,** 998–1008.

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.

Farrell, B., 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.,* **47,** 2409–2416.

Fisher, M., and P. Courtier, 1995: Estimating the covariance matrices of analysis and forecast difference in variational data assimilation. ECMWF Research Department Tech. Memo. 220, 27 pp. [Available from European Centre for Medium-Range Weather Forecasts, Shefield Park, Reading, Berkshire R62 9AX, United Kingdom.]

Gelaro, R., E. Klinker, and F. Rabier, 1996: Real and near-real time corrections to forecast initial conditions using adjoint methods: A feasibility study. Preprints, *11th Conf. on Numerical Weather Prediction,* Norfolk, VA, Amer. Meteor. Soc., J58–J60.

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

Holton, J. R., 1979: *An Introduction to Dynamic Meteorology.* 2d ed. Academic Press, 391 pp.

Klinker, E., F. Rabier, and R. Gelaro, 1998: Estimation of key analysis errors using the adjoint technique. *Quart. J. Roy. Meteor. Soc.,* in press.

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

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

Ménard, R., and R. Daley, 1996: The application of Kalman smoother theory to the estimation of 4DVAR error statistics. *Tellus,* **48A,** 221–237.

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.

Mureau, R., F. Molteni, and T. N. Palmer, 1993: Ensemble prediction using dynamically conditioned perturbations. *Quart. J. Roy. Meteor. Soc.*, **119,** 299–323.

Palmer, T., C. Brankovic, F. Molteni, and S. Tibaldi, 1990: Extended-range predictions with ECMWF models: Interannual variability in operational model integrations. *Quart. J. Roy. Meteor. Soc.*, **116,** 799–834.

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

Pires, C., R. Vautard, and O. Talagrand, 1996: On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus*, **48A,** 96–121.

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

Rabier, F., P. Courtier, J. Pailleux, O. Talagrand, and V. Deasko, 1993: A comparison between four-dimensional variational assimilation and simplified sequential assimilation relying on three-dimensional variational analysis. *Quart. J. Roy. Meteor. Soc.*, **119,** 845–880.

——, F., J.-N. Thépaut, and P. Courtier, 1996a: Four dimensional variational data assimilation at ECMWF. *ECMWF Seminar on Data Assimilation*, Reading, United Kingdom, ECMWF, 213–250.

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

Reynolds, C., P. Webster, and E. Kalnay, 1994: Random error growth in NMC's global forecasts. *Mon. Wea. Rev.*, **122,** 1281–1305.

Saha, S., 1992: Response of the NMC MRF model to systematic-error correction within integration. *Mon. Wea. Rev.*, **120,** 345–360.

Sardeshmukh, P. D., M. Newman, and M. D. Borges, 1997: Free barotropic Rossby wave dynamics of the wintertime low-frequency flow. *J. Atmos. Sci.*, **54,** 5–23

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

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,** 535–561.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of initial perturbations. *Bull. Amer. Meteor. Soc.*, **74,** 2317–2330.

Zupanski, M., 1993: Regional four-dimensional variational data assimilation in a quasi-operational forecasting environment. *Mon. Wea. Rev.*, **121,** 2396–2408.