Q. J. R. Meteorol. Soc. (1999), **125**, pp. 2333–2351

3D-Var Hessian singular vectors and their potential use in the ECMWF Ensemble Prediction System

By J. BARKMEIJER*, R. BUIZZA and T. N. PALMER
European Centre for Medium-Range Weather Forecasts, UK

(Received 6 May 1998; revised 28 October 1998)

SUMMARY

Singular vectors are computed which are consistent with 3D-Var (three-dimensional variational) estimates of analysis error statistics. This is achieved by defining the norm at initial time in terms of the full Hessian of the 3D-Var cost function. At final time the total energy norm is used. The properties of these Hessian singular vectors (HSVs) differ considerably from total energy singular vectors (TESVs) in such aspects as energy spectrum and growth rate. Despite these differences, the leading 25 TESVs and HSVs explain nearly the same part of the 2-day forecast error.

Two experimental ensemble configurations are studied. One configuration uses perturbations based on HSVs in the computation of initial perturbation, the other uses TESVs and 2-day linearly evolved singular vectors (ESVs) of two days before. The latter approach provides a way to include more stable and large-scale structures in the perturbations. Ten pairs of ensembles are compared to the operational European Centre for Medium-Range Weather Forecasts Ensemble Prediction System. The ensembles using ESVs perform slightly better. The ensembles based on HSVs show a slightly worse performance and are lacking some spread in the medium range. Possible directions to improve the computation of HSVs are discussed.

KEYWORDS: Ensemble forecasting Hessian singular vectors Medium-range forecasting Numerical weather prediction

1. INTRODUCTION

Since December 1996, the operational Ensemble Prediction System (EPS) at the European Centre for Medium-Range Weather Forecasts (ECMWF) comprises 50 non-linear integrations of the $T_L 159$ operational model version (using linear grid), with as initial conditions the 1200 UTC analysis perturbed along growing directions. The development of the EPS is documented in Palmer *et al.* (1992), Molteni *et al.* (1996) and Buizza *et al.* (1998a); recent findings indicate that its performance is useful (Palmer *et al.* 1998b). Three other numerical weather prediction (NWP) centres run an EPS on a daily basis, namely the US National Centers for Environmental Prediction (NCEP), the Canadian Meteorological Centre (CMC) and the Fleet Numerical Meteorological and Oceanographic Center (FNMOC) of the US Navy. The approaches followed by the various NWP centres to define their EPS differ considerably. The breeding method is used at NCEP (Toth and Kalnay 1997) and FNMOC (Rennick 1995) to create initial perturbations, while CMC generates initial perturbations by running various assimilation schemes using perturbed observations, but also takes into account model errors by allowing different model configurations in the ensemble (Houtekamer *et al.* 1996).

At ECMWF singular vectors (SVs) are used to define initial perturbations for the EPS (Buizza and Palmer 1995), as they are believed to sample the unstable linear subspace as efficiently as possible. Two sets of SVs at horizontal resolution T42 and vertical resolution 31 (T42L31) are computed, targeted respectively at the northern and southern hemisphere extratropics (the tropics being defined as 30°S to 30°N) and with an optimization time of 2 days. From these, 25 perturbations are generated for each hemisphere separately (Molteni *et al.* 1996), and then the two sets of perturbations are added to yield extratropical perturbations. By adding and subtracting the extratropical perturbations to and from the 1200 UTC analysis the 50 perturbed initial conditions for the EPS are defined.

* Corresponding author: European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, UK.

In order to obtain the forecast probability distribution function (PDF) under the assumption of a perfect forecast model, ideally, one would like to integrate the appropriate Liouville equation (Epstein 1969; Ehrendorfer 1994). However, the large dimension of NWP models makes this impossible. The SV approach provides a possibility of searching for directions in phase space where the errors in the initial condition will amplify rapidly. Nevertheless, the interpretation of results as derived from the EPS is not straightforward. The large discrepancy between the size of the ensemble (50) and the dimension of NWP models (10^6 to 10^7) may lead to sampling errors in describing the forecast PDF: initial directions in phase space which result in improved forecasts may easily be missed. In fact, estimates of the dimension of the linear unstable subspace for a 1449-variable T21L3 QG (quasi-geostrophic) model are of order 10^2 to 10^3 (Palmer *et al.* 1998a). Another difficulty is that the SVs are computed to produce large growth in the first 2 days of the forecast. It is possible that slower growing SVs become more important in the medium range when error growth has become nonlinear. This was indeed the case for perturbations which were specifically defined to trigger the onset of weather regimes at forecast day 5 in the context of a 3-level QG model (Oortwijn and Barkmeijer 1995). In addition to this, medium-range forecast errors cannot solely be attributed to errors in the initial condition. Experiments show that model errors can become equally important in causing forecast errors (Harrison *et al.* 1995; Richardson 1998). Research is under way to complement the SV approach by allowing the model perturbations during the actual time integration of each ensemble member (Buizza *et al.* 1998b).

In this paper we assume a perfect-model approach and we shall focus on defining SVs in accordance with analysis error statistics. The specification of the initial and, to a lesser extent, the final norm plays a crucial role in this. In the ECMWF operational EPS, SVs are computed with the so-called total energy norm at initial and final time (see section 2 for more details). It can be shown that among simple norms, the total energy norm provides SVs which agree best with analysis error statistics (Palmer *et al.* 1998a). In Barkmeijer *et al.* (1998) a method is proposed to incorporate analysis error statistics directly in the SV computation. This is done by taking the full Hessian of the 3D-Var (three-dimensional variational) cost function as an approximation to the inverse of the analysis error covariance matrix, and using it to define a norm at initial time. The inverse of the analysis error covariance matrix $\mathbf{A}$ is not explicitly known. It suffices to be able to compute $\mathbf{y} = \mathbf{A}^{-1}\mathbf{x}$ for a given input vector $\mathbf{x}$. By doing so the SV computation becomes consistent with the 3D-Var procedure to determine the analysed state. The total energy norm is still used at optimization time. We call SVs calculated in this way Hessian singular vectors (HSVs). Ehrendorfer and Tribbia (1997) state that such an approach to determine SVs provides an efficient way to describe the forecast error covariance matrix when only a limited number of linear integrations are possible. In their SV computations, however, explicit knowledge of the analysis error covariance matrix is used, which is unavailable in an operational data-assimilation system. It is this complication which requires an efficient generalized eigenvalue problem solver to compute SVs.

The purpose of this paper is to report on properties of these HSVs defined with an initial norm given by the full 3D-Var Hessian, and to describe their impact on the performance of the ECMWF EPS. To that end 10 ensembles starting from days in December, January, February 1996/97 are compared, where initial perturbations are created with either total energy singular vectors (TESVs) or HSVs. Parallel to this we also investigate an easy way to complement perturbations based on initial TESVs with more stable and large-scale directions. For this third set of ensembles, the 2-day linearly evolved initial TESVs computed from two days before were used. The perturbations

generated by the evolved singular vectors (ESVs) were simply added to perturbations based on initial SVs.

The organization of the paper is as follows. In the next section we introduce the computation of the HSVs. The comparison of the HSVs and TESVs is given in section 3. Results from the three sets of ensembles are presented in section 4. The paper concludes with a brief summary of the results and possible areas for future work. In the appendix one of the statistical verification methods, the relative operating characteristic (ROC), is briefly described.

2. Use of the 3D-Var Hessian

The computation of SVs requires the specification of a norm at initial time t_0 and at optimization time t_1. In this paper we consider SVs, $\boldsymbol{e}$, which maximize the ratio

$$\frac{\langle \mathbf{P}\boldsymbol{e}(t_1), \mathbf{EP}\boldsymbol{e}(t_1)\rangle}{\langle \boldsymbol{e}(t_0), \mathbf{C}\boldsymbol{e}(t_0)\rangle}. \tag{1}$$

Here $\langle\ ,\ \rangle$ denotes the Euclidean inner product $\langle \mathbf{x}, \mathbf{y}\rangle = \sum x_i y_i$. The positive definite and symmetric operators **C** and **E** induce a norm at initial and optimization time respectively. The projection operator **P** sets a vector to zero outside a given domain, e.g. south of 30°N as in this paper. The first singular vector SV1 maximizes the ratio (1), the second singular vector SV2 maximizes (1) in the subspace **C** orthogonal to SV1, and so forth. The evolved SVs $\boldsymbol{e}(t_1) = \mathbf{M}\boldsymbol{e}(t_0)$ form an **E**-orthogonal set at optimization time, where **M** is the propagator of the tangent linear model. Alternatively, the SVs defined by (1) are solutions of the following generalized eigenvalue problem

$$\mathbf{M}^*\mathbf{P}^*\mathbf{EPMx} = \lambda\mathbf{Cx}, \tag{2}$$

where λ is the eigenvalue corresponding to **x**. The adjoint operators $\mathbf{M}^*$ and $\mathbf{P}^*$ are determined with respect to the Euclidean inner product.

In the computation of the TESVs, the total energy metric is used at initial and optimization time, i.e. **E** and **C** are identical and:

$$\langle \mathbf{x}, \mathbf{Ey}\rangle = \frac{1}{2}\int_0^1 \int_\Sigma \left(\nabla\Delta^{-1}\zeta_x \cdot \nabla\Delta^{-1}\zeta_y + \nabla\Delta^{-1}D_x \cdot \nabla\Delta^{-1}D_y + \frac{c_p}{T_\mathrm{r}}T_x T_y\right) \mathrm{d}\Sigma \left(\frac{\partial p}{\partial \eta}\right) \mathrm{d}\eta + \frac{1}{2}\int_\Sigma R_\mathrm{d} T_\mathrm{r} P_\mathrm{r} \ln \pi_x \cdot \ln \pi_y \,\mathrm{d}\Sigma, \tag{3}$$

with $(\zeta_x, D_x, T_x, \ln \pi_x)$ being the vorticity, divergence temperature and logarithm of the surface pressure components of the state vector **x**, Σ denotes integration over a sphere, $p(\eta)$ is pressure at level η ($0 =$ the surface; $1 =$ top of atmosphere), c_p is the specific heat of dry air at constant pressure, R_d is the gas constant for dry air, $T_\mathrm{r} = 300$ K is a reference temperature and $P_\mathrm{r} = 800$ hPa is a reference pressure.

In the TESV case **C** has the form of a diagonal matrix and the square root of **C** can be readily determined. Multiplying both sides of (2) with the inverse of the square root of **C**, yields an equation which can be solved using the Lanczos algorithm (Partlett 1980). Palmer *et al.* (1998a) study the impact on SVs of choosing different simple metrics at initial time, while keeping the total energy metric at optimization time. They were interested in initial metrics so that the spectra of the dominant SVs was not inconsistent with the spectra of an estimate of analysis error variance. Inconsistency is meant in the sense that spectra of SVs are dominant in a wave-number band for which the analysis

error variance is small. It turns out that, of the simple metrics considered, the total energy metric satisfies this consistency condition. In Barkmeijer *et al.* (1998) a method is proposed to make the singular vector computation fully consistent with analysis error statistics by employing the Hessian of the 3D-Var cost function.

In the incremental formulation of 3D-Var a cost function $\mathcal{J}$ of the form

$$\mathcal{J}(\delta\mathbf{x}) = \tfrac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \tfrac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d})^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}) \tag{4}$$

is minimized (Courtier *et al.* 1998). The increment $\delta\mathbf{x}^a$, which minimizes $\mathcal{J}$, provides the analysis $\mathbf{x}^a$ which is defined by adding $\delta\mathbf{x}^a$ to the background $\mathbf{x}^b$

$$\mathbf{x}^a = \mathbf{x}^b + \delta\mathbf{x}^a. \tag{5}$$

The matrices $\mathbf{B}$ and $\mathbf{R}$ are covariance matrices for background errors $(\mathbf{x}^b - \mathbf{x}^t)$ and observation errors $(\mathbf{y}^o - H(\mathbf{x}^t))$ respectively with $\mathbf{y}^o$ the observation vector and $\mathbf{x}^t$ the true state of the atmosphere, $\mathbf{H}$ is the linear approximation of the observation operator H in the vicinity of $\mathbf{x}^b$ and $\mathbf{d}$ is the innovation vector

$$\mathbf{d} = \mathbf{y}^o - H\mathbf{x}^b \tag{6}$$

The Hessian $\nabla\nabla\mathcal{J}$ of the cost function is given by

$$\nabla\nabla\mathcal{J} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \tag{7}$$

Provided that the background error and observation errors are uncorrelated, it can be shown that the Hessian of the cost function is equal to the inverse of the analysis error covariance matrix (see Rabier and Courtier 1992; Fisher and Courtier 1995).

In the HSV computation the inverse of the analysis error covariance matrix is used to define the norm at initial time. The operator $\mathbf{C}$ is specified to be equal to the full Hessian of the 3D-Var cost function. The operator $\mathbf{C} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ is not known in matrix form and determining its square root is not feasible. In order to solve (2), a generalized eigenvalue problem solver, called generalized Davidson algorithm, is used (Barkmeijer *et al.* 1998). This algorithm can solve (2) efficiently and requires only the ability to calculate $\mathbf{y} = \mathbf{S}\mathbf{x}$, where $\mathbf{S}$ is any of the operators appearing in (2). No explicit knowledge of any operator is needed. In the following we assume that the total energy metric is always used at optimization time.

To improve the convergence of the generalized Davidson algorithm a coordinate transformation $\boldsymbol{\chi} = \mathbf{L}^{-1}\mathbf{x}$ is carried out with $\mathbf{L}\mathbf{L}^T = \mathbf{B}$. Applying the transformation $\mathbf{L}$, the Hessian becomes equal to the sum of the identity and a matrix of rank less than or equal to the dimension of the vector of observations (see also Fisher and Courtier 1995). Thus, when no observations are used in the cost function, the transformed operator $(\mathbf{L}^{-1})^T\mathbf{C}\mathbf{L}^{-1}$ is the identity and the generalized Davidson algorithm becomes equivalent to the Lanczos algorithm (Barkmeijer *et al.* 1998).

3. HESSIAN SINGULAR VECTORS

For 10 days from winter 1996/97, T42L31 SVs were computed using an initial norm derived from the full Hessian of the 3D-Var cost function and with an optimization period of 2 days. Since the SVs will also be used to create EPS perturbations, we selected an equal number of initial days with below as with above 1996/97 winter average ensemble spread and ensemble control forecast error at day 5 over the northern hemisphere (NH). The final norm in the defining equation (2) is the total energy

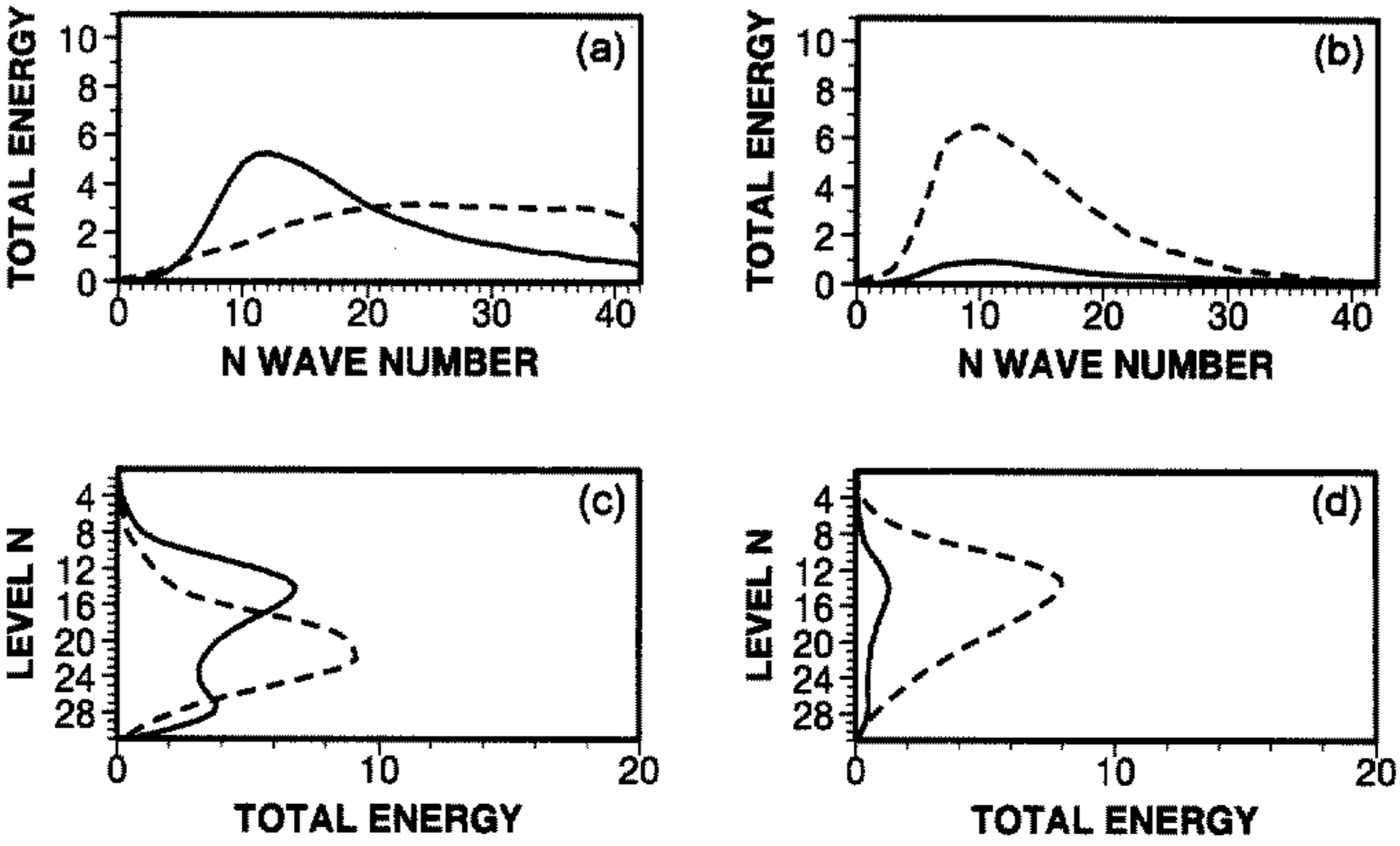


Figure 1. (a) and (b) The total energy spectrum, and (c) and (d) the vertical distribution of the total energy spectrum, of total energy singular vectors and Hessian singular vectors respectively. Values at initial (final) time are given by dashed (solid) lines. At initial time the total energy (m^2s^{-2}) has been multiplied by a factor of 100.

norm (3). The solutions of this generalized eigenvalue problem will be referred to as HSVs. In evaluating the Hessian, we used the new formulation of the background error covariances described by Bouttier *et al.* (1997), together with most of the conventional observations (SYNOP, AIREP, SATOB, DRIBU, TEMP, PILOT, SATEM and PAOB) for NH. The necessary gradient computations of the cost function and the slower convergence of the generalized Davidson algorithm make the HSVs around 5 times computationally more expensive to determine than TESVs.

The HSVs have properties considerably different from TESVs which use the total energy norm also at initial time. Figures 1(a) and (b) give the TESV and HSV spectrum in terms of total energy averaged over the 10 cases, each consisting of 25 SVs. Clearly, the TESVs are initially (dashed line) more small-scale than HSVs, with a dominant wave number around 30 compared with 10. The spectra at final time (solid line) for both sets of SVs peak around wave number 10. Decreasing the horizontal diffusion for vorticity, divergence and temperature by a factor of 4 (damping times are 60 h, 24 h and 60 h respectively) slowed little impact on the spectra, in contrast to findings for TESVs (Buizza 1998). Also the vertical structures of HSVs and TESVs are quite different, see Fig. 1(c) and (d). Most of the HSV total energy is initially confined to the levels around 300 hPa, whereas TESVs have most of the energy near the jet steering level (700 hPa). At optimization time the vertical energy distributions of TESVs and HSVs are comparable. The total energy amplification of HSVs is smaller, as indicated by the area under the solid curves in Figs. 1(a) and (b).

On average TESVs grow approximately twice as fast as HSVs in terms of total energy. Figure 2 shows the total energy amplification of the 15 leading HSVs and TESVs for a particular day with a 2-day optimization period. Observe that, as a consequence of using different norms at initial and final time, the ordering of the HSVs no longer corresponds with their total energy growth.

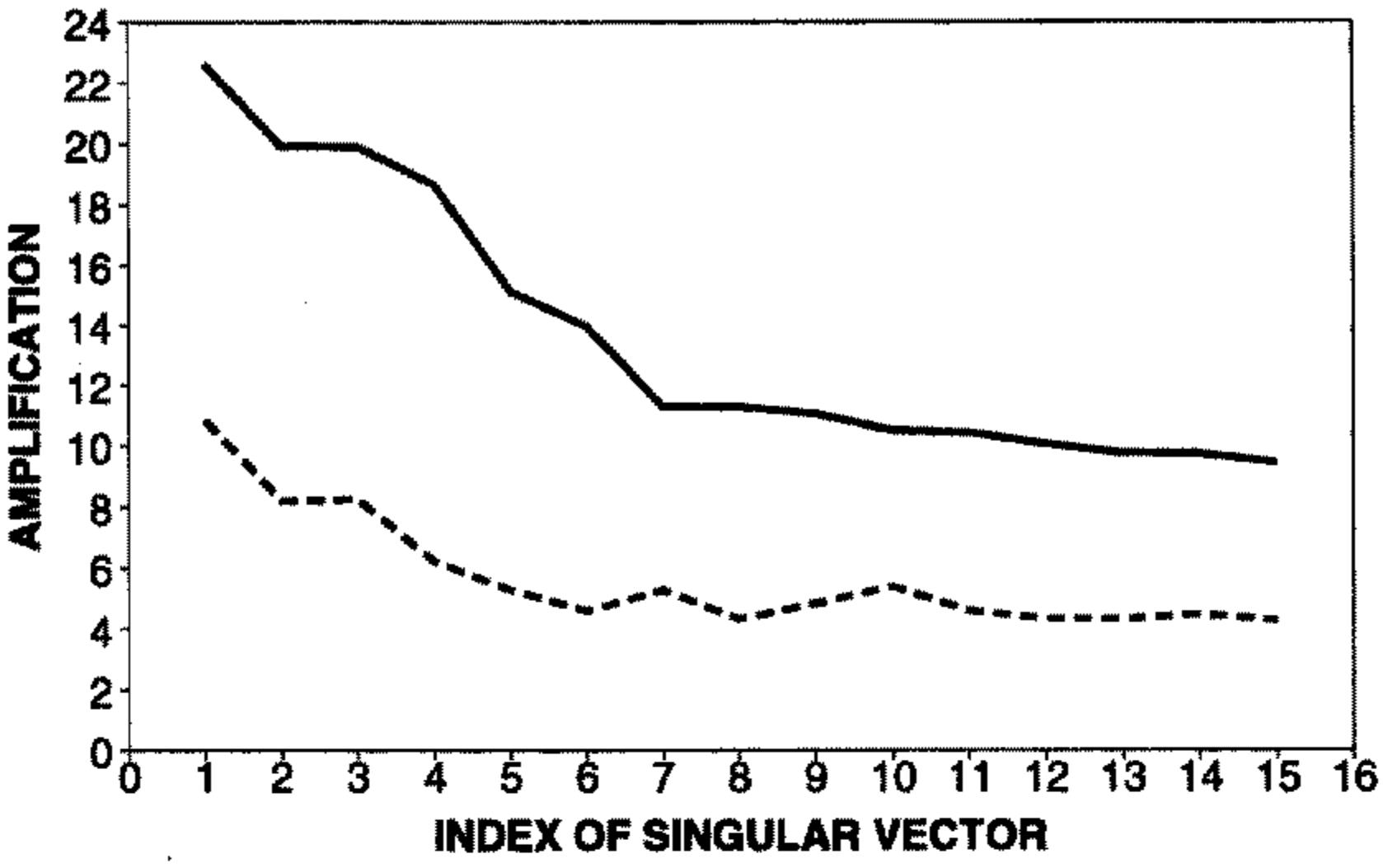


Figure 2. Typical total energy amplification of a total energy singular vector (solid) and Hessian singular vector (dashed).

The large-scale structure and energy distribution of HSVs is to a large extent determined by the formulation of the background error covariance matrix **B**. Essential in defining the **B** matrix are error statistics derived from the difference between 2-day and 1-day forecasts verifying at the same day (commonly known as the NMC method; Parrish and Derber 1992). The statistics obtained for a period of 90 consecutive days extending from December 1992 to February 1993 are described in Rabier *et al.* (1998). Power spectra of the error statistics indicate that the energy spectrum peaks around total wave number 10. The vertical distribution of the background error variance reaches a maximum near the jet level. Both these properties are reflected in the metric defined by the Hessian. Also, the background error covariance matrix is specified to have broad horizontal and vertical correlations, and thus penalizes the occurance of baroclinic structures in the analysis error. Hence we may suspect that the 3D-Var Hessian metric penalizes the small-scale and baroclinic perturbation patterns in the areas which are picked up by the TESV computation. Thépaut *et al.* (1996) have compared power spectra of 3D-Var and 4D-Var analysis increments in the vicinity of an extratropical storm. Although their results must be regarded as indicative and depend on, for instance, the background error covariance matrix, the 4D-Var approach clearly allowed for more energy in larger total wave numbers, peaking between wave number 7 and 15. Also the slope of the 4D-Var increments' power spectrum increased to 6, as compared to 3 for 3D-Var. The flat spectrum that occurs for initial TESVs up to the truncation limit, see Fig. 1(a), is not observed. The NMC method is well suited for statistical data assimilation in that it provides a good estimate of the time averaged global background error covariances. However, it is likely that the average covariance structures are not optimal in dynamically unstable areas.

The differences between HSVs and TESVs can also be exhibited by using a similarity index (Buizza 1998), which measures how parallel subspaces are spanned by the leading HSVs and TESVs. Values of the similarity index range from 0 to 1 and increasing values mean that the subspaces become increasingly parallel. Figure 3 shows for each of the 10 cases the similarity index between the unstable HSV and TESV subspaces at initial and optimization time when 10 or 25 SVs are used to span the

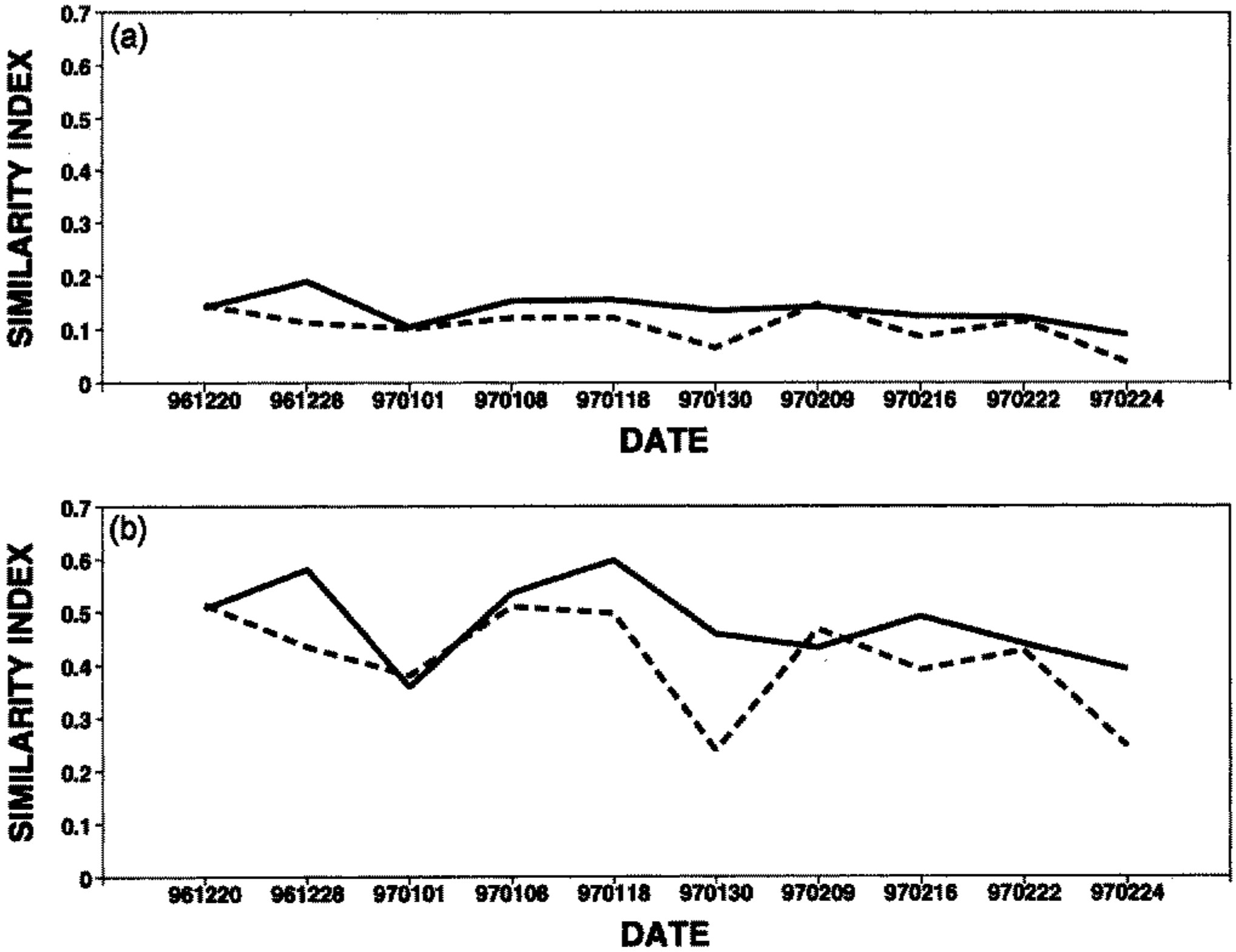


Figure 3. Similarity index between total energy singular vector and Hessian singular vector unstable subspaces at (a) initial and (b) final time. Dashed (solid) lines indicate that 10(25) singular vectors are used in spanning the subspace.

unstable subspace. It is clear from Fig. 3(a) that the unstable TESV and HSV subspaces are almost orthogonal at initial time. At final time the subspaces have become more parallel although the similarity index is still quite low.

Given such small similarity indices, one wonders whether the two sets of SVs describe different parts of the forecast error. To investigate this, the operational 2-day NH forecast error $\boldsymbol{\varepsilon}(48)$ was projected onto the 2-day linearly evolved HSVs and TESVs for each of the 10 cases. In the projection 25 SVs were used. We denote the portion of the forecast error thus explained by $\widetilde{\boldsymbol{\varepsilon}}(48)$ and the associated so-called pseudo analysis error by $\widetilde{\boldsymbol{\varepsilon}}(0)$ (see also Buizza *et al.* 1997):

$$\widetilde{\boldsymbol{\varepsilon}}(48) = \sum_{i=1}^{25} a_i \mathbf{M}(SV_i) = \mathbf{M}\widetilde{\boldsymbol{\varepsilon}}(0), \tag{8}$$

where $\mathbf{M}$ is the tangent model. The percentage of the total energy of $\boldsymbol{\varepsilon}(48)$ as explained by $\widetilde{\boldsymbol{\varepsilon}}(48)$ is given in Fig. 4 for TESVs and HSVs. Both types of SVs describe nearly the same fraction of $\boldsymbol{\varepsilon}(48)$ in terms of total energy.

Despite their very different structures, TESVs and HSVs also describe similar geographical patterns of the 2-day forecast error when 25 SVs are used in the expansion (8). Figure 5(a) and (b) shows for 1200 UTC 18 January 1997 the pseudo analysis error $\widetilde{\boldsymbol{\varepsilon}}(0)$ at 500 hPa in geopotential height using 25 TESVs and HSVs respectively; the corresponding $\widetilde{\boldsymbol{\varepsilon}}(48)$ is given in Fig. 5(c) and (d). The actual 2-day forecast error is given in Fig. 5(e). For both sets of SVs the projected forecast error $\widetilde{\boldsymbol{\varepsilon}}(48)$ is almost

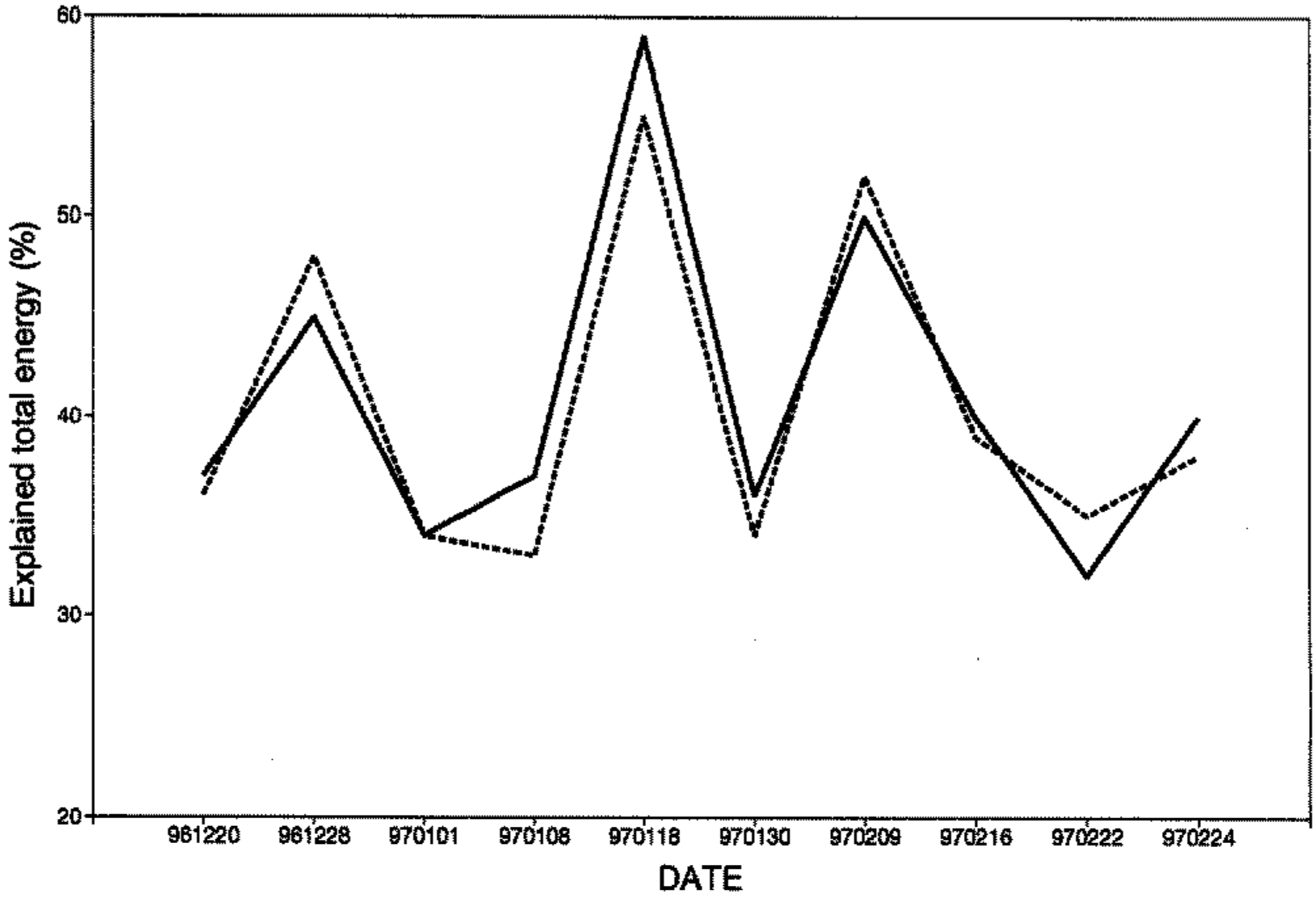


Figure 4. Explained part of the 2-day northern hemisphere forecast error in terms of total energy. The dashed (solid) line indicates that the 25 leading total energy singular vectors (Hessian singular vectors) are used.

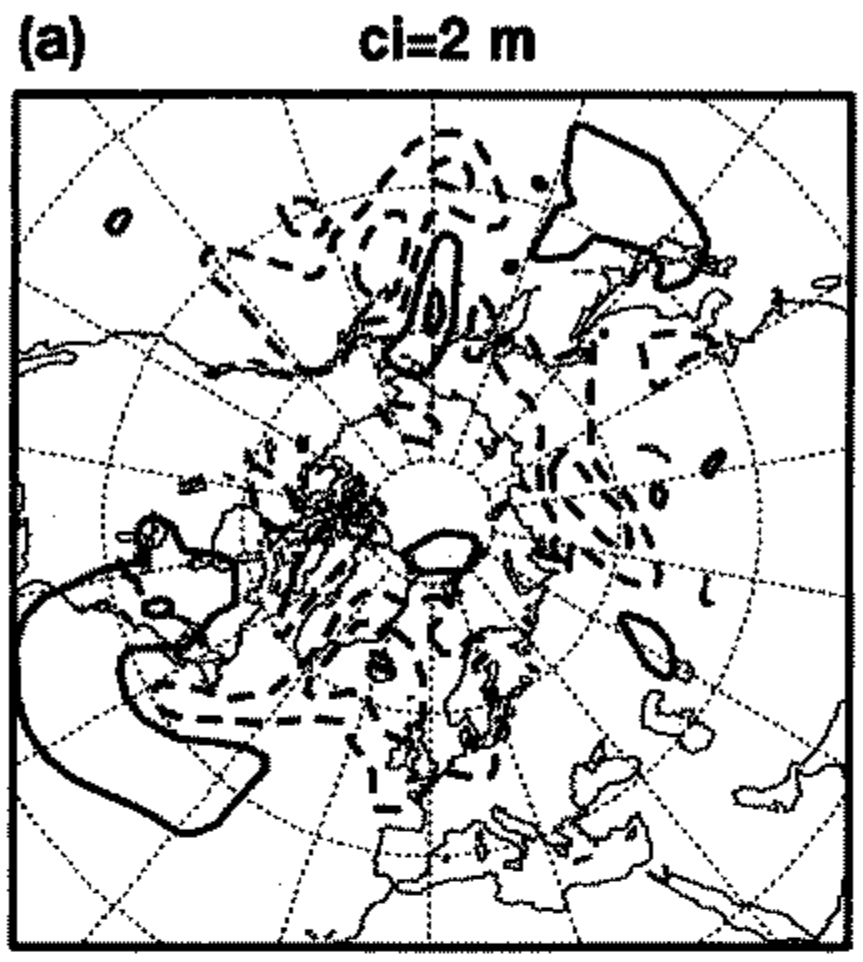
indistinguishable (the correlation coefficient is 0.93). The patterns of $\tilde{\boldsymbol{\varepsilon}}(0)$ are more different (correlation 0.59), although the centres of most of the maxima are located at the same positions. The amplitude of $\tilde{\boldsymbol{\varepsilon}}(0)$ when using HSVs in the expansion is larger than $\tilde{\boldsymbol{\varepsilon}}(0)$ obtained with TESVs (note the different contour interval). The corresponding $\tilde{\boldsymbol{\varepsilon}}(0)$ and $\tilde{\boldsymbol{\varepsilon}}(48)$ for temperature yield similar results, with correlations of 0.62 and 0.95 at initial and final time respectively, see Fig. 6. This result holds for all the cases and shows that, despite their different structures, TESVs and HSVs explain the same part of the forecast error.

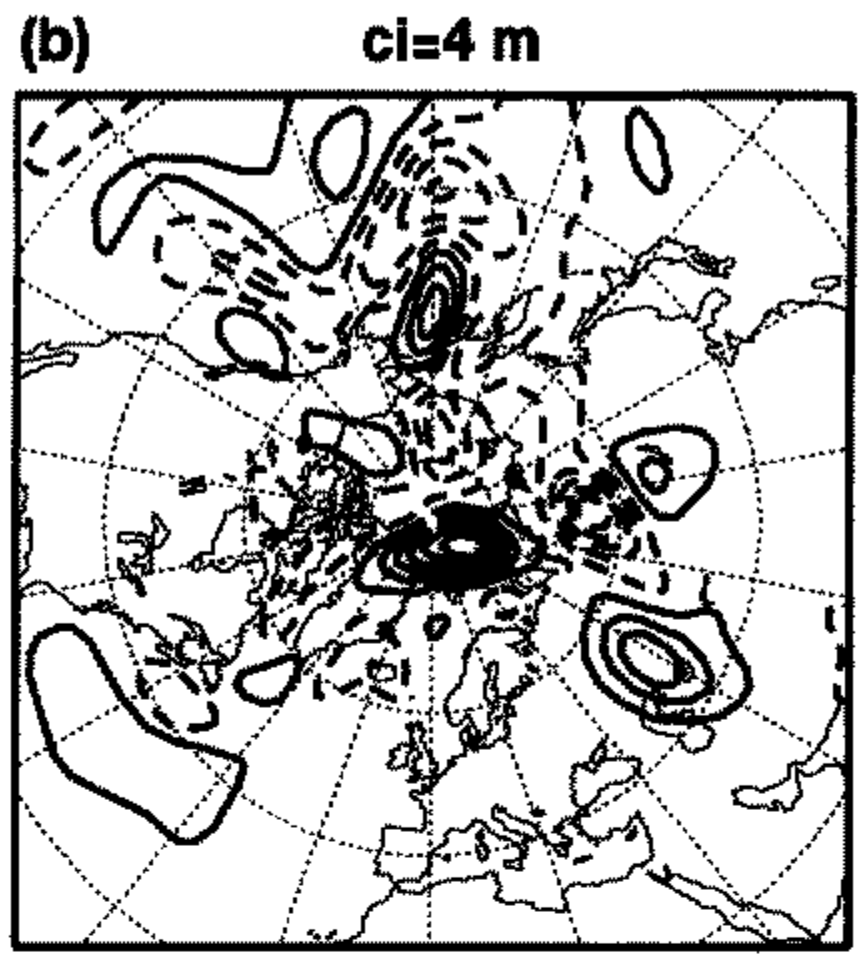
When the number of SVs is decreased, differences between the explained part of the forecast error become visible, see Fig. 7. Here the 2-day forecast error from 1200 UTC 30 January 1997 is projected onto 10 evolved SVs. The TESVs explain better the forecast error over the Atlantic, whilst the HSVs describe some of the errors over Europe.

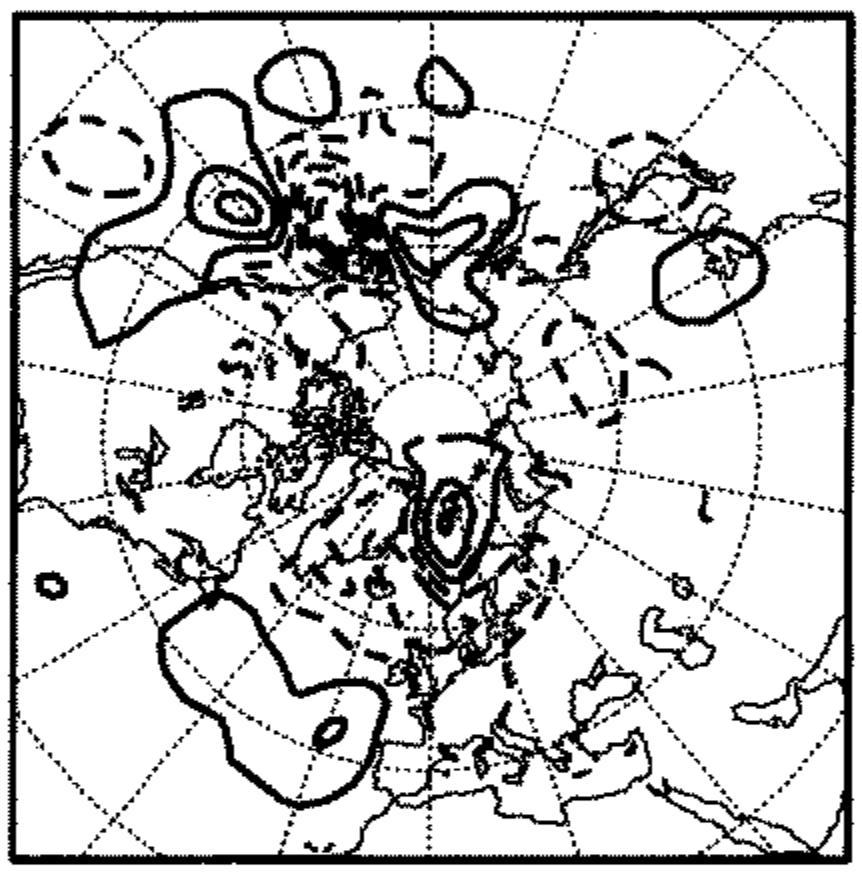
Note that in Fig. 6 the temperature amplitude of the pseudo analysis error $\tilde{\boldsymbol{\varepsilon}}(0)$ computed with TESVs is larger than for HSVs. It is a direct consequence of the different distribution of total energy over the four components of the SV state vector: vorticity, divergence, temperature and logarithmic surface pressure. For TESVs most of the total energy is in the temperature component, while HSVs have a dominant vorticity component, see Fig. 8.

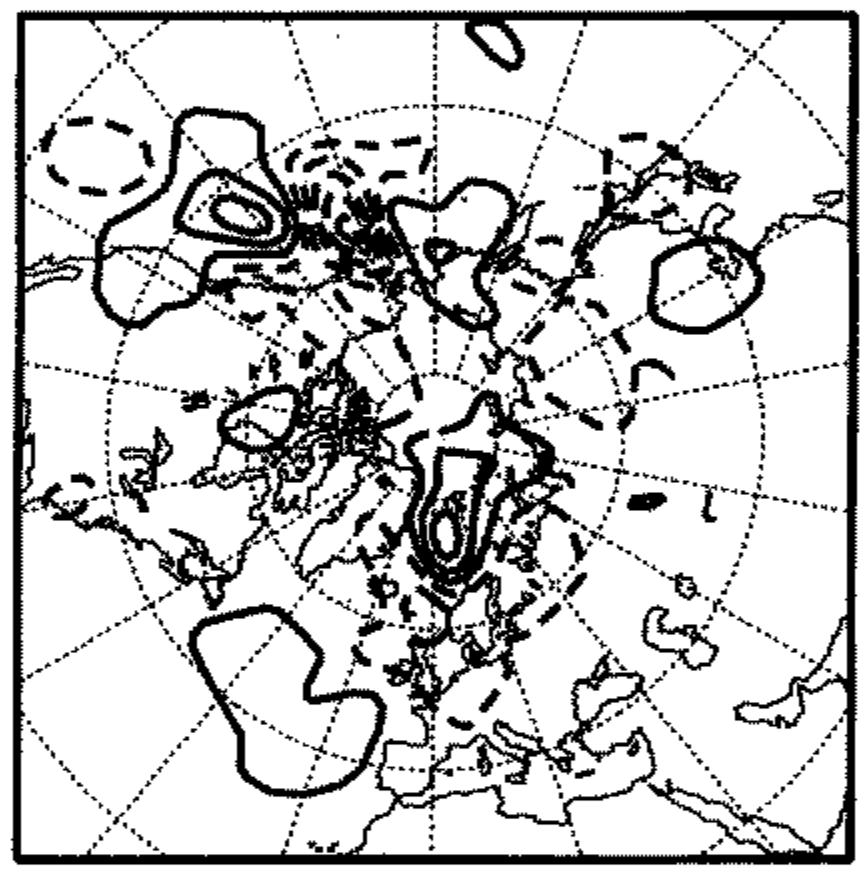
4. HESSIAN AND EVOLVED SINGULAR VECTORS IN THE ECMWF EPS

For the 10 days for which HSVs were available, alternative ensembles have been integrated using the same setting as the operational ensemble: 50 + 1 (control) 10-day T_L159L31 nonlinear integrations. In creating the 50 initial perturbations, now based on HSVs, the same methodology was followed as for the operational ensemble, see Molteni *et al.* (1996), except that the initial amplitude was set so that both TESV and HSV









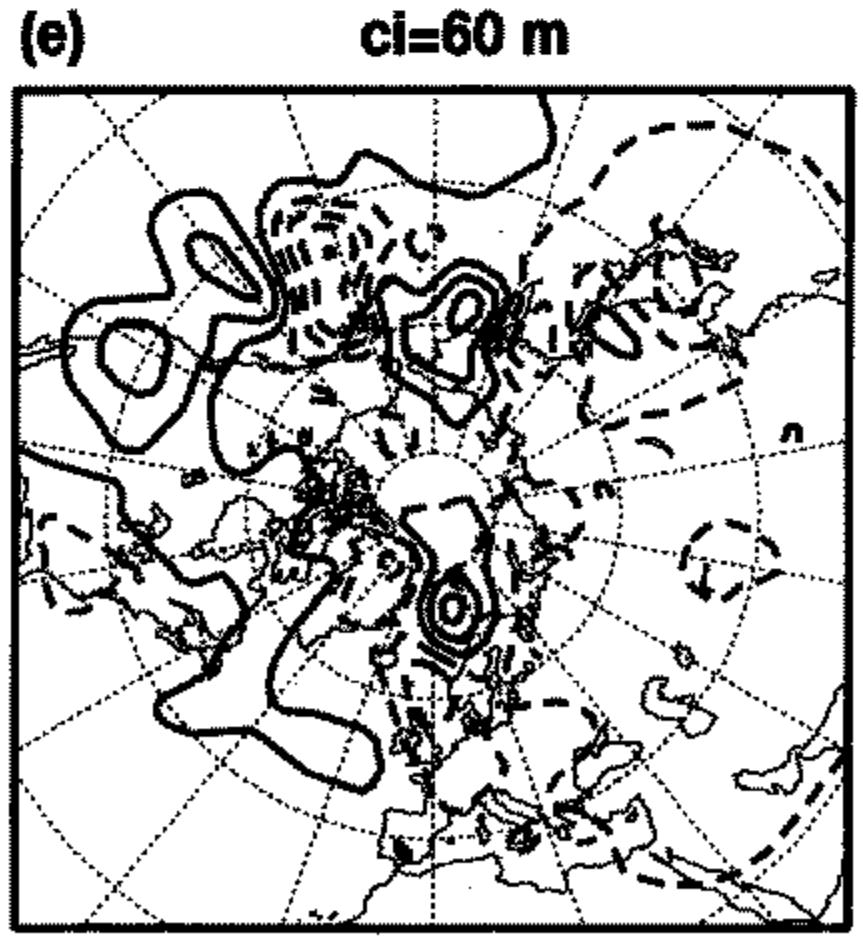


Figure 5. (a) Pseudo analysis $\tilde{\boldsymbol{\varepsilon}}(0)$ in geopotential height at 500 hPa for total energy singular vectors. (b) As (a) but for Hessian singular vectors. (c) As (a) but for corresponding $\tilde{\boldsymbol{\varepsilon}}(48)$. (d) As (b) but for corresponding $\tilde{\boldsymbol{\varepsilon}}(48)$. (e) 2-day geopotential height forecast error (m) from 1200 UTC 18 January 1997. Solid (dashed) lines denote positive (negative) values. Contour interval (ci) is given above each panel.

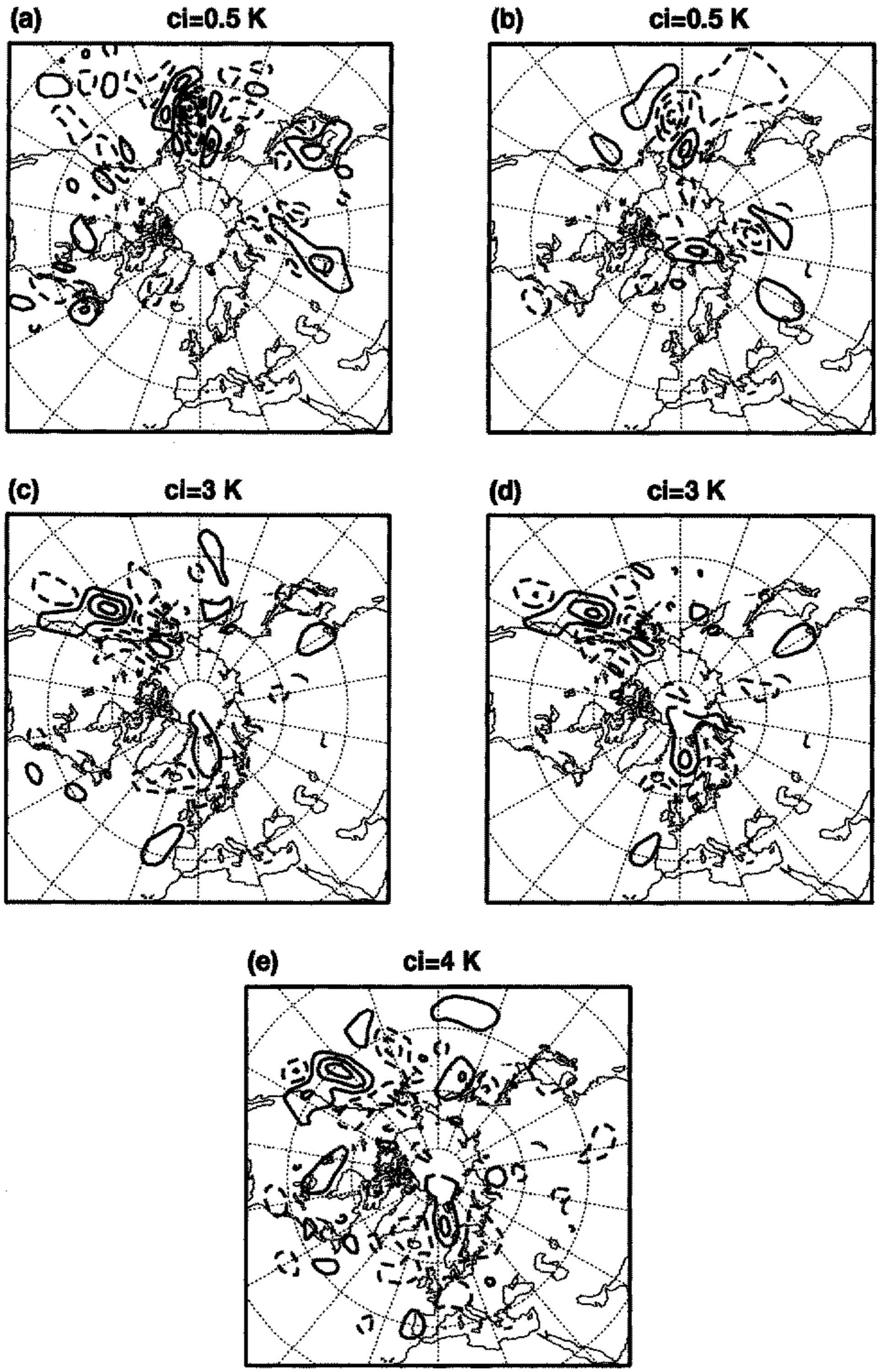
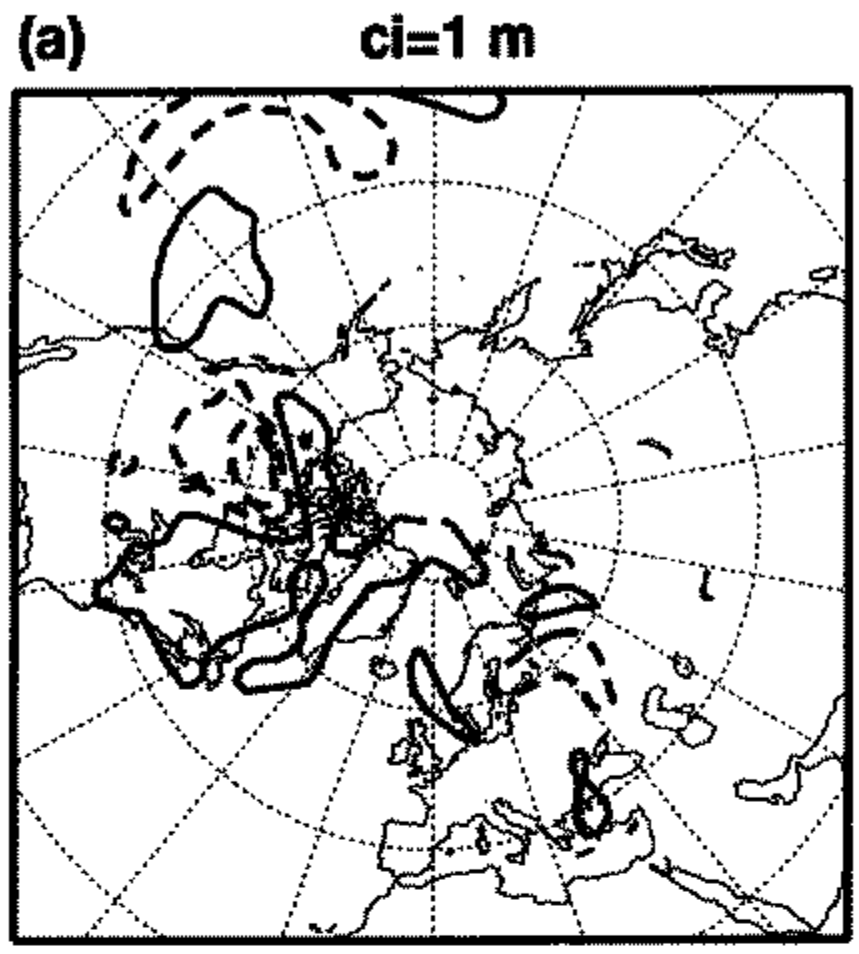
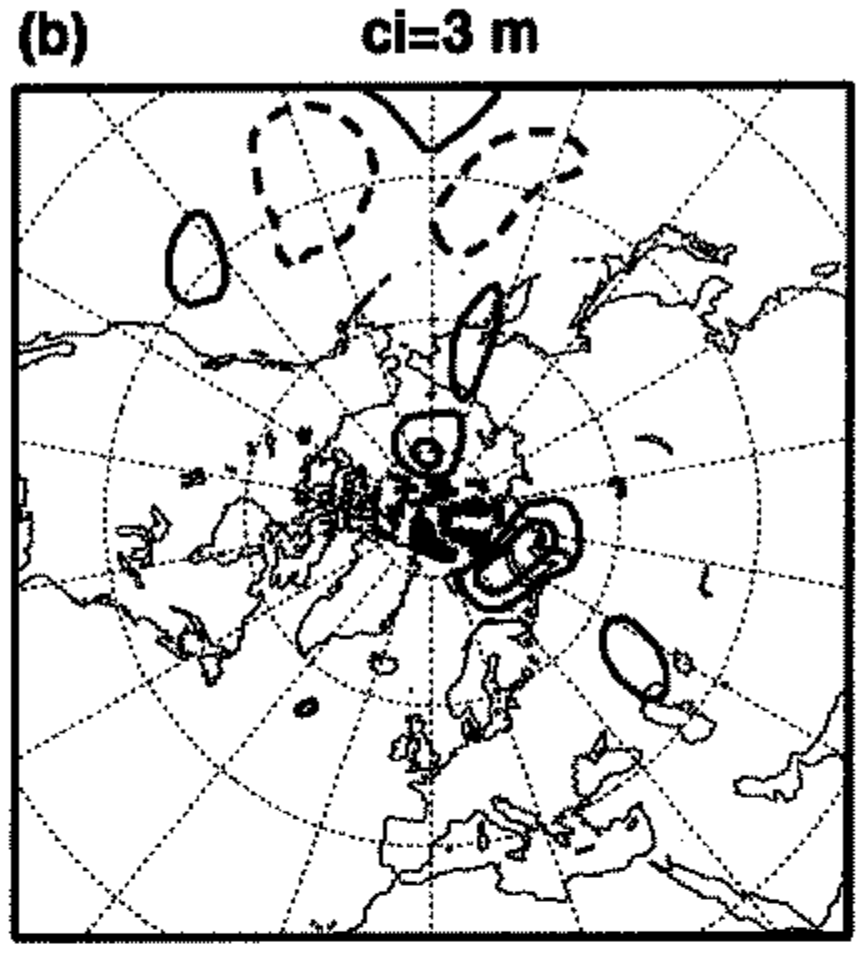
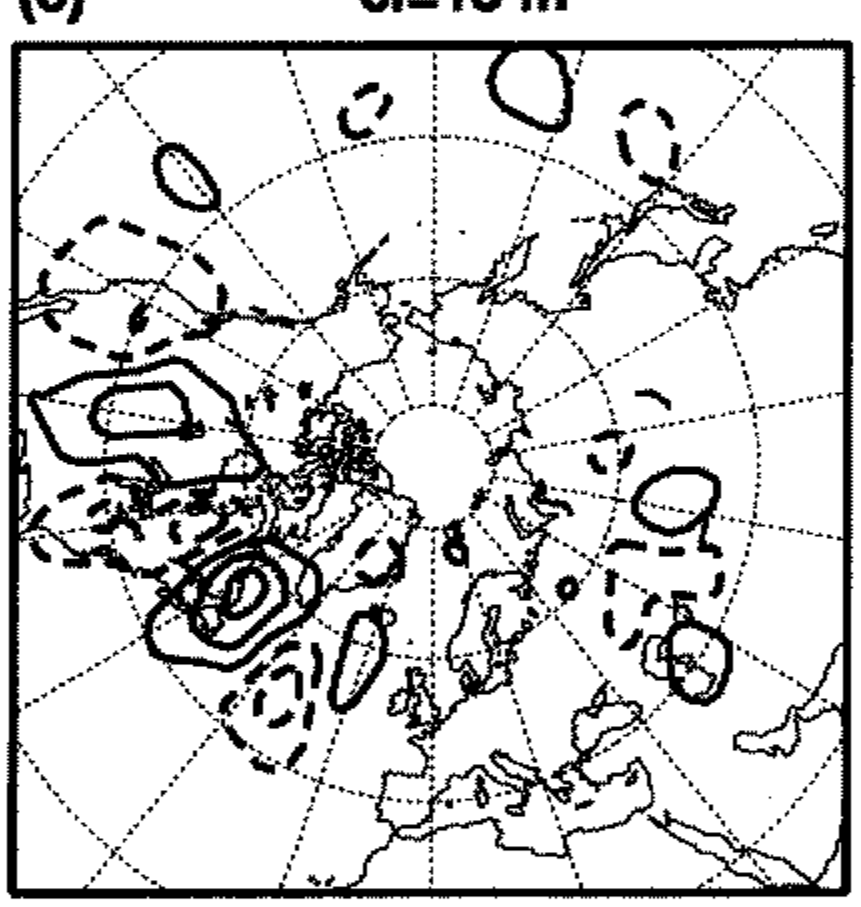
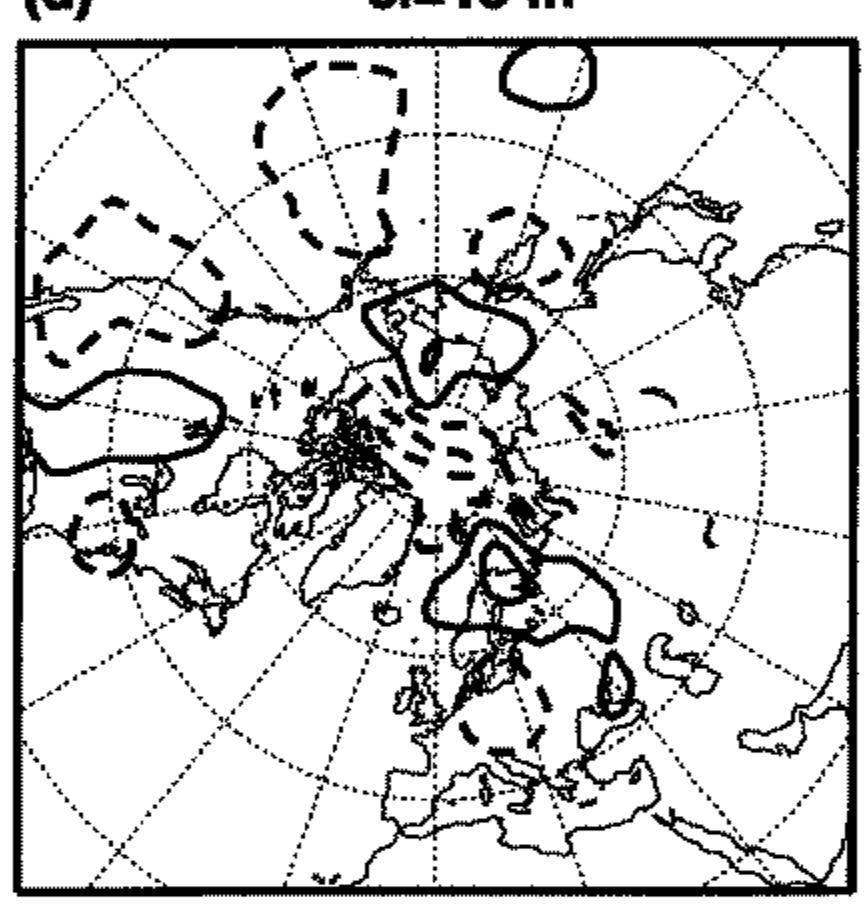


Figure 6. As Fig. 5 but for temperature forecast error (K) at 500 hPa.









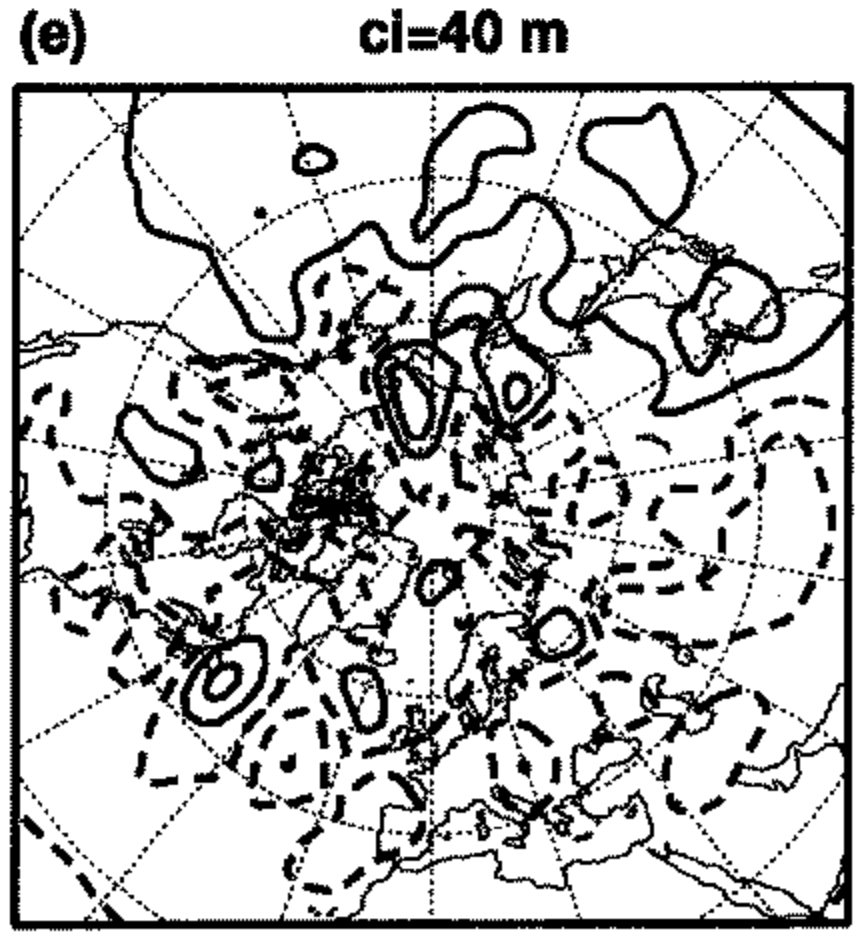


Figure 7. As Fig. 5 but for 1200 UTC 30 January 1997 and using 10 singular vectors.

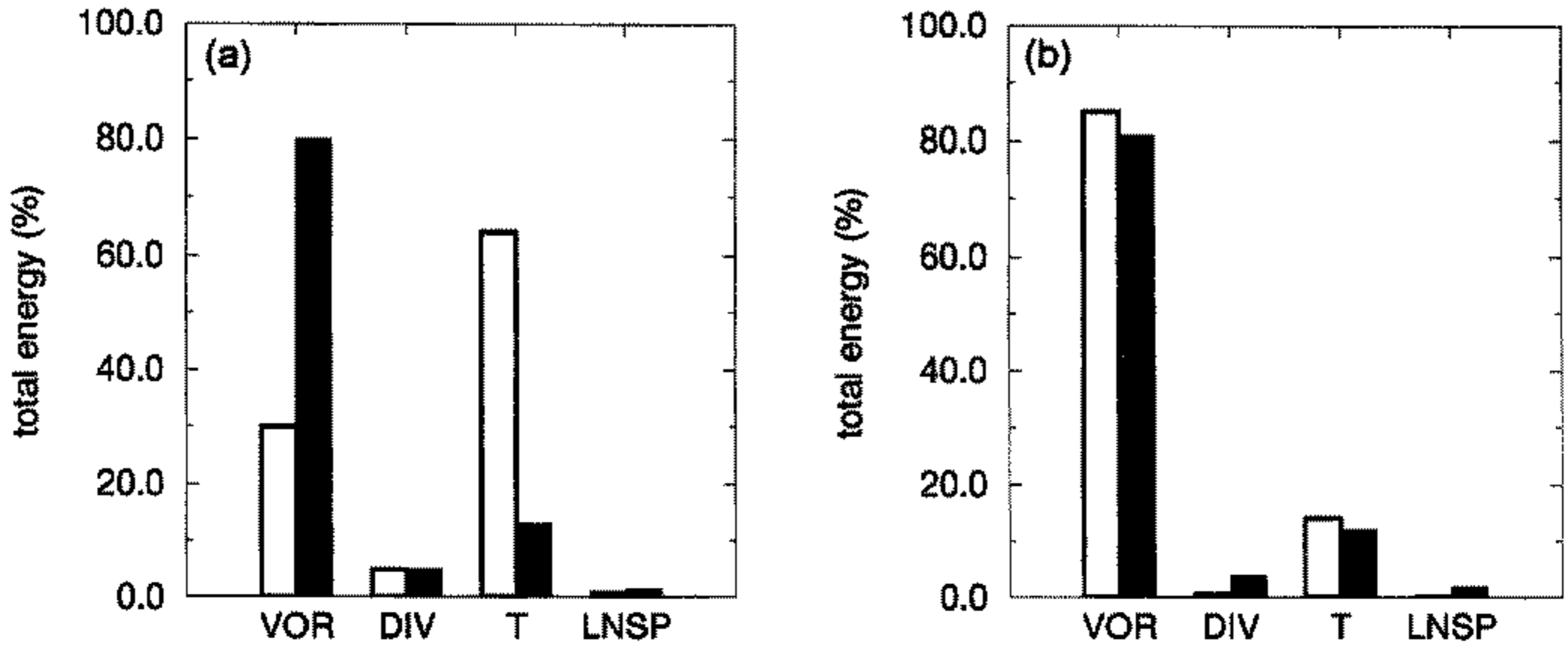


Figure 8. Distribution of total energy in percentages over the singular vector state vector components (vorticity, divergence, temperature and (log) surface pressure) at (a) initial and (b) final time. White (black) bars correspond to total energy singular vector (Hessian singular vector).

based ensemble configurations have a nearly equal spread with respect to the control over NH at day 2. This was done by tuning a parameter α that sets the amplitude of the initial perturbations. Here α represents the maximal acceptable ratio between the initial perturbations and an estimate of analysis error variances. Perturbations are compared to analysis error variances at 5 pressure levels and for zonal and meridional wind and temperature. In practice, the analysis error variances of temperature put the strongest constraint on the perturbation amplitude. Because of the completely different growth rate for HSVs and the different distribution of total energy over the state vector components, in particular temperature, this parameter α had to be changed significantly. It was reduced from 0.6 for the operational TESV perturbations to 0.25 for perturbations based on HSVs. As a third EPS configuration we exploited the use of evolved TESVs in creating initial perturbations. In addition to the perturbations, $\mathbf{p}_i$, of the operational EPS, 50 perturbations, $\mathbf{ep}_i$, are computed in a similar manner (and using the same initial amplitude) but now based on the 2-day linearly evolved singular vectors computed two days before. The initial perturbations, $\mathbf{pert}_i$, are defined by adding the two sets:

$$\mathbf{pert}_i = \mathbf{p}_i + \mathbf{ep}_i, \quad i = 1, \ldots, 50. \tag{9}$$

The use of evolved SVs provides an easy way to include more stable and large-scale directions in the generation of EPS perturbations. See Fig. 1(a) for the spectra of initial and evolved TESVs. In the following we refer to this configuration as the ESV ensemble.

Figure 9(a) shows, as a function of forecast time, the root-mean-square (r.m.s.) error of the control forecast for NH and the r.m.s. spread of the ensemble with respect to the control for the three EPS configurations, averaged over the 10 cases. Notice that the perturbations based on HSVs cannot create as much spread in the medium range as is present in the operational ensemble. Other verification areas show the same deficiency in spread in the medium range for the HSV ensemble. Although evolved singular vectors are in the more stable directions of the analysis error, they lead to an increased spread in the ESV ensemble not only during the quasi-linear stage but up to day 10. There is little difference between the skill of the ensemble mean for all three configurations in terms of the r.m.s. error as can be seen in Fig. 9(b). The ensemble mean of the HSV (ESV) ensemble is on average slightly worse (better) for the medium range.

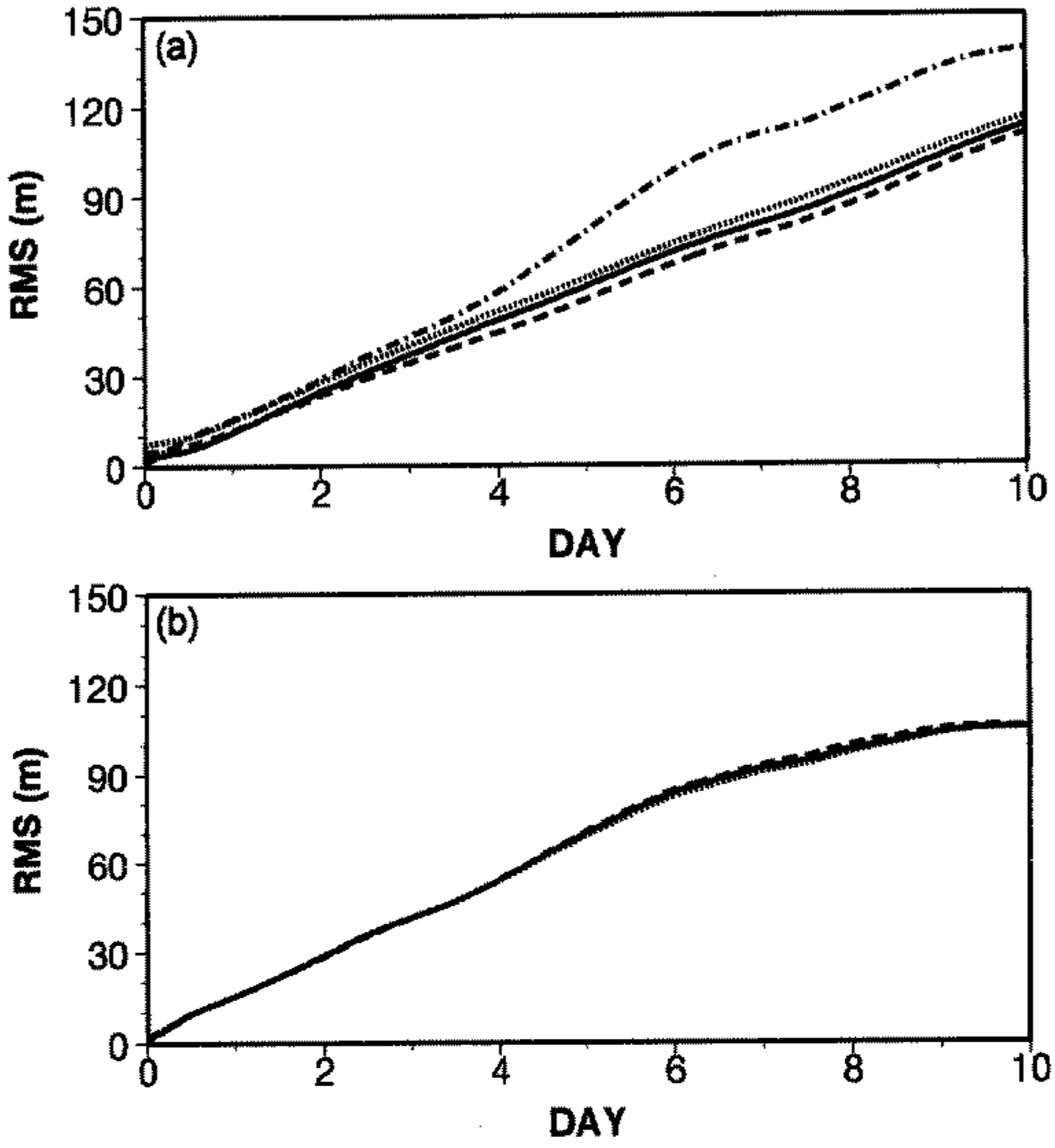


Figure 9. (a) Average root mean square (r.m.s.) height error (m) of the T_L159L31 (see text) ensemble control forecasts (chain-dashed), the average r.m.s. spread of the operational (solid), the Hessian singular vector (HSV; dashed) ensembles and the evolved singular vector (ESV; dotted) ensembles. (b) Average r.m.s. error of the ensemble mean for the operational (solid), HSV (dashed) and ESV (dotted) ensembles. All values are relative to 500 hPa geopotential height over the northern hemisphere.

In fact, ensemble means of the three configurations reveal a quite similar spatial pattern for each individual case. Figure 10 shows the difference between the ensemble mean and the control forecast at forecast day 5 for an arbitrarily chosen HSV and operational ensemble. For a more detailed discussion on this issue see Hersbach *et al.* (1998).

(*a*) *Brier skill score and relative operating characteristics*

Brier skill scores (Brier 1950; Stanski *et al.* 1989) have been computed for probability predictions of geopotential height anomalies exceeding certain thresholds (25 and 50 m positive/negative anomalies at 500 hPa) and analogously for temperature (4 and 8 K high/low anomalies at 850 hPa). Figure 11 gives the Brier skill score as a function of forecast time for a 50 m positive anomaly threshold at 500 hPa. The Brier skill score compares the skill of the probabilistic forecast to climatology: it is 1 for a perfect forecast, 0 when the probabilistic forecast does not perform better than climatology, and negative for even worse forecasts. All three ensemble configurations give indistinguishable results up to day 4, after which the ESV ensemble performs slightly better. The same conclusion holds for the other thresholds and for the Brier skill scores for temperature at 850 hPa.

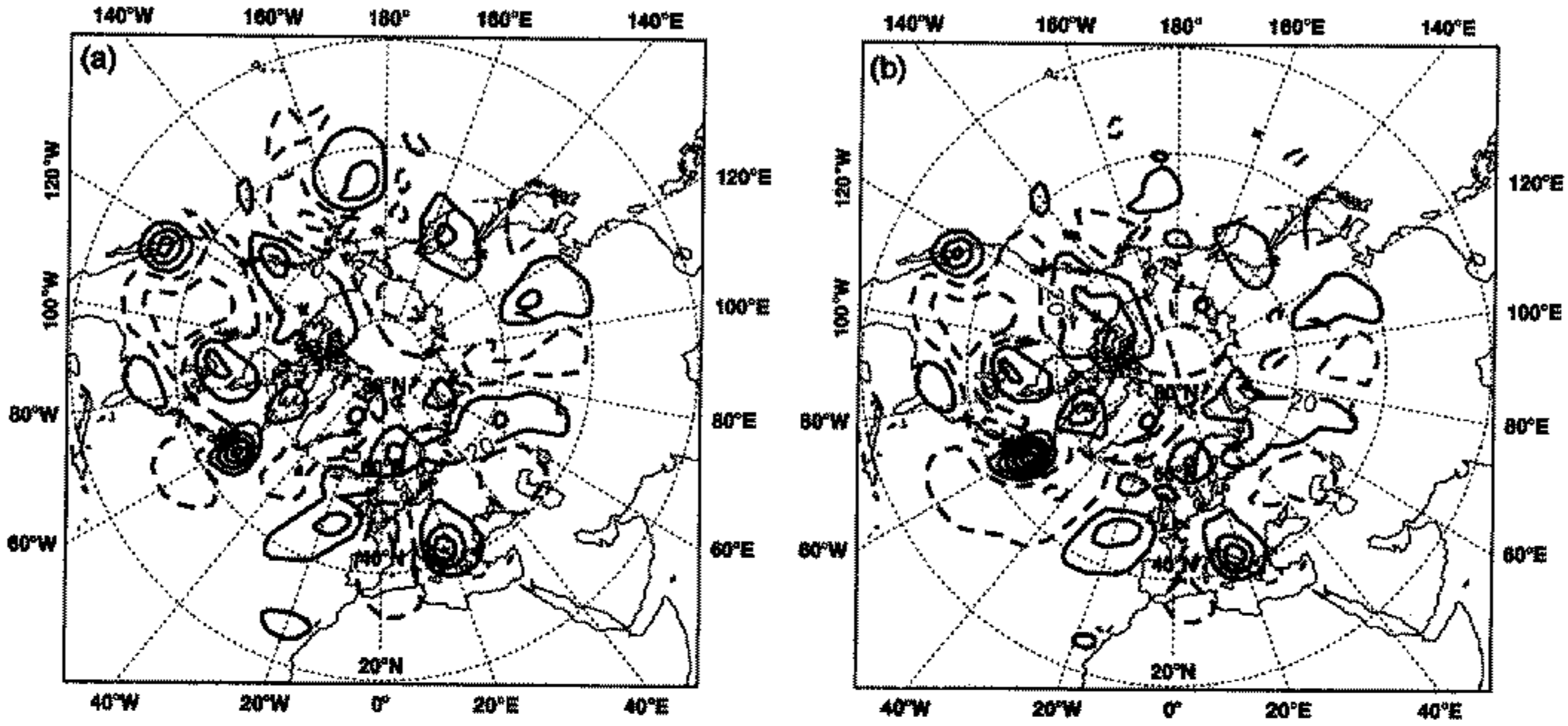


Figure 10. Geopotential height difference (m) at 500 hPa between the ensemble mean and the control forecast at day 5 for the (a) operational and (b) Hessian singular vector ensemble starting from 1200 UTC 9 February 1997. Solid (dashed) lines denote positive (negative) values and the contour interval is 20 m.

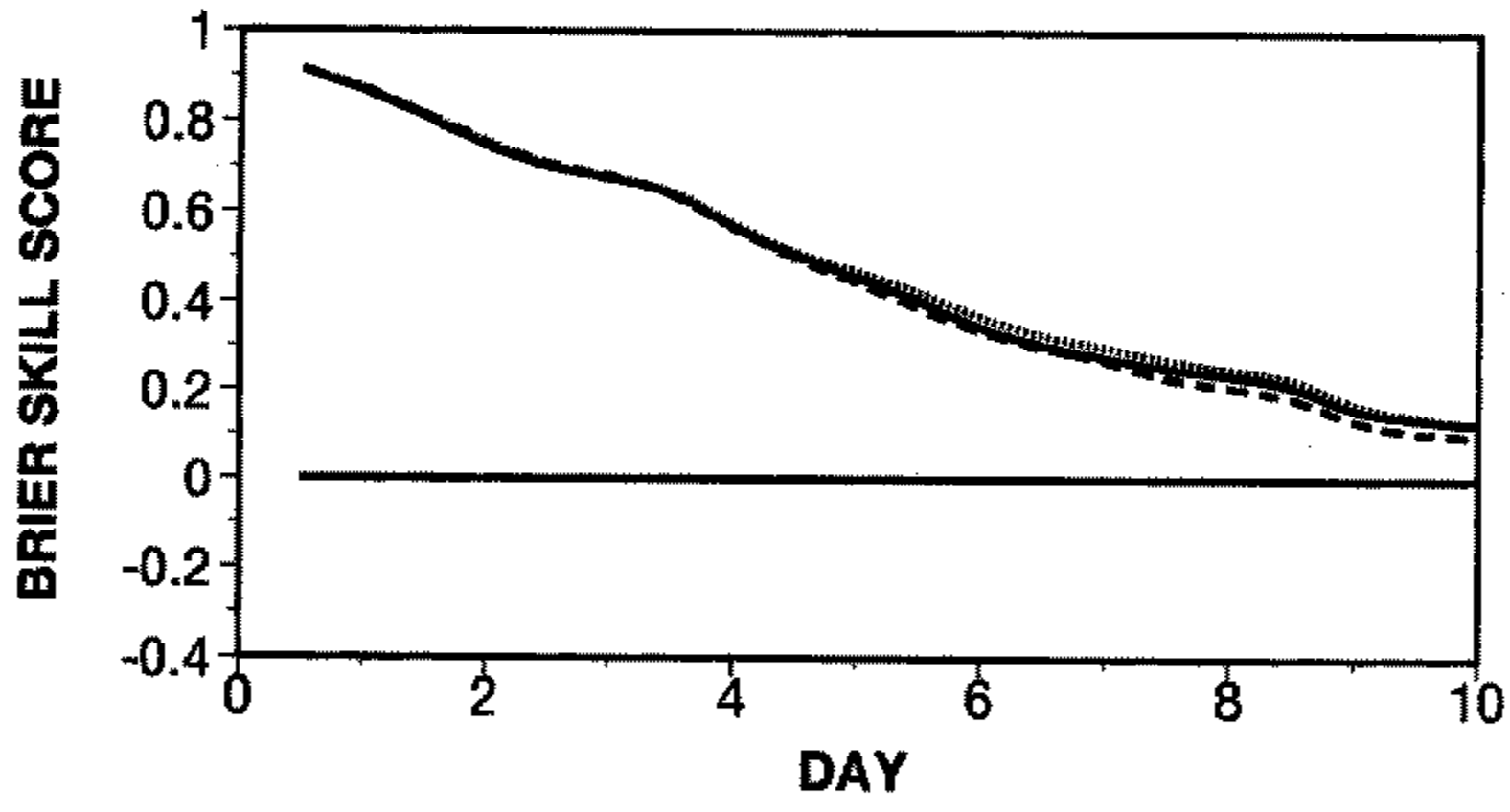


Figure 11. Brier skill score of the operational (solid), Hessian singular vector (dashed) and evolved singular vector (dotted) ensemble at different forecast times for a geopotential height anomaly threshold of 50 m at 500 hPa over the northern hemisphere.

From signal-detection theory (Mason 1982; Stanski *et al.* 1989) so-called relative operating characteristics (ROC) have been computed for the same variables and thresholds as the Brier skill score; see also the appendix. A convenient measure associated with the ROC is the area under the curve. It ranges from 1 for a perfect forecast system (i.e. a forecast system with zero false alarms) to 0. A value of 0.5 is produced by a useless forecast system which cannot discriminate between occurrences and non-occurrences of events. Figure 12 gives the ROC area for a negative geopotential height anomaly of 50 m at 500 hPa over NH. Other thresholds and ROC areas for temperature at 850 hPa give similar results, a slightly better performance of the experimental ensembles up to day 3. After day 4 the ESV ensembles continue to give the best results; the HSV ensembles show a deterioration in the performance compared to the operational ensembles. This confirms the results obtained with the Brier skill score.

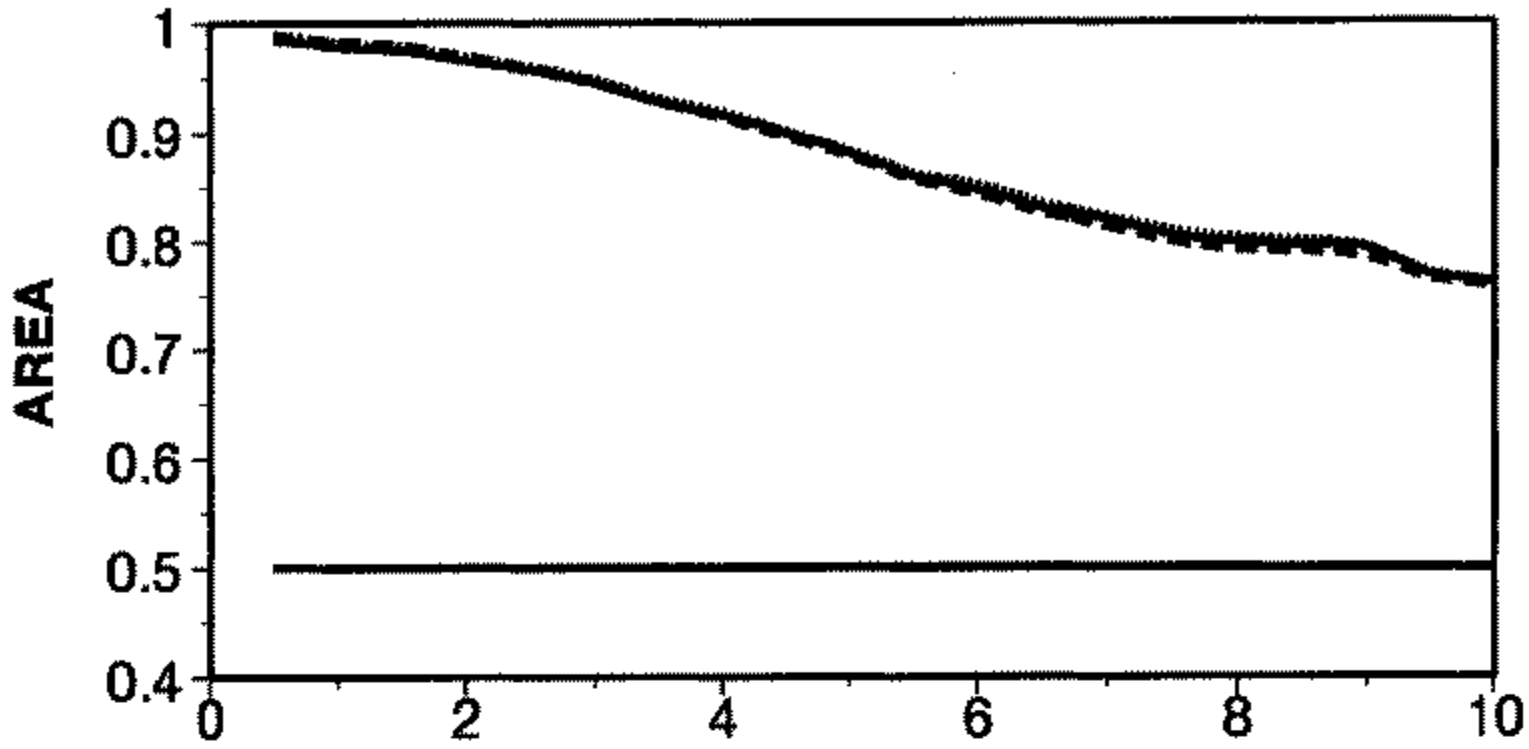


Figure 12. Relative operating characteristics area of the operational (solid), Hessian singular vector (dashed) and evolved singular vector (dotted) ensemble for a negative geopotential height anomaly of 50 m at 500 hPa over the northern hemisphere at different forecast times.

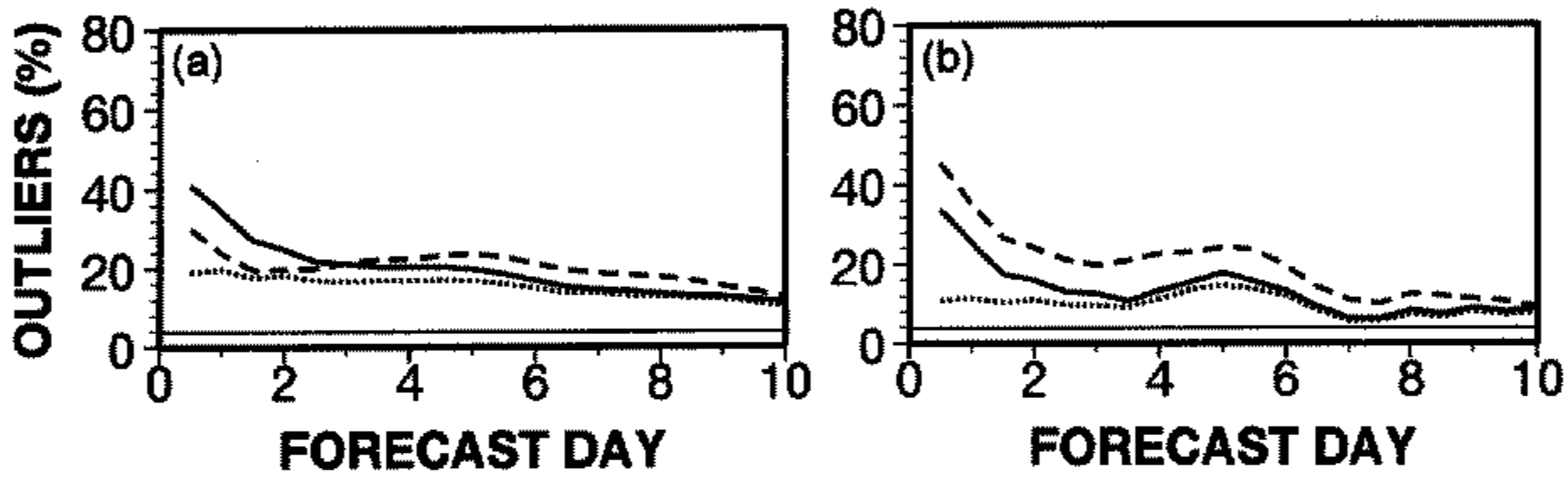


Figure 13. Percentage of analysis outliers for the operational (solid), Hessian singular vector (dashed) and evolved singular vector (dotted) ensembles as a function of forecast time for geopotential height at 500 hPa for: (a) the northern hemisphere, and (b) Europe.

(*b*) *Percentage of analysis outliers*

At each grid point the 50 ensemble values partition the real line into 51 intervals. Under the assumption of no model errors and a random sampling of the analysis error PDF, each interval is equally likely to contain the analysis value (when averaged over a verification area). Figure 13 gives the percentage of analysis values lying outside the ensemble forecast range for geopotential height at 500 hPa over NH and Europe (averaged over the 10 cases). When all ensemble members are equally likely, then the percentage of analysis outliers is $\frac{2}{51} \times 100\%$. Model errors and wrong sampling of the initial perturbations may cause the actual percentage of analysis outliers to be considerably larger than the expected value. The experimental ensembles have less outliers than the operational ensemble for the short range with respect to NH. This coincides with the larger spread in the experimental ensembles during the first two days, see Fig. 9(a). Areas where ensembles produce less spread than the operational ensemble, such as Europe (not shown) for the HSV ensembles, yield larger percentages of analysis outliers. The ESV ensembles gave the smallest percentage of outliers. This is because the orthogonality of initial and evolved SVs results in a more uniform perturbation coverage, especially during the first few days.

5. FINAL REMARKS

In this paper so-called Hessian singular vectors (HSVs) are computed which, at initial time, are constrained by an estimate of the analysis error covariance metric. Up to now the calculation of SVs as used in the ECMWF EPS is based on an energy metric at initial time which may be considered as a first approximation of the analysis error covariance metric (Palmer *et al.* 1998a). In computing HSVs the full Hessian of the cost function of the variational data assimilation is used as an approximation to the analysis error covariance matrix. In this way the calculation of singular vectors and the analysed state in 3D/4D-Var become consistent.

The HSVs are solutions of a generalized eigenvalue problem, and by using a generalization of the Davidson algorithm (Davidson 1975; Sleijpen and Van der Vorst 1996) the leading SVs can be determined. It only requires that the propagators of the linear and adjoint model and the Hessian of the 3D-Var cost function are available in operator form, i.e. $\mathbf{y} = \mathbf{S}\mathbf{x}$ can be computed, where $\mathbf{S}$ is any of these operators and $\mathbf{x}$ is an input vector. The computation of 25 HSVs, as needed for the ensemble perturbations, is of the order of 5 times more expensive than the computations of the same number of TESVs.

Earlier results obtained with a T21L5 PE model (Barkmeijer *et al.* 1998) already indicated significant differences between HSVs and TESVs. In the present study these results are confirmed for SVs with a resolution at T42L31. At initial time the horizontal structure of HSVs is more large scale than TESVs, with energy spectra attaining their maximum at wave numbers 10 and 30 respectively. The energy spectra at optimization time are comparable. Also the vertical structure of HSVs and TESVs show a striking difference in terms of the energy distribution. Most of the HSV energy is at initial time confined to the jet level, instead of peaking around the baroclinic steering level as TESV energy. At final time both type of SVs show the same vertical energy distribution. The growth rate of TESVs in terms of total energy is typically twice as large as for HSVs.

The large-scale structure and the vertical energy distribution of HSVs are to a large extent determined by the formulation of the background error covariance matrix $\mathbf{B}$. The first-guess error statistics are based on the difference between the 2-day and 1-day forecast valid for the same day (the so-called NMC method; Parrish and Derber 1992). The $\mathbf{B}$ matrix defined in this way lacks a realistic description of flow-dependent small-scale error structures. Also the broad horizontal and vertical correlations will penalize the occurrence of baroclinic structures in the analysis error. From this it is clear that an SV computation using the 3D-Var Hessian is not optimal in dynamically unstable areas.

A first improvement in making $\mathbf{B}$ more realistic would be to relax its static character by including some flow-dependent error covariances. This approach is attempted at ECMWF by experimenting with a simplified Kalman filter in an operational environment (Fisher 1998; Rabier *et al.* 1997). Here the $\mathbf{B}$ matrix is modified for each 6 h analysis cycle in the unstable subspace spanned by the leading HSVs. Parallel to this, a full Kalman filter is being developed in the context of a T21L3 quasi-geostrophic model (Ehrendorfer 1998; Ehrendorfer and Bouttier 1998).

Finally, the impact of HSVs in the ECMWF EPS was investigated. Perturbations based on HSVs were determined in a similar way as for the operational ensemble. Both the operational and HSV ensemble configurations use perturbations that give comparable spreads at day 2 for NH relative to geopotential height at 500 hPa. In addition to this, the use of evolved SVs was investigated. These 2-day linearly evolved SVs (ESVs), computed two days before, provide an easy way to add large-scale structures to the EPS perturbations. The initial perturbations for the ESV ensembles are

defined by adding the operational perturbations to perturbations based on the evolved SVs (using the same initial amplitude as used for the operational perturbations).

Results show that the HSV perturbations give slightly less spread over NH in the medium range compared to the operational ensemble. This is in contrast to the ESV ensembles where the spread is slightly larger up to day 10. The percentage of analysis outliers over NH has decreased for the experimental ensembles up to day 3. For longer lead times the ESV ensembles have the smallest percentage of outliers. Statistical verification methods, such as ROC and Brier skill score, for geopotential height at 500 hPa and temperature at 850 hPa do not show large differences between the three ensemble configurations. The impact of the HSV or ESV perturbations is neutral or positive for forecasts up to day 3; after day 4 the ESV ensembles show a slightly better performance. Additional experimentation with ESV perturbations confirmed a small but consistent improvement using the Brier skill score and ROC, and a substantial decrease in the percentage of analysis outliers. From 25 March 1998 onwards the ESV perturbations are used in the operational ECMWF EPS (the parameter α was slightly reduced from 0.6 to 0.5).

ACKNOWLEDGEMENTS

We thank F. Bouttier and M. Fisher for the stimulating discussions during the preparation of this paper. Also thanks are due to A. Hollingsworth and A. Simmons for reading the manuscript carefully. The reviews of two referees provided helpful comments and suggestions.

APPENDIX

The relative operating characteristic (ROC) originating from signal-detection theory is briefly described here. The reader is referred to Stanski *et al.* (1989) for a more detailed definition. Consider a two-category contingency table:

	Forecast = Yes	Forecast = No	Total Observed
Observed = Yes	X	Y	X + Y
Observed = No	Z	W	Z + W
Total Forecast	X + Z	Y + W	

The two entries X and Z can be referred to as 'hits' and 'false alarms' respectively. The hit rate is then given by X/(X + Y) (percentage of correct forecasts) and the false alarm rate by Z/(Z + W) (percentage of forecasts of the event, given that the event did not occur).

Signal detection theory generalizes the concept of hit and false alarm rates to probability forecasts. Suppose a forecast distribution is stratified into probability categories 10% wide, and non-occurences, a_i, and occurences, b_i, of an event are tabulated for each category. Here the j-th category is related to a forecast probability between $(j-1) \times 10\%$ and $j \times 10\%$. For a certain probability threshold, $j \times 10\%$, the entries a_i and b_i can be summed to give the four entries of the two-by-two contingency table, the hit and false alarm rate calculated and a point plotted on a graph. The four entries are given by $W = \sum_{i=1}^{j} a_i$, $Y = \sum_{i=1}^{j} b_i$, $Z = \sum_{i=j+1}^{10} a_i$ and $X = \sum_{i=j+1}^{10} b_i$. By repeating this process for all thresholds $j \times 10\%$, $j = 1, \ldots, 10$, a curve is obtained called the ROC.

A convenient measure associated with the ROC is the area under the curve, which decreases from 1 to 0 as the false-alarm rate increases. In this approach a useless probability forecast has an area of 0.5, because such a system cannot discriminate between the occurrences and non-occurrences of events.

REFERENCES

Barkmeijer, J., Gijzen, Van M. and Bouttier, F. 1998 Singular vectors and estimates of the analysis error covariance metric. *Q. J. R. Meteorol. Soc.*, **124,** 1695–1713

Bouttier, F., Derber, J. and Fisher, M. 1997 'The 1997 revision of the J_b term in 3D/4D-Var'. ECMWF Technical Memorandum No. 238. Available from ECMWF, Shinfield Park, Reading, UK

Brier, G. W. 1950 Verification of forecasts in terms of probability. *Mon. Weather Rev.*, **78,** 1–3

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

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

Buizza, R., Gelaro, R., Molteni, F. and Palmer, T. N. 1997 The impact of increased resolution on predictability studies with singular vectors. *Q. J. R. Meteorol. Soc.*, **123,** 1007–1033

Buizza, R., Petroliagis, T., Palmer, T. N., Barkmeijer, J., Hamrud, M., Hollingsworth, A., Simmons, A. and Wedi, N. 1998a Impact of model resolution and ensemble size on the performance of an ensemble prediction system. *Q. J. R. Meteorol. Soc.*, **124,** 1935–1960

Buizza, R., Miller, M. and Palmer, T. N. 1998b 'Stochastic simulation of model uncertainties in the ECMWF ensemble prediction system'. In Proceedings of the ECMWF workshop on predictability, 20–22 October 1997. ECMWF, Shinfield Park, Reading, UK

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljević, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M. 1998 The ECMWF implementation of three dimensional variational assimilation (3D-Var). I: Formulation. *Q. J. R. Meteorol. Soc.*, **124,** 1783–1807

Davidson, E. R. 1975 The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices. *J. Comput. Phys.*, **17,** 87–94

Ehrendorfer, M. 1994 The Liouville equation and its potential usefulness for the prediction of forecast skill. Part I: Theory. *Mon. Weather Rev.*, **122,** 703–713

1998 'Prediction of the uncertainty of numerical weather forecasts; problems and approaches'. In Proceedings of the ECMWF workshop on predictability, 20–22 October, 1997. ECMWF, Shinfield Park, Reading, UK

Ehrendorfer, M. and Bouttier, F. 1998 'An explicit low-resolution extended Kalman filter: Implementation and preliminary results'. ECMWF Technical Memorandum No. 259. Available from ECMWF, Shinfield Park, Reading, UK

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

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

Fisher, M. 1998 'Development of a simplified Kalman filter'. ECMWF Technical Memorandum No. 260. Available from ECMWF, Shinfield Park, Reading, UK

Fisher, M. and Courtier, P. 1995 'Estimating the covariance matrices of analysis and forecast error in variational data assimilation'. ECMWF Technical Memorandum No. 220. Available from ECMWF, Shinfield Park, Reading, UK

Harrison, M. S. J., Palmer, T. N., Richardson, D. S., Buizza, R. and Petroliagis, T. 1995 'Joint ensembles from the UKMO and ECMWF models'. Proceedings of the ECMWF seminar on predictability: Vol. II, 4–8 September 1995. ECMWF, Shinfield Park, Reading, UK

Hersbach, H., Mureau, R., Opsteegh, J. D. and Barkmeijer, J. 1998 'An EPS for the short and early medium range'. In Proceedings of the ECMWF workshop on predictability, 20–22 October 1997. ECMWF, Shinfield Park, Reading, UK

Houtekamer, P. L., Lefaivre, L., Derome, J., Ritchie, H. and Mitchell, H. L. 1996 A system simulation approach to ensemble prediction. *Mon. Weather Rev.*, **124,** 1225–1242

Mason, I. 1982 A model for assessment of weather forecasts. *Aust. Meteorol. Mag.*, **30,** 291–303

Molteni, F., Buizza, R., Palmer, T. N. and Petroliagis, T. 1996 The ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc.*, **122,** 269–298

Oortwijn, J. and Barkmeijer, J. 1995 Perturbations that optimally trigger weather regimes. *J. Atmos. Sci.*, **52,** 3932–3944

Palmer, T. N., Molteni, F., Mureau, R., Buizza, R., Chapalet, P. and Tribbia, J. 1992 'Ensemble prediction'. Proceedings of the ECMWF seminar on validation of models over Europe: Vol I, 7–11 September 1992. ECMWF, Shinfield Park, Reading, UK

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

Palmer, T. N., Buizza, R. and Lalaurette, F. 1998b Performance of the ECMWF Ensemble Prediction System. In Proceedings of the ECMWF workshop on predictability, 20–22 October 1997. ECMWF, Shinfield Park, Reading, UK

Parlett, P. 1980 *The symmetric eigenvalue problem*. Series in computational mathematics. Prentice Hall, New Jersey, USA

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

Rabier, F. and Courtier, P. 1992 Four-dimensional assimilation in the presence of baroclinic instability. *Q. J. R. Meteorol. Soc.*, **118,** 649–672

Rabier, F., Mahfouf, J.-F., Fisher, M., Järvinen, H., Simmons, A., Andersson, E., Bouttier, F., Courtier, P., Hamrud, M., Haseler, J., Hollingsworth, A., Isaksen, L., Klinker, E., Saarinen, S., Temperton, C., Thépaut, J.-N., Undén, P. and Vasiljeviç, D. 1997 'Recent experimentation on 4D-Var and first results from a simplified Kalman filter'. ECMWF Technical Memorandum No. 240. Available from ECMWF, Shinfield Park, Reading, UK

Rabier, F., McNally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A. and Bouttier, F. 1998 The ECMWF implementation of three dimensional variational assimilation (3D-Var). II: Structure functions. *Q. J. R. Meteorol. Soc.*, **124,** 1809–1829

Rennick, M. A. 1995 'The ensemble forecast system (EFS)'. Technical Note 2–95, Fleet Numerical Meteorology and Oceanography Center. Available from Models Department, FLENUMMETOCCEN, 1 Grace Hopper Avenue, Monterey, CA 93943, USA

Richardson, D. 1998 'The relative effect of model and analysis differences on ECMWF and UKMO operational forecasts'. In Proceedings of the ECMWF workshop on predictability, 20–22 October 1997. ECMWF, Shinfield Park, Reading, UK

Sleijpen, G. L. G. and van der Vorst, H. A. 1996 A Jacobi–Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, **17,** 401–425

Stanski, H. R., Wilson, L. J. and Burrows, W. R. 1989 'Survey of common verification methods in meteorology'. Research report No. 89–5, Atmospheric Environment Service, Forecast Research Division, 4905 Dufferin Street, Downsview, Ontario, Canada

Thépaut, J.-N., Courtier, P., Belaud, G. and Lemaître, G. 1996 Dynamical structure functions in a four-dimensional variational assimilation: A case study. *Q. J. R. Meteorol. Soc.*, **122,** 535–561

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