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

J. Barkmeijer, R. Buizza and T.N. Palmer

European Centre for Medium-Range Weather Forecasts

Abstract

Singular vectors are computed which are consistent with 3D-Var 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 differ considerably from total energy singular vectors in such aspects as energy spectrum and growth rate. Despite these differences, the leading 25 total energy and Hessian singular vectors explain nearly the same part of the 2-day forecast error.

Two experimental ensemble configurations are studied. One configuration uses perturbations based on Hessian singular vectors, the other uses total energy singular vectors and 2-day linearly evolved singular vectors of two days before in the computation of initial perturbation. 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 ECMWF Ensemble Prediction System. The ensembles using evolved singular vectors perform slightly better. The Hessian singular vectors based ensembles show a slightly worse performance and are lacking spread in the medium range. Possible directions to improve the computation of Hessian singular vectors are discussed.

1 Introduction

The current operational Ensemble Prediction System (EPS) at the European Centre for Medium-Range Weather Forecasts (ECMWF) comprises 50 non-linear integrations of the T_L159 operational model version, with as initial condition the 12 UTC analysis perturbed along growing

directions. The development of the EPS is documented in (Palmer et al, 1992; Molteni et al, 1996; Buizza et al, 1998a) and recent findings indicate that its performance is quite useful (Palmer et al, 1998). Two other numerical weather prediction (NWP) centres run an EPS on a daily basis, namely the US National Centers for Environmental Prediction (NCEP) and the Canadian Meteorological Centre (CMC). The approaches followed by the various NWP centres to define their EPS differs considerably. At NCEP the breeding method (Toth and Kalnay, 1997) is used 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 (SV) are used to obtain initial perturbations for the EPS (Buizza and Palmer, 1995), as they are believed to sample the unstable linear subspace as efficiently as possible. Currently, two sets of SVs at resolution T42L31 are computed, targeted respectively for the Northern and Southern Hemisphere extra tropics and with an optimization time of 2 days. From these, 25 perturbations are generated for each hemisphere seperately (Molteni et al, 1996) and then the two sets of perturbations are added to yield global perturbations. By adding and subtracting the global perturbations to the 12 UTC analysis the 50 perturbed initial conditions for the EPS are defined.

In order to obtain the forecast probability distribution function (PDF) under the assumption of a perfect forecast model, one, ideally, would like to integrate the appropriate Liouville equation (Epstein, 1969; Ehrendorfer, 1994). However, the large dimension of the current NWP models makes this impossible. The SV approach provides a possibility to search 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 (50) of the ensemble and the dimension of NWP models (10^6–10^7) may lead to sampling errors in describing the forecast probability distribution function: initial directions in phase space which result in erroneous forecasts may easily be missed. In fact, estimates of the dimension of the linear unstable subspace for a 1449-variable T21L3 QG model are of the order 10^2–10^3, (Palmer et al, 1997). 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 quasi-geostrophic model (Oortwijn and Barkmeijer, 1996). Weather regime transitions may be associated with large spread in the ensemble. A study by Trevisan et al (1997) in a 2-level quasi-geostrophic model indicated

that the use of singular vectors are preferable to Lyapunov vectors in ensemble forecasting to detect the enhanced spread associated with regime transitions. 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 for model perturbations during the actual time integration of each ensemble member (Buizza et al, 1998b).

In this paper we assume a perfect model approch 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, 1997). 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 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 analyzed state. The total energy norm is still used at optimization time. We call singular vectors calculated in this way Hessian singular vectors. 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 Hessian singular vectors 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 two sets of 10 ensembles starting from days in DJF 1996/97 will be compared where initial perturbations are created with total energy or Hessian singular vectors.

The organization of the paper is as follows. In the next section we introduce the computation of the Hessian singular vectors. The comparison of the Hessian and total energy singular vectors

is given in section 3. Results from the 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 tests, the relative operating characteristic, is briefly described.

2 Use of the 3D-Var Hessian

The computation of singular vectors requires the specification of a norm at initial time t_0 and at optimization time t_1. In this paper we consider singular vectors ε which maximize the ratio

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

Here $\langle\ ,\ \rangle$ denotes the Euclidean inner product $\langle \mathbf{x}, \mathbf{y}\rangle = \sum \mathbf{x}_i \mathbf{y}_i$. The positive definite and symmetric operators $\mathbf{C}$ and $\mathbf{E}$ induce a norm at initial and optimization time respectively. The projection operator $\mathbf{P}$ 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 singular vectors $\varepsilon(t_1) = \mathbf{M}\varepsilon(t_0)$ form an $\mathbf{E}$-orthogonal set at optimization time. Alternatively, the singular vectors defined by (1) are solutions of the following generalized eigenvalue problem

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

The adjoint operators $\mathbf{M}^\star$ and $\mathbf{P}^\star$ are determined with respect to the Euclidean inner product. In the computation of the total energy singular vectors (TESV), the total energy metric is used at initial and optimization time, i.e. $\mathbf{E}$ and $\mathbf{C}$ are identical and:

$$\begin{aligned}\langle \mathbf{x}, \mathbf{Ey}\rangle = \frac{1}{2}\int_0^1 \int_\Sigma (\nabla\Delta^{-1}\zeta_x \cdot \nabla\Delta^{-1}\zeta_y + \nabla\Delta^{-1}\mathbf{D}_x \cdot \nabla\Delta^{-1}\mathbf{D}_y \\ + \frac{c_p}{T_r}\mathbf{T}_x\mathbf{T}_y)\mathrm{d}\Sigma(\frac{\partial p}{\partial \eta})\mathrm{d}\eta + \frac{1}{2}\int_\Sigma R_d T_r P_r \ln\pi_x \cdot \ln\pi_y \mathrm{d}\Sigma\end{aligned} \tag{3}$$

with $(\zeta_x, \mathbf{D}_x, \mathbf{T}_x, \ln\pi_x)$ being the vorticity, divergence temperature and logarithm of the surface pressure components of the state vector $\mathbf{x}$ and c_p is the specific heat of dry air at constant pressure, R_d is the gas constant for dry air, $T_r = 300$ K is a reference temperature and $P_r = 800$ hPa is a reference pressure.

In this TESV case $\mathbf{C}$ has the form of a diagonal matrix and the square root of $\mathbf{C}$ can be readily determined. Multiplying both sides of (2) to the left and right with the inverse of the square root of $\mathbf{C}$, yields an equation which can be solved using the Lanczos algorithm (Strang, 1986). Palmer et al (1997) study the impact of choosing different simple metrics at initial time, keeping

the total energy metric at optimization time. It turns out that of the simple metrics considered, the total energy metric is the most consistent with the analysis error statistics. In Barkmeijer et al (1998) a method is proposed to make the singular vector computation 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

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

is minimized. 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 Hessian singular vector 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 covergence of the generalized Davidson algorithm a coordinate transformation $\chi = \mathbf{L}^{-1}\mathbf{x}$ 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 singular vectors 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 initial days for which the ensemble spread and the ensemble control forecast at day 5 over the Northern Hemisphere either have values that lie below or above their 1996/97 winter average. The final norm in the defining equation (2) is the total energy norm (3). The solutions of this generalized eigenvalue problem will be referred to as Hessian singular vectors (HSV). 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 the Northern Hemisphere.

The HSVs have properties considerably different from the total energy singular vectors (TESV) which use the total energy norm also at initial time. Figures 1a,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 wavenumber around 30 compared with 10. The spectra at final time (solid line) for both sets of SVs peak around wavenumber 10. Also the vertical structure of HSVs and TESVs are quite different, see Fig 1c,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 Figures 1a,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. Observe that, as a consequence of using different norms at initial and final time, the ordering of the HSVs does not correspond anymore with their total energy growth.

The large scale structure and energy distribution of HSVs is to a large extent determined by the formulation of the background error covariance matrix $\mathbf{B}$. Essential in defining the $\mathbf{B}$

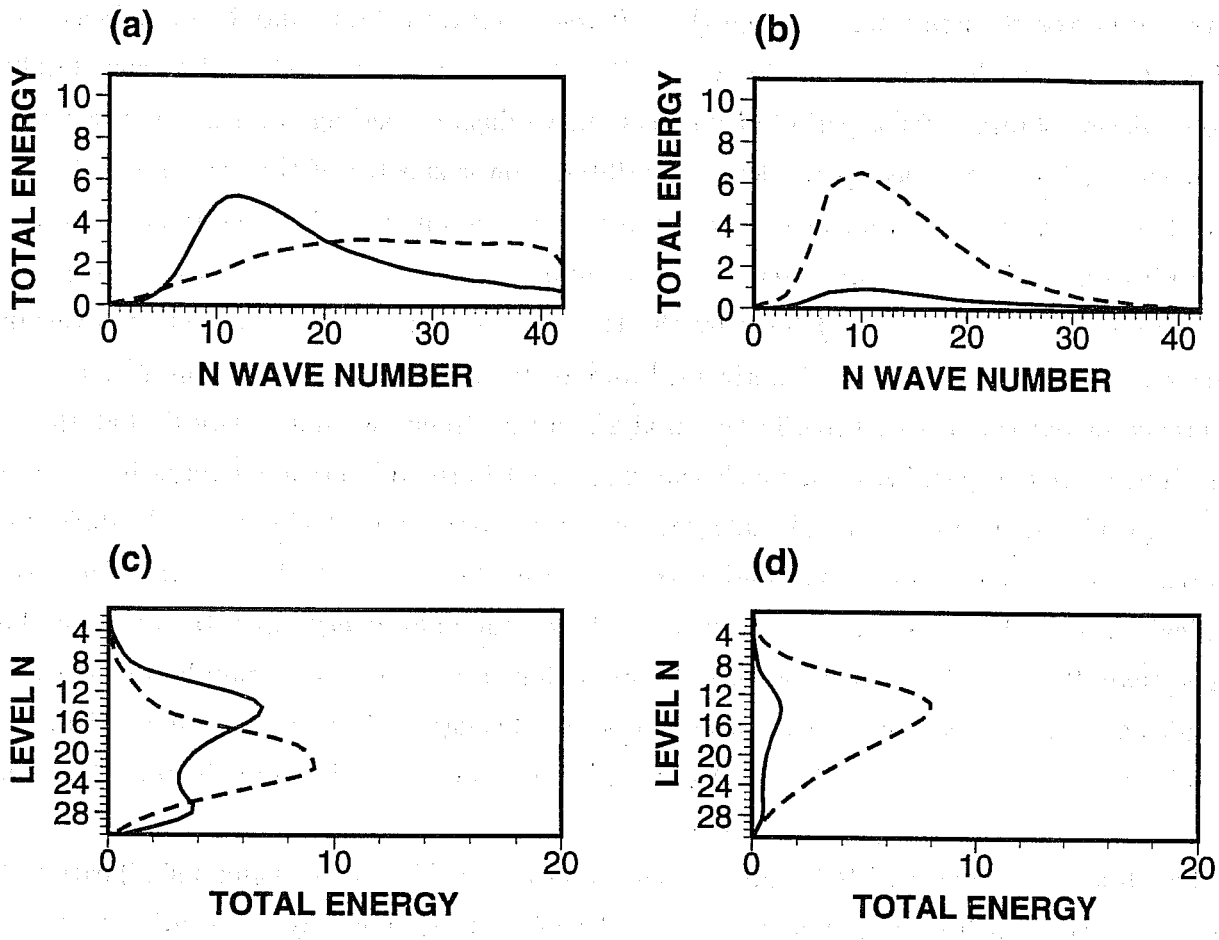


Figure 1: *(a-b) total energy spectrum and (c-d) vertical distribution of the total energy spectrum of TESVs and HSVs respectively. Values at initial (final) time are given by dashed (solid) lines. At initial time the total energy ($m^2 s^{-2}$) has been multiplied by a factor of 100.*

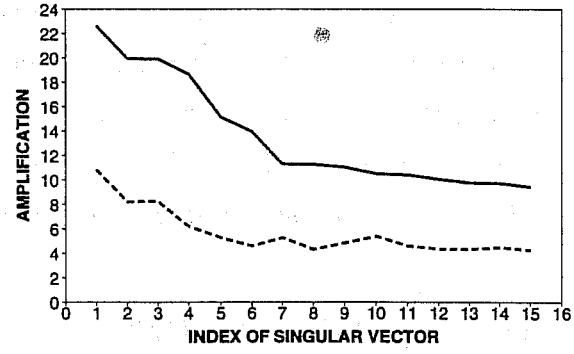


Figure 2: *Typical total energy amplification of TESVs (solid) and HSVs (dashed).*

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 occurence of baroclinic structures in the analysis error. Hence we may suspect that the 3D-Var Hessian metric penalizes too much the sharp and baroclinic error patterns in the areas which are picked up by the TESV singular vector computation. Thépaut et al (1995) have compared power spectra of 3D-Var and 4D-var analysis increments in the vicinity of an extratropical storm. The 4D-Var approach clearly allowed for more energy in total wave numbers larger than 10. 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, 1994), 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 beween the unstable HSV and TESV subspaces at initial and optimization time when 10 or 25 SVs are used to span the unstable subspace. It is clear from Figure 3a that the unstable TESV and HSV subspaces are almost orthogonal at initial time. At final time the subpaces 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 Northern Hemisphere forecast error $\varepsilon(48)$ was projected onto the 2-day linearly evolved HSVs and TESVs for each of the 10 cases. In the projection 25 SVs were used. Denote by $\tilde{\varepsilon}(48)$ the portion of the forecast error thus explained and the associated so-called pseudo analysis error by $\tilde{\varepsilon}(0)$ (see also Buizza et al, 1997):

$$\tilde{\varepsilon}(48) = \sum_{i=1}^{25} a_i \mathbf{M}(\mathrm{SV}_i) = \mathbf{M}\tilde{\varepsilon}(0) \tag{8}$$

where $\mathbf{M}$ is the tangent model. The percentage of the total energy of $\varepsilon(48)$ as explained by $\tilde{\varepsilon}(48)$ is given in Fig. 4 for TESV (solid line) and HSV (dashed line). Both types of SVs describe

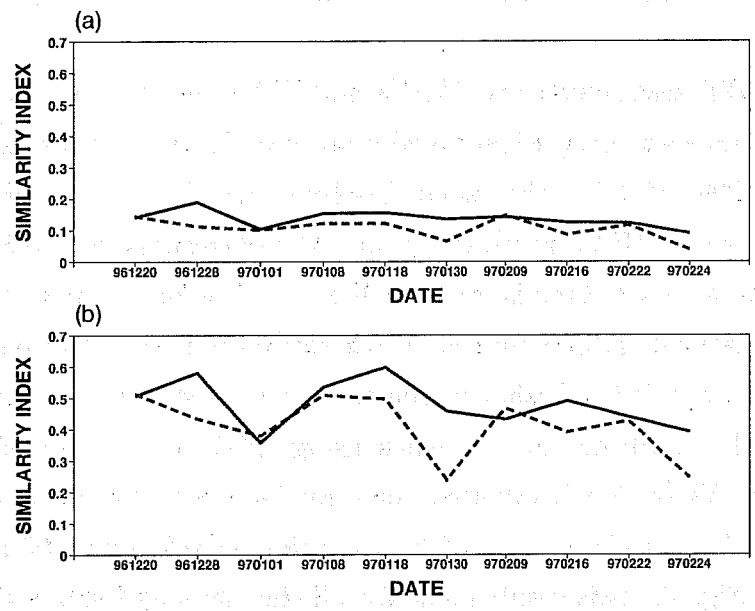


Figure 3: *Similarity index between TESV and HSV unstable subspaces at (a) initial and (b) final time. Dashed (solid) lines indicate that 10(25) singular vectors are used in spanning the subspace.*

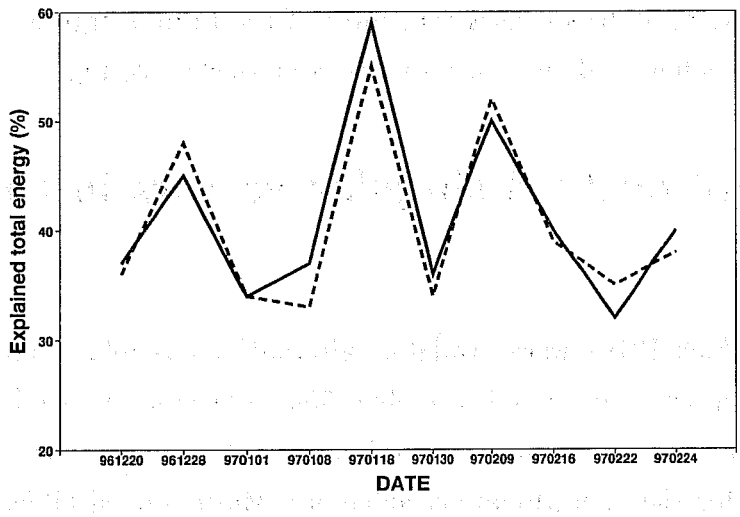


Figure 4: *Explained part of the 2-day Northern Hemisphere forecast error in terms of total energy. The dashed (solid) line indicates that the 25 leading TESVs (HSVs) are used.*

nearly the same fraction of $\varepsilon(48)$ in terms of total energy.

Despite their very different structures, TESVs and HSVs describe also similar geographical patterns of the 2-day forecast error when 25 SVs are used in the expansion (8). Figure 5a,b shows for 18 January 1997 12 UTC the pseudo analysis error $\tilde{\varepsilon}(0)$ at 500 hPa in geopotential height using 25 TESVs and HSVs respectively and the corresponding $\tilde{\varepsilon}(48)$ is given in Fig. 5c,d. The actual 2-day forecast error is given in Fig. 5e. For both sets of SVs the projected forecast error $\tilde{\varepsilon}(48)$ is almost indistinguishable (correlation is 0.93). The patterns of $\tilde{\varepsilon}(0)$ are more different (correlation is 0.59), although the centres of most of the maxima are located at the same positions. The amplitude of $\tilde{\varepsilon}(0)$ when using HSVs in the expansion is larger than $\tilde{\varepsilon}(0)$ obtained with TESVs (note the different countour interval). The corresponding $\tilde{\varepsilon}(0)$ and $\tilde{\varepsilon}(48)$ 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 stuctures, 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 30 January 1997 12 UTC 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 pseudo anlysis error $\tilde{\varepsilon}(0)$ computed with TESVs has larger amplitudes. 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. The baroclinic structure of TESVs causes that most of the total energy is in the temperature component, whilst 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 based ensemble configurations have a nearly equal spread with respect to the control over the Northern Hemisphere (NH) at day 2. This was done by tuning the parameter that sets the amplitude of the initial perturbations. Because of the completely different growth rate and the different distribution of total energy this parameter had to be changed significantly. It was reduced from 0.6 to 0.25. As a third EPS

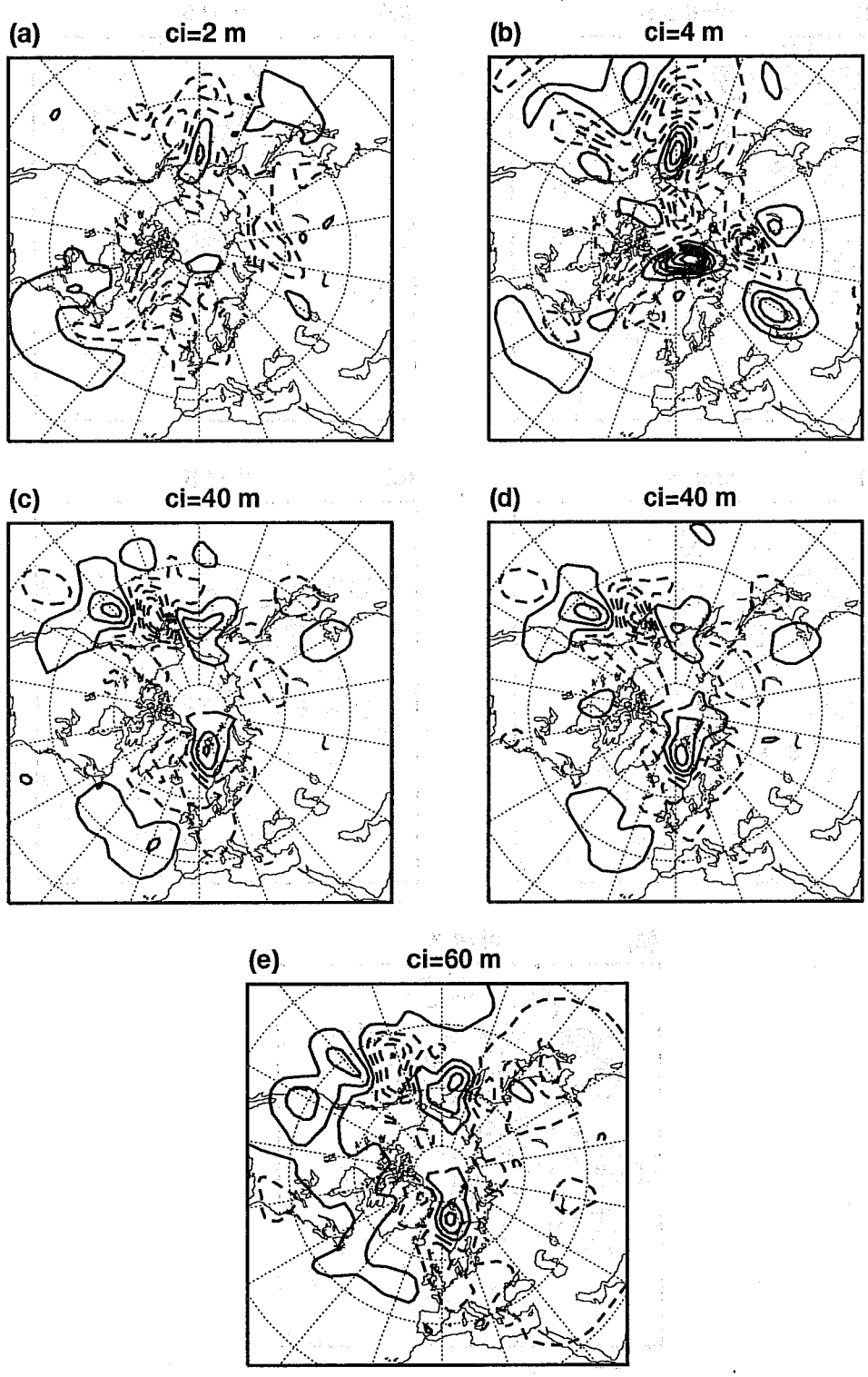


Figure 5: *(a-b) pseudo analysis* $\tilde{\varepsilon}(0)$ *in geopotential height at 500 hPa and (c-d) corresponding* $\tilde{\varepsilon}(48)$*, (e) 2-day forecast error from 18 January 1997 12 UTC. Contour interval is given above each panel.*

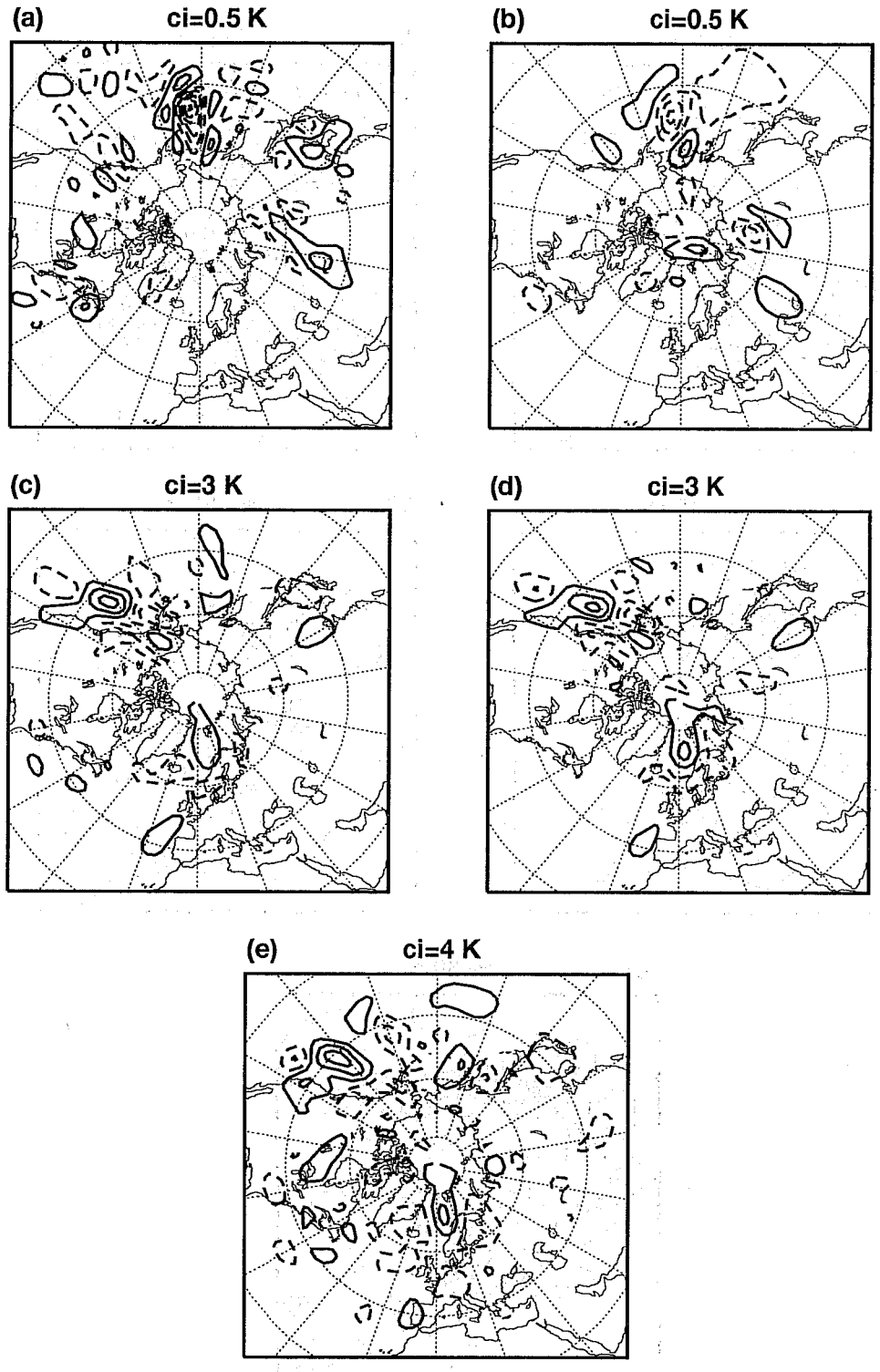


Figure 6: *Same as fig. 5 but for temperature at 500 hPa.*

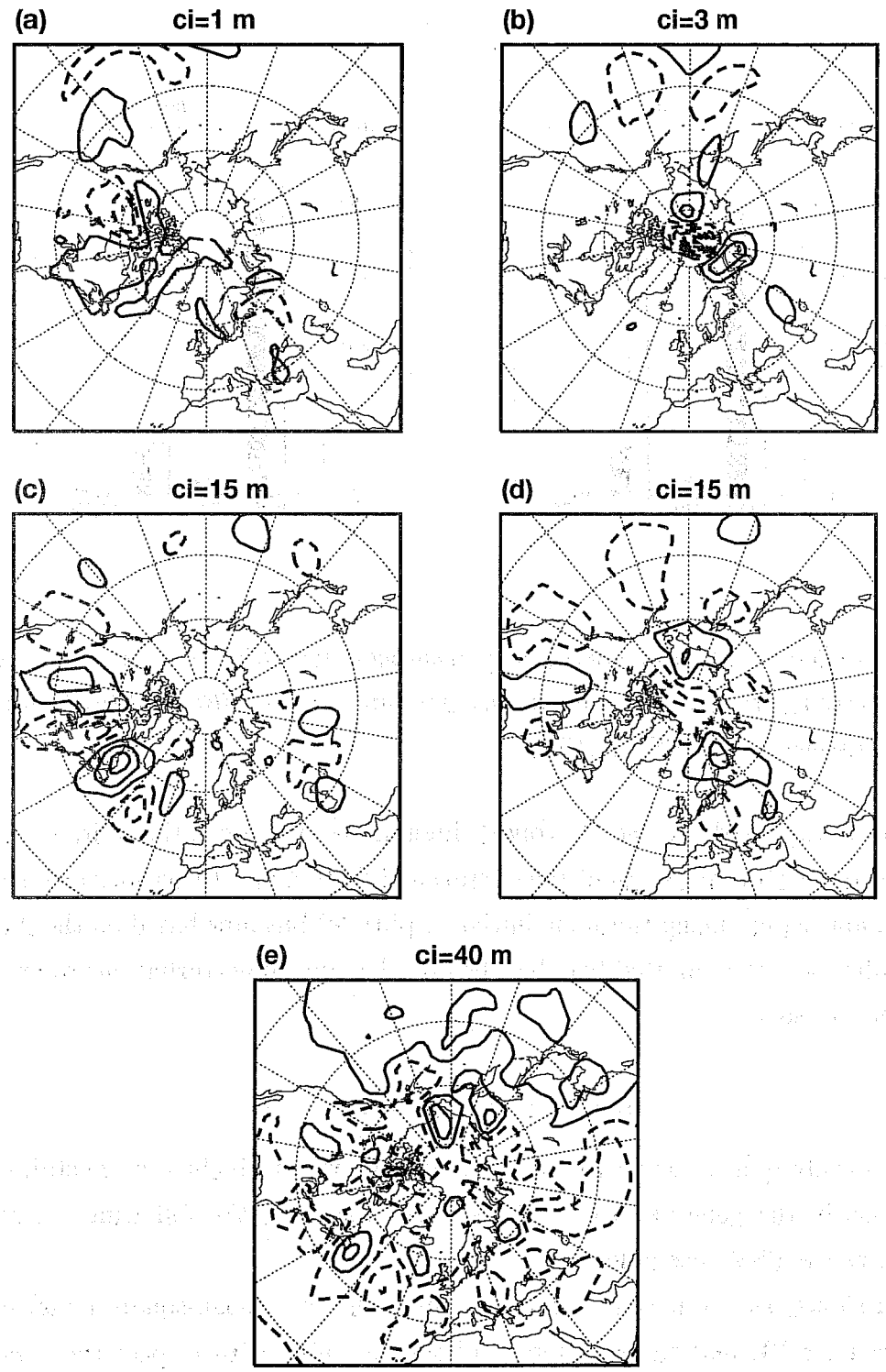


Figure 7: *Same as fig. 5 but for 30 January 1997 12UTC and using 10 singular vectors.*

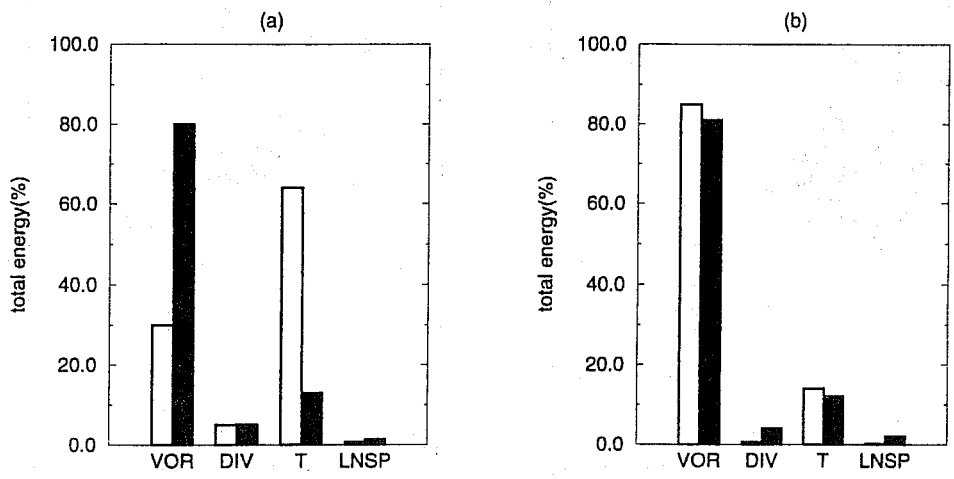


Figure 8: *Distribution of total energy in percentages over the SV state vector components (vorticity, divergence, temperature and surface pressure) at (a) initial and (b) final time. White (black) bar corresponds to TESV (HSV).*

configuration we exploited the use of evolved singular vectors in creating initial perturbations. Additional 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:

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

The use of evolved singular vector provides an easy way to include more stable and large-scale directions in the generation of EPS perturbations. In the following we refer to this configuration as the ESV ensemble.

Figure 9a shows, as a function of forecast time, the root-mean-square (rms) error of the control forecast for NH and the rms spread of the ensemble with respect the control for the three EPS configurations, averaged over the 10 cases. Notice that the perturbations based on HSVs cannot sustain the same 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 rms error as can be seen in Fig. 9b. The ensemble mean of the HSV (ESV) ensemble is on average slightly worse (better) for the medium range.

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

4.1 Brier skill score and relative operating characteristic

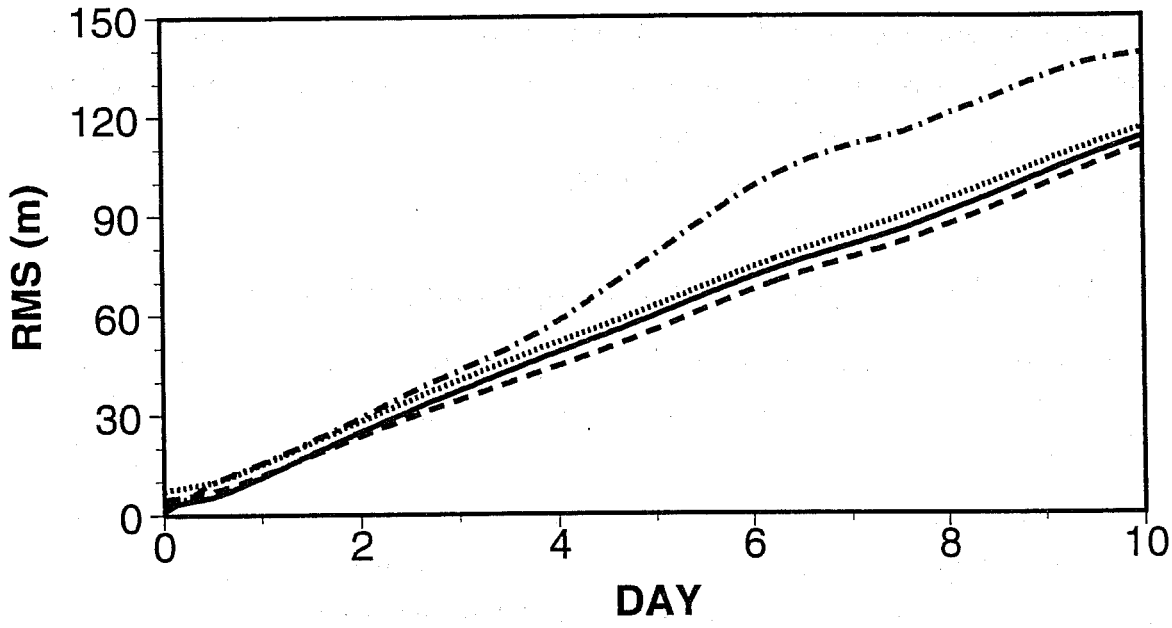
Brier skill scores (Brier, 1950; Stanski et al, 1989) have been computed for probability predictions of geopotential height anomalies exceeding a certain threshold (selected thresholds are geopotential height 25 and 50 m positive/negative anomalies at 500 hPa) and analogously for temperature (selected thresholds are temperature 4 and 8 K warm/cold anomalies at 850 hPa). Figure 11 gives the Brier skill score as a function of forecast time for the 50 m positive anomaly threshold at 500 hPa. In the Brier skill score the skill of the probabilistic forecast is compared 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.

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. for a forecast system with zero false alarms) to 0. A value of 0.5 is produced by an useless forecast system which cannot discriminate between occurrences and non-occurrences of an event. 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.

4.2 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

(a)



(b)

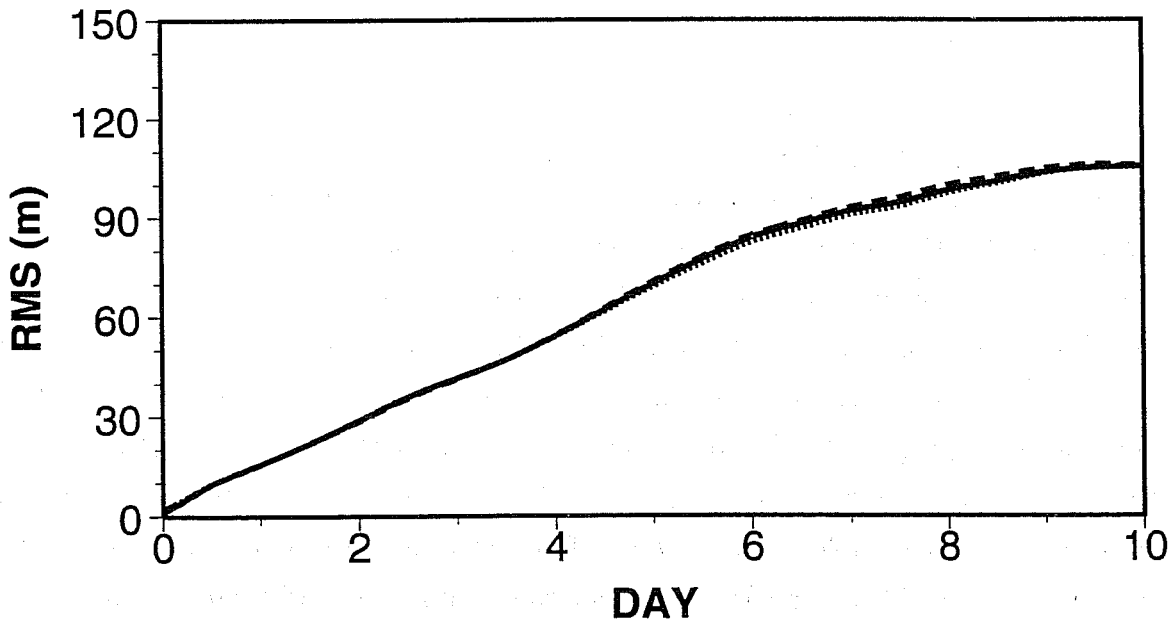


Figure 9: *(a) average rms error of the $T_L159L31$ ensemble control forecasts (chain-dashed) and the average rms spread of the operational (solid), the HSV (dashed) ensembles and the ESV (dotted) ensembles. (b) average rms error of the ensemble mean for the operational (solid), HSV (dashed) and ESV (dotted) ensembles. All values are relative to 500 hPa geopotential height over NH*

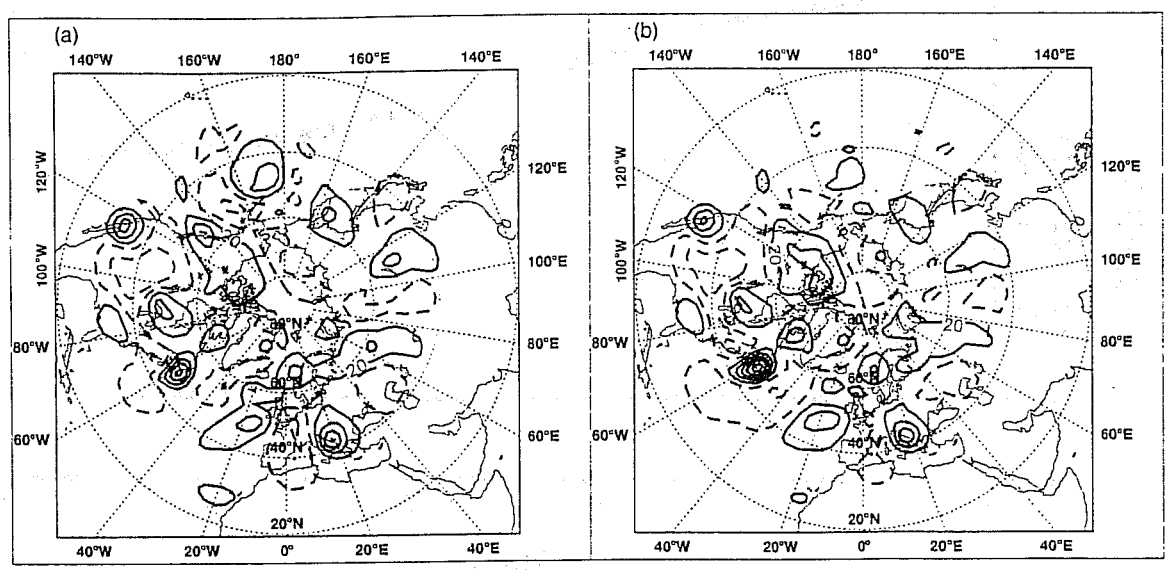


Figure 10: *Geopotential height difference at 500 hPa between the ensemble mean and the control forecast at day 5 for the (a) operational and (b) HSV ensemble starting from 9 February 1997 12 UTC. Solid (dashed) lines denote positive (negative) values and contour interval is 20 m.*

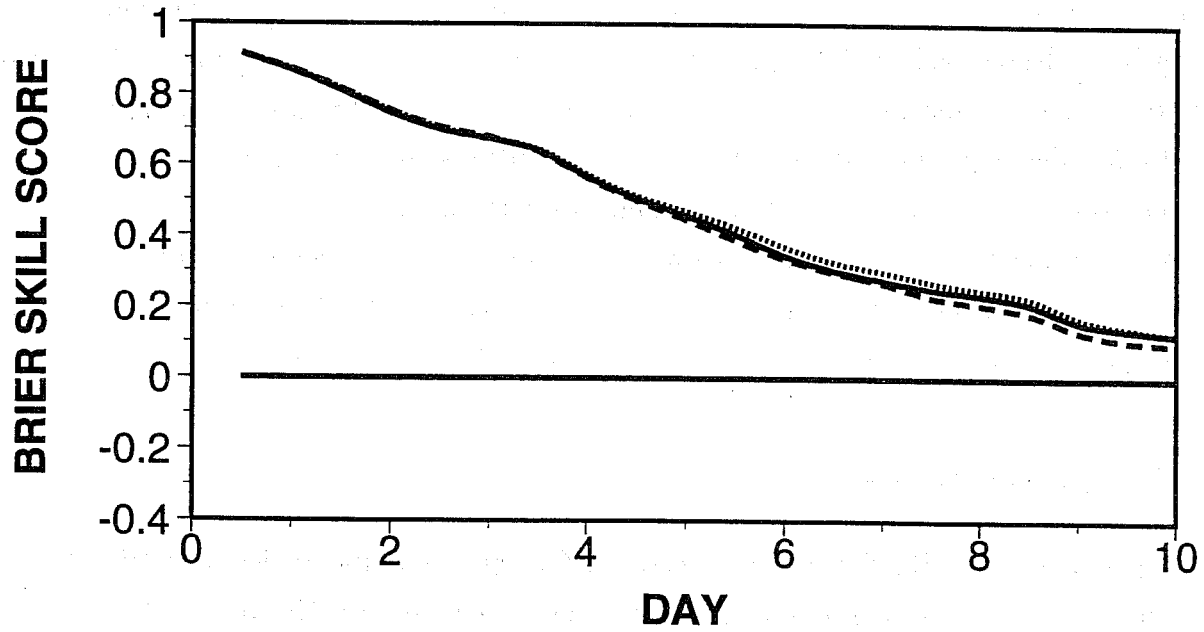


Figure 11: *Brier skill score of the operational (solid), HSV (dashed) and ESV(dotted) ensemble at different forecast times for a geopotential height anomaly threshold of 50 m at 500 hPa over NH.*

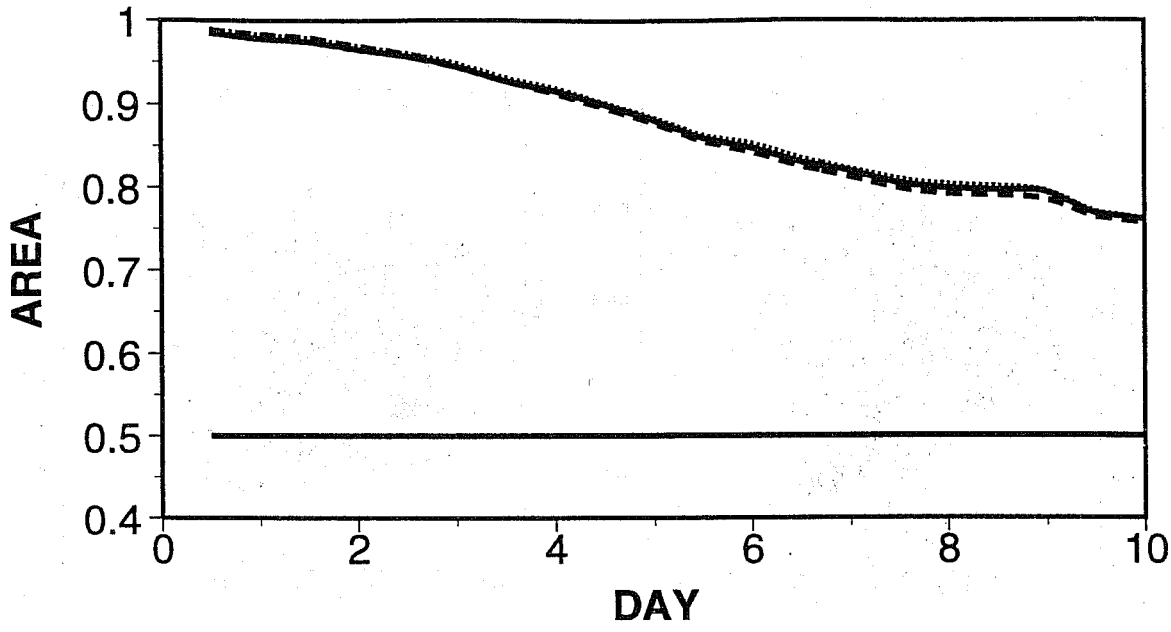


Figure 12: *ROC area of the operational (solid), HSV (dashed) and ESV (dotted) ensemble for a negative geopotential height anomaly of -50 m at 500 hPa over NH at different forecast times.*

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 in case of NH and Europe (averaged over the 10 cases). Because model errors cannot be neglected and the initial perturbations do not form a random sample but are fixed-amplitude perturbations, the percentage of analysis outliers differs from the expected value $\frac{2}{51} \times 100\%$. 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. 9a. 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 result in the smallest percentage of outliers.

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 are 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, 1997). 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

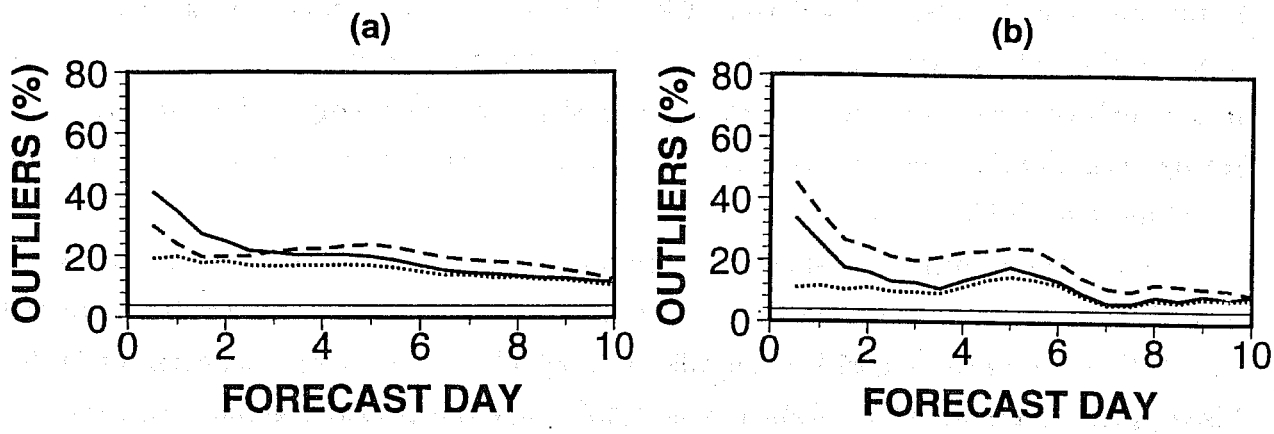


Figure 13: *Percentage of analysis outliers for the operational (solid), HSV (dashed) and ESV (dotted) ensemble as a function of forecast time for geopotential height at 500 hPa in case of (a) Northern Hemisphere and (b) Europe*

of both SVs 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) 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 wavenumber 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, in stead of peaking around the baroclinic steering level as is the case of TESVs. At final time both type of SVs show the same vertical energy distribution.

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 $\mathbf{B}$. Currently 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 occurence of baroclinic structures in the analysis error. From this it is clear that a singular vector computation using the 3D-Var Hessian is not optimal in dynamical unstable areas and will not produce small-scale baroclinic structures like TESVs.

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 currently attempted at ECMWF by experimenting with a simplified Kalman filter in an operational environment(Mike Fisher, pers. commun. 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 (Martin Ehrendorfer, 1998). It will certainly provide a guideline for devising a strategy to implement the simplified Kalman filter.

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 spread at day 2 for NH relative to geopotential height at 500 hPa. In addition to this, the use of evolved singular vector was investigated. These 2-day linearly evolved singular vectors (ESV) 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 singular vectors (using the same initial amplitude as used for the operational perturbations). Results show that the HSV perturbations do not produce the same spread over NH in the medium range as occurs in the operational ensemble. This in contrast to the ESV ensembles where the spread is 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. The statistical tests 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/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 of 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 initial amplitude was slightly reduced from 0.6 to 0.5)

6 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	Tot Observed
Observed=Yes	**X**	**Y**	**X+Y**
Observed=NO	**Z**	**W**	**Z+W**
Tot Forecast	**X+Z**	**Y+W**	

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

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

7 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 helpful comments improving the manuscript.

8 References

Barkmeijer, J., Gijzen, Van M., and F. Bouttier, 1998: Singular vectors and estimates of the analysis error covariance metric. *Quart. J. Roy. Meteor. Soc.* in press.

Bouttier, F., 1997: The 1997 revision of the J_b term in 3D/4D-Var. ECMWF, Technical Memorandum No 238, 54pp. Available from ECMWF, Shinfield Park, Reading RG2 9AX, UK.

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

Buizza, R., 1994: Sensitivity of optimal unstable structures. *Quart. J. Roy. Meteor. Soc.*, 120, 429-451.

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

Buizza, R., T. Petroliagis, T. Palmer, J. Barkmeijer, M. Hamrud, A. Hollingsworth, A. Simmons and N. Wedi, 1998a: Impact of model resolution and ensemble size on the performance of an ensemble prediction system. *Quart. J. Roy. Meteor. Soc.* to appear.

Buizza, R., Miller, M., and T.N. Palmer, 1998b: Stochastic simulation of model uncertainties in the ECMWF ensemble prediction system. In ECMWF Workshop Proceedings of the ECMWF Workshop on Predictability, 20-22 October, 1997, ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljević, M. Hamrud, A. Hollingsworth, F. Rabier and M. Fisher, 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Quart. J. Roy. Meteor. Soc.* to appear.

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

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

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

Ehrendorfer, M., 1998: Prediction of the uncertainty of numerical weather forecasts: problems and approaches. In ECMWF Workshop Proceedings of the ECMWF Workshop on Predictability, 20-22 October, 1997, ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

Fisher, M. and P. Courtier, 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. ECMWF, Technical Memorandum No 220, 27pp. Available from ECMWF, Shinfield Park, Reading RG2 9AX, UK.

Hersbach, H., Mureau, R., Opsteegh, J.D., and Barkmeijer, J., 1998: An EPS for the short and early medium range. In ECMWF Workshop Proceedings of the ECMWF Workshop on Predictability, 20-22 October, 1997, ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

Harrison, M.S.J., Palmer, T.N., Richardson, D.S., Buizza, R., and T. Petroliagis, 1995: Joint ensembles form the UKMO and ECMWF models. ECMWF Seminar Proceedings of the ECMWF Seminar on Predictability, Vol II, 4-8 September, 1995, ECMWF, Shinfield Park, Reading, RG2 9AX, UK, 259pp.

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

Molteni, F. R. Buizza, T.N. Palmer and T. Petroliagis, 1996: The ECMWF ensemble prediction system: methodology and validation. *Quart. J. Roy. Meteor. Soc.*, 119, 269-298.

Palmer, T.N., R. Gelaro, J. Barkmeijer and R. Buizza, 1998: Singular vectors, metrics and adaptive observations. *J. Atmos. Sci.*, 54, 633-653.

Palmer, T.N., F. Molteni, R. Mureau, R. Buizza, P. Chapalet, and J. Tribbia, 1992: Ensemble prediction. ECMWF Seminar Proceedings of the ECMWF Seminar on Validation of models over Europe, Vol I, 7-11 September, 1992, ECMWF, Shinfield Park, Reading, RG2 9AX, UK, 285pp.

Palmer, T.N., Buizza, R., and F. Lalaurette, 1998: Performance of the ECMWF Ensemble Prediction System. In ECMWF Workshop Proceedings of the ECMWF Workshop on Predictability, 20-22 October, 1997, ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

Parrish, D.I. and J.C. Derber, 1992: The National Meteorological Centre's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, 120, 1740-1763.

Rabier, F. and P. Courtier, 1992: Four-dimensional assimilation in the presence of baroclinic instability. *Quart. J. Roy. Meteor. Soc.*, 118, 649-672.

Rabier, F., A. Mc Nally, E. Anderson, P. Courtier, P. Undén, J. Eyre, A. Hollingsworth and F. Bouttier, 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part II: Structure functions. *Quart. J. Roy. Meteor. Soc.* to appear.

Rabier, F., Mahfouf, J.-F., Fisher, M., Järvinen, H., Simmons, A., Anderson, E., Bouttier, F., Courtier, P., Hamrud, M., Haseler, J., Hollingsworth, A., Isaksen, L., Klinker, E., Saarinen, S., Temperton, C., Thépaut, J.-N., Undén, p., Vasiljeviç, D., 1997: Recent experimentation on 4D-Var and first results form a simplified Kalman filter. ECMWF, Technical Memorandum No 240, 42pp. Available from ECMWF, Shinfield Park, Reading RG2 9AX, UK.

Richardson, D., 1998: The relative effect of model and analysis differences on ECMWF and UKMO operational forecasts. In ECMWF Workshop Proceedings of the ECMWF Workshop on Predictability, 20-22 October, 1997, ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

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

Strang, G, 1986: Introduction to applied mathematics. Wellesley-Cambridge Press, pp758.

Thépaut, J.-N., P. Courtier, G. Belaud and G. Lemaître, 1995: 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, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, 125, 3297-3319.

Trevisan, A., F. Pancotti and F. Molteni, 1997: Ensemble prediction in a model with flow regimes. Submitted to *J. Atmos. Sci.*.