TECHNICAL MEMORANDUM

219

Predictability studies with high resolution singular vectors

R. Buizza, R. Gelaro, F. Molteni
and T.N. Palmer

Research Department

August 1995

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

ABSTRACT

The dominant singular vectors of the tangent propagator of the ECMWF numerical weather prediction model are an essential component of the ECMWF ensemble prediction system. These singular vectors describe the principal finite-time linear instabilities of the northern extratropical atmospheric circulation. The impact of increasing the horizontal resolution of the tangent model from T21 to T42 on three different types of initial perturbation, which make use of these singular vectors, is considered. The effect of this increase in resolution is to allow the possibility of describing more accurately instabilities with an upscale cascade of energy from sub-synoptic to synoptic scales. Two of the perturbations are referred to as the pseudo-inverse and sensitivity vectors. These are both diagnostic, and involve estimating from the short-range forecast error, the component of initial error in the unstable subspace. The third type of perturbation is used to construct the set of initial states for the ensemble prediction scheme. Linear and non-linear integrations are described using these different types of perturbation.

All the results point to the conclusion that the higher resolution calculations lead to more accurate results. This is found to be especially true in a number of cases where the control forecast was particularly poor. As a consequence the predictability of synoptic-scale disturbances in the atmosphere on timescales of a few days is likely to be determined by errors in the initial state on sub-synoptic scales. In addition to these conclusions, the relationship between the amplitude of the ensemble perturbations and the pseudo-inverse perturbations is discussed.

1. INTRODUCTION

The concepts of instability and predictability are intimately linked; broadly speaking, the more unstable a dynamical system is, the less predictable it will be. The patterns of instability which dominate the energy-spectrum of the atmosphere are associated with baroclinic modes, with typical scales of a few thousand km. This leads to the commonly-held view (e.g. *Eady*, 1949) that the predictability of the synoptic-scale flow is determined by the (inverse of the) growth rate of these baroclinic instabilities.

However, the extra-tropical atmosphere also shares many characteristics with those of a quasi two dimensional turbulent fluid containing a spectrum of interacting scales (*Pedlosky*, 1979). As discussed e.g. by *Lilly* (1983), energy injection at sub-synoptic scales in a 2-dimensional turbulent flow would lead to an upscale cascade of energy to the synoptic scales. More recently, *Hartmann et al* (1995) found that synoptic waves that become fully developed in three days can arise from perturbations initially contained in the sub-synoptic scales. An uncertainty in these sub-synoptic scales could then influence the predictability of the synoptic-scale flow, irrespective of uncertainty in the synoptic scales themselves.

Consider a time interval $t_0 \leq t \leq t_1$, characteristic of the inverse growth rate of a baroclinic wave. The question arises as to whether, over this time interval, the amplification of energy at the scale associated with the baroclinic wave, exceeds that associated with the upscale energy cascade to the baroclinic scale from initially sub-baroclinic scales. This question is central in determining the precise relationship between the instability and predictability of synoptic-scale weather patterns. Uncertainties in the initial conditions of a weather prediction occur at all scales. On the scale of the weather patterns themselves, uncertainties arise particularly from the paucity of conventional data over the oceans. However, irrespective of location, uncertainties inevitably arise on scales unresolved by the measurement network. Our question can be reformulated as follows: is the predictability of weather determined by these synoptic or sub-synoptic scale errors?

The ECMWF (European Centre for Medium-Range Weather Forecasts) Ensemble Prediction System (EPS) is a practical tool for determining atmospheric predictability in the medium range (*Molteni et al*, 1995). At the time of writing, the EPS comprises 32 medium-range forecasts run from initial conditions which are slightly perturbed from the operational analysis. These 32 perturbations are themselves constructed from a 16-dimensional space that, in a precise sense is estimated to contain a selection of the dominant instabilities of the atmospheric circulation. This unstable sub-space can be calculated from the singular vectors of the forward tangent propagator of the primitive equations of motion (*Buizza and Palmer*, 1995; see also section 2). Until 14 of March 1995, operational EPS perturbations were made using this propagator truncated at horizontal spectral truncation T21 with 19 vertical levels (T21L19). Whilst this choice was dictated by technical constraints, it can be argued that all of the dominant synoptic-scale instabilities were captured at this resolution.

In this paper we compare estimates of the initial error in the unstable sub-space based on singular vectors made with both T21 and T42 resolution tangent models, and corresponding ensemble forecasts. We will take the view that scales smaller than that resolved by a T21 truncation are certainly sub-synoptic. The comparative results, therefore, determine the relative importance of synoptic and sub-synoptic scales.

The two estimates of the component of initial error in the unstable sub-space are obtained by techniques referred to as "pseudo-inverse" and "sensitivity" analysis. Of course, if the forward tangent propagator were invertible, then the full initial error could be obtained from the forecast error. Non-invertibility is associated with the existence of singular values close to or equal to zero. The pseudo-inverse defined in this paper equals the familiar Moore-Penrose pseudo-inverse (*Penrose*, 1955; *Golub and van Loan*, 1983) in the approximation that the singular values outside the unstable subspace are negligible.

Recent studies with adjoint models in numerical weather prediction (*Rabier et al,* 1994) have shown that the gradient of the short-range forecast error taken with respect to the initial conditions, commonly referred to as a 'sensitivity pattern', can be an effective means of identifying structures in the initial conditions that might cause large forecast errors. The formal difference between the pseudo-inverse and the sensitivity pattern is described in the main body of this paper. *Rabier et al* (1994) showed that, in cases where the operational forecast was poor, the *a posteriori* introduction of an analysis perturbation based on the (scaled) sensitivity pattern led to a significant reduction in the forecast error.

The results of this paper are based on seven case studies in which we consider singular vectors computed at T21 and T42 horizontal spectral truncations growing over a 36-hour optimisation period. The control forecasts in all seven cases can be classified as having average or below-average skill in the medium range (forecast days 5-7, see Appendix A). However, three of the cases (94.01.03, 94.03.15 and 94.04.06), for which the control forecasts were well below average, showed particularly strong sensitivity to the horizontal resolution of the singular vector perturbations, especially over Europe, and are analyzed in greater detail. For reference, a brief description of the synoptic situations during these sensitive cases is given in Appendix A.

The paper is organized as follows. In section 2, we briefly review the mathematical formalism of the singular vector calculation, and define the (unstable sub-space) projection operators used to conduct the pseudo-inverse and sensitivity analyses in section 3. In section 3, we compare the T21 and T42 unstable sub-spaces in terms of the pseudo-inverse and sensitivity patterns. In section 4, we examine the impact of the singular vector resolution on EPS performance. Conclusions are drawn in section 5.

2. MATHEMATICAL FORMALISM

The generation of the unstable sub-space of singular vectors is the first component in the calculation of initial perturbations for the EPS. The mathematical development and dynamical interpretations of the singular vectors have been described in detail by *Buizza and Palmer* (1995) but see also earlier work e.g. by (*Farrell,*

1982: *Lacarra and Talagrand*, 1988; *Borges and Hartmann*, 1992; *Molteni and Palmer*, 1993). In this section, we review the definition of the singular vectors, and define their relationship to various measures of forecast error used for conducting the pseudo-inverse and sensitivity analysis in section 3.

a. Singular vectors and the probability distribution of error

Let $\bar{x}(t)$ be a portion of a phase-space trajectory defined by (an M-dimensional truncation of) the nonlinear primitive equations of motion over an interval $t_0 \le t \le t_1$. Let $L \equiv L(t_0,t_1)$ denote the forward propagator of the dynamical system linearised about $\bar{x}(t)$ so that

$$\boldsymbol{x}(t_1) = \boldsymbol{L}\ \boldsymbol{x}(t_0) \tag{1}$$

defines a mapping of the M-dimensional tangent space $T_{\bar{x}(t_0)}$ at $\bar{x}(t_0)$ to the tangent space $T_{\bar{x}(t_1)}$ at $\bar{x}(t_1)$.

Finally, let $\boldsymbol{L}^*$ denote the adjoint operator of $\boldsymbol{L}$, defined with respect to a total energy inner product $(..;..)$ [see *Buizza and Palmer*, 1995, Eqs (2.4) and (2.5)]. Then the total energy of a perturbation $\boldsymbol{x}(t)$ is

$$\|\boldsymbol{x}(t)\|^2 = (\boldsymbol{L}^*\boldsymbol{L}\ \boldsymbol{x}(t_0);\boldsymbol{x}(t_0)), \tag{2}$$

where $\|..\|$ is the norm associated to the inner product $(..;..)$.

The eigenvectors $\boldsymbol{v}_i$ of the compound operator $\boldsymbol{L}^*\boldsymbol{L}$, with associated eigenvalues $\sigma_i^2 \ge 0$, given by

$$\boldsymbol{L}^*\boldsymbol{L}\ \boldsymbol{v}_i = \sigma_i^2\ \boldsymbol{v}_i, \tag{3}$$

form a complete orthonormal basis in $T_{\bar{x}(t_0)}$. In linear algebra terminology, the $\boldsymbol{v}_i$ and σ_i^2 are referred to as the singular vectors and singular values, respectively, of $\boldsymbol{L}$ itself. The spectrum of $\boldsymbol{v}_i$ determines the amplification of an element of $T_{\bar{x}(t_0)}$ as it is mapped by $\boldsymbol{L}$ along the trajectory to $T_{\bar{x}(t_1)}$. At time t_1, the singular vectors evolve to

$$\hat{\boldsymbol{u}}_i = \boldsymbol{L}\boldsymbol{v}_i, \tag{4}$$

$$\|\boldsymbol{v}_i\| = 1\ ,\ \ \|\hat{\boldsymbol{u}}_i\| = \sigma_i. \tag{5}$$

The σ_i, ranked in terms of magnitude, are the amplification factors of the singular vectors over the interval $t_1 - t_0$. Maximum energy growth over this interval is associated with the evolution of the dominant singular vector $\boldsymbol{v}_1 \rightarrow \hat{\boldsymbol{u}}_1$. The sub-spaces $P^n_{\bar{x}(t_0)} \subset T_{\bar{x}(t_0)}$ and $P^n_{\bar{x}(t_1)} \subset T_{\bar{x}(t_1)}$ spanned by the first n singular vectors of the spectrum will be referred to as the n-dimensional unstable sub-spaces at $\bar{x}(t_0)$ and $\bar{x}(t_1)$.

The choice of inner product is not arbitrary. It is easily shown that the singular vectors $\hat{\boldsymbol{u}}_i$ are equal to the dominant eigenvectors of the forecast error covariance matrix, in the inner product $(...;..)_C$ = <..;*C*..>, where <..;..> denotes the Euclidean inner product, and C is the analysis error covariance matrix. As discussed in *Molteni et al* (1995), the energy inner product (..;..) is a fair approximation to $(..;..)_C$, especially compared with other standard inner products based on perturbation enstrophy or streamfunction squared.

In essence, then, the dominant singular vectors at optimisation time (with energy inner product) approximate the major axes of the forecast error covariance matrix (the second moment of the forecast error probability distribution function). The phase space directions associated with the initial state, which evolve into these major axes, are given by the dominant singular vectors at initial time. In terms of the $(..;..)_C$ inner product, the probability of making an initial error along one of the dominant singular vector directions is as likely as making an error along any other direction. To reiterate, this condition appears to be reasonably well approximated using the energy inner product.

b. Relationship to pseudo-inverse and forecast sensitivity

Since the singular vectors used in this study are computed with an optimisation time interval $t_1 - t_0$ equal to 36 hours, let us first consider a 36-hour forecast error $\boldsymbol{e}_{36}$. It is convenient to introduce a normalized set of evolved singular vectors $\boldsymbol{u}_i$ with unit total energy norm, such that

$$\boldsymbol{u}_i = \sigma_i^{-1} \hat{\boldsymbol{u}}_i, \tag{6}$$

$$\|\boldsymbol{u}_i\| = 1. \tag{7}$$

The projection of $\boldsymbol{e}_{36}$ into the sub-space $\mathrm{P}^n_{\vec{x}(t_1)}$ spanned by n selected singular vectors $\boldsymbol{u}_i$ can thus be written

$$\tilde{\boldsymbol{e}}_{36} = \boldsymbol{U}\boldsymbol{U}^T\boldsymbol{e}_{36}, \tag{8}$$

where $\boldsymbol{U}$ is an $m \times n$ matrix with columns $\boldsymbol{u}_i$, and the superscript T denotes a transpose. The dimension m depends on the horizontal resolution of the model.

Using Eqs (4) and (6), we can approximate the projection of the forward propagator $\boldsymbol{L}$ into the unstable sub-space by

$$\tilde{\boldsymbol{L}} = \boldsymbol{U}\Sigma\boldsymbol{V}^T, \tag{9}$$

where $\boldsymbol{V}$ is an $m \times n$ matrix with columns $\boldsymbol{v}_i$, and Σ is an $n \times n$ diagonal matrix with elements σ_i.

From Eqs (8) and (9) it follows that $\tilde{\boldsymbol{e}}_{36}$ is the linear evolution of the initial perturbation

$$\tilde{\boldsymbol{e}}_0 = \tilde{\boldsymbol{L}}^{-1}\boldsymbol{e}_{36} = \boldsymbol{V}\Sigma^{-1}\boldsymbol{U}^T\boldsymbol{e}_{36}, \tag{10}$$

Equation (10) is the analogue, in the unstable sub-space, of inverting the forward propagator $\boldsymbol{L}$ to obtain the initial perturbation $\boldsymbol{e}_0$. With respect to the full m dimensional space, $\tilde{\boldsymbol{L}}^{-1}$ is the Moore-Penrose pseudo-inverse (*Penrose*, 1955; *Golub and van Loan*, 1983) of $\boldsymbol{L}$ in the approximation where all the singular values outside the unstable subspace are small compared with those inside the unstable subspace. In a perfect model, $\boldsymbol{e}_0$ is the analysis error, so that a perturbation $\boldsymbol{p}_0 = -\boldsymbol{e}_0$ would cancel $\boldsymbol{e}_{36}$ exactly in a linear sense. In contrast, a perturbation $\boldsymbol{p}_0 = -\tilde{\boldsymbol{e}}_0$ describes the linear combination of n selected singular vectors that minimizes the cost function

$$F = \|\boldsymbol{L}\boldsymbol{p}_0 + \tilde{\boldsymbol{e}}_{36}\|^2 + \alpha^2\|\boldsymbol{p}_0\|^2, \tag{11}$$

with $\alpha = 0$ (see Appendix B). The computation of $\tilde{\boldsymbol{e}}_0$ is straightforward for fixed n. However, as n increases, this computation will eventually be divergent owing to the factors $1/\sigma_i$, where $\sigma_1 > \sigma_2 > \ldots > \sigma_n$ associated with the non-invertibility of the full operator $\boldsymbol{L}$.

Closely related to the pseudo-inverse is the gradient of the forecast error with respect to the initial conditions

$$\nabla_0 J = \boldsymbol{L}^* \boldsymbol{e}_{36}, \tag{12}$$

where J is a diagnostic function defined, in this case, as the difference, with respect to the total energy norm, between the 36-h forecast and verifying analysis. The gradient $\nabla_0 J$ is often referred to as a sensitivity pattern (following e.g.*Marchuk*, 1974; *Cacuci*, 1981) which, for convenience, we denote hereafter by $\boldsymbol{s}$.

The relationship between the sensitivity $\boldsymbol{s}$ and the analysis error $\boldsymbol{e}_0$ is most readily seen by expressing Eq (12) in the form

$$\boldsymbol{s} = \boldsymbol{L}^* \boldsymbol{L} \boldsymbol{e}_0. \tag{13}$$

Expanding $\boldsymbol{e}_0$ in terms of singular vectors and using Eq (3), we obtain the projection of the sensitivity pattern onto the unstable sub-space

$$\tilde{\boldsymbol{s}} = \boldsymbol{V}\boldsymbol{\Sigma}^2\boldsymbol{V}^T\boldsymbol{e}_0. \tag{14}$$

It is clear from Eq (14) that the sensitivity pattern is dominated by the most unstable singular vectors. The weights σ_i^2 given to the leading components of $\tilde{\boldsymbol{s}}$ are typically of order 100 or more at T42 resolution, corresponding to amplification factors of 10 or more for the fastest growing singular vectors over a 36-h optimization interval (see Fig 1). As demonstrated by *Rabier et al* (1994), the introduction of a perturbation proportional to $-\boldsymbol{s}$ can thus lead to a substantial reduction in the forecast error in cases where the error is dominated by the growth of these rapidly growing components. The relation between the sensitivity vector

and the pseudo-inverse can be obtained by noting that a perturbation $\boldsymbol{p}_0 = -\tilde{\boldsymbol{s}}$ minimizes Eq (11) when $\alpha = \infty$ (see Appendix B).

Using Eq (10) to substitute for $\boldsymbol{e}_0$, we obtain the more practical form

$$\tilde{\boldsymbol{s}} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{U}^T\boldsymbol{e}_{36}. \tag{15}$$

The comparison of Eqs (10) and (15) shows that $\tilde{\boldsymbol{e}}_0$ and $\tilde{\boldsymbol{s}}$ differ only in the weighting of the singular vectors used in the projection. However, note that in contrast with $\tilde{\boldsymbol{e}}_0$, $\tilde{\boldsymbol{s}}$ converges for increasing values of n owing to the factors σ_i.

3. COMPARISON OF T21 AND T42 UNSTABLE SUB-SPACES

In this section, we examine singular vector growth in the T21 and T42 unstable sub-spaces, and the impact of resolution on the pseudo-inverse and sensitivity patterns.

a. Singular vector growth

Figure 1 shows the singular values for the first 25 singular vectors at T21 (solid line) and T42 (dashed line) resolution, averaged over the seven cases examined in this study. The impact of the higher resolution model is fairly uniform across the spectrum, with the singular values at both resolutions showing a rather gradual decrease with singular vector index. The dotted line in Fig 1 shows the ratio between the T42 and T21 singular values. Overall, growth rates are approximately 60% larger at the higher resolution.

Figure 2 shows the energy spectrum for the T21 and T42 calculations, respectively, averaged over all seven cases, and over the first 16 singular vectors. At optimisation time, both the T21 and T42 singular vectors have maximum energy at synoptic scales. However, at initial time, the energy is spread over a broader range of scales at T42 than at T21. Although the energy appears to be increasing to a maximum at the truncation limit at both resolutions, the flatter spectrum at T42 suggests that a further increase in resolution would have less impact than the increase from T21 to T42 (see also *Hartmann et al*, 1995).

These results show that the unstable sub-space is strongly non-modal. Energy cascades upscale as the singular vectors amplify, and an artificial truncation that restricts the number of octaves over which upscale growth can occur also restricts the amplification of the perturbation. *(Molteni et al* (1995) showed how the use of an enstrophy norm would reverse the spectrum characteristics at initial and final time, giving perturbations whose time evolution is characterized by a downscale energy cascade. However, it was argued in that paper that the analysis error probability distribution would not be isotropic with respect to an enstrophy inner product.)

Figure 3 shows the geographical distribution of the first 16 singular vectors at T21 and T42 resolution for all seven cases. The location of each singular vector is marked by a single black square positioned where the singular vector vorticity at initial time is a maximum. The two distributions are fundamentally quite similar. For example, both show a relative maximum in singular vector distribution over the regions of mean baroclinity near the eastern seaboards of Asia and North America. There are, however, some detailed differences between the two distributions. At T42 resolution, there is a reduction in the number of singular vectors associated with the sub-tropical jet over the middle east, and a corresponding increase in the number of singular vectors over the western Pacific. In general, the T42 singular vectors are somewhat more concentrated in the jet stream regions. These differences are in part due to the fact that the jet streams are more localized and characterized by stronger gradients at T42 than at T21. However, the larger number of degrees of freedom associated with the T42 singular vectors will automatically imply a relative concentration of energy near the most unstable regions.

Figure 4 shows the vertical distribution of total energy for the first 16 singular vectors averaged over all seven cases. The vertical structure is little changed between T21 and T42. Both unstable sub-spaces show initial singular vector structure with maxima in the lower troposphere, and final singular vector structure near the tropopause. These characteristics were discussed by *Buizza and Palmer* (1995) and can be understood qualitatively in terms of wave-action concepts. In general, if a linear perturbation is launched close to the baroclinic steering level, and the perturbation propagates vertically towards the jet level, then conservation of wave-action can imply substantial energy growth. Such a process has been modelled in a WKBJ approximation by *Zeng* (1983).

b. **Pseudo-inverse and forecast error analysis**

As described in section 2b, the projected 36-hour forecast error, $\tilde{e}_{36}$, is readily computed using Eq (8). In Fig 5 the 500 hPa height of the forecast error $\tilde{e}_{36}$ is compared with the projections $\tilde{e}_{36}^{T21}$ and $\tilde{e}_{36}^{T42}$ using 16 singular vectors at T21 and T42 resolution for the two cases, 94.03.15 (left panels) and 94.04.06 (right panels).

The characteristic spatial scale of the e_{36} field appears to be somewhat better reproduced in the T42 projection than in the T21 projection. However, despite the scale difference, the T21 projections have comparable amplitude to the T42 projections, namely about 30%-50% of the amplitude of e_{36}. As such, the smaller-scale features of the T42 singular vectors at optimisation time do not contribute much to the overall proportion of forecast error explained by the singular vector projections (though as will be discussed, the same conclusion is not true for the pseudo-inverse error field). Note that the spatial features of the forecast error that are best reproduced in the singular vector projections are baroclinic wave trains in the oceanic storm tracks (e.g.the Pacific wave pattern that stretches from approximately 35°N-165°E to the Gulf of Alaska in the 94.03.15 case).

A more quantitative comparison between e_{36} and its projections into the T21 and T42 unstable sub-spaces can be made by looking at the energy distribution as a function of height (model level) and total wave number. Figure 6 shows these distributions for the 94.04.06 case, in which the energy values for the projected fields (Figs 6c-f) have been multiplied by 9 (equivalent to a factor of 3 in the field amplitude) to make them visually comparable with the full fields. With this scaling factor, the vertical energy profiles of e_{36}, $\tilde{e}_{36}^{T21}$, and $\tilde{e}_{36}^{T42}$ appear very similar, apart from the error at the model's upper boundary. The energy peaks at level 8 (approximately 250 hPa) in e_{36}, and at level 9 (approximately 320 hPa) in the two projections. Even the secondary maximum around level 16 (approximately 900 hPa) has a counterpart in the curvature of the $\tilde{e}_{36}$ profiles.

The analysis of the total wave number spectra reveals that e_{36} has a broader spectrum than either $\tilde{e}_{36}^{T21}$ or $\tilde{e}_{36}^{T42}$. However, both projections reproduce the double maxima at n=12-13 and n=17-18, which have been documented in the singular vector spectra analyzed by *Buizza and Palmer* (1995). One can see that a large proportion of the energy of $\tilde{e}_{36}^{T42}$ is accounted for by total wave numbers $n \leq 21$, consistent with the fact that the norm of $\tilde{e}_{36}^{T21}$ is only marginally smaller than the norm of $\tilde{e}_{36}^{T42}$.

Because the initial perturbation $\tilde{e}_0$ that linearly evolves into $\tilde{e}_{36}$ is a linear combination of the dominant singular vectors, it is, by construction, a linear combination of the initial perturbations used by the EPS (see Section 2b). Figure 7 shows $\tilde{e}_0^{T21}$ and $\tilde{e}_0^{T42}$ for the cases 94.03.15 and 94.04.06, as computed from Eq (10) using the 16 singular vectors selected by the EPS for each date. The impact of the higher resolution singular vectors on these perturbations is quite dramatic.

By construction, an initial perturbation that is equal to $\tilde{e}_0$ will cancel the component $\tilde{e}_{36}$ of the forecast error at 36 hours (in the linear approximation). However, it is not guaranteed that perturbation will also continue to cancel a significant part of the forecast error beyond the linear range. Below we examine the results of forecast experiments in which the initial perturbation is larger than, but proportional to, the pseudo-inverse fields. In particular, Table 1 shows the skill of a series of T63 forecasts for the 94.04.06 case, in which the initial analysis was perturbed by a pattern equal to $k\tilde{e}_0$. It compares the geopotential height errors at 500 hPa for the control forecast with those of the perturbed forecasts using T21 and T42 singular vectors with different values of k. The RMS errors for the northern hemisphere extra-tropics (i.e. the extra-tropical region $\phi \geq 30°N$, hereafter NHET) are listed for forecast days 1, 3, 5 and 7, with the smallest RMS error for each forecast day underlined.

		day 1	day 3	day 5	day 7
Control		13.1	36.6	73.7	78.9
$k = 1$	T21	<u>12.9</u>	34.6	69.4	76.7
	T42	<u>12.9</u>	32.2	62.2	76.0
$k = \sqrt{2}$	T21	13.0	34.0	67.6	77.2
	T42	12.9	30.7	57.1	72.7
$k = 2$	T21	13.1	33.2	65.8	77.2
	T42	13.0	29.7	53.2	65.5
$k = 2\sqrt{2}$	T21	13.5	32.7	65.0	77.0
	T42	13.4	<u>29.4</u>	<u>50.5</u>	<u>63.4</u>
$k = 4$	T21	14.5	33.1	65.0	78.3
	T42	14.3	31.8	52.7	65.8

Table 1 NHET RMS error of control and perturbed forecasts for 500 hPa geopotential height, with perturbations equal to $k\tilde{e}_0$, for the 94.04.06 case. For each forecast day, the smallest error is underlined.

At forecast day 1, which lies within the optimization time interval, the smallest error is obtained for k=1 for both the T21 and T42 perturbations, as expected from theory. In this case, the two resolutions give the same result, which is only slightly better than the score of the control forecast. This is consistent with the fact that both $\tilde{e}_{36}^{T21}$ and $\tilde{e}_{36}^{T42}$ explain about 1/9 of the energy of e_{36}. After forecast day 1, however, the most skilful forecasts are obtained for larger values of k, and (for any k) the T42 perturbations provide a greater improvement over the control than their T21 counterparts. The T42 perturbation with k=$2\sqrt{2}$ is the most effective beyond the linear optimization time. For example, at forecast day 5 it gives an RMS error of 51 m, compared with 74 m for the control and 65 m for the T21 perturbation with the same amplitude. It is intriguing to note that the value $2\sqrt{2}$ is very close to the factor that would make the energy of the 36-hour perturbation, $\tilde{e}_{36}$, equal to the energy of the 36-hour forecast error, e_{36}.

Results for the other sensitive cases produce similar conclusions, namely, that from forecast day 2 onwards, the greatest improvement over the control forecast is obtained with a T42 perturbation with k>1. The optimal value may be $\sqrt{2}$<k<2 for these cases, depending on the forecast.

There are two possible explanations of these results. Firstly a larger amplitude may be needed in order to account for nonlinear error growth. In this respect it may be important at the end of the period of linear growth to have achieved a perturbation amplitude comparable with the total forecast error amplitude. However, it is also possible that the larger initial amplitude may be necessary to compensate for the effects of random model error (which themselves could project onto the unstable subspace, see the conclusions section). The implications of these results on EPS perturbations are discussed in section 4.

There is another reason why the T42 tangent model may be preferable for calculating initial perturbations for T63 nonlinear integrations. Specifically, the T42 linear propagator L can be expected to be closer to the actual evolution with the nonlinear T63 model, than the equivalent T21 propagator. To verify this, we can compare e_{36} (which represents the linear evolution of $\tilde{e}_0$ according to L) with the 36-hour perturbations originated from $\tilde{e}_0$ and $-\tilde{e}_0$ in the T63 model. The fields $\tilde{e}_{36}^{T21}$ and $\tilde{e}_{36}^{T42}$ were shown in Figs 5d,f, respectively, while the corresponding nonlinear 36-hour evolutions are shown in Fig 8.

It can be seen that, for this time range and amplitude, T63 perturbations started from $\tilde{e}_0$ and $-\tilde{e}_0$ are very closely anti-correlated, indicating that the perturbation growth is still linear in the T63 model. This is in agreement with the results of *Buizza* (1995) who estimated that non-linearity becomes important after 2 days for perturbations generated as in the T21 configurations. However, while the T63 nonlinear evolution of $\tilde{e}_0^{T42}$ is very close to $\tilde{e}_{36}^{T42}$, larger differences can be seen for the T21 perturbation. For example, the Pacific wave pattern is very well reproduced at T42 resolution, but has a different phase and meridional shape at T21 (especially along the dateline). Differences are also evident over the western Atlantic and the Mediterranean. A perturbation 'optimized' along a linear T21 trajectory may therefore behave in a non-optimal way when superimposed to a T63 trajectory, even in the time range during which the evolution is linear. At T42 this discrepancy is much less evident, providing a strong argument for the use of T42 singular vectors in the construction of ensemble perturbations.

c. Sensitivity analysis

As described in section 2.b, the projected sensitivity field, $\tilde{s}$, is easily derived from e_{36} using Eq (14). Here, as in section 4, we focus our analysis on the two sensitive forecasts cases 94.03.15 and 94.04.06.

Figures 9a,b show the sensitivity s for these cases in terms of geopotential height at approximately 500 hPa computed using the adjoint operator of the ECMWF forecast model at T63 resolution. For presentation purposes, these fields have been truncated to T42 resolution. Figures 9c-f show the projections $\tilde{s}^{T21}$ and $\tilde{s}^{T42}$ of these fields using 16 singular vectors computed at T21 and T42 resolution, respectively. As in Fig 5, the panels on the left (right) correspond to the 94.03.15 (94.04.06) case. There are several interesting points of comparison between the sensitivity projections shown in Fig 9 and the forecast error projection and pseudo-inverse shown in Figs 5 and 7.

First, it should be noted that the contour intervals in Fig 9 are many times larger than in Figs 5 and 7 since the sensitivity fields in Fig 9 represent a *gradient* field rather than the geopotential height field itself. [These fields have units of geopotential height as a result of using the total energy norm to define the adjoint operator L^*; see *Rabier et al* (1994).] An appropriate scaling of the gradient is thus required before a

meaningful comparison can be made between, say, the RMS amplitudes of s and $\tilde{e}_0$. *Rabier et al* (1994) have shown by empirical means that scale factors on the order of 10^{-2} give reasonable perturbation amplitudes for the gradient of the 2-day forecast error computed with a T63 adjoint model. Although scale factors of this order are in fact consistent with the differences in magnitude between $\tilde{e}_0$ and s as shown in Figs 7 and 9, we have made no attempt to derive an appropriate perturbation scaling of s for the purposes of this study. This topic will be discussed in a subsequent paper that examines more thoroughly the relationship between sensitivity perturbations and singular vectors.

Figure 9 shows that, as in the case of $\tilde{e}_0$, the increase in horizontal resolution has a significant impact on $\tilde{s}$. Clearly, the typical spatial scale of s is far better reproduced in the T42 projections (Figs 9e,f) than in the T21 projections (Figs 9c,d). Note, for example, the close correspondence between the major sensitivity features in Figs 9b,f for the 94.04.06 case, especially over North America, central Europe, and the western Pacific. The positions, spatial scales, and orientations of these features are well reproduced by $\tilde{s}^{T42}$ in Fig 9f. The "L" shaped double maximum in the sensitivity pattern over North America is particularly well represented in the projected field.

We can compare these sensitivity projections with the pseudo-inverse $\tilde{e}_0$, shown in Fig 7. Overall, there is a strong resemblance between these fields. This result is consistent with the fact that $\tilde{s}$ and $\tilde{e}_0$ differ only in the reciprocal weighting of each term in their expansions by the appropriate growth rate σ_i [see Eqs (10) and (14)], where the spectrum of σ_i tends to decrease gradually as shown in Fig 1. Thus, $\tilde{e}_0$ and $\tilde{s}$ are similar in appearance except that, for example, in the case of $\tilde{s}$, those structures that project onto the fastest growing singular vectors (corresponding to the largest values of σ_i), have greater relative amplitude than the corresponding structures in $\tilde{e}_0$. Note, for example, the difference between the relative amplitude of the "L" shaped double maximum in $\tilde{s}^{T42}$ over North America (Fig 9f), which projects strongly onto the first and second fastest growing singular vectors for this date, and the corresponding feature in $\tilde{e}_0^{T42}$ (Fig 7d).

Similar conclusions can be drawn for the case 94.03.15, namely, that $\tilde{s}^{T42}$ is superior to $\tilde{s}^{T21}$ in capturing the full sensitivity pattern and that, in general, $\tilde{s}$ and $\tilde{e}_0$ are similar except for the relative amplitudes of individual features. In this case, there is a striking difference between the relative amplitude of the western Pacific wave train in $\tilde{s}^{T42}$ and $\tilde{e}_0^{T42}$ (Figs 9e and 7c, respectively). Again, this feature projects strongly onto two of the fastest growing singular vectors for this date. In summary, the similarities and differences between Figs 7 and 9 are clearly consistent with the properties of $\tilde{e}_0$ and $\tilde{s}$ described in section 2b.

Figure 10 shows the energy distribution of s and its projections into the T21 and T42 unstable sub-spaces as a function of height (using model level) and total wave number for the 94.04.06 case. These plots are analogous to those shown in Fig 6 for e_{36}. Similar energy distributions were computed for $\tilde{e}_0$ (not shown), and these show little qualitative differences from those for $\tilde{s}$. Thus, given the strong qualitative similarity between $\tilde{e}_0$ and $\tilde{s}$, as also revealed, for example, in Figs 7 and 9, and in view of the fact that e_0 itself is unknown, we restrict our discussion for now to s and e_{36} and their projections. We refer again to the fact that the relatively large values in Fig 10, compared with those in Fig 6 (note the different axes scaling) are, in part, a result of the sensitivity fields being representative of the unscaled gradient of the meteorological fields. Also note that the projected fields in Fig 10 were plotted using the same scaling as for the full fields, whereas the projected fields in Figs 6 were multiplied by a factor of 9 in order to make them visually comparable with the full fields.

The vertical energy profile for s in Fig 10a shows a single significant peak near model level 10 (approximately 410 hPa), which is somewhat lower in height than the peak at level 8 (approximately 250 hPa) for e_{36} in Fig 6a. This is consistent with the fact that the sensitivity pattern tends to have its maximum amplitude in the lower to middle troposphere, while the forecast error tends to have its maximum amplitude near the upper-tropospheric jet (*Rabier et al*, 1994). In this case, the energy peak for s occurs at a higher level than the typically observed level of maximum sensitivity between 850 hPa and 600 hPa. The vertical energy profiles for $\tilde{s}^{T21}$ and $\tilde{s}^{T42}$ in Figs 10c,e also show a single peak with similar structure as the full field, but with less amplitude and at a slightly lower level of 600 hPa. Roughly speaking, we can deduce from these profiles that the energy of s is nearly four times larger than that of $\tilde{s}^{T21}$, but only slightly more than two times larger than that of $\tilde{s}^{T42}$. These ratios can be compared with those deduced from the vertical energy profiles in Fig 6, which indicate that the energy of e_{36} is roughly nine times larger than that of both $\tilde{e}_{36}^{T21}$ and $\tilde{e}_{36}^{T42}$. Thus, we again draw the conclusion that the increase in horizontal resolution has a significant positive impact on the projection of the sensitivity pattern into the unstable sub-space of singular vectors.

The total wave number spectrum in Fig 10b shows that s has substantial energy at scales $n \geq 10$, with a very broad peak near n=20. The spectrum is consistent with those observed by *Rabier et al* (1994) for sensitivity perturbations based on 2-day forecast errors. As in the case of $\tilde{e}_0^{T21}$ (see Fig 6c), the spectrum for $\tilde{s}^{T21}$ in Fig 10d shows that the projection onto T21 singular vectors is clearly inadequate for capturing the small scale features of s, and that most of the energy is confined near the truncation limit. In contrast, the spectrum for $\tilde{s}^{T42}$ in Fig 10f shows a more or less even distribution of energy at scales $n \geq 15$, and no evidence of energy being confined near the truncation limit. Again, the total energy of s appears to be just over two times larger

than that of $\tilde{s}^{T42}$. Note that the spectra for both projections have very little energy at large scales $n<10$, indicative of the relatively small scale structure of the singular vectors at initial time.

The relatively large impact of the T42 singular vectors on the sensitivity projections is clearly revealed by comparing the sensitivity spectra in Figs 10d,f with the forecast error spectra in Figs 6d-f. In Figs 6d,f the overall character of the spectrum is changed very little by increasing the singular vector resolution from T21 to T42 since most of the energy in e_{36} is in the range $n<20$. In Figs 10d,f the character of the spectrum is changed dramatically with the T42 singular vectors. The large impact of the T42 singular vectors on $\tilde{s}$ compared with $\tilde{e}_{36}$ is indicative of the fact that the unstable sub-space is characterized by small spatial scales at initial time and larger spatial scales at optimization time.

As mentioned earlier, energy distributions for $\tilde{e}_0$ were found to be qualitatively similar to those for $\tilde{s}$ shown in Figs 10c-f, except that, for $\tilde{e}_0$, the total energy was greater for the T21 projections. This is explained by the more localized structure of the T42 singular vectors, and by the reciprocal weighting by the growth rates σ_i in the case of $\tilde{e}_0$, which tends to reduce the amplitudes of the fastest growing structures. These effects become especially noticeable at T42 resolution since the growth rates are larger than at T21 (cf. Figs 7c,d and 9e,f). In addition, the smaller growth rates for the T21 singular vectors imply that a given forecast error at, say 36 hours, must necessarily have started with larger initial amplitude at T21 than at T42.

4. ENSEMBLE PREDICTION

As a final test of the impact of horizontal resolution on singular vector estimates, we discuss results from ensemble forecasts with perturbed initial conditions generated using T21 and T42 singular vectors. For each case study, ensembles have been run in four different configurations using the T63L19 version of the ECMWF operational model. Configurations "T21" and "2T21" use as initial conditions the same set of singular vectors computed at T21, the latter with an amplitude $\sqrt{2}$ larger than the former. Configurations "T42" and "2T42" use the same set of T42 singular vectors, the latter with an amplitude $\sqrt{2}$ larger than the former. Note that the ECMWF operational configuration has been "T21" until 23 August 1994 (see *Molteni et al*, 1995), and "2T21" until 13 March 1995 (see *Buizza*, 1995).

For reference, the norm of a typical "T42" ensemble perturbation is about 0.8 m^2s^{-2}. This can be compared with a typical amplitude of about 0.5 m^2s^{-2} for the pseudo-inverse perturbation, as discussed in section 3b. The amplitude of the EPS perturbation is determined in part by the requirement that the spread of the EPS in the medium range should not be much smaller than the error of the control forecast. The fact that the EPS perturbation amplitude is larger than the pseudo-inverse, is consistent with the 10-day forecast experiments discussed in section 3b where it was shown that a minimisation of medium-range forecast error was obtained with perturbations proportional to, but larger than, the pseudo-inverse. The extent to which this is

predominantly associated with nonlinear effects or model error effects is not known at present (see conclusions).

In sub-section 4.a we analyze in detail ensembles run in the four configurations for one of the sensitive cases; in sub-section 4.b. we discuss results averaged over the seven cases, and over the three sensitive cases.

a. 94.04.06: ensemble performances

Figure 11 shows the average (computed over the 16 perturbations) RMS amplitude of perturbation temperature at about 700 hPa, close to the level of maximum amplitude (see section 2) for configurations 2T21 and 2T42. Due to their scale, T42 perturbations tend to be more localized in the regions of instability, and this explains why larger local maxima can be detected in the average RMS perturbation amplitude field.

Figure 12 shows the RMS error and spread of the ensembles run in the four different configurations, computed over the NHET (left panels) and over Europe (right panels). For the T21 ensembles (Fig 12a-d), we can see that 2T21 has larger spread throughout most of the forecast period, and this leads to a slightly better ensemble-mean performance after forecast day 5, especially over the European area. Note that, during most of the forecast period, the mean spread of the T42 ensemble is comparable to the spread of the 2T21 configuration (Fig 12c-f), due to the faster growth of the initially smaller T42 perturbations. Comparing ensembles run in the 2T21 and T42 configurations (i.e. comparing the two configurations with similar spread), ensembles with the higher resolution singular vectors have more skilful ensemble-mean fields. The comparison of the T42 and 2T42 configurations shows that the further increase of the average spread does not produce a significant improvement of the ensemble skill.

For each configuration, the skill of the control and most skilful ensemble member (between day 5 to 8) is listed in Table 2. The most skilful member (number 4) of the 2T21 ensemble is more skilful than the most skilful member (also number 4) of the T21 ensemble. On the other hand, the best member of the T42 ensemble performs better than the best member of the 2T42 ensemble at forecast days 5-6, and worse at days 7-8.

Configuration	Ens-num	ACC skill Eur at fc-day				RMS err Eur at fc-day 7
		5	6	7	8	
	Control	.05	-.09	.12	.17	129.8
T21	4	.74	.74	.71	.58	76.0
2T21	4	.83	.83	.79	.70	70.8
T42	19	.85	.85	.77	.61	70.6
2T42	19	.80	.80	.78	.82	78.2

Table 2 ACC skill of the control forecast, and of the best members of ensembles run in all configurations for the 94.04.06 case.

Figure 13 shows the initial perturbations associated with the most skilful member of the 2T21 and the 2T42 ensembles. The region where both perturbations have significant amplitude is over North America. Figure 14 shows the verifying analysis, the control forecast, and the best member of the T42 ensemble at forecast day 7. The perturbed forecast performs better than the control since it succeeds in predicting the cut-off low over eastern Europe.

b. Composite results

Figure 15a shows the average over the seven cases of the ensemble-mean NHET RMS error for configurations T21 (solid), 2T21 (dash), T42 (dotted) and 2T42 (chain-dash), while Fig 15b shows the NHET RMS error but averaged over the three 'sensitive' cases. In a similar way, panels (c-d) show the RMS spread associated with the four ensemble configurations.

Comparing the 2T21 and T42 ensembles (which, on average, have similar NHET spread), Figs 15a-b show that configuration T42 is slightly more skilful. Configuration 2T42 has a slightly more skilful ensemble-mean after forecast day 5, but has the least skilful ensemble-mean between forecast days 1-4. Figures 15c-d show the impact of increasing perturbation amplitude throughout the forecast range, and illustrates the fact that to obtain a given level of spread in the medium range, a larger initial amplitude is needed using T21 perturbations.

Figure 16 shows European RMS error and RMS spread averaged over the three sensitive cases. These results indicate that the T42 ensembles perform notably better than either the T21 or the 2T21 ensembles. Again, Fig 16 shows that while the skill of 2T42 ensemble mean is the best after forecast day 7, it is the poorest between forecast days 1-4.

In addition to these scores, a cluster analysis has been performed on the 500 hPa geopotential height at forecast day 7 over the European area. As in *Molteni et al* (1995), the Ward hierarchical clustering algorithm (e.g. *Anderberg*, 1973) has been used, with the internal variance of each cluster set equal to the monthly-average forecast error variance at day 3.

Table 3 lists, for each case and for each ensemble configuration, the ACC skill of the best cluster and the number of ensemble members in that cluster (in brackets). The clusters are ordered by population, the first number in the super-script is the cluster order number and the second number is the total number of computed clusters. When the best cluster coincides with the most-populated cluster its ACC skill is underlined (e.g. the T42 ensemble for the 94.04.06 case). For example, for the 94.04.06 case, we can see that, for the two T42 configurations, the best cluster coincides with the most-populated and has a higher level of skill than the equivalent T21 clusters. On average the most skilful of the T42 clusters are more skilful than the equivalent T21 values. The average number of members in the most skilful cluster is larger with higher resolution, with the largest average size occurring for the T42 configuration.

	T21	2T21	T42	2T42
92.12.28	77 $^{3/6}$ (7)	76 $^{2/6}$ (6)	<u>78</u> $^{1/6}$ (11)	75 $^{3/6}$ (6)
93.02.14	77 $^{4/5}$ (4)	81 $^{2/6}$ (7)	85 $^{4/6}$ (4)	<u>78</u> $^{1/6}$ (8)
93.09.06	80 $^{2/3}$ (11)	78 $^{4/6}$ (6)	74 $^{2/3}$ (11)	76 $^{2/5}$ (7)
94.01.03	61 $^{5/5}$ (3)	69 $^{2/6}$ (6)	72 $^{3/6}$ (6)	77 $^{4/6}$ (4)
94.02.20	71 $^{2/4}$ (9)	76 $^{4/6}$ (5)	72 $^{4/4}$ (4)	74 $^{4/6}$ (5)
94.03.15	26 $^{2/2}$ (6)	<u>39</u> $^{1/4}$ (14)	74 $^{5/6}$ (3)	<u>72</u> $^{1/6}$ (7)
94.04.06	50 $^{5/5}$ (2)	57 $^{3/6}$ (4)	<u>65</u> $^{1/6}$ (11)	<u>71</u> $^{1/6}$ (12)
Average	63 (6.0)	68 (6.8)	74 (7.1)	74 (7.0)

Table 3 Characteristics, at forecast day 7, of the cluster with the best ACC skill. For each case and for each configuration, the first number represents the ACC skill (in %), the superscript fraction represents the cluster ordering number (ordered according to the population) with respect to the total number of clusters, and the value in brackets represents the number of ensemble members in the cluster (e.g., for the 94.01.03 T42 ensemble, the best cluster is number 5 of the 6 generated clusters, it has 74% ACC skill, and it is composed of 3 members). Underlined ACC skill highlighted the cases when the best cluster coincides with the most-populated cluster.

7. DISCUSSION AND CONCLUSIONS

In the introduction we posed the question as to whether the predictability of a synoptic-scale weather system in the short or medium range was determined by uncertainties in the initial state on the scale of the weather system itself, or by uncertainties in the initial state on scales much smaller than the weather system. To test this, we have studied the impact of various relevant initial perturbations estimated using forward and adjoint tangent models with either T21 or T42 resolution. If the synoptic-scale flow was only dependent on synoptic-scale uncertainties, then results should be largely independent of horizontal resolution.

The initial perturbations studied in this paper lie in the unstable subspace defined by the dominant singular vectors of the forward tangent propagator of the ECMWF numerical weather prediction model. These singular vectors have been calculated using a total energy inner product. It has been argued that the analysis error covariance matrix is approximately isotropic in this inner product space.

The perturbations considered are of three types. The first two types of initial perturbations are diagnostic; they are based on a projection of a given 36-hour forecast error field onto the set of singular vectors. The first perturbation ($\tilde{\boldsymbol{e}}_0$) is essentially the pseudo-inverse of the forecast error field, the second ($\tilde{s}$) gives the gradient of the forecast error with respect to the initial state. The third perturbation is prognostic; it is a perturbation that would be used in the ECMWF Ensemble Prediction System.

All the results highlight the sensitivity of the $\tilde{\boldsymbol{e}}_0$ and $\tilde{s}$ perturbations to the resolution of the tangent model, and show that significant amounts of energy in the $\tilde{\boldsymbol{e}}_0$ and $\tilde{s}$ perturbations resided in scales smaller than T21. This was particularly evident from the pseudo-inverse and sensitivity projections using the higher resolution singular vectors.

Moreover, by studying the skill of ensemble forecasts, it was shown that the skill of the resulting medium-range forecast could, in general, be improved more by using perturbations from the T42 tangent model than using the T21 tangent model. The implication is that the quality of the initial conditions between wavenumbers 21 and 42 can play an important role in determining the skill of the subsequent forecast of forecast activity.

These results are consistent with studies of upscale energy cascades in turbulent fluids studied theoretically and experimentally (e.g.*Lilly*, 1983, *Metais et al*, 1994). In these studies it is suggested that if atmospheric motions can be modelled by quasi-two dimensional turbulence, then energy associated with mesoscale forcing can propagate in spectral space, forming the -5/3 spectrum observed on scales up to about 1000 km. Energy could then be expected to disperse further up the spectrum to synoptic scales. Shutts (personal communication, 1995) has modelled this process in a large-eddy model with explicit convective forcing.

However, the benefit of the high resolution singular vectors was not felt uniformly over all cases studied. Specifically the impact of the T42 singular vectors was largest for cases of low predictability in which both the sensitivity pattern and the pseudo-inverse patterns had relatively large amplitude.

As a result of these studies, together with a more extensive set of quasi-operational tests (not reported here), the real-time configuration of the ECMWF Ensemble Prediction System was modified (on 14 March 1995) to calculate T42 singular vectors. At the same time, the optimisation interval was increased to 48 hours to be consistent with routine real-time sensitivity studies. As a result, the amplitude of the ensemble perturbations was set to a value smaller than the T42 configuration studied here (by a factor $\tilde{e}_0$).

It should be mentioned that, in this paper, we did not actually confirm the optimality of the evolved sensitivity perturbations using the full nonlinear model. To perform this evolution with the full nonlinear model requires a re-scaling of the sensitivity pattern so as to produce an appropriate perturbation amplitude. As stated earlier, further investigation is required to determine a method for obtaining these scale factors since there is no *a priori* guarantee that for each case the perturbation will be comparable in magnitude to, say, a typical analysis error. This and other aspects of the relationship between singular vectors and sensitivity perturbations will be investigated in detail in a subsequent paper.

In addition, we have not considered the role of model error itself in estimates of predictability. As well as model systematic error, forecast skill can be expected to be affected significantly by random model error. For example, within any grid box at any time step, the parametrized diabatic tendencies will have a random error associated with the fact that sub-grid scale processes are not in strict quasi-equilibrium with the large-scale flow. An example of the breakdown of quasi-equilibrium would be associated with convectively driven circulations which occur on scales comparable with the model grid. Hence, in addition to the random forcing associated with initial error, one can imagine a continuous background stochastic noise contributing to the

error field. Just as with the initial conditions, the atmospheric response to this forcing would be felt by components which project onto the dominant singular vectors. It would seem reasonable to suspect that such stochastic forcing would occur on relatively small scales, and that the T42 energy-norm singular vectors are appropriate quantities to characterise the atmospheric response.

A corollary of this is that the initial amplitude given to the singular vectors in the ensemble system might also have to reflect both initial and model random error. In this way, perturbations which are significantly larger than that given by the pseudo-inverse could be justified in the EPS. This would be consistent with results which showed that the maximum impact in the medium range of the pseudo-inverse perturbation occurred when the amplitude of the perturbation was increased.

REFERENCES

Anderberg, M R, 1973: Cluster analysis for applications. Academic Press, New York.

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

Buizza, R, 1995: Optimal perturbation time evolution and sensitivity of ensemble prediction to perturbation amplitude. Q J Roy Meteor Soc, in press.

Buizza, R and T N Palmer, 1995: The singular vector structure of the atmospheric general circulation. J Atmos Sci, 52, 9, 1434-1456.

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

Eady, E T, 1949: Long waves and cyclone waves. Tellus, 1, 33-52.

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

Golub, G H and C F van Loan, 1983: Matrix computations. North Oxford Academic. 476 pp.

Hartmann, D L, R Buizza and T N Palmer, 1995: Singular vectors: the effect of spatial scale on linear growth of disturbances. J Atmos Sci, in press.

Lacarra, J-F and O Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. Tellus, 17, 321-333.

Lilly, D K, 1983: Stratified turbulence and the mesoscale variability of the atmosphere. J Atmos Sci, 40, 749-761.

Marchuk, G I, 1974: Osnovnye i soprazhennye uravneniya dinamiki atmosfery i okeana. Meteor Gidrol, 2, 9-37.

Metais, O, J J Riley and M Lesieur, 1995: Numerical simulations of stably-stratified rotating turbulence. From "Stably stratified flow and dispersion over topography." Eds I P Castro and N J Rockliff. Clarendon Press, Oxford, 1994.

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

Molteni, F, R Buizza, T Palmer and T Petroliagis, 1995: The ECMWF ensemble prediction system: methodology and validation. Q J Roy Meteor Soc, to appear.

Pedlosky, J, 1979: Geophysical fluid dynamics. Spinger-Verlag New York Inc, 626 pp.

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

Rabier, F, E Klinker, P Courtier and A Hollingsworth, 1994: Sensitivity of two-day forecast errors over the Northern Hemisphere to initial conditions. ECMWF Research Department Tech Mem No. 203, ECMWF, Shinfield Park, Reading RG9-2AX, UK.

Zeng, Q-C, 1983: The evolution of a Rossby-wave packet in a three dimensional baroclinic atmosphere. J Atmos Sci, 40, 73-84.

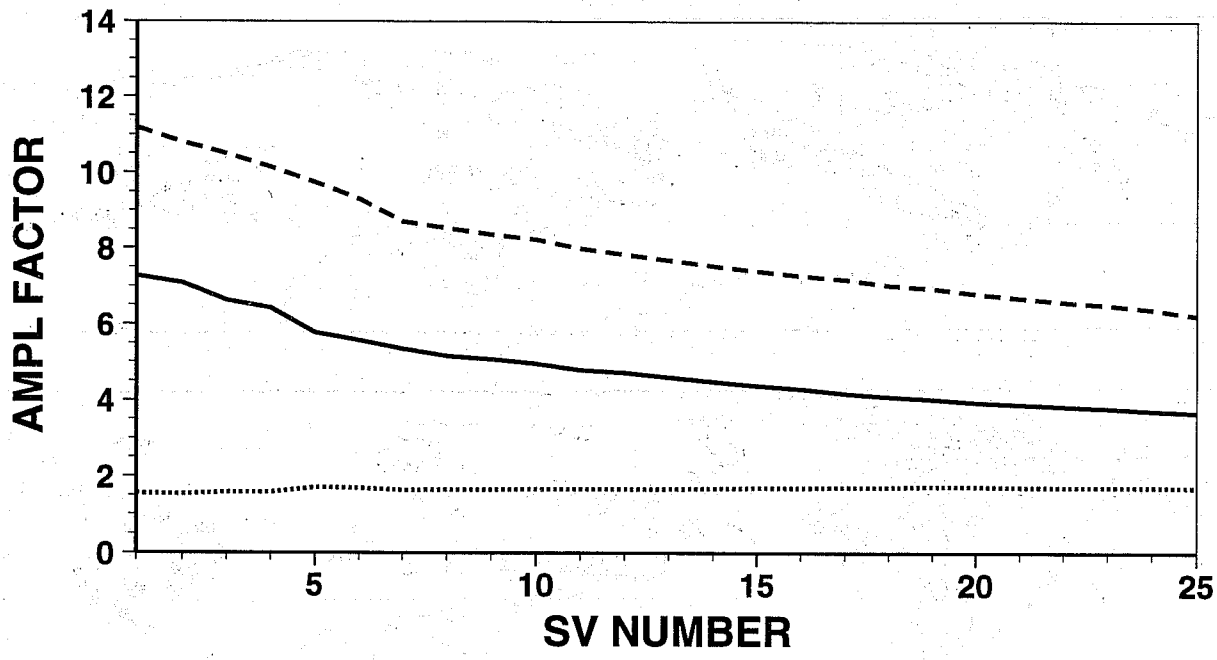

Fig 1 Average singular values of the first 25 unstable singular vectors computed at horizontal resolution T21 (solid) and T42 (dash), and ratio between the singular values factors of T42 and T21 singular vectors with the same sorting number (dot).

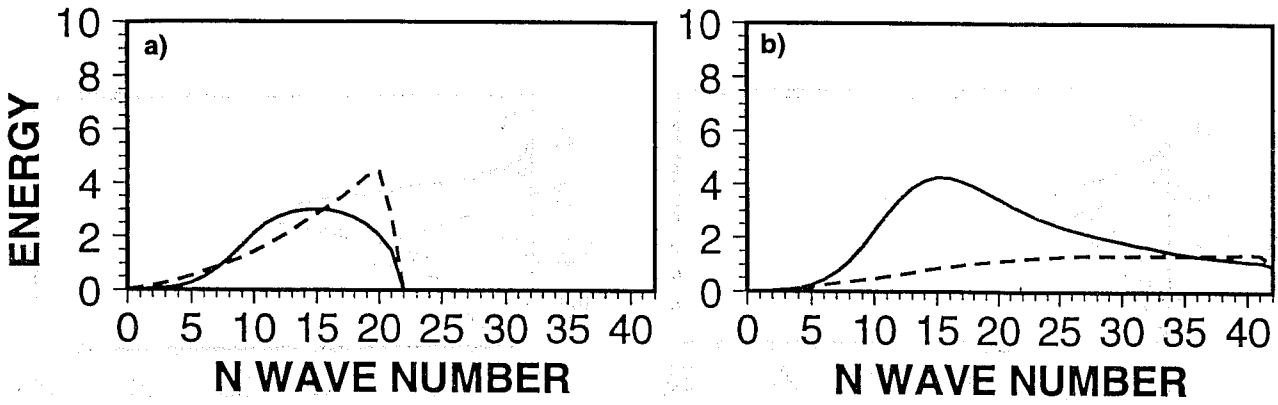

Fig 2 Energy spectrum (m^2, s^{-2}) of the (a) T21 and (b) T42 singular vectors, averaged over all the cases and over the first 16 singular vectors, at initial (dashed, x40) and final time (solid).

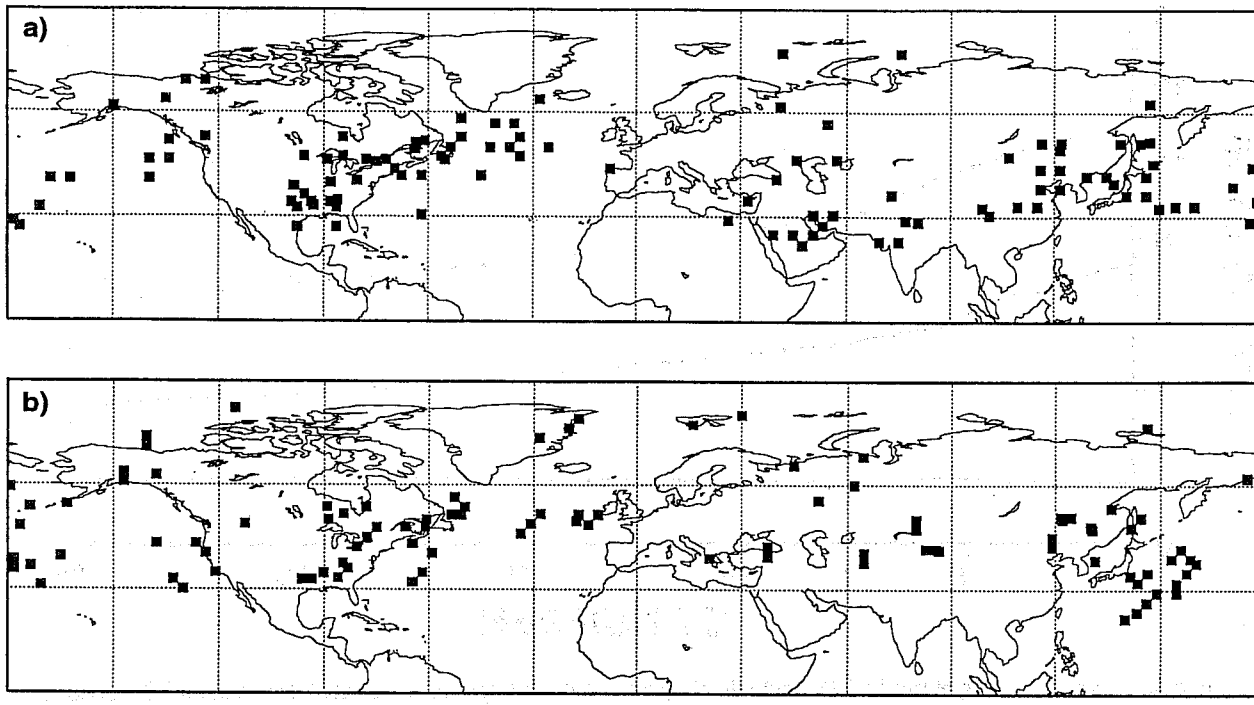

Fig 3 Location of vorticity maxima of the first 16 (a) T21 and (b) T42 singular vectors for all seven cases.

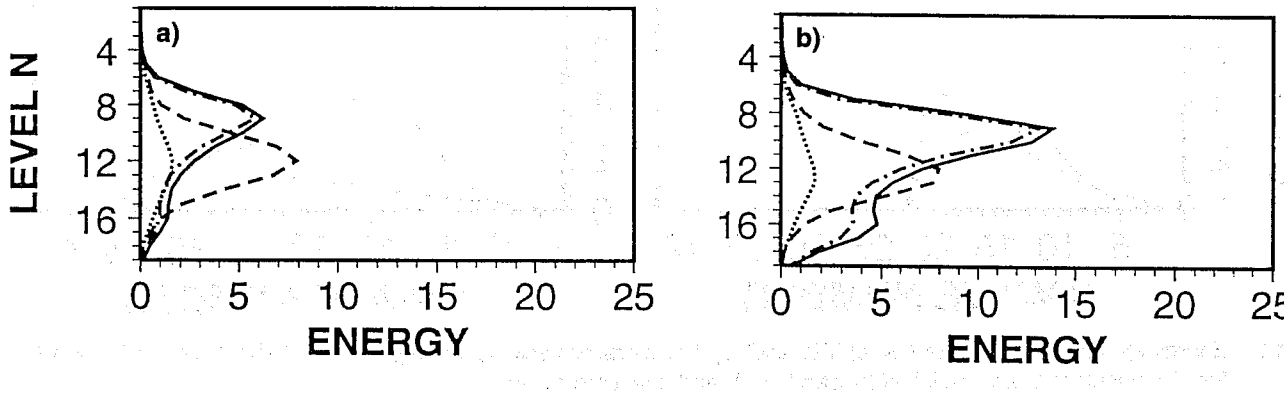

Fig 4 Vertical cross section of the total energy (m^2, s^{-2}, dash, x40) and the kinetic energy (dot, x40) at initial time, and of the total energy (solid) and the kinetic energy (chain-dash) at optimisation time, of the (a) T21 and (b) T42 singular vectors, averaged over all the cases and over the first 16 singular vectors.

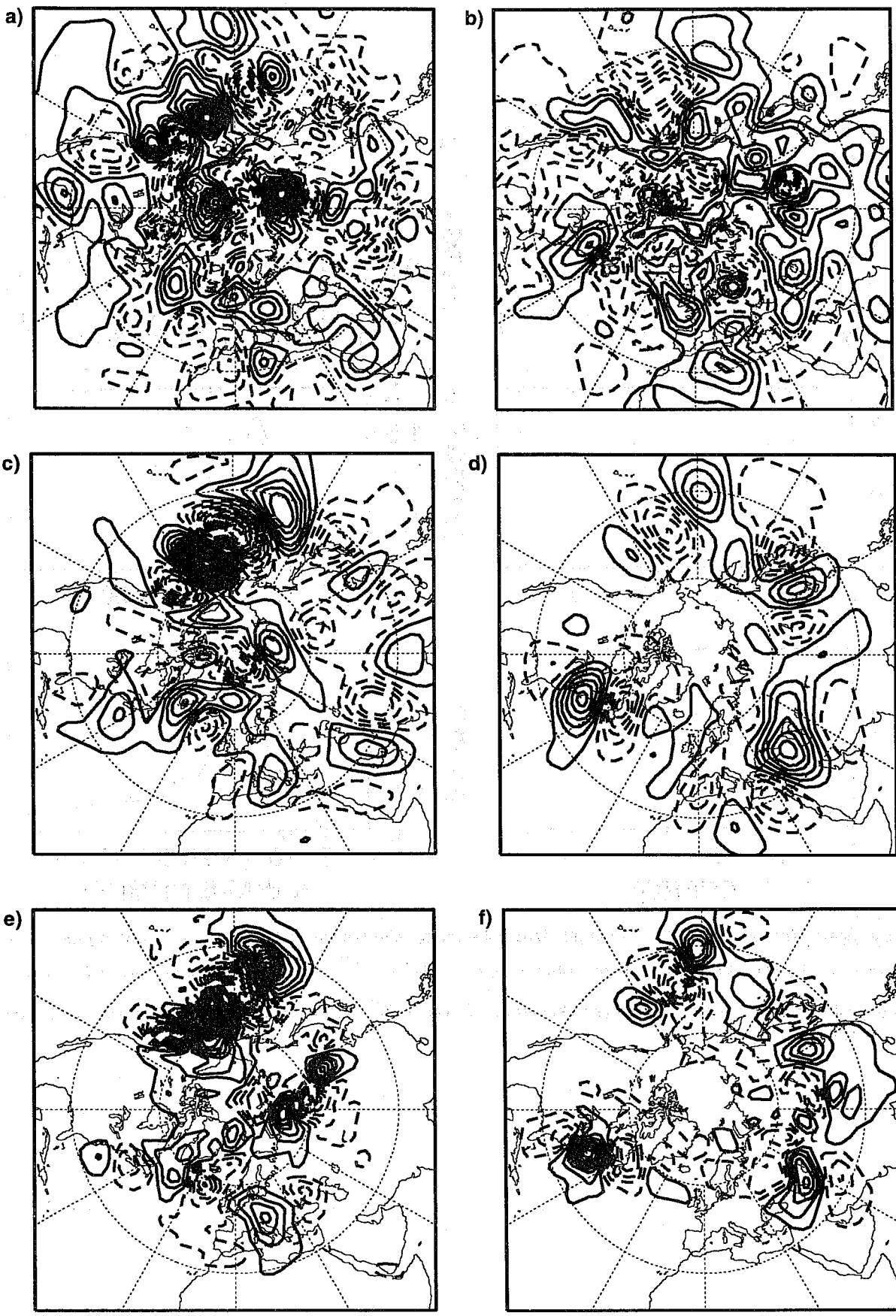

Fig 5 36h forecast error e_{36}, in terms of 500 hPa geopotential height, for the cases (a) 94.03.15 and (b) 94.04.06. (c-d): as (a-b) but for the projection $\tilde{\boldsymbol{e}}_{36}^{T21}$ of the forecast error on the T21 singular vectors. (e-f): as (a-b) but for the projection $\tilde{\boldsymbol{e}}_{36}^{T42}$ of the forecast error on the T42 singular. Contour interval 10 m in (a-b) and 4 m in (c-f), with negative values dashed.

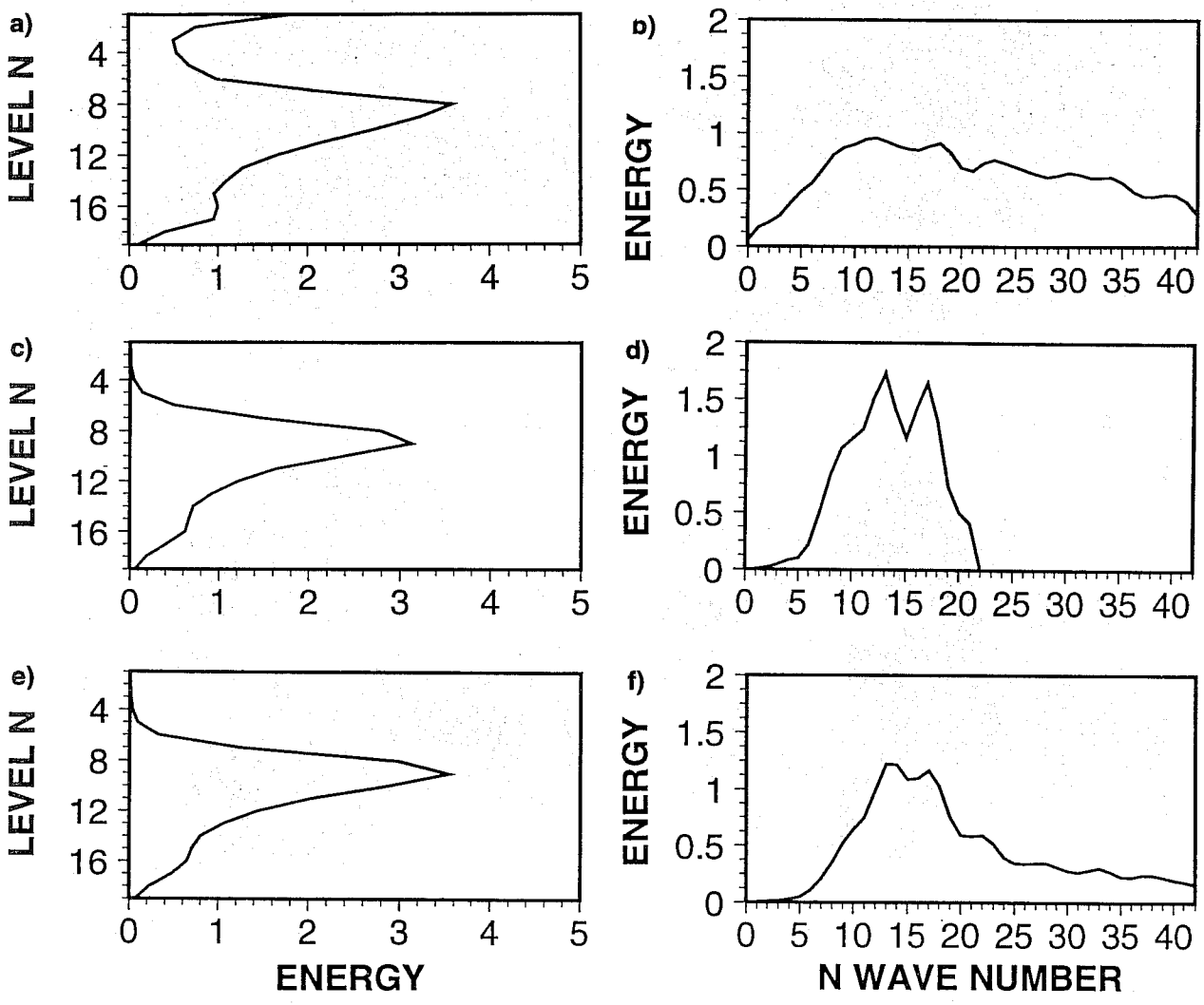

Fig 6 Total energy (m^2, s^{-2}) distribution of the 36h forecast error e_{36} with respect to (a) vertical level and (b) total wave number for the case 94.04.06. (c-d): as (a-b) but for the projection $\tilde{e}_{36}^{T21}$ using 16 singular vectors at T21 (values are multiplied by a factor 9). (e-f): as (a-b) but for the projection $\tilde{e}_{36}^{T42}$ using 16 singular vectors at T42 (values are multiplied by a factor 9).

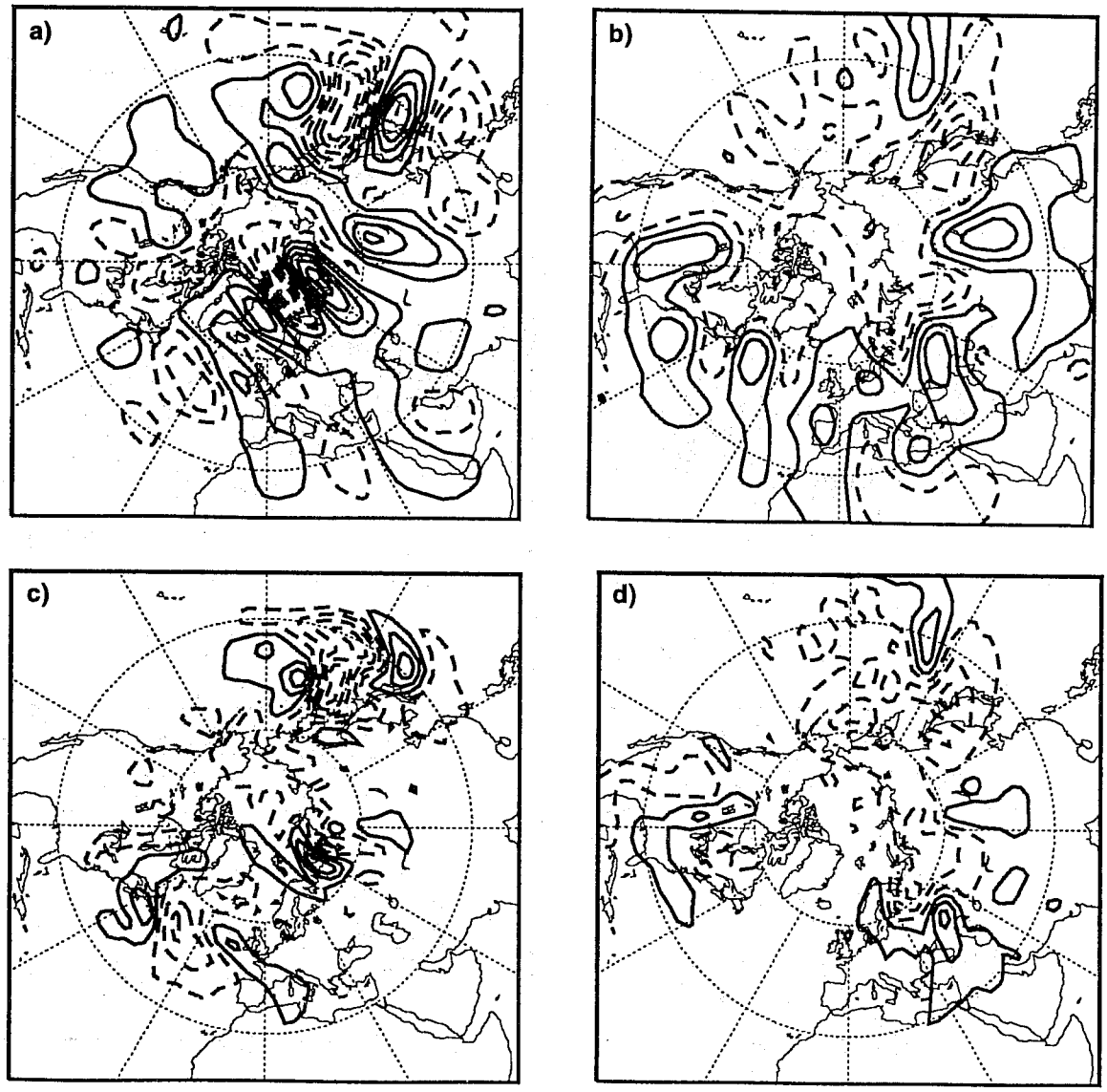

Fig 7 Projection of the initial error, in terms of 500 hPa geopotential height, on the T21 singular vector (i.e. $\tilde{e}_0^{T21}$) for (a) the 94.03.15 case, and (b) the 94.04.06 case. (c-d): as (a-b) but for the projection on the T42 singular vectors (i.e. $\tilde{e}_0^{T42}$). Contour intervals 0.5 m.

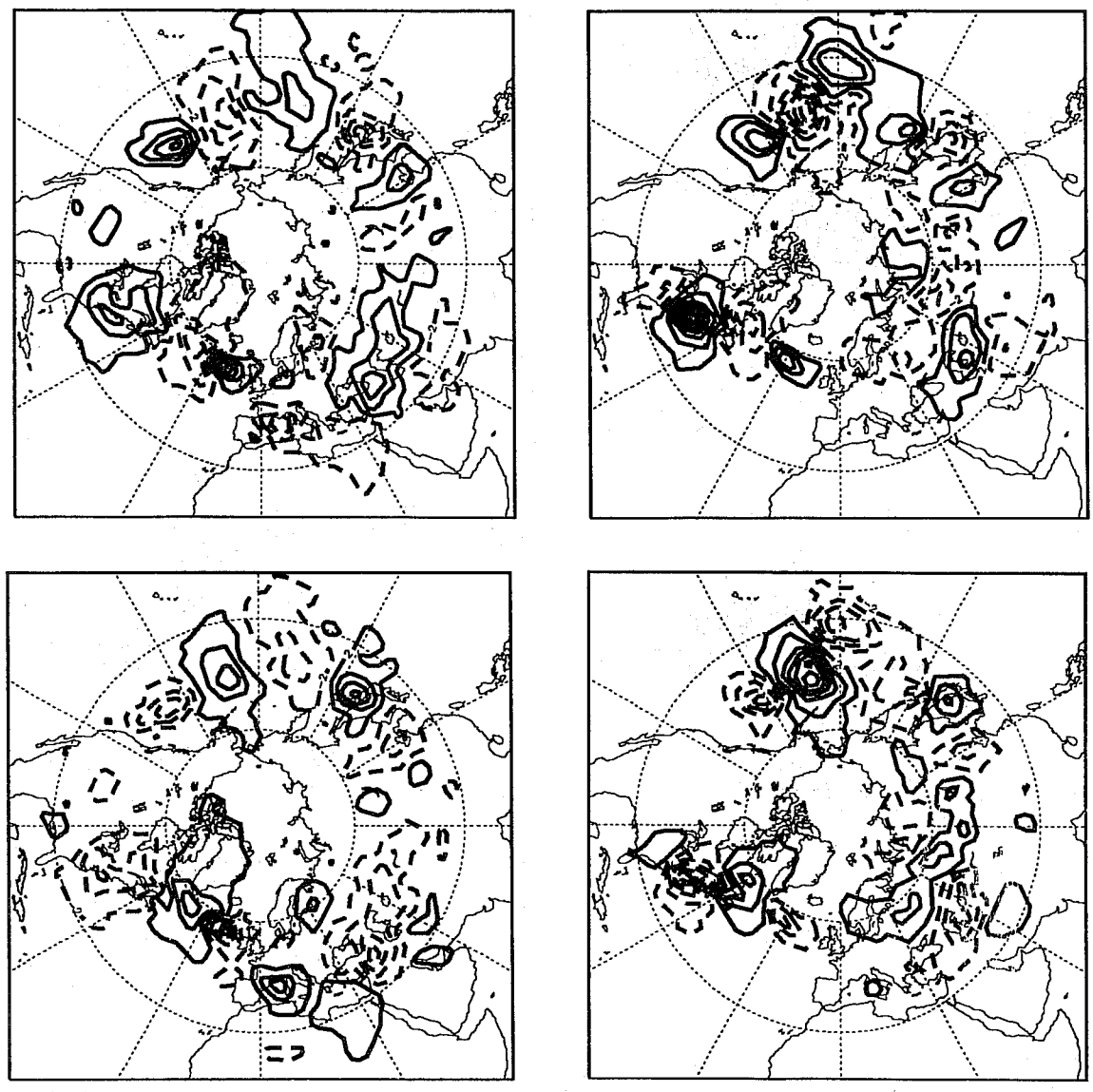

Fig 8 94.04.06, 500 hPa geopotential height: 36h non-linear time evolution, computed using the T63L19 model, of (a) $\tilde{e}_0^{T21}$ and (b) $\tilde{e}_0^{T42}$; 36h non-linear time evolution, computed using the T63L19 model, of (c) $-\tilde{e}_0^{T21}$ and (d) $-\tilde{e}_0^{T42}$. Contour interval 4 m.

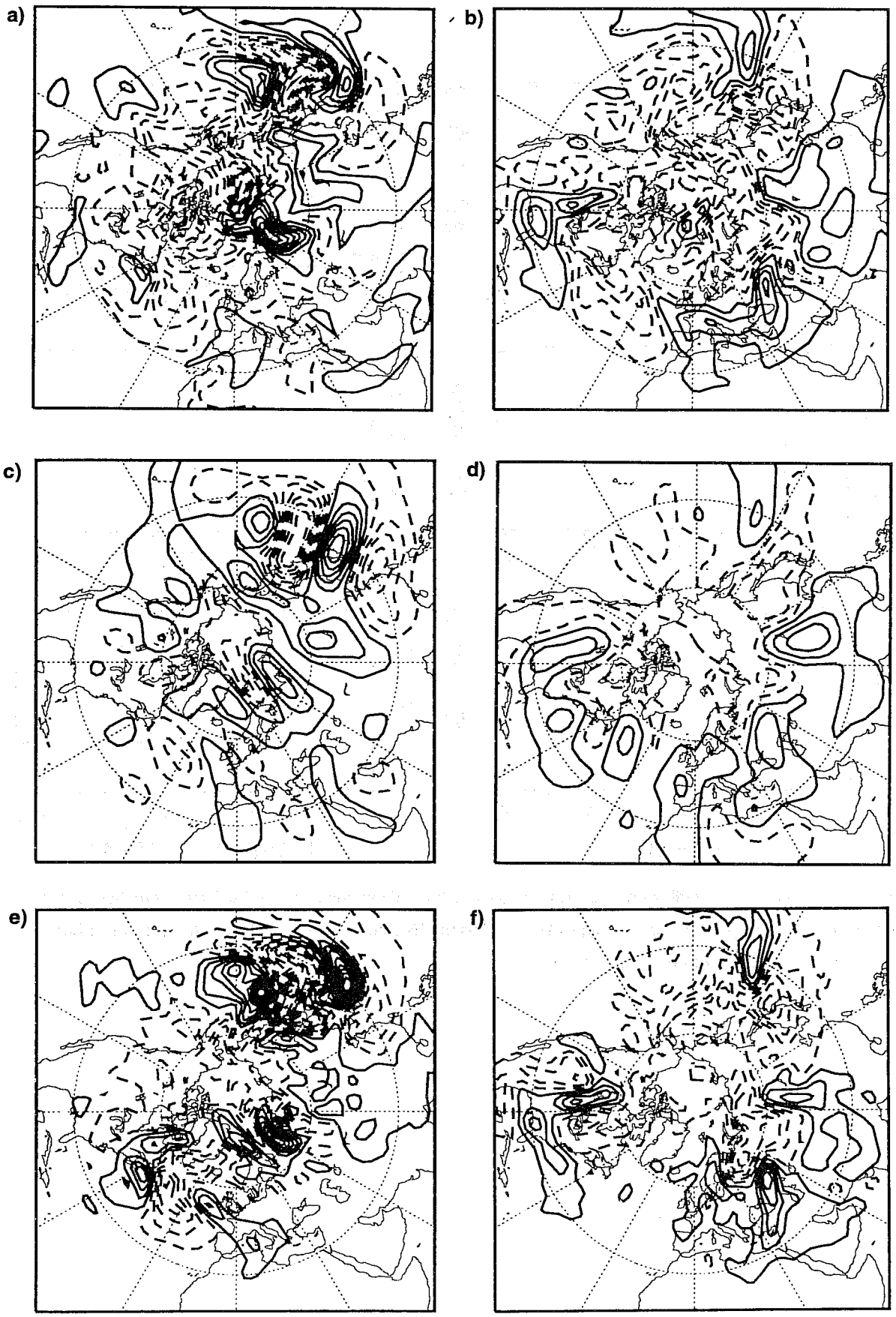

Fig 9 Sensitivity, s, in terms of 500 hPa geopotential height, for the cases (a) 94.03.15 and (b) 94.04.06. (c-d): as (a-b) but for the projection $\tilde{s}^{T21}$ using 16 singular vectors at T21. (e-f): as (a-b) but for the projection $\tilde{s}^{T42}$ using 16 singular vectors at T42. Contour interval 40 m in (a-b) and 20 m in (c-f), with negative values dashed.

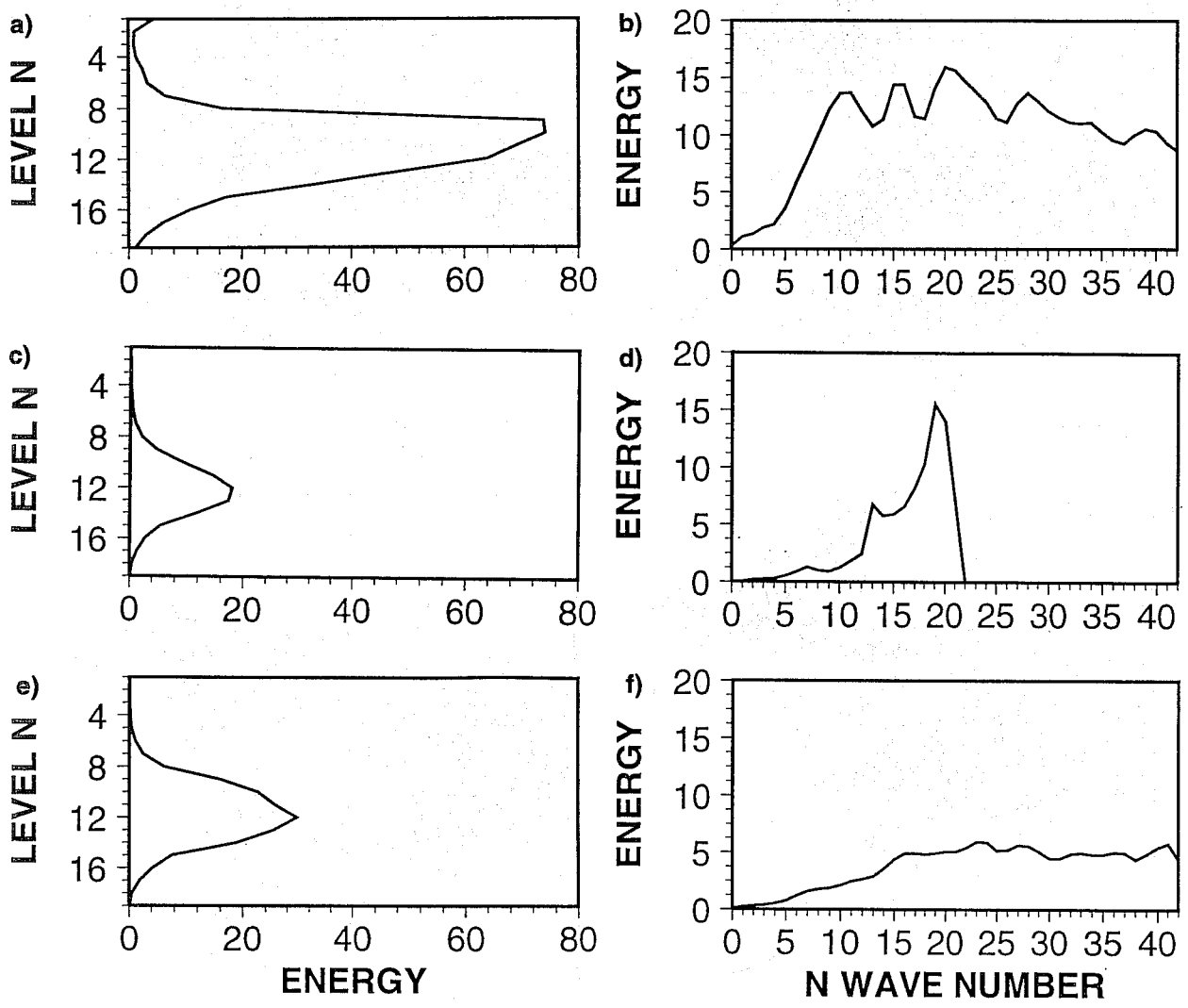

Fig 10 Total energy ($m^2 s^{-2}$) distribution of s with respect to (a) vertical level and (b) total wave number for the case 94.04.06. (c-d): as (a-b) but for the projection $\tilde{s}^{T21}$ using 16 singular vectors at T21. (e-f): as (a-b) but for the projection $\tilde{s}^{T42}$ using 16 singular vectors at T42.

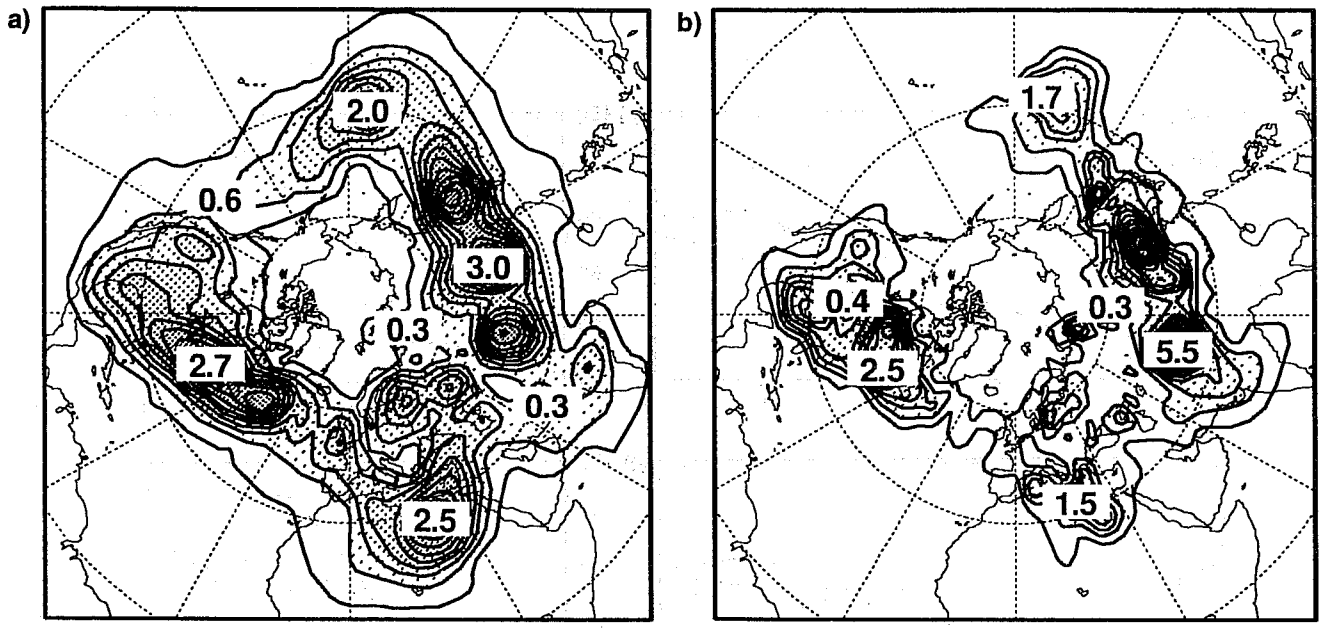

Fig 11 Average RMS amplitude of the perturbations (temperature at model level 13) added to the control initial conditions for ensemble configurations (a) 2T21 and (b) 2T42, for the 94.04.06 case. Contour interval 0.25°K, starting from 0.125°K.

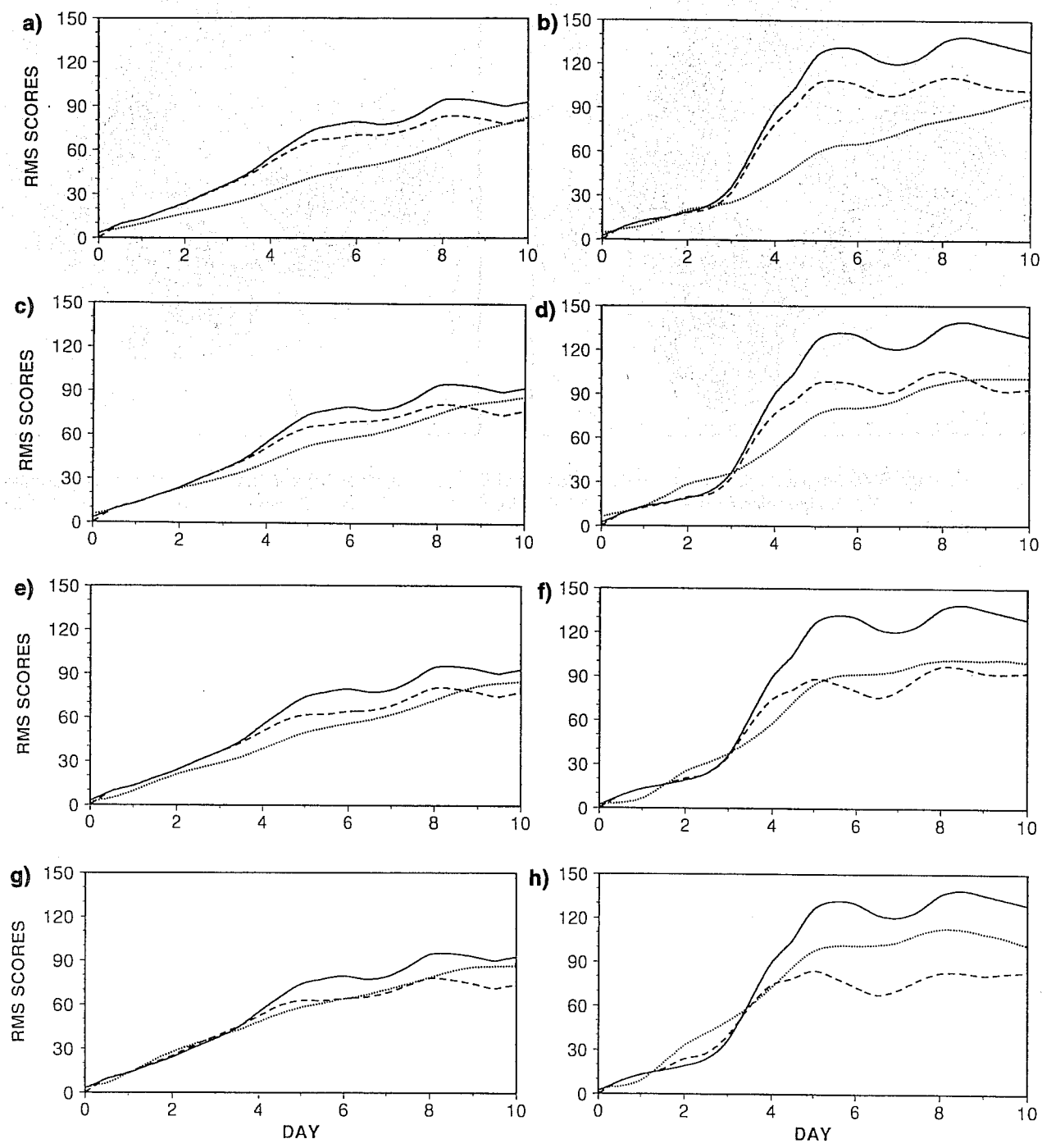

Fig 12 Control RMS error (solid), ensemble-mean RMS error (dash), and average of the ensemble members' RM spread (dot) of ensembles run for the 94.04.06 case, in configuration (a) T21, (b) 2T21, (c) T42, and (d) 2T42. All RMS values refer to the NHET. (e-h): as (a-d) but for Europe.

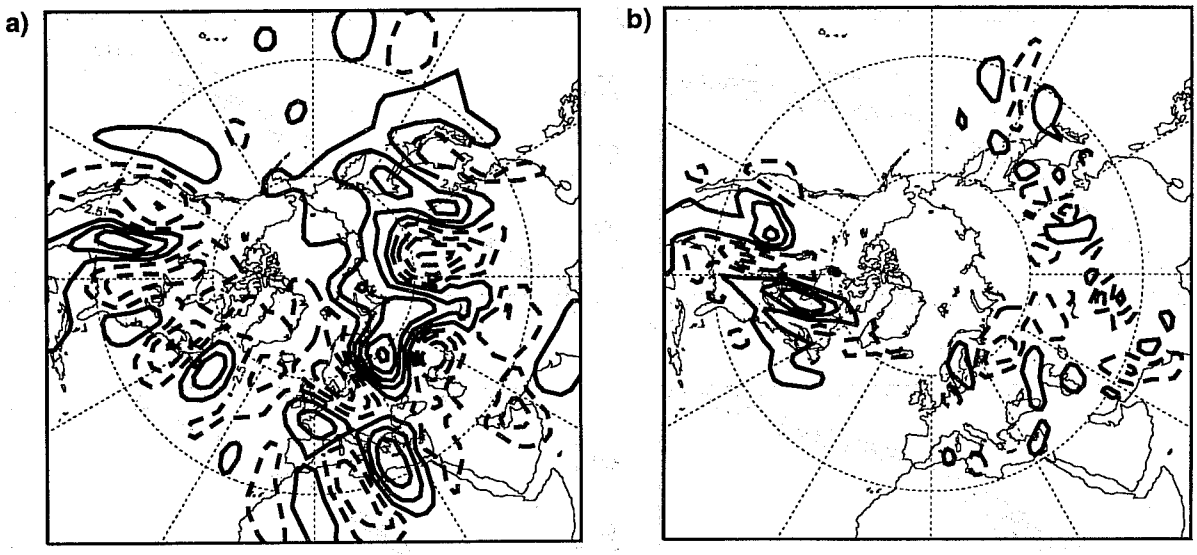

Fig 13 Initial perturbations, in terms of 500 hPa geopotential height, that added to the control initial condition gave the best prediction over Europe (in terms of 500 hPa geopotential height ACC), of ensembles run for the 94.04.06 case, in configuration (a) 2T21 (ensemble member number 4) and (b) 2T42 (ensemble member number 19). Contour interval 5 *m*, starting from 2.5 *m*, with negative values dashed.

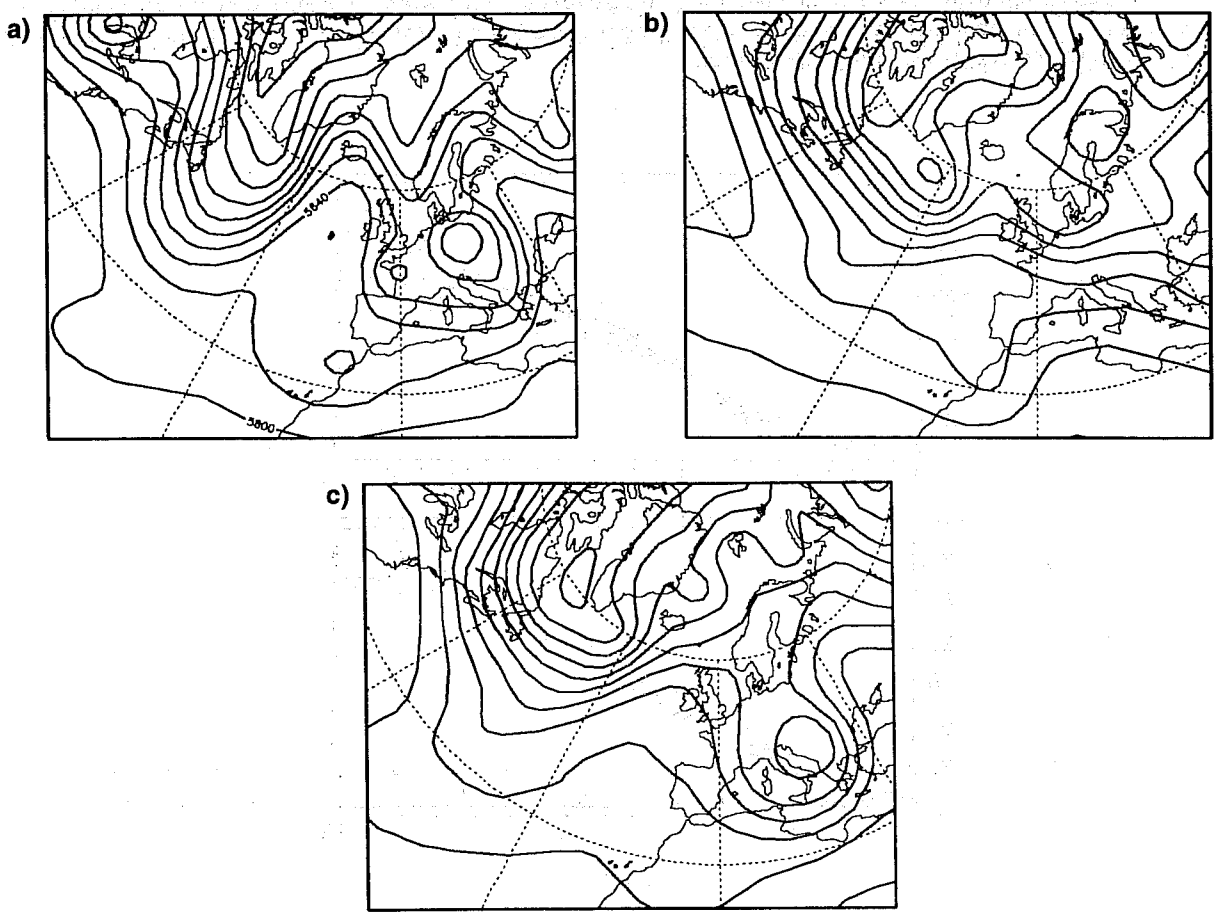

Fig 14 500 hPa geopotential height over Europe of (a) the analysis for 94.04.13, i.e. corresponding to forecast day 7 for ensembles started on 94.04.06, (b) control, (c) best T42 7-day forecast over Europe. Contour interval 80*m*.

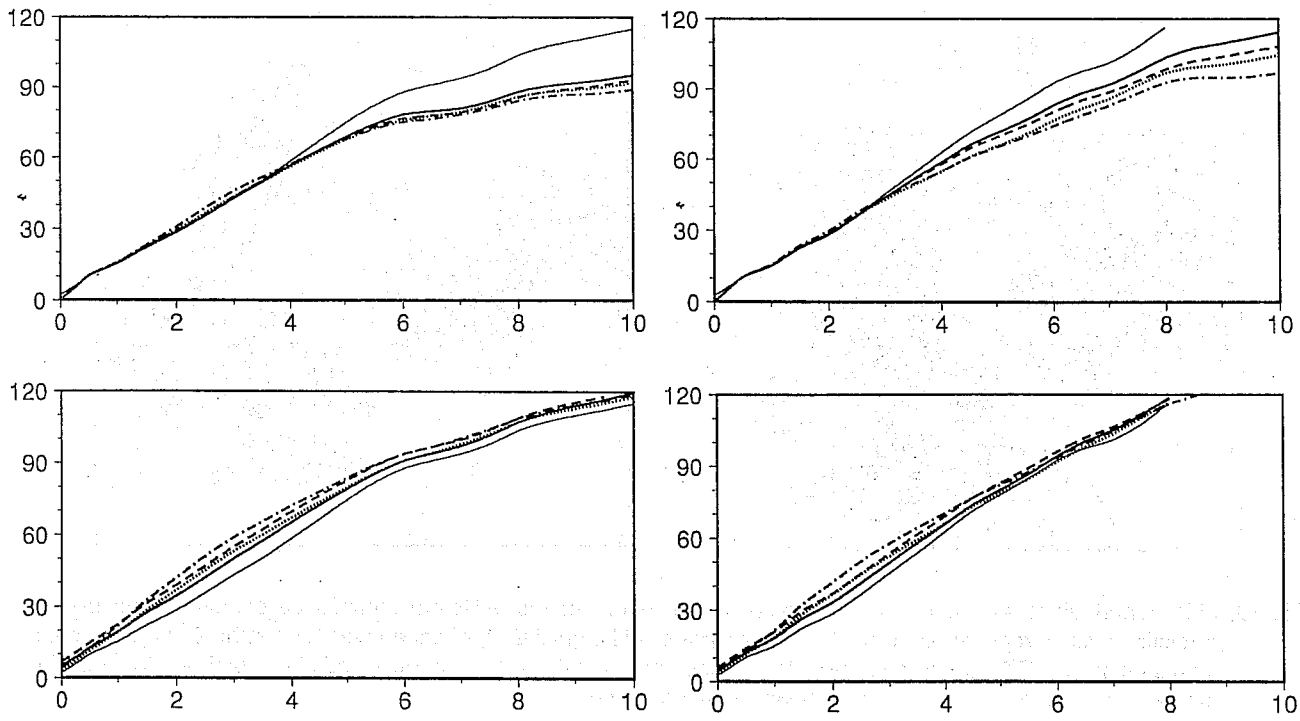

Fig 15 (a): average computed among the seven cases of the ensemble-mean RMS error for configurations T21 (solid), 2T21 (dashed), T42 (dotted) and 2T42 (chain-dashed) over the NHET. (b): as (a) but considering only the three sensitive cases (94.01.03, 94.03.15 and 94.04.06). (c): as (a) but for the average of the mean ensemble members' RMS spread over the NHET. (d): as (c) but considering only the three sensitive cases. As reference in each panel the thin-solid line shows the control RMS error.

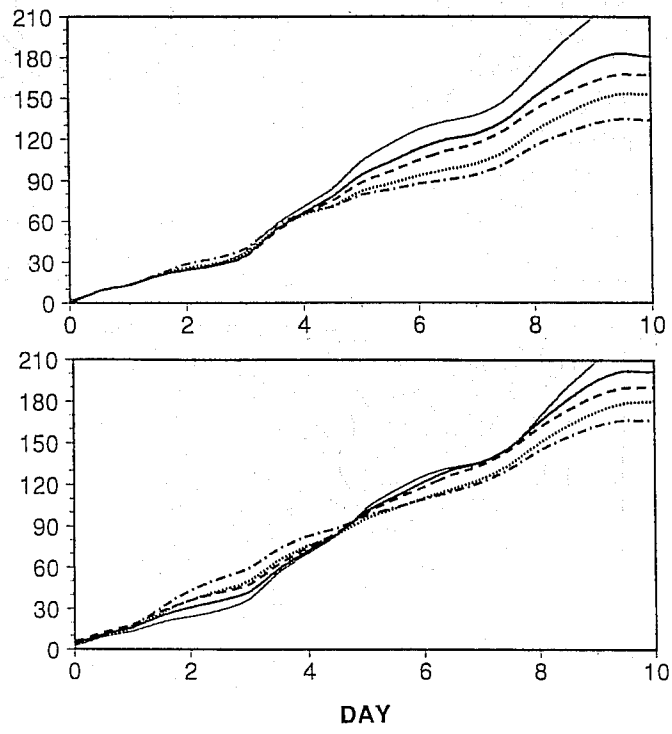

Fig 16 (a): average computed among the three sensitive cases (94.01.03, 94.03.15 and 94.04.06) of the ensemble-mean RMS error for configurations T21 (solid), 2T21 (dashed), T42 (dotted) and 2T42 (chain-dashed) over Europe. (b): as (a) but for the average of the mean ensemble members' RMS spread. As reference in each panel the thin-solid line shows the control RMS error.

APPENDIX A: SYNOPTIC SITUATIONS

Figure A1 shows the skill scores for the control forecasts in terms of anomaly correlation coefficient (ACC) and root mean square (RMS) error for geopotential height at 500 hPa over Europe. For days 5-7, the forecasts started on 92.12.28 and 93.02.14 can be classified as having average skill, forecasts started on 93.09.06, 94.01.03 and 94.03.15 as poor, and forecasts started on 94.02.20 and 94.04.06 as very poor. The synoptic situations during the three poor or very poor cases (94.01.03, 94.03.15 and 94.04.06) are described briefly.

a. 94.01.03

At initial time, a trough is located over the northeastern Atlantic, whilst continental Europe is characterized by a weak anti-cyclonic circulation and southern Europe by a northwesterly flow (Fig A2.a). Two days later, a cyclone starts developing over the Atlantic, and in the following days a deep trough appears over Europe. By forecast day 7, the flow over Europe is split, with a cut-off low over Greece and an anti-cyclonic circulation over northern Europe (Fig A2.b). After this, the flow over northern Europe becomes more zonal, while a deeper cut-off low develops over the southern Mediterranean. The poor performance of the control forecast after forecast day 6 over Europe is due to the erroneous prediction of a mature block over Europe at forecast day 7.

b. 94.03.15

At initial time, the flow over the Atlantic-European sector is zonal north of $45°N$, while two deep troughs are positioned over the eastern Pacific and eastern North America (Fig A2.c). During the following days, the flow over the British Isles becomes more northwesterly, associated with the arrival from the Atlantic of a ridge. At forecast day 7, this ridge moves towards The Netherlands, while the flow over Scotland becomes strongly westerly (Fig A2.d). The flow over southern Europe is weakly anticyclonic. The control forecast is very good up to forecast day 4, though later it predicts an excessive ridge west of the British Isles and an excessively deep and incorrectly tilted trough over central Europe. This leads to a completely erroneous prediction of a ridge over Spain and of a cut-off low over southern Italy at forecast day 7.

c. 94.04.06

At initial time, the flow over the Pacific and over the Atlantic is zonal, with troughs positioned over the central United States and over Europe; the latter having a northwest to southeast tilt (Fig A2.e). During the following days, a small cyclonic perturbation developing in the Atlantic moves toward Europe, and by forecast day 4, a deep cut-off low is observed over central Europe, with a ridge west of the British Isles. This situation persists until forecast day 7 (Fig A2.f). During the following days, a ridge over the northeastern Atlantic intensifies, a cut-off low develops to its south, and a cyclonic circulation persists over central Europe. The control prediction looses skill rapidly after forecast day 3, when an excessively weak and incorrectly tilted trough is forecast over central Europe. This leads, at forecast day 7, to a completely erroneous prediction of almost zonal flow over Europe.

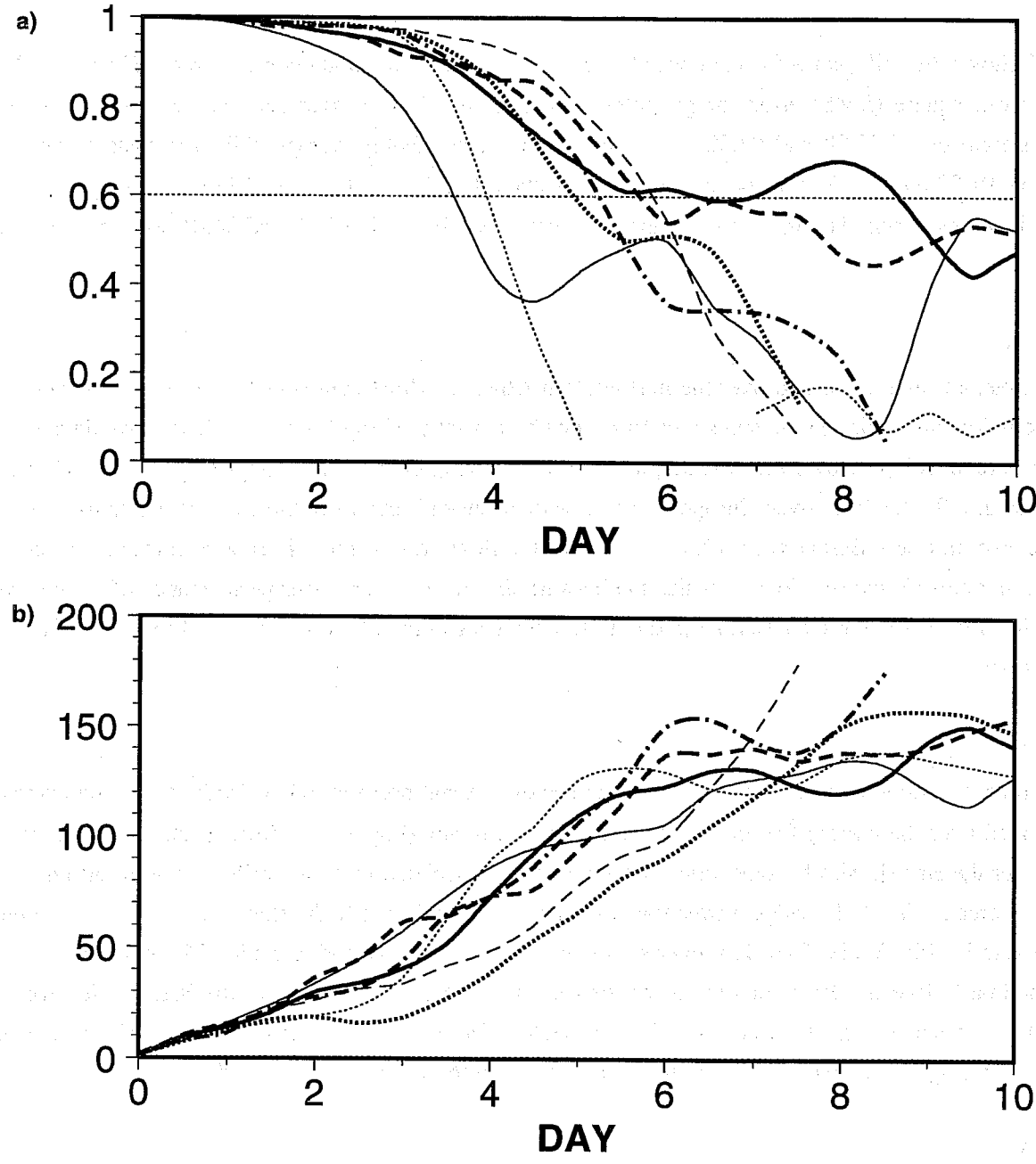

Fig A1 (a) ACC skill and (b) RMS error, computed over Europe, of the 10-day T63L19 control with initial dates: 92.12.28 (thick solid), 93.02.14 (thick dash), 93.09.06 (thick dot), 94.01.03 (chain-dash), 94.02.20 (thin solid), 94.03.15 (thin dash) and 94.04.06 (thin dot).

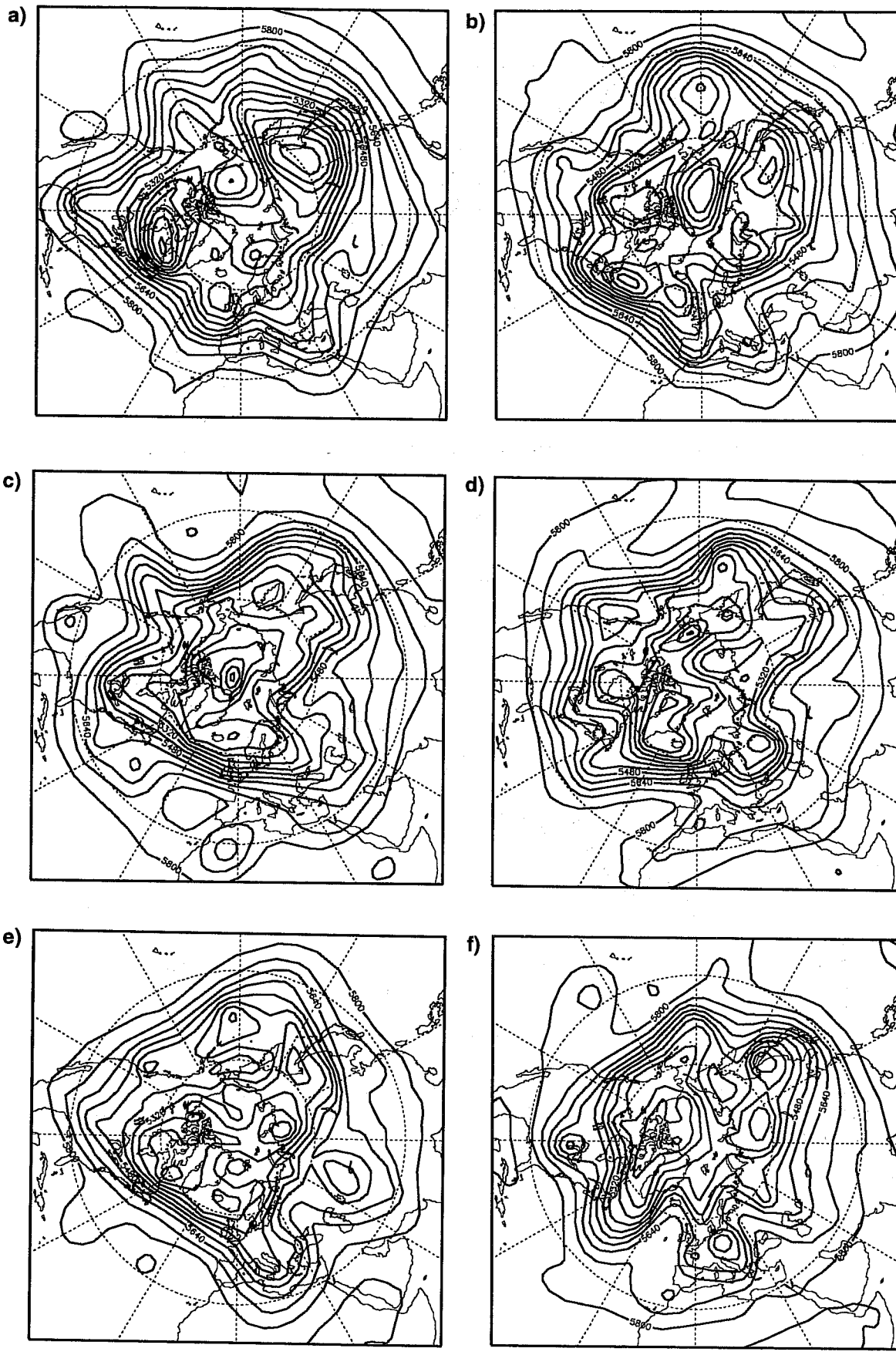

Fig A2 Analysis corresponding to the initial condition (left panels) and to the 7 day forecast, relative to (a-b) 94.01.03, (c-d) 94.03.15 and (e-f) 94.04.06. Contour interval 80 m.

APPENDIX B: AN OPTIMIZED ESTIMATE OF ANALYSIS ERROR.

In this appendix, we will demonstrate that the pseudo-inverse estimate $\tilde{e}_0$ of the analysis error and the sensitivity pattern $\tilde{s}$ can be viewed as solutions of the same optimization problem, corresponding to different values of an adjustable parameter.

The problem can be formulated as follows. Suppose that a 3-dimensional data assimilation system provides two independent analyses a_0 and a_1 for the state of the atmosphere at time t_0 and t_1 respectively. Let

$$f_1 = T(a_0, t_0, t_1) \tag{B.1}$$

denote the forecast obtained by integrating a dynamical model T from t_0 to t_1 with a_0 as initial condition. The mean-square error of this forecast with respect to a_1 is

$$E = \| T(a_0, t_0, t_1) - a_1 \|^2 \tag{B.2}$$

where a suitable norm has been assumed.

One may argue that a better estimate of the true state of the atmosphere at time t_0 is given by a corrected analysis a'_0 which minimizes the function

$$F = w_1 \| T(a_0', t_0, t_1) - a_1 \|^2 + w_0 \| a_0' - a_0 \|^2. \tag{B.3}$$

F represents the weighted sum of the mean-square forecast error at time t_1, computed with respect to a_1, and the mean-square analysis difference.

The relative amplitude of the weights w_0 and w_1 should reflect the expected accuracy of the assimilation and the forecast model. The ratio

$$\alpha = w_0 / w_1 \tag{B.4}$$

could be interpreted as the ratio between the expected mean-square errors of the forecast and assimilation systems. However, even assuming a perfect model, α should not be less than one, since the forecast error term in (B.3) is computed with respect to the a_1 analysis, which is itself affected by an error likely to be as large as the error in a_0.

If we define the analysis error of a_0 as $e_0 = a_0 - a'_0$ and assume that its norm is small compared to the norm of the full analysis, we can write

$$T(a_0', t_0, t_1) = T(a_0 - e_0, t_0, t_1) \simeq T(a_0, t_0, t_1) - L e_0 \tag{B.5}$$

where L is the linear propagator obtained by linearizing the model equations around the forecast trajectory started from a_0. With this approximation, and dividing all terms by w_1, (B.3) becomes:

$$F' = F / w_1 = \| e_1 - L e_0 \|^2 + \alpha \| e_0 \|^2 \tag{B.6}$$

where $e_1 = f_1 - a_1$ is the forecast error at time t_1.

As in Section 2, let us write the singular value decomposition of L as

$$L = U \Sigma V^t \tag{B.7}$$

where V is the matrix of the orthonormal singular vectors v_i at time t_0, U is the matrix of the orthonormal singular vectors u_i at time t_1, and $\sum$ is the diagonal matrix of the singular values. Expanding e_0 and e_1 onto the appropriate singular vector basis:

$$e_0 = \sum_i c_i v_i \tag{B.8a}$$

$$e_1 = \sum_i d_i u_i \tag{B.8b}$$

and writing

$$L e_0 = \sum_i c_i L v_i = \sum_i c_i \sigma_i u_i \tag{B.9}$$

the cost function F' can be expressed as

$$F' = \| \sum_i (d_i - c_i \sigma_i) u_i \|^2 + \alpha \| \sum_i c_i v_i \|^2 \tag{B.10}$$

and from the orthonormality of the singular vector it follows:

$$F' = \sum_i (d_i - c_i \sigma_i)^2 + \alpha \sum_i c_i^2. \tag{B.11}$$

Given the e_1 forecast error, and therefore the d_i coefficients, the c_i values that minimize F' can be found by setting the partial derivative of F' with respect to each c_i equal to zero. One obtains:

$$c_i = \frac{\sigma_i}{\sigma_i^2 + \alpha} d_i. \tag{B.12}$$

For singular values such that $\sigma_i^2 > \alpha$, then $c_i \simeq \sigma_i^{-1} d_i$. Therefore, for small values of α, the projection of e_1 onto the fastest growing singular vectors is well approximated by the pseudo-inverse estimate of analysis error:

$$e_0 \simeq V \Sigma^{-1} U^t e_1. \tag{B.13}$$

Conversely, when $\sigma_i^2 < \alpha$, then $c_i \simeq \alpha^{-1} \sigma_i d_i$. In the subspace of such singular vectors, e_0 becomes parallel to the sensitivity vector:

$$e_0 \simeq \alpha^{-1} s = \alpha^{-1} V \Sigma U^t e_1. \tag{B.14}$$

In practice, the full sensitivity pattern s becomes a good approximation to the solution of the optimization problem given by (B.12) when α is of the order of the largest squared singular value.