TECHNICAL MEMORANDUM

203

Sensitivity of two day forecast errors over the Northern Hemisphere to initial conditions

F. Rabier, E. Klinker, P. Courtier
and A. Hollingsworth

Research Department

September 1994 (SAC paper)

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

SENSITIVITY OF TWO-DAY FORECAST ERRORS OVER THE NORTHERN HEMISPHERE TO INITIAL CONDITIONS

F Rabier, E Klinker, P Courtier and A Hollingsworth
ECMWF
Reading, UK

1. INTRODUCTION

Large variations in daily forecast performance are a common problem of numerical forecasting systems at different NWP centres. ECMWF experience with data assimilation experiments and forecast experiments suggests that the largest forecast errors usually arise from errors in the initial conditions rather than from errors in the model formulation. However, single cases have to be examined carefully before a final conclusion can be drawn regarding error sources, as large sensitivity to changes in the parametrized forcing has occasionally been found as well.

There are several ways of separating analysis errors and model errors. The most commonly used approaches are either to run one model with a number of initial conditions generated by different NWP centres or to run different models with the same analysis (*Hollingsworth et al*, 1985; *Arpe et al*, 1985). The adjoint model, developed mainly for data assimilation applications at ECMWF, can also be used to identify sensitive structures in the initial conditions that might cause forecast errors. The adjoint equations, as explained by *Le Dimet and Talagrand* (1986), allow efficient computation of the gradient (or vector of first derivatives) of one output parameter of a numerical weather prediction model with respect to all input parameters. Applications of adjoint equations have mainly been to data assimilation, sensitivity problems and predictability. Meteorological applications up to 1992 are listed in *Courtier et al* (1993).

The four-dimensional variational data assimilation problem can be seen as finding the starting point of a model trajectory which will best fit the available data. Once an objective criterion J measuring the misfit between the model trajectory and the observations has been defined, the adjoint model gives the gradient of J with respect to the starting point x. This gradient can then be provided to an iterative minimization algorithm which will determine the x minimizing the objective function J. The theory and first applications of the method can be found in *Lewis and Derber* (1986), *Talagrand and Courtier* (1987) and *Courtier and Talagrand* (1987). For latest results, see *Zou et al* (1992), *Thépaut et al* (1993) or *Rabier et al* (1993a).

Adjoint equations have been used for a wide range of sensitivity problems. The sensitivity to model parameters has been investigated for instance by *Hall et al* (1982), *Hall* (1986), *Courtier* (1987), *Marais and Musson-Genon* (1992) and *Rinne and Järvinnen* (1993). The sensitivity of one aspect of the forecast to initial conditions has been the subject of the study of *Errico and Vukicevic* (1992) on real data cases. Their results were confirmed by *Rabier et al* (1992), in which the sensitivity of cyclogenesis to initial conditions was tackled in the context of an idealized flow, and a method was presented to filter out the effect of gravity waves on the gradient of the diagnostic function. It involves using the adjoint of nonlinear normal mode initialization at the end of the integration of the adjoint model.

Preliminary results on the sensitivity of forecast errors to initial conditions were obtained by *Rabier et al* (1993b) in which the diagnostic function was the square norm of the difference between the operational 48 hr forecast and the verifying analysis. This "energy" norm is derived from the quadratic invariant of the primitive equations linearized in the vicinity of a state of rest

$$\|x\|_W^2 = 1/2\int_0^1\int\int_\Sigma (u^2 + v^2 + R_a T_r(\ln\pi)^2 + (C_p/T_r)T^2)d\Sigma(\partial p_r/\partial\eta)d\eta$$

In this definition of the norm, x is the vector representing a perturbation of the atmospheric flow: its components (u, v, T, π) stand for horizontal components of the wind, temperature and surface pressure. W is the matrix of weights defining the norm. The weights are a function of T_r and P_r, which are a reference temperature and pressure, and of the constants R_a (gas constant for dry air) and C_p (specific heat at constant pressure for dry air). $\sum$ represents the integration domain.

For their experimentation *Rabier et al* (1993b) used a T63L31 near-adiabatic model (with horizontal and vertical diffusion and a surface drag, as in *Buizza*, 1993). The adjoint integrations were performed in the vicinity of a trajectory derived from a model employing the same simple physics starting from the ECMWF analysis valid 48 hours before the verification time. The limitation of the adjoint equations is that they are linear and can only describe the sensitivity to small initial perturbations for which the time evolution can be described by the tangent linear model. For this reason, the integrations of the adjoint model were not extended beyond 48 hrs since this has been shown to be a reasonable time limit for the validity of the tangent linear hypothesis (*Lacarra and Talagrand*, 1988; *Vukicevic*, 1991 or *Rabier and Courtier*, 1992).

The results of *Rabier et al* (1993b) can be summarised as follows. Firstly, the adjoint technique was shown to give reasonable results in a statistical sense, when studying the sensitivity of the 48 hr forecast error over the Northern Hemisphere with respect to initial conditions, during the month of January 1993. The sensitivity was found to be particularly large at the lower model levels, and over the oceanic mid-latitudes. In the mean, the analysis was too anticyclonic over the Atlantic ocean and too cyclonic over the Pacific

ocean. When applied to individual cases and more localised forecast errors, the approach helped to focus attention directly on those sensitive areas that were independently found to be the source of problems by synopticians, by more subjective error-tracking techniques.

On the basis of the encouraging results from the experimentation in January 1993, it was decided to implement the sensitivity calculation into routine operation starting mid-February 1994. Each day, adjoint integrations are performed to investigate the sensitivity of 24 hr and 48 hr forecast errors to the atmospheric state one and two days before. Forecast errors are defined as the differences between the operational forecasts and the corresponding analyses, both truncated at T63L31. The diagnostic functions are taken as the square norms of the forecast errors, computed over either the extra-tropical Northern Hemisphere or Europe. Both the trajectories needed for the adjoint integrations and the adjoint integrations themselves are performed at T63L31. An improvement has been made to the preliminary results of *Rabier et al* (1993b): the trajectories are now computed using the full ECMWF physics package in the forward nonlinear model. The adjoint integrations are performed with the simple physics consisting of horizontal and vertical diffusion and of surface drag. In this study, we have mainly studied the sensitivity of the 48 hr forecast error over the extra-tropical Northern Hemisphere (latitudes higher than 30N). The gradients of the 48 hr forecast error with respect to the initial conditions are studied as sensitivity patterns. They can also be used to generate small perturbations to be added to the initial conditions in order to improve the two-day forecast. The mathematics are detailed in section 2. Results are presented in section 3 for individual cases and in section 4 for the entire month of April 1994. Concluding remarks are given in section 5.

2. MATHEMATICAL FORMULATION OF THE PROBLEM

2.1 Expression of the gradient fields

The diagnostic function is

$$J = \|x_T - x_T^{ref}\|_W^2$$

where the norm $\|.\|_W$ is the "energy" norm as defined in the introduction, with the integration domain $\sum$ taken as the area (30N:90N). x_T represents the two-day forecast valid at time T. x_T results from the integration of the operational model M over 48 hours, starting from the operational analysis x_o: $x_T = M(x_o)$, which makes J a function of the analysis x_o. x_T^{ref} stands for the verifying analysis at time T.

Let us introduce the inner-product $<.,.>$ such that $<x,y> = x^t(2W)y$ where x^t stands for x transpose and W is the matrix of weights defined in the previous section, and the projection operator P projecting any global field x onto the integration domain $\sum$. J can be rewritten as

$$J = \tfrac{1}{2}<P(x_T - x_T^{ref}),\ P(x_T - x_T^{ref})>$$

A perturbation δx_o in the initial conditions would lead to a perturbation δx_T in the final time, with $\delta x_T = M(x_o+\delta x_o) - M(x_o)$. This would create a change in J

$$\begin{aligned}\delta J = J(x_o+\delta x_o) - J(x_o) &= \tfrac{1}{2}<P(x_T+\delta x_T - x_T^{ref}),\ P(x_T+\delta x_T - x_T^{ref})> - \tfrac{1}{2}<P(x_T - x_T^{ref}),\ P(x_T - x_T^{ref})> \\ &= <P(x_T - x_T^{ref}),\ P\delta x_T> + O(\delta x_T^2)\end{aligned}$$

P is a linear operator, its adjoint P^* is defined by

$<Px,y> = <x,P^*y>$, for any (x,y).

We can use this to transform δJ

$$\delta J = <P^*P(x_T - x_T^{ref}),\ \delta x_T> + O(\delta x_T^2).$$

i) By definition, the gradient of J at time T is such that, to first order approximation $\delta J = <\nabla J_T,\ \delta x_T>$. Identifying with the expression for δJ, we have

$$\nabla J_T = P^*P(x_T - x_T^{ref}).$$

This can be simplified further noting that P is a projection operator orthogonal for $< >$, thus $P^*P = P^2 = P$

$$\nabla J_T = P(x_T - x_T^{ref})$$

The gradient of J at time T is simply equal to the projection of the error field onto the domain $\sum = (30N{:}90N)$.

Remark: In practice, a technical difficulty is introduced since x is represented in spectral space. Thus P is defined as first going from spectral space to grid-point space, then resetting the grid-point values outside $\sum$ to 0, and finally going back to spectral space. Thanks to the Parseval quality, P is orthogonal for the inner-product and the previous discussion applies.

ii) Similarly the gradient of J at initial time is defined by

$\delta J = <\nabla J_o, \delta x_o> + O(\delta x_o^2)$.

It is now necessary to introduce a linearisation in the time evolution of the perturbation:

$\delta x_T = M(x_o+\delta x_o) - M(x_o) \simeq R\delta x_o$ where R is the resolvent of the tangent linear model.

We then have

$\delta J \simeq <P^*P(x_T-x_T^{ref}), \delta x_T> \simeq <R^*P^*P(x_T-x_T^{ref}), \delta x_o>$

$$\nabla J_o = R^*P(x_T-x_T^{ref})$$

with R^* the adjoint of the resolvent of the tangent linear model

By definition, a change δx_o in the initial conditions x_o will lead to a change in the forecast error J given by $\delta J = <\nabla J_o, \delta x_o>$. One can say that, in regions where the gradient ∇J_o is large, a change in the initial conditions would have created a large impact on the forecast error. Similarly, in regions where the gradient is small, such a change in the initial conditions would have affected the subsequent forecast error very little. Consequently, the pattern of the gradient of J at time 0 shows the sensitivity of the day 2 forecast error to the initial conditions. We remark that the advantage of using the inner-product <.,.> used in the definition of J is that the gradients of J at any time have the same units as meteorological fields. It should be noted that the experimental results depend on the choice of inner-product. In particular, the gradient structures discussed in the following sections would have different scales with an enstrophy norm, for instance.

2.2 Interpretation of the sensitivity field

i) If we assume that the forecast model is perfect and the linear approximation is valid for two days, then we can suppose that the forecast error at day 2 is due to the growth of a small analysis error δx_o evolving according to the tangent linear model. Thus

$Mx_o - x_T^{ref} = R\delta x_o$

and so $\nabla J_o = R^*R\delta x_o$.

The operator R^*R is real symmetric and operates in finite dimension and so it has a complete set of orthogonal eigenvectors v_i and associated (non-negative) eigenvalues λ_i^2, $\lambda_1^2 \geq \lambda_2^2 \geq \ldots \geq \lambda_N^2 \geq 0$.

Thus if we expand δx_o as

$$\delta x_o = \sum_{i-1}^{N} e_i v_i$$

then $\nabla J_o = \sum_{i-1}^{N} \lambda_i^2 e_i v_i$.

If the analysis error has a white spectrum in the expansion (i.e. e_i=1), the sensitivity pattern is dominated by the eigenvectors with the largest amplification factors λ_i. The eigenvectors of the operator R^*R or so-called singular vectors play a key role in the Ensemble Prediction System described by *Molteni et al* (1994). The fastest growing eigenvectors can amplify by a factor of 10 to 15 in a period of two days giving a value of λ_i^2 of order 100 to 225. Typically there are about 50-100 eigenvectors with amplification factors greater than 2. It is therefore reasonable to speculate that the introduction of an initial perturbation of the form $\delta x_o = -\alpha \nabla J_o$ where $\alpha \approx \frac{1}{\lambda_1^2}$ would lead

to a substantial reduction in forecast error at day 2, at least in those cases where the forecast error is dominated by rapidly growing eigenvectors. The use of a vector of the form $\delta x_o = -\alpha \nabla J_o$ as a perturbation of the initial conditions can be considered as a first step in the process of minimizing the diagnostic function J. This is typically what would be done by a conjugate-gradient minimization algorithm, with α a computed optimal step-size (for example, in a one-dimensional problem, this optimal step-size is given by the second derivative of J: $\delta x_o = -(J'')^{-1}\nabla J$). It would bring a change $\delta J = \langle \nabla J_o, \delta x_o \rangle = -\alpha \langle \nabla J_o, \nabla J_o \rangle$ which, for a given value of the norm $\|\delta x_o\|$ of the initial small perturbation, is maximum. For a larger initial perturbation one can still use $\delta x_o = -\alpha \nabla J_o$ to bring a decrease in the diagnostic function J but one cannot ensure that there is no other perturbation of same norm which is more effective in decreasing J, as second order terms $(o(\delta x_o^2))$ cannot be neglected any more in $J(x_o+\delta x_o) - J(x_o)$.

ii) If, however, one actually wants to cancel the forecast error at time T in the tangent linear context, one has to invert the tangent linear model. If R is singular, one can use the pseudo-inverse of R. As a matter of fact, taking as initial perturbation $\delta x_o = -R^{-1}(M(x_o)-X_T^{ref})$ will create $\delta x_T = R\delta x_o = -(M(x_o)-X_T^{ref}) = -(x_T-x_T^{ref})$. One would then manage to minimize $J = \langle x_T-x_T^{ref}, x_T-x_T^{ref} \rangle$ by having $x_T - x_T^{ref} = 0$.

iii) There is no equivalence between these two different approaches to perturbing the initial conditions so as to reduce the function J. Only under the special circumstances would one get the same answer from the two approaches. For example, if $R^*R = \alpha I$, then one would have $\delta x_o = -R^{-1}(M(x_o)-x_T^{ref})$ proportional to $\nabla J_o = R^*(M(x_o)-x_T^{ref})$. The relation $R^*R = \alpha I$ implies that all components of the initial error grow at the same rate. Let us suppose that δx_o is a small initial error, evolving to give $R\delta x_o$ at time T. One has for any δx_o, $<R\delta x_o, R\delta x_o> = <R^*R\delta x_o, \delta x_o> = \alpha<\delta x_o, \delta x_o>$ which means that the energy of the final perturbation $R\delta x_o$ is proportional to the energy of the initial perturbation δx_o. This does not happen in practice with real meteorological models for which it is known that the growth rate of error is very variable with phase-space direction (*Buizza and Palmer*, 1994).

One can illustrate the difference between the sensitivity field ∇J_o and the inversion of the tangent linear model in a simple example. Let us imagine that the atmospheric state x can be split into N components $(x_i)_{1 \le i \le N}$ with different amplification error rates such that the tangent linear model writes $R = \begin{pmatrix} \lambda_1 & 0 \\ & \ddots \\ 0 & \lambda_N \end{pmatrix}$

with $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_N > 0$. This poses the problem in the particular case of a real, diagonal R. The initial error $e = x_o - x_o^{ref} = (e_i)_{1 \le i \le N}$ is directly decomposed in singular vectors. It evolves in time as $x_T - x_T^{ref} = (\lambda_i e_i)_{1 \le i \le N}$. The diagnostic function measuring the forecast error at time T can be written $J = 1/2\ (x_T-x_T^{ref})^t(x_T-x_T^{ref})$. The canonical inner product will also be used to define the gradient. Inverting the tangent linear model would consist of building

$$\delta x_o = -R^{-1}(x_T-x_T^{ref}) = -\begin{pmatrix} \frac{1}{\lambda_1} & 0 \\ & \ddots \\ 0 & \frac{1}{\lambda_N} \end{pmatrix} (x_T-x_T^{ref}) = -(e_i)_{1 \le i \le N}$$

which is the negative of the initial error in the analysis. As a comparison, going along the gradient $\nabla J_o = R^t(x_T - x_T^{ref}) = \begin{pmatrix} \lambda_1^2 e_1 \\ \cdot \\ \cdot \\ \lambda_N^2 e_N \end{pmatrix}$ means starting from $\delta x_o = -\alpha \nabla_o J = -\alpha \begin{pmatrix} \lambda_1^2 e_1 \\ \cdot \\ \cdot \\ \lambda_N^2 e_N \end{pmatrix}$. In this simple example the optimal α is easy to find analytically. It is the α bringing the largest decrease in diagnostic function $J(x_o+\delta x_o) - J(x_o) = (x_T - x_T^{ref})^t R\delta x_o + \frac{1}{2}(R\delta x_o)^t R\delta x_o$ replacing δx_o by $-\alpha \nabla_o J$

$$\delta J(\alpha) = -\alpha (x_T - x_T^{ref})^t R\nabla_o J + \frac{\alpha^2}{2}(R\nabla_o J)^t R\nabla_o J.$$

Differentiating with respect to α

$$0 = -(x_T - x_T^{ref})^t R\nabla_o J + \alpha (R\nabla_o J)^t R\nabla_o J$$

$$\alpha = \frac{(x_T - x_T^{ref})^t R\nabla_o J}{(R\nabla_o J)^t R\nabla_o J}$$

We then have $\alpha = \dfrac{\sum_{i-1}^{N} \lambda_i^4 e_i^2}{\sum_{i-1}^{N} \lambda_i^6 e_i^2}$ The optimal perturbation along the gradient is

$$\delta x_o = -\frac{\sum_{i-1}^{N} \lambda_i^4 e_i^2}{\sum_{i-1}^{N} \lambda_i^6 e_i^2} \times \begin{pmatrix} \lambda_1^2 e_1 \\ \cdot \\ \cdot \\ \lambda_N^2 e_N \end{pmatrix}.$$

One can see that the weight λ_1^2 given to e_1, the fastest growing component of the error, is larger than the one given to the second eigenvector, λ_2^2, etc ... as $\lambda_1^2 > \lambda_2^2 > \ldots > \lambda_N^2$. This shows that the gradient enhances the unstable components of the initial analysis error because these are the directions along which the forecast error is likely to vary the most. This particular case illustrates the connection between the gradient-related perturbations and the singular vectors (fastest growing modes of the Ensemble Prediction System).

2.3 **Experimentation**

It should be noted that, in the previous mathematical deviation, ***M*** and ***R*** stand for the nonlinear and tangent linear models, with full physics and at high resolution T213L31. Actually, in the experimentation, only a

simplified version M_s of M was used in the computation of the trajectory leading from time 0 to time T = 48 hours. M_s is the nonlinear model at lower resolution T63L31. It was checked that this approximation is reasonable, leading to a few differences in the 48 hr forecasts, mainly over orography. A simplified version R_s of R was also used, at truncation T63L31, and with simplified physics as in *Buizza* (1993). R_s is then not exactly the tangent linear of M_s. The validity of this approximation was investigated in a case discussed further below, for which both the nonlinear evolution of a perturbation δx_o at T106L31, then truncated at T63L31, and its linear evolution at T63L31 according to R_s were produced. δx_o is the perturbation of the initial analysis that was chosen to give the best improvement of the two-day forecast. Its nonlinear evolution and linear evolution after two days can be seen in Fig 1, panels a and b respectively. The agreement between those 2 panels is good, with differences in phase and intensity but it reproduces all the main features. This result gives us confidence in the validity of the approach.

The forecast experimentation consists of a "control" forecast started from operational analysis x_o truncated at T106 and of a "perturbed" forecast started from $x_o - \alpha \nabla J_o$ where ∇J_o is the gradient of the 2-day forecast error over the Northern Hemisphere (computed at T63L31), and α is a scaling factor that was chosen to give a good improvement in the 2-day forecast as above. It is not possible to find the best factor α analytically in the full complexity of J. It could be found by a linear search in the direction of ∇J_o. However, the following values of α have been derived in this study by trial and error. The order of magnitude of this scaling factor is around 1/100. This value seems reasonable if we interpret it in terms of $\frac{1}{\lambda_1^2}$ as in section 2.1. It then corresponds to a mode growing by a factor of 10 in 48 hours. The perturbed forecast will be called the "sensitivity" forecast.

3. INDIVIDUAL CASES

3.1 **Winter case**

Comparing the forecast performance from different initial conditions helped to identify an important analysis problem for 3 January 1994 when the operational ECMWF forecast produced a spurious ridge over western Europe by day-5 (Fig 2b). When the low resolution version of the ECMWF model (horizontal resolution T106 and 19 levels in the vertical) was run from the DWD analysis, the forecast improved the positioning and the intensity of the upper level trough over the western Mediterranean and North-Africa (Fig 2c). Over Ireland and Britain the spurious upper level ridge was replaced by a more correct cyclonic westerly flow as compared to the verifying analysis (Fig 2a).

The differences between the ECMWF and DWD analyses are fairly large in this case (Fig 3). Upper air differences over the Pacific and over the south-eastern parts of North America seem to be most important for the divergent evolution of the two forecast runs from the two analyses. Large forecast differences propagate eastward and reach Europe by day-4. By transplanting only a small area of the Pacific DWD analysis into the ECMWF analysis it could be shown that for the flow configuration at the beginning of January, analysis errors over the Pacific alone affected Europe by day-5 in the forecast and described a large part of the forecast differences shown in the last panel of Fig 3.

The objective scores in terms of anomaly correlation and rms errors (Fig 4) show the dependence of the forecast performance on the initial conditions very clearly. Whereas the operational ECMWF forecast loses skill very rapidly after day-3 (solid line), the forecast from the DWD analysis (dotted line) stays above 60% close to the end of the operational forecast range. A noticeable reduction of errors was already achieved by transplanting part of the Pacific DWD analysis into the ECMWF analysis (dashed line).

In contrast to fairly large differences between the DWD and ECMWF analysis, with maximum values of 60 metres close to the date line, the perturbations based on the sensitivity calculations are almost an order of magnitude smaller, maximum values in the Pacific being 10 metres (Fig 5). However, the unstable character of the sensitivity perturbations becomes evident by their relatively large growth rate. Within two days they have grown by a factor of ten and reached almost the same magnitude as the differences between integrations from the DWD and ECMWF analyses. The effect of the sensitivity perturbations on the flow pattern in the day-5 forecast over Europe is quite remarkable. The trough over the western Mediterranean has now almost the correct intensity and position (Fig 2d). Similar to the forecast from the DWD analysis the flow pattern over most parts of Western Europe is now much closer to the analysis.

The improvement of the forecast run from an analysis perturbed with the sensitivity perturbations is also reflected in substantially improved skill scores. The anomaly correlation stays above 60% up to the end of the forecast period of 10 days (Fig 4 dashed-dotted line).

Both the sensitivity calculations and the forecast experiments run from two different analyses identified in the case of 3 January 1994 the same area in the central Pacific where modifications to the initial conditions caused large changes in the forecast further downstream. The sensitivity calculation gives a strong indication of the three-dimensional structure of the initial error. However, the data coverage over the Pacific is not good enough to verify that the analysis is indeed in error in the manner indicated.

3.2 Spring case

Further cases of poor forecast performance (particularly over Europe) have been investigated by using the sensitivity calculations to modify the initial conditions. The consistent improvement of forecast skill in all cases proves the usefulness of this approach at least as a diagnostic tool to identify possible regions of analysis problems that cause major forecast failures. Of particular interest are cases like the one from 6 April 1994 where a large sensitivity of two day forecast errors to initial conditions was found over central parts of North America where the data density is relatively high. The perturbations derived from the sensitivity calculations produce a much sharper upper air trough to the south-west of the Great Lakes which seems to be supported by data (Fig 6). The largest modifications to the analysis in the region of 40N and 96W improves the fit of the initial conditions to the available observations at 400 hPa in that area. For instance, the standard-deviation of the difference between the analyses and the radiosonde height data at 400 hPa in the region (50N:30N, 110W:80W) displayed in Fig 6 is reduced from 20 to 19.5 m.

The vertical east-west cross section at 40N of the eddy height field shows the baroclinic nature of a deep tropospheric trough (Fig 7a) that is highly sensitive in its further growth to baroclinic perturbations at the initial time. The eastward moving trough tilts backward by around 1500 km from the surface to 250 hPa. For the present OI scheme used in data assimilation a baroclinic system like this is difficult to analyse. Indeed the OI analysis increments (Fig 7b) suggest that the baroclinic structure of this system is mainly produced by the model. In contrast, the perturbations derived from the sensitivity calculations have been able to contribute to a vertically consistent baroclinic change of the trough (Fig 7c). The sensitivity perturbations have a rather small horizontal wavelength of around 1200 km compared to the OI analysis increments and a strong tilt following fairly closely the tilt of the trough in the initialized analysis.

The time evolution of the sensitivity perturbations in terms of stream function (Fig 8), shown as east-west cross sections in 12-hourly intervals, exhibits the typical features of a most unstable baroclinic wave (*Farrell*, 1990). The strong tilt in the vertical is also typical of the fastest growing perturbations of the Ensemble Prediction System. During the first 36 hours the growth is very much restricted to low levels and the westward tilt with height gradually decreases with time. However, with increasing horizontal and vertical scale the vertical tilt disappears and the growth shifts to higher levels. Substantial downstream development produces a large impact of the disturbances (initially located around 60W) close to Europe at 60 hours.

The analysis problem identified by the sensitivity calculations in this case has a major effect on the forecast performance over Europe from day-4 onwards. By day-5 the operational forecast misses an intense cyclogenesis over the Mediterranean and consequently has an upper level ridge over Western Europe which has moved too far to the east (Fig 9). The modification of the analysis from the sensitivity calculations

contributes to a substantial improvement of the upper air flow pattern in the forecast, with the cut-off cyclone in the right position and a correct prediction of the ridge over Ireland and the UK.

Similar to the January case discussed before the forecast scores improve substantially from the modification of the analysis (Fig 10). The sharp decline of the anomaly correlation and the simultaneous increase of rms-errors changes to a more gradual error growth with a gain of predictive skill of more than 3 days at 500 hPa.

3.3 **Floyd case**

Another interesting case regarding the performance of the operational model occurred in September 1993. A disturbance which could be identified as a remainder of the hurricane Floyd developed into an intensive extra-tropical cyclone (middle top panel of Fig 11) and affected the south-west coast of France and South-West England with gales and heavy rain. As in similar cases where a tropical air mass was involved in the development of a mid-latitude cyclone the operational forecast performance was rather poor. The left hand side columns of Fig 11 show the operational forecasts at different time ranges all verifying on 12 September 1993. Even in the short range forecast of two days the intensity of the cyclone had been completely underestimated. All other forecast ranges had a similar problem in underestimating the intensity. The day-3 forecast did not have a low at all close to the west coast of France, whereas the day-4 to day-6 forecasts at least had a cyclonic flow pattern in the region but still not of the right intensity and not at the right place.

The forecast failures with mid-latitude disturbances of tropical origin suggest that errors in the diabatic forcing could be responsible for the underestimation of the cyclone development. Indeed experimentation in the Floyd case with a different convection scheme showed a definite sensitivity to diabatic forcing. However, the improvement of the forecast was not consistent throughout the forecast ranges. The use of an adjustment scheme for convection improved mainly the forecast at medium but not at short range.

For this case the gradient of two-day forecast errors has been used to modify the initial conditions for five different cases ranging from two to six days before the development of the strong cyclone occurs. The integrations from the modified analyses produce a deep cyclone for all forecast ranges from day-2 to day-6 (right hand side panels of Fig 11). Though the location and the intensity could still be improved, a forecast series like this would have given the forecaster a clear basis for issuing the appropriate warnings of strong winds in the coastal areas of France and the South-West of England.

In these cases it is clear that there is a marked sensitivity to initial conditions. There is also a sensitivity to the choice of the convection scheme. Until we quantify the impact of the convection scheme on the important unstable structures in the initial error field it is not possible to say whether the analysis or the physics has a greater impact on forecast error.

4. STATISTICS OVER APRIL 1994

Each day over the entire month of April 1994, four sets of T106 L31 forecasts were produced. The "control" forecast and a set of three perturbed forecasts started from $x_o - \alpha_i \nabla J_o$ where ∇J_o, and α_i is a scaling factor taken as $\alpha = 1/50$, $\alpha_2 = 1/100$, $\alpha_3 = 1/200$. Each day, the perturbed forecast giving the best improvement in the 2-day forecast is chosen (sensitivity forecast). In most cases, it corresponds to the scaling factor $\alpha = 1/100$. For 30 days, statistics are computed from the two sets of forecasts: control and sensitivity with the optimum scaling factor. In particular, we are interested in the shape of the sensitivity patterns shown by the initial perturbations. The evolution in time of the differences between control and sensitivity forecasts is also investigated.

4.1 **Time evolution of sensitivity perturbations**

The geographical distribution of the sensitivity patterns for the basic variables temperature and vorticity are displayed in Fig 12, which shows the rms at model level 22 corresponding approximately to a pressure level of 700 hPa. As expected, the sensitive areas are the mid-latitudes from 40 to 70 N, with maximum sensitivity over the Pacific Ocean and East Canada. It should be noted that the exact location of the maxima of sensitivity and their relative amplitudes depends on the choice of norm. For instance if, instead of taking the L^2 norm $\frac{1}{N}\sum_{i=1}^{N} x_i^2$ as in the rms computation, the L^1 norm $\frac{1}{N}\sum_{i=1}^{N} |x_i|$ is used, the impact of a few outliers is lessened. Comparing Fig 12, panel a with Fig 13 which represents the same sensitivity fields combined with the L^1 norm, one can notice a change in the sensitivity pattern over the Eastern part of North America for instance. The absolute maximum close to the Hudson Bay in Fig 12a, which corresponds to a couple of extreme sensitivity values, does not appear as the absolute maximum any longer in Fig 13 in which the Nova Scotia area is at least as sensitive. But the gross patterns are preserved. The two main sensitive areas, Pacific Ocean and North-East of the American continent are well-known regions of cyclogenesis in the Northern Hemisphere. They also correspond to regions identified for significant singular vector growth in *Buizza and Palmer* (1994).

As far as the vertical distribution of the sensitivity is concerned, cross-sections of rms differences between the control and sensitivity runs, averaged over the latitudes 40 to 65 N, are presented in Figs 14 and 15 for the basic variables temperature and vorticity. Panel a represents the initial perturbations, and the subsequent panels b to d their time evolution from day 1 to day 5. The initial perturbations exhibit maximum values around 400 to 700 hPa for temperature and 500 to 850 hPa for vorticity - these perturbations are then mainly in the mid-low troposphere. They are tilted in the vertical similar to baroclinic instability-type structures, although this is more noticeable on the spring case presented in section 3.2 than on these averaged cross-sections. The maximum of sensitivity in the mid-low troposphere was already found in *Rabier et al* (1993)

over the month of January 1993. For an idealized baroclinic instability problem (*Rabier et al*, 1992) the cyclogenesis was also found to be the most sensitive to initial values around 700 hPa which corresponded to the steering level. This fits well with the study of *Farrell* (1989) where the author relates the optimal excitation of unstable Charney modes to waves concentrated near this significant level. It is also consistent with the structure of the first singular vectors (*Buizza and Palmer*, 1994). When looking at the time evolution of these perturbations for the temperature field in Fig 14, one sees that the mid-troposphere structures initially located between 400 hPa and 700 hPa rapidly evolve to give maxima at two levels around 250 hPa and 500 hPa. As for the vorticity perturbations (Fig 15), the region of largest amplitude moves upwards rapidly in the first 12 to 24 hours from 850 hPa to 300-400 hPa. One can notice that the perturbations located over North-East America are growing much faster than any other local perturbations during the first two days, leading to large differences from the control run.

4.2 Spectral distribution of sensitivity perturbations compared with forecast errors

Fig 16 shows averaged spectra of perturbations in geopotential height at a few selected levels. These are log-log figures for which the abscissa represents log(n) with n the horizontal wavenumber from n=1 to n=63. The bottom set of curves in Fig 16, panel a, represents the initial perturbations, the middle one represents the perturbations after evolution up to 48 hours, and the top one is the 48 hour forecast error (2-day forecast minus analysis). In Fig 16, panel b, the bottom set of curves is the perturbation spectra after 5 days and the top one is the 5-day forecast error.

The initial perturbations exhibit rather flat spectra, with quite a lot of energy in small scales. The 1000 hPa level is even more heavily weighted towards the small scales than the other levels. At 250 hPa, the peak is somewhere around x=1.3 which corresponds to n=20. After 48 hours of evolution, the perturbations can be compared to the "true" forecast error. The overall shape of the spectra is rather similar, with a peak around x=1.0 (n=10). However, at large scales (n<10), the amplitude of the sensitivity perturbations is much less than the actual error, by a factor of **1/50** at n=1 for example. The peak at n=2 is also missing. For synoptic and small scales (n≥10), the amplitude of the perturbations is about a tenth of the amplitude of the forecast error. For synoptic scales (n around 10 to 20), both sets show higher amplitudes for higher atmospheric level, whereas for small scales (n greater than 30), the near surface represents a big proportion of total amplitude in both cases. After five days, the same comparison applies for Fig 16, panel b. For both sets of curves, the amplitudes have been growing a lot around n=10, whereas small scales and, to a lesser extent, large scales tend to saturate. This saturation is particularly true for the actual forecast error (top curve), which leads to a better agreement between perturbations and errors at those scales than after 48 hours. There is now a factor of **1/10** at n=1, and **1/5** at n=63 compared with **1/50** and **1/10** before. This spectral analysis shows that although the initial perturbations are small-scale patterns, they evolve in such a way as to reduce mainly the synoptic part of the forecast error from day two onwards, up to about 20 percent.

4.3 Improvements in forecast quality from the sensitivity perturbations

One can actually investigate which parts of the 2-day and 5-day errors have been reduced from a geographical point of view in Fig 17, representing the initial difference between the sensitivity and the control initial conditions and the subsequent difference between the rms errors of the sensitivity runs and the rms errors of the control runs at different ranges for the geopotential height at 500 hPa. The picture is clear: the sensitivity run is better than the control one all over the Northern Hemisphere with a few spots of large improvement. By day 5, two main large areas seem to benefit from the introduction of the initial perturbation. These are Canada and Europe. Very small initial perturbations (maximum of 3 m) lead to considerable improvement in the short range (more than 30 m locally at day 2) and in the medium range (broad areas for which the forecast error has been reduced by more than 20 m at day 5).

This method does not tell the whole story of the causes of forecast error. This can be investigated by comparing Fig 18 representing the two-day forecast error of the control run and Figs 14 and 15 panels c showing the evolution of the perturbation after two days. The sensitivity pattern has evolved in a manner which leads to changes in the 200 hPa-700 hPa layer, but misses completely the large forecast errors located close to the ground or at the very top of the model. These are not coming mainly from analysis errors being amplified by dynamical instabilities, but almost certainly from problems in physical parametrizations or lack of stratospheric resolution and are therefore impossible to describe with this method, which describes the sensitivity to initial conditions only and not to model formulation.

Operational performance in April 1993 showed typical day-to-day variations of forecast skill. The top panel of Fig 19 shows the anomaly correlation for the T106 control experiments (dashed line) and the sensitivity experiments (solid line) for Europe. Fairly good operational forecasts at the beginning of the month are followed by a complete failure on the 6th when the anomaly correlation drops to negative values. In the discussion of single cases it has already been shown that this day was very sensitive to initial conditions and the sensitivity run improved the forecast skill over Europe from -25% to 85%. But not only cases with low forecast skill have been improved, in fact most forecasts from the modified analysis show an increase in forecast skill. An important feature of the sensitivity experiments is that their forecast skill is much more consistent throughout the month, suggesting that day-to-day forecast inconsistencies are closely related to analysis problems. A similar improved performance can be seen over the Northern Hemisphere (Fig 19 bottom panel), where the control forecasts show three major periods of deterioration in skill with a drop of anomaly correlation close to 60% whereas the sensitivity experiments follow a smoother line with anomaly correlations staying above 73%.

The average error development and the evolution of the initial perturbations for April is shown in Fig 20 for Europe in the top panel and for the Northern Hemisphere in the bottom panel. The rms-difference between

the control runs and the sensitivity runs (solid line) gives an indication of the growth of the initial perturbations with forecast time. In contrast to an initial linear growth of both the control forecast errors (dashed line) and errors of the sensitivity runs (dotted line), the perturbations show an exponential growth up to day 2 to 3, a clear sign that mainly the most unstable modes contributing to day-2 forecast errors have contributed to the calculations of the initial perturbations. After day-1 the perturbations have reached large enough values to produce an increasing reduction of rms-errors with forecast time. In particular the nonlinear increase of rms-errors in the control forecasts between day-2 and day-6 has been reduced in the sensitivity runs. The fact that the perturbations have been derived to reduce errors that grow fastest within the first 2 days can also be seen from the Northern Hemisphere scores. By modifying the analysis, the error growth has been reduced mainly between day-1 and day-3, whereas after that time the two rms-error graphs are almost parallel.

5. CONCLUDING REMARKS

In this study, the adjoint method has been used to provide the sensitivity of short-range forecast errors to the initial data. The gradient of the day-two forecast error with respect to the initial conditions identifies which small changes could be made to the analysis to improve most effectively the short-range forecast. If we assume that the resolvent of the tangent linear model is a real diagonal matrix, this gradient is equal to $(\lambda_i^2 e_i)_{1 \le i \le N}$ where the $(e_i)_{1 \le i \le N}$ are the components of the initial error growing respectively by factors of $(\lambda_i)_{1 \le i \le N}$ during the two-day period. It thus identifies the fast growing components of the analysis error. When scaled by $\alpha \sim \frac{1}{\lambda_{max}^2}$ and subtracted from the analysis, it provides a "sensitivity run" which can be used as a diagnostic tool.

In cases when the analysis error is dominated by a limited number of unstable modes, the forecast errors are likely to be substantially reduced in the sensitivity run. In this case, the sensitivity pattern and the first few singular vectors of the ECMWF Ensemble Prediction Scheme, if they were computed at truncation T63, should compare well.

Statistics over the month of April 1994 have characterised the sensitivity patterns as small-scale, mid-low tropospheric structures tilted in the vertical. Because of this predominance of small scales, it seems that at least T42 is necessary to get the patterns reasonably close to the operational T63 structures for the sensitivity calculations. The general pattern of these structures is known to be associated with the fastest possible growth of forecast error. The amplitude of the initial perturbations is kept well within the error bars of the analysis. When inserted as initial perturbations, they evolve rapidly into synoptic-scales structures, propagating into higher atmospheric levels.

On average, the sensitivity run corrects for about a tenth of the day two forecast error, which indicates that indeed not all of the error is in the fastest amplifying modes. But the fraction of the error corrected at day 2 is important for an improvement in the medium-range, as it continues to grow substantially in the nonlinear regime. The region which seems to benefit most from these sensitivity runs is Europe. For the April period, changes over the North-East of the American continent and the North-West of the Atlantic Ocean are particularly beneficial to the quality of the forecast over the European region. On the other hand, in a winter case, the Pacific was found to be the most sensitive area for medium range forecast errors over Europe. The fact that the modification of the initial conditions showed a much reduced day-to-day variability of forecast skill suggests that one of the major problems of operational forecasting, the inconsistency of consecutive forecasts, is to a large extent due to analysis errors.

These results have proven that there is still scope for great improvement in the forecast in the medium range, particularly over Europe, by better describing the initial conditions. The sensitivity experimentation suggests that many cases of major forecast errors may be explained by weaknesses in the analysis. A small change in the analysis is enough to improve the forecast quality. However, to analyse correctly the key structures which lead to serious forecast errors, the assimilation and observing systems must be able to deal with strongly tilted small scale (both on the horizontal and the vertical) structures. Advanced assimilation methods such as 4D-Var and Kalman filter are expected to be required for this purpose.

Future work will concentrate on developing tools for an objective verification of the modified initial conditions against observations. This will help us to diagnose cases of analysis shortcomings. The improvement of the sensitivity calculations currently under development is to evaluate objectively the optimal scaling factor for the perturbations by using a minimization algorithm. In a few cases completed so far the optimal factor proved to be consistent with the subjectively chosen factor in the April experimentation.

ACKNOWLEDGEMENTS

Werner Wergen, of Deutscher Wetterdienst, kindly provided the initialised analysis for 3 January 1994. Martin Miller is acknowledged for his contribution to the Floyd case. The authors would like to thank Tim Palmer, Adrian Simmons, Dave Burridge, Jean-Noël Thépaut and Per Undén for their useful comments. Bernard Strauss and Anders Persson are acknowledged for their enthusiastic support.

REFERENCES

Arpe, K, A Hollingsworth, M S Tracton, A C Lorenc, S Uppala and P Kållberg, 1985: The response of numerical weather prediction systems to FGGE level IIb data. Part II: Forecast verifications and implications for predictability. Q J R Meteorol Soc, 111, 67-101.

Buizza, R, 1993: Impact of a simple vertical diffusion scheme and of the optimization time interval on optimal unstable structures. ECMWF Tech Memo No. 192, February 1993.

Buizza, R, and T N Palmer, 1994: The Singular-Vector structure of the Atmospheric General Circulation. Submitted to J Atmos Sci.

Courtier, P, 1987: Application du contrôle optimal à la prévision numérique en météorologie. Thèse de doctorat de l'université de Paris VI.

Courtier, P and O Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation- Part II. Numerical results. Q J R Meteorol Soc, 113, 1329-1347.

Courtier, P, J Derber, R Errico, J-F Louis and T Vukićević, 1993: Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology. Tellus, 45A, 342-357.

Errico, R M and T Vukicevic, 1992: Sensitivity analysis using an adjoint of the PSU/NCAR mesoscale model. Mon Wea Rev, 120, 1644-1660.

Farrell, B F, 1989: Optimal excitation of baroclinic waves. J Atmos Sci, 46, 1193-1206.

Farrell, B F, 1990: Small error dynamics and the predicatbility of atmospheric flows. J Atmos Sci, 47, 2409-2416.

Hall, M C G, 1986: Application of adjoint sensitivity theory to an atmospheric general circulation model. J Atmos Sci, 43, 2644-2651.

Hall, M C G, D G Cacuci and M R Schlesinger, 1982: Sensitivity analysis of a radiative-convective model by the adjoint method. J Atmos Sci, 39, 2038-2050.

Hollingsworth, A, A C Lorenc, M S Tracton, K Arpe, G Cats, S Uppala and P Kållberg 1985: The response of numerical weather prediction systems to FGGE level IIb data. Part I: Analyses. Q J R Meteorol Soc, 111, 1-66.

Lacarra, J-F and O Talagrand, 1988: Short range evolution of small perturbations in a barotropic model. Tellus, 40A, 81-95.

Le Dimet, F-X and O Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations. Tellus, 38A, 97-110.

Lewis, J M and J C Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. Tellus, 37A, 309-322.

Marais, C and L Musson-Genon, 1992: Forecasting the surface weather elements with a local dynamical-adaptation method using a variational technique. Mon Wea Rev, 120, 1035-1049.

Molteni, F, R Buizza, T N Palmer and T Petroliagis, 1994: The ECMWF ensemble prediction system: methodology and validation. Submitted to Q J R Meteorol Soc.

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

Rabier, F, P Courtier and O Talagrand, 1992: An application of adjoint models to sensitivity analysis. Beitr Phys Atmosph, 65, 177-192.

Rabier, F, P Courtier, J Pailleux, O Talagrand and D Vasiljevic, 1993a: A comparison between four-dimensional variational assimilation and simplified sequential assimilation relying on three-dimensional variational analysis. Q J R Meteorol Soc, 119, 845-880.

Rabier, F, P Courtier, M Hervéou, B Strauss and A Persson, 1993b: Sensitivity of forecast error to initial conditions using the adjoint model. Technical Memorandum No 197, October 1993.

Rinne, J and H Järvinnen, 1993: Estimation of the Cressman term for a barotropic model through optimisation with use of the adjoint model. Mon Wea Rev, Vol 121, No 3, 826-833.

Talagrand, O and P Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation- Part I Theory. Q J R Meteorol Soc, 113, 1311-1328.

Thépaut, J-N, D Vasiljevic, P Courtier and J Pailleux, 1993: Variational assimilation of conventional observations with a multilevel primitive equation model. Q J R Meteorol Soc, 119, 153-186.

Vukicevic, T, 1991: Nonlinear and linear evolution of initial forecast errors. Mon Wea Rev, 119, 1603-1611.

Zou, X, I M Navon and F-X Le Dimet, 1992: Incomplete observations and control of gravity waves in variational data assimilation. Tellus, 44A, 273-296.

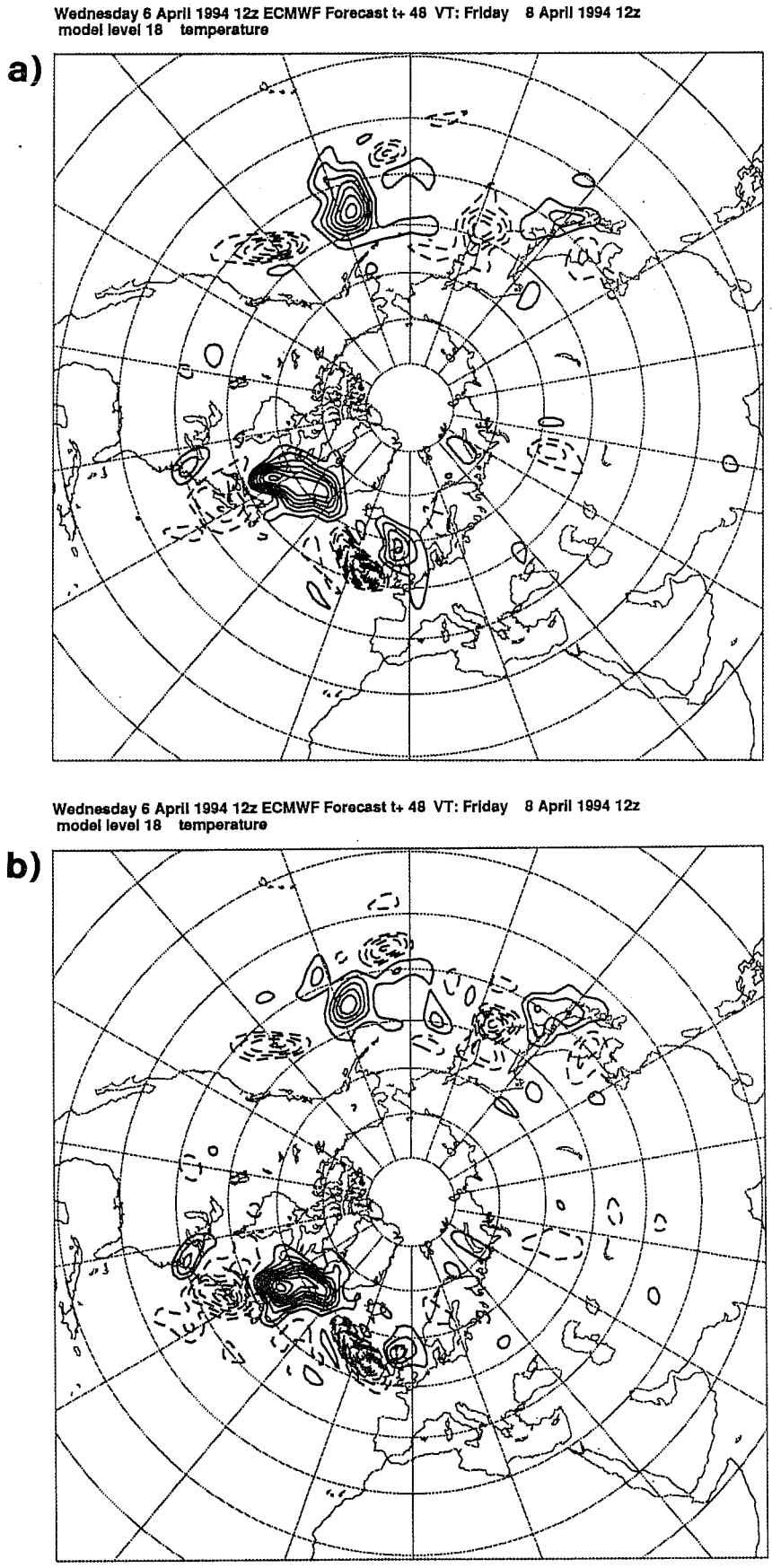


Fig 1 Evolution of a temperature perturbation δx_o in the analysis of 6/4/94 up to 48 hours at model level 18 (~ 500 hPa), using (a) the nonlinear T106L31 model with full physics and (b) the simplified tangent linear model at T63L31. In (a) the field is truncated at T63 for comparison with (b). Isolines are every 1 K. Dashed contours are used for negative values.

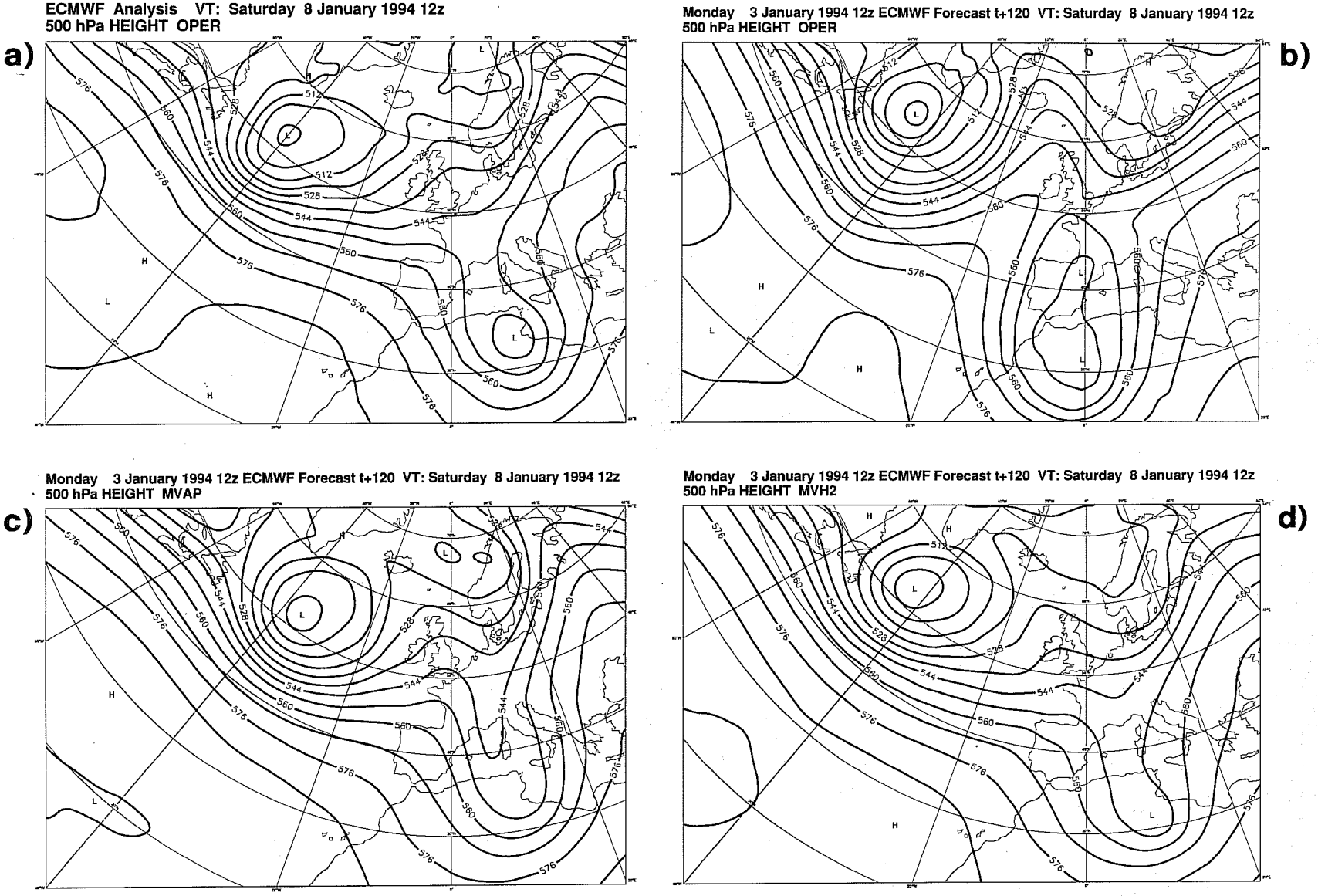


Fig 2 Maps of geopotential height at 500 hPa for 3 January 1994, (a) analysis, (b) operational forecast, (c) forecast from DWD analysis, (d) sensitivity forecast. Units: dam.

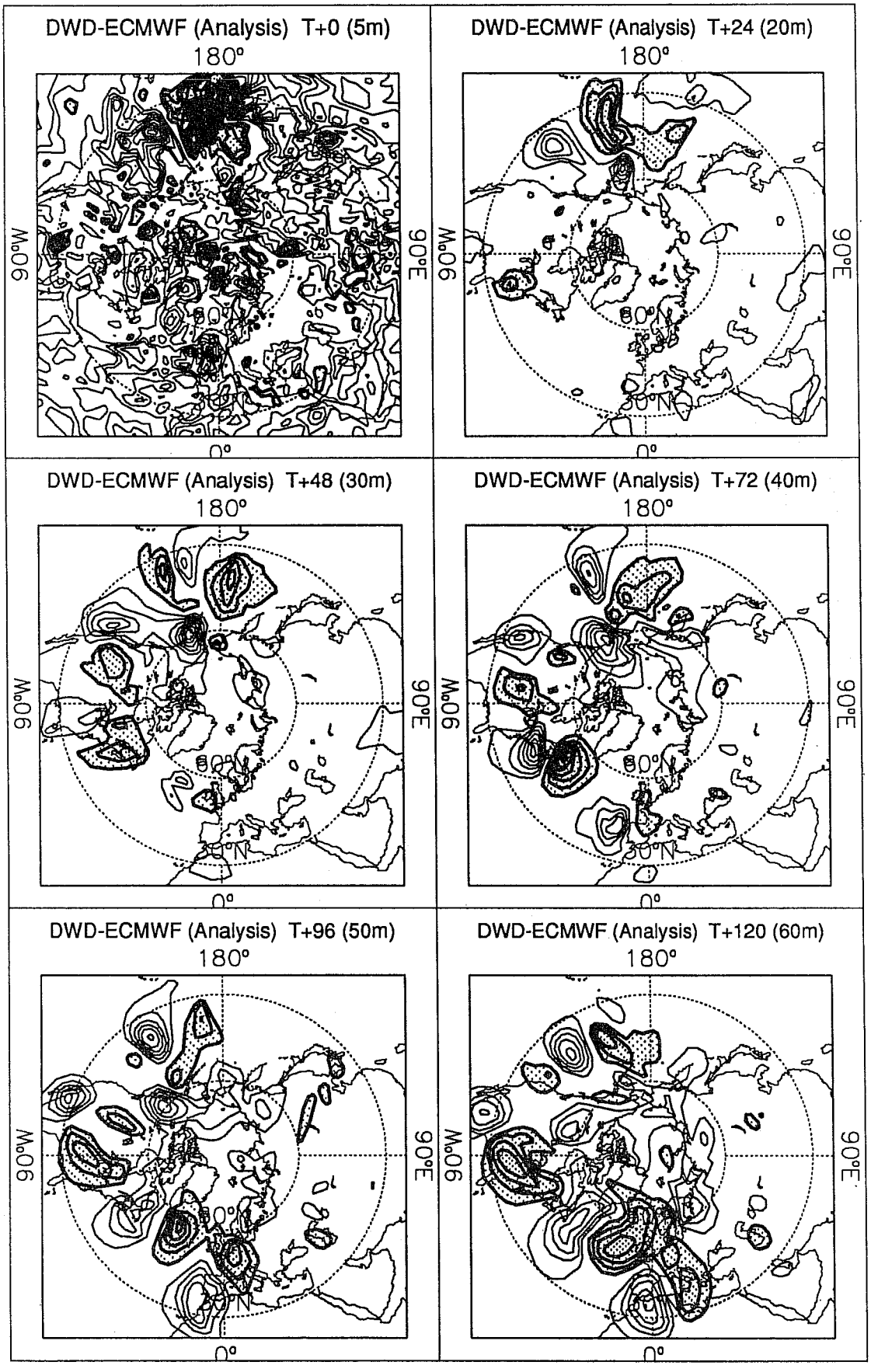


Fig 3 Height difference between the forecasts form the DWD analysis and the ECMWF analysis for forecast steps 0,24,,,,120 hours for 3 January 1994. Units: m.

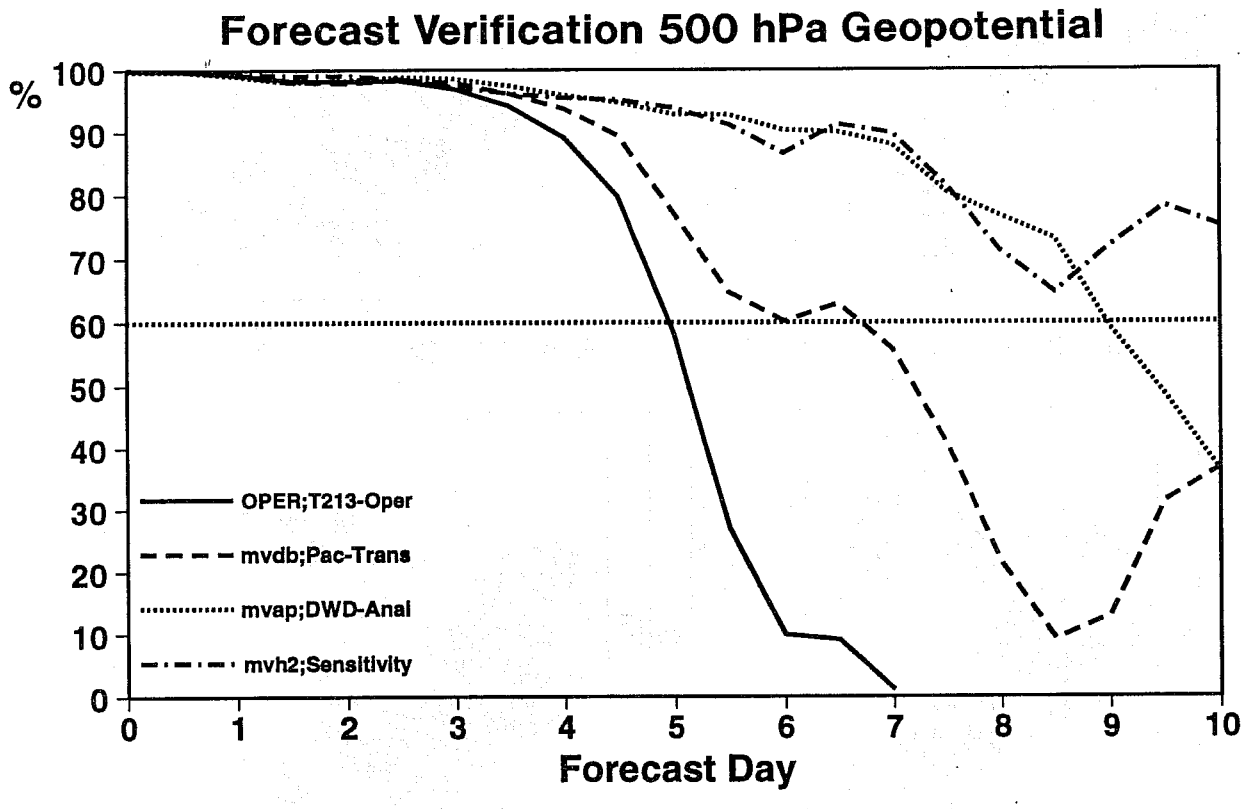


Fig 4 Forecast scores for 3 January 1994 (anomaly correlation). Scores are shown for the operational forecast (solid line), for the forecast from the DWD-analysis (dotted line), for the forecast from a combination of the ECMWF analysis with a DWD analysis transplant in the Central Pacific (dashed line), and for the sensitivity forecast (dashed-dotted line).

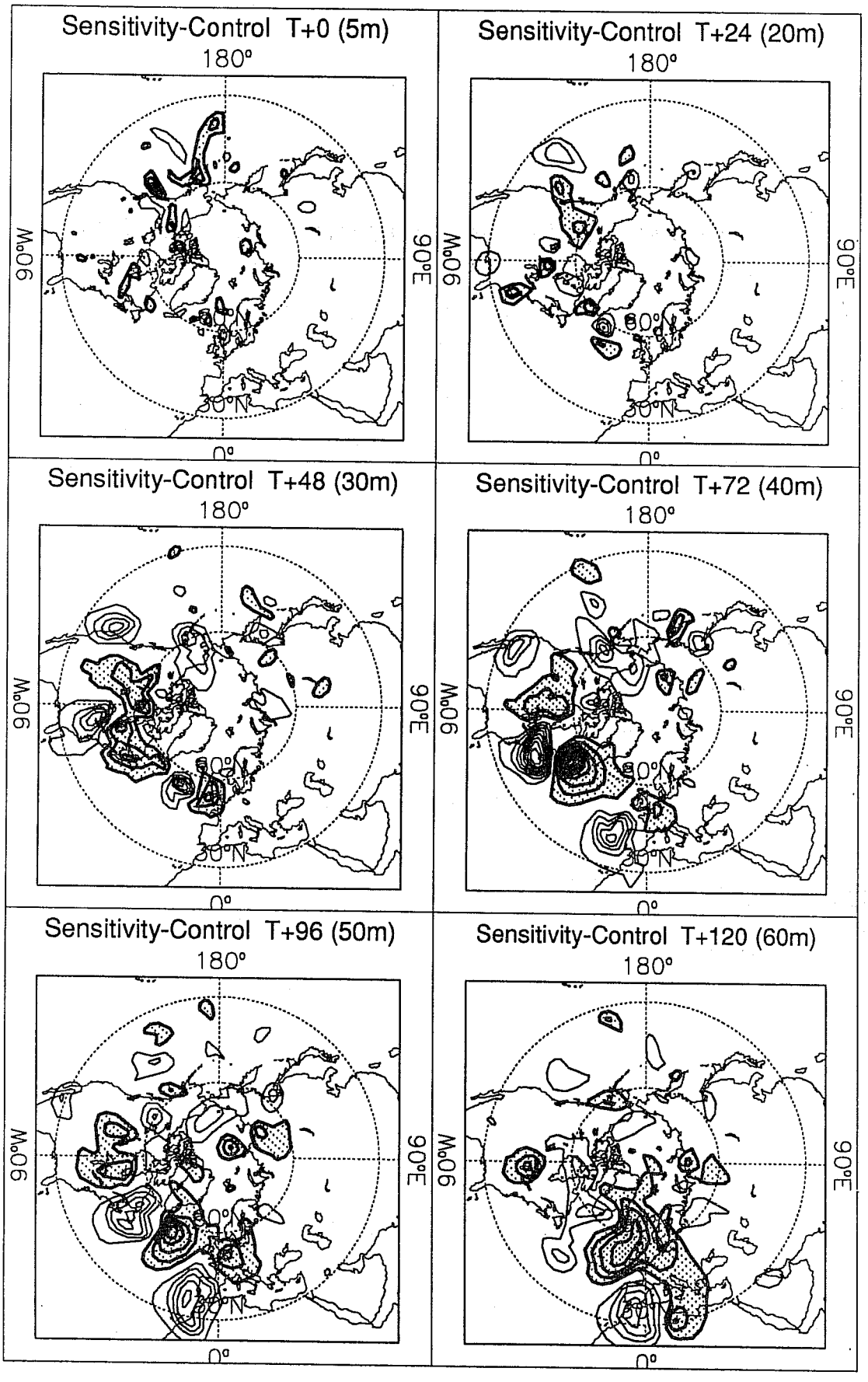


Fig 5 Height difference between the sensitivity and the control forecast for steps 0,24,,,,120 hours for 3 January 1994. Units: m.

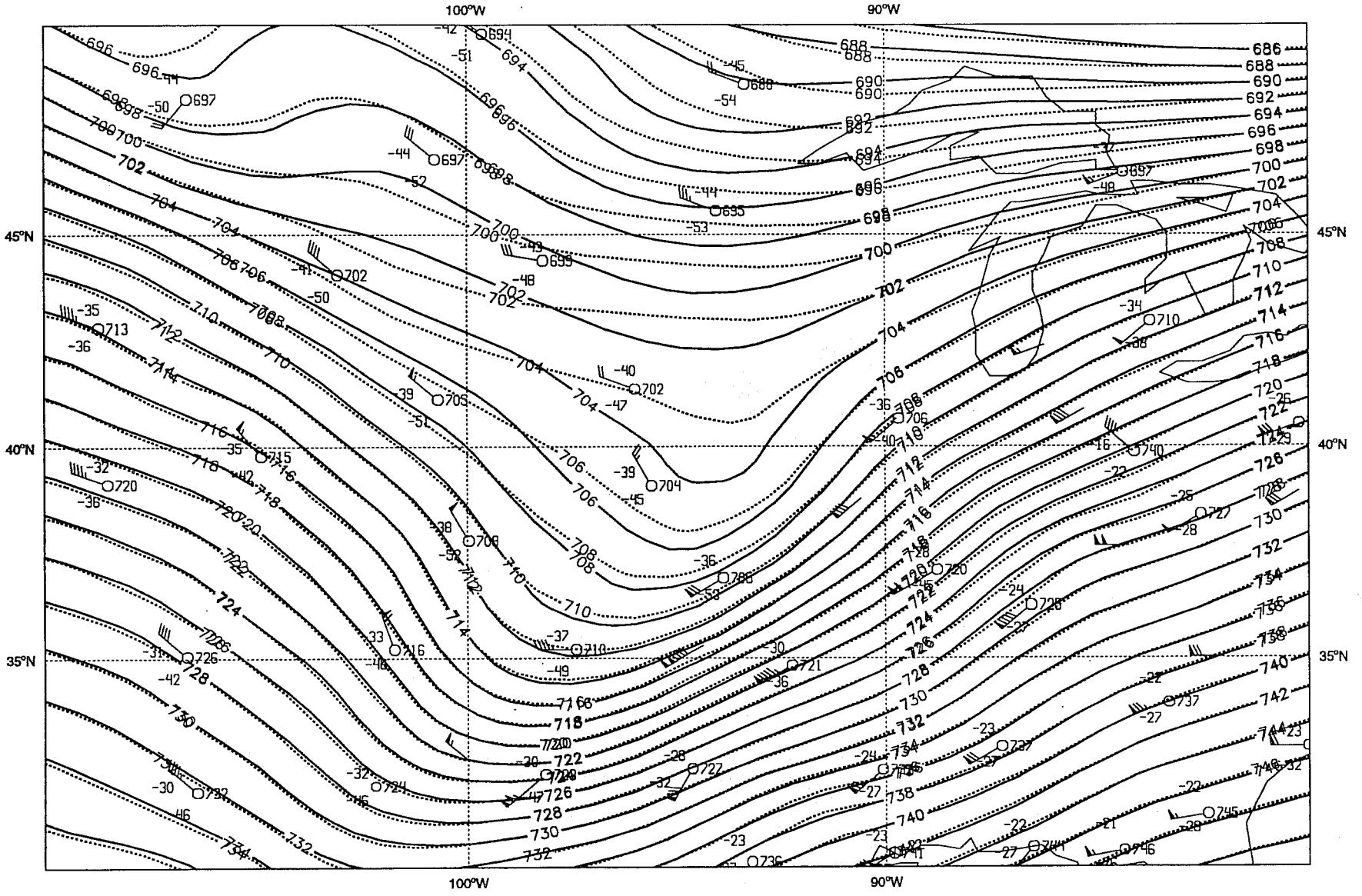


Fig 6 Operational initialized analysis (dotted contours), modified initial condition proportional to the gradient of the day-2 forecast error (solid contours) and observations at 400 hPa for 6 April 1994.

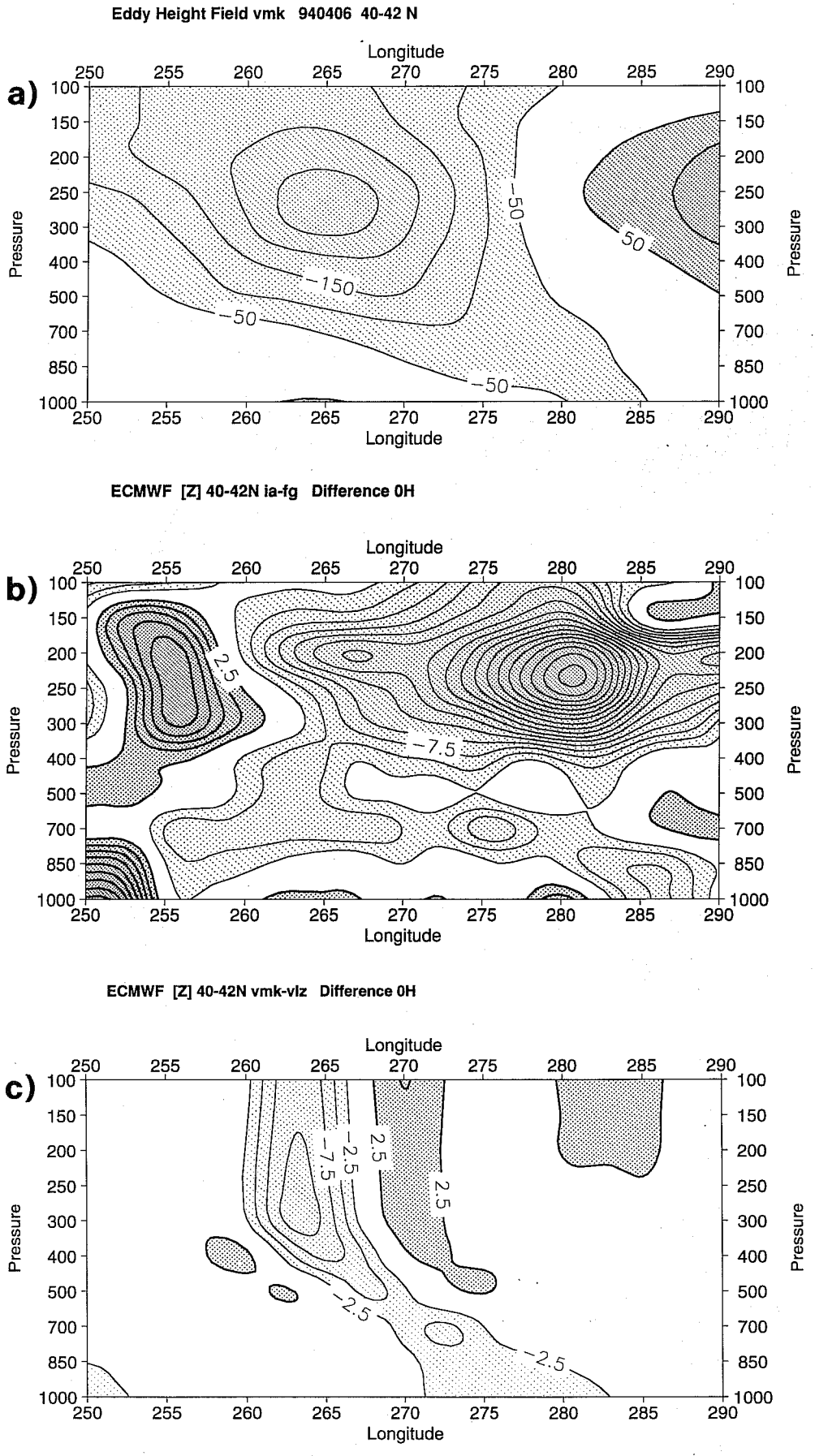


Fig 7 Vertical west-east cross-section of the geopotential height at 41 degrees north for (a) departure of analyzed values from the zonal mean, (b) analysis increments, (c) sensitivity perturbations for 6 April 1994. Units: m.

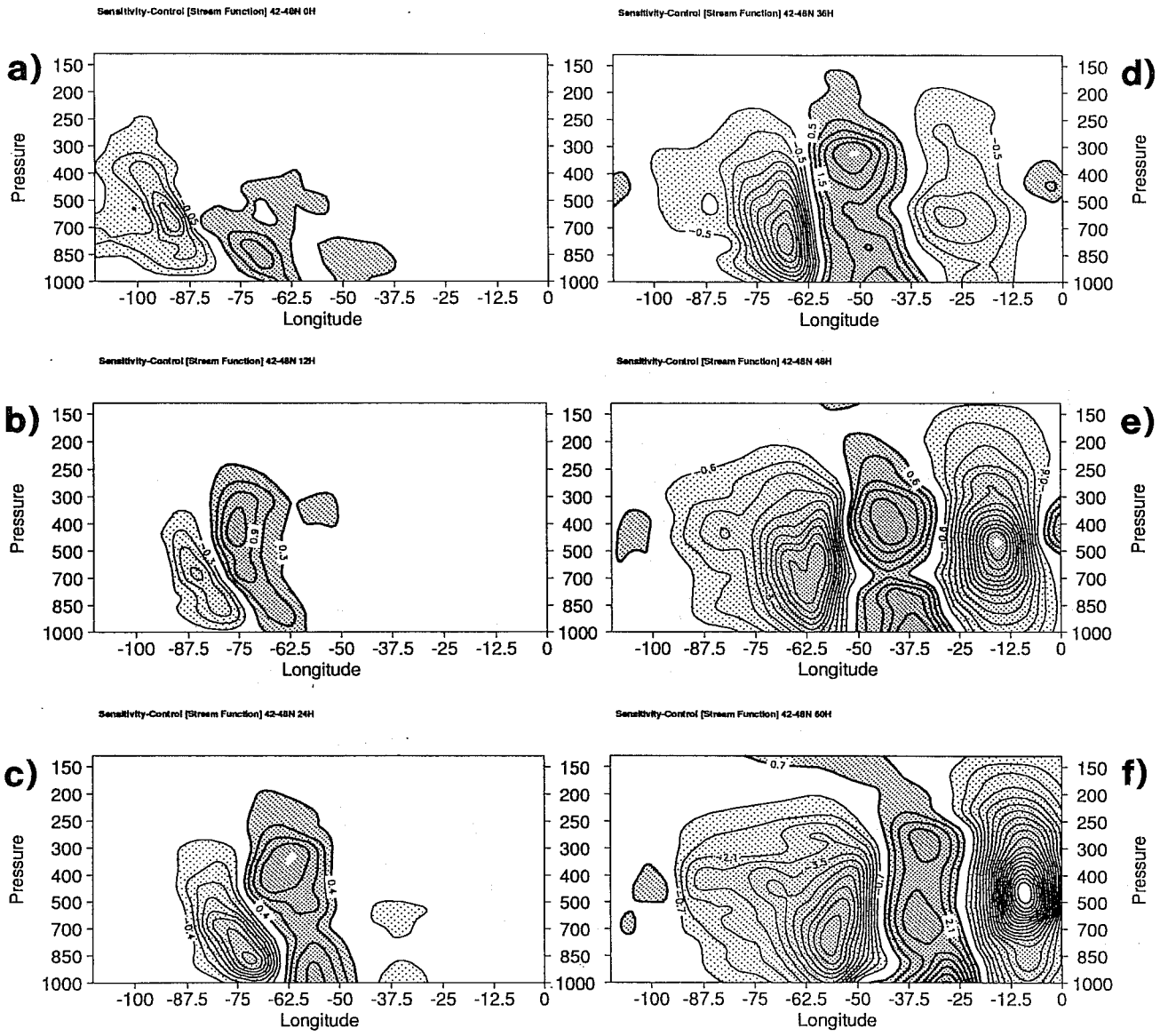
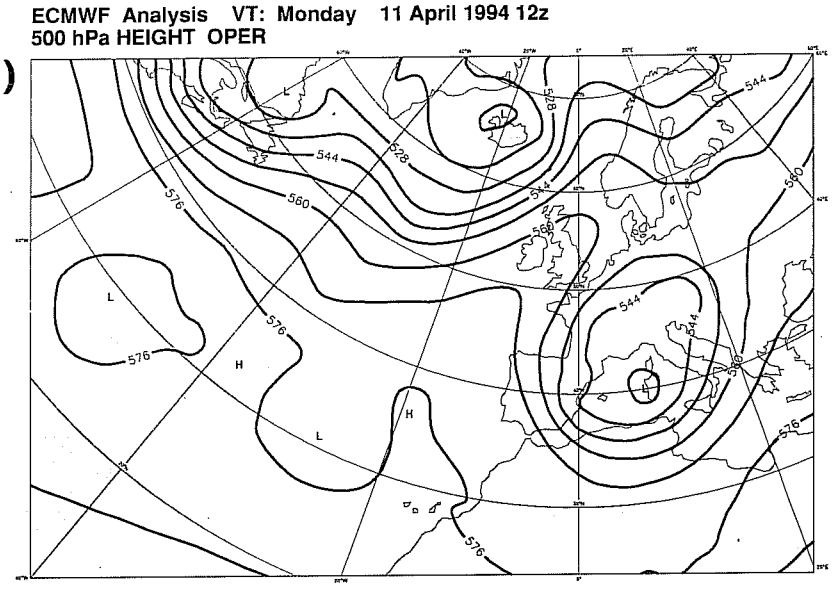
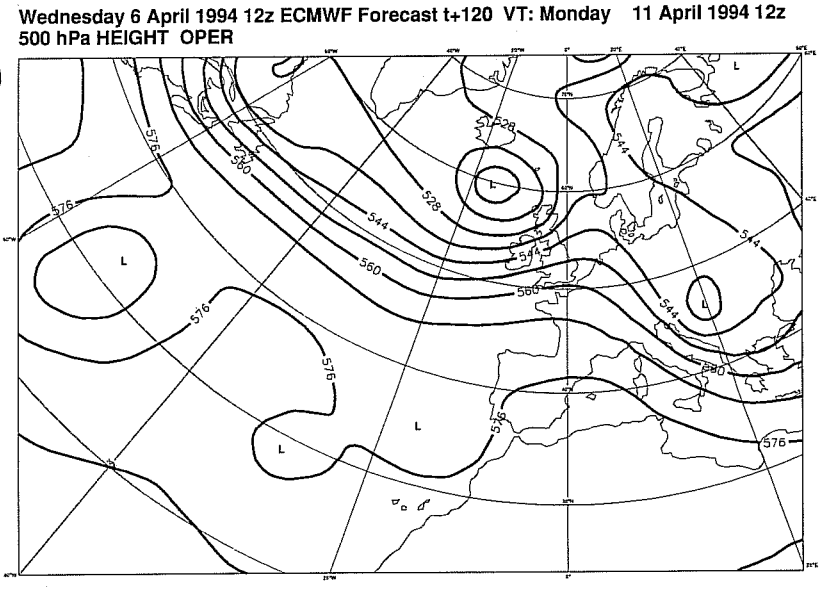


Fig 8 Vertical west-east cross-section of the sensitivity perturbation as time evolution following the latitude of the maximum amplitude for steps; 0,12,24,36,48,60. The parameter shown is the stream function perturbation for the forecast from 6 April 1994. Units: $10^6 m^2 s^{-1}$.

a)



b)



c)

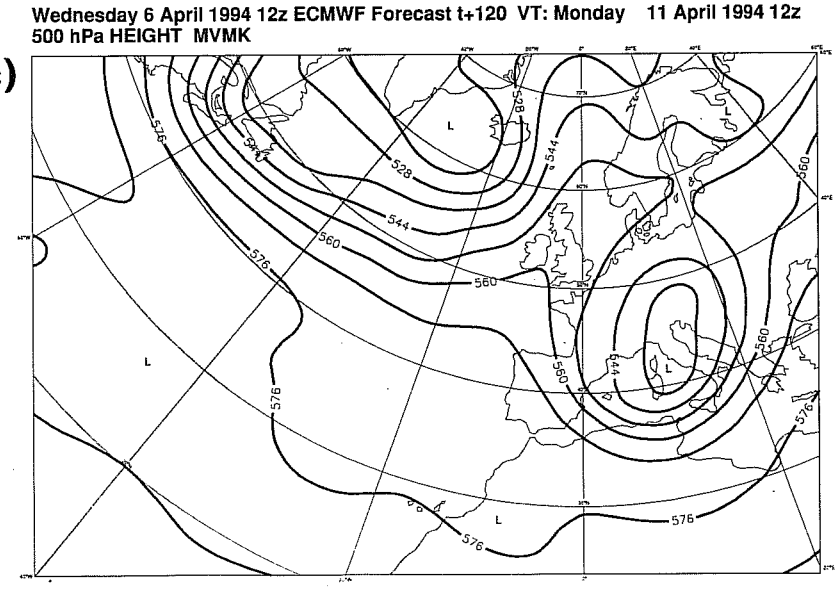


Fig 9 Maps of geopotential height at 500 hPa for 11 April 1994, (a) analysis, (b) operational forecast, (d) sensitivity forecast. Forecast range 120 hours. Units: dam.

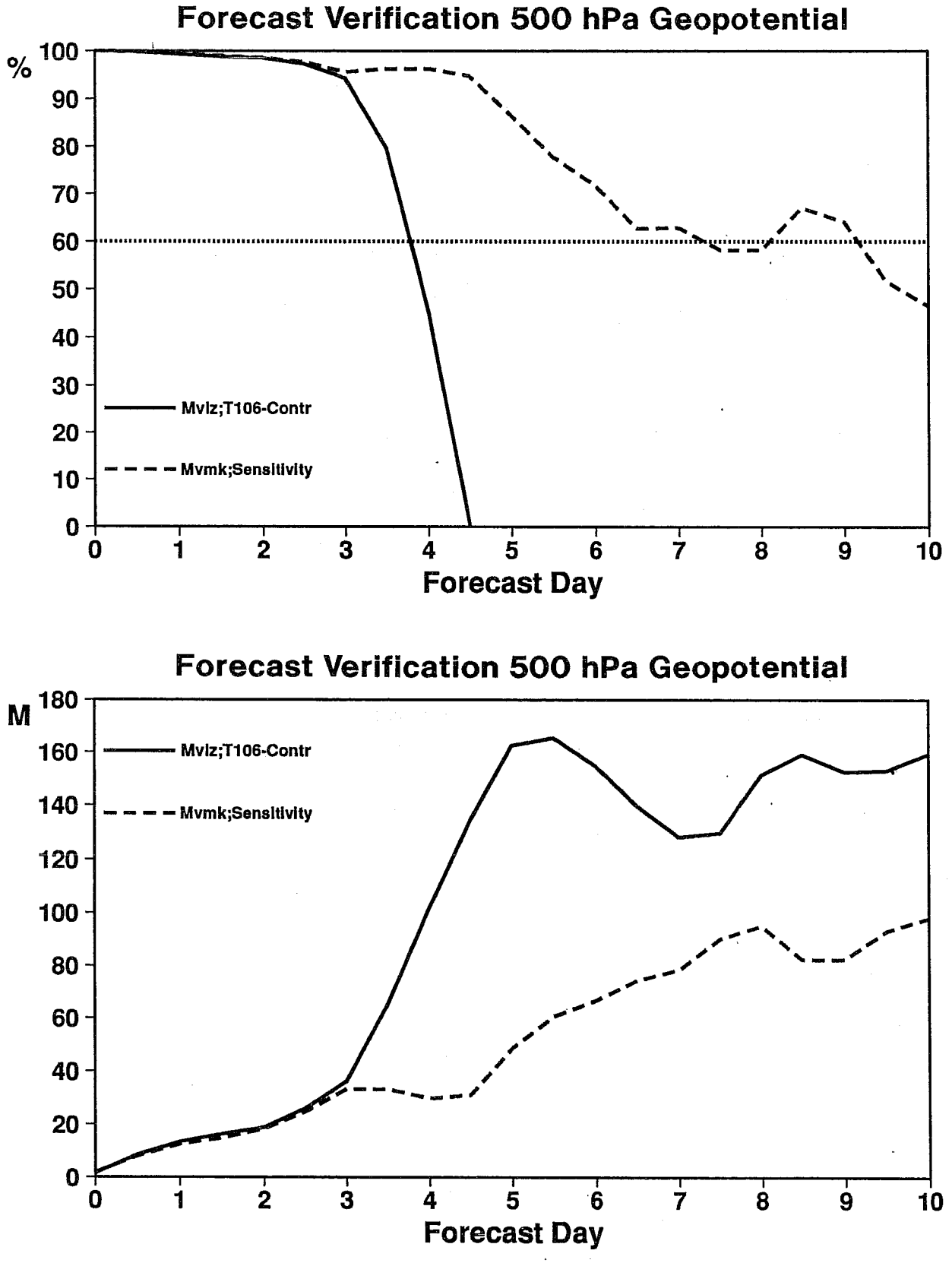


Fig 10 Forecast scores for 6 April 1994, upper panel anomaly correlation, lower panel rms-errors. Scores are shown for the control forecast (solid line) and the sensitivity forecast (dashed line)

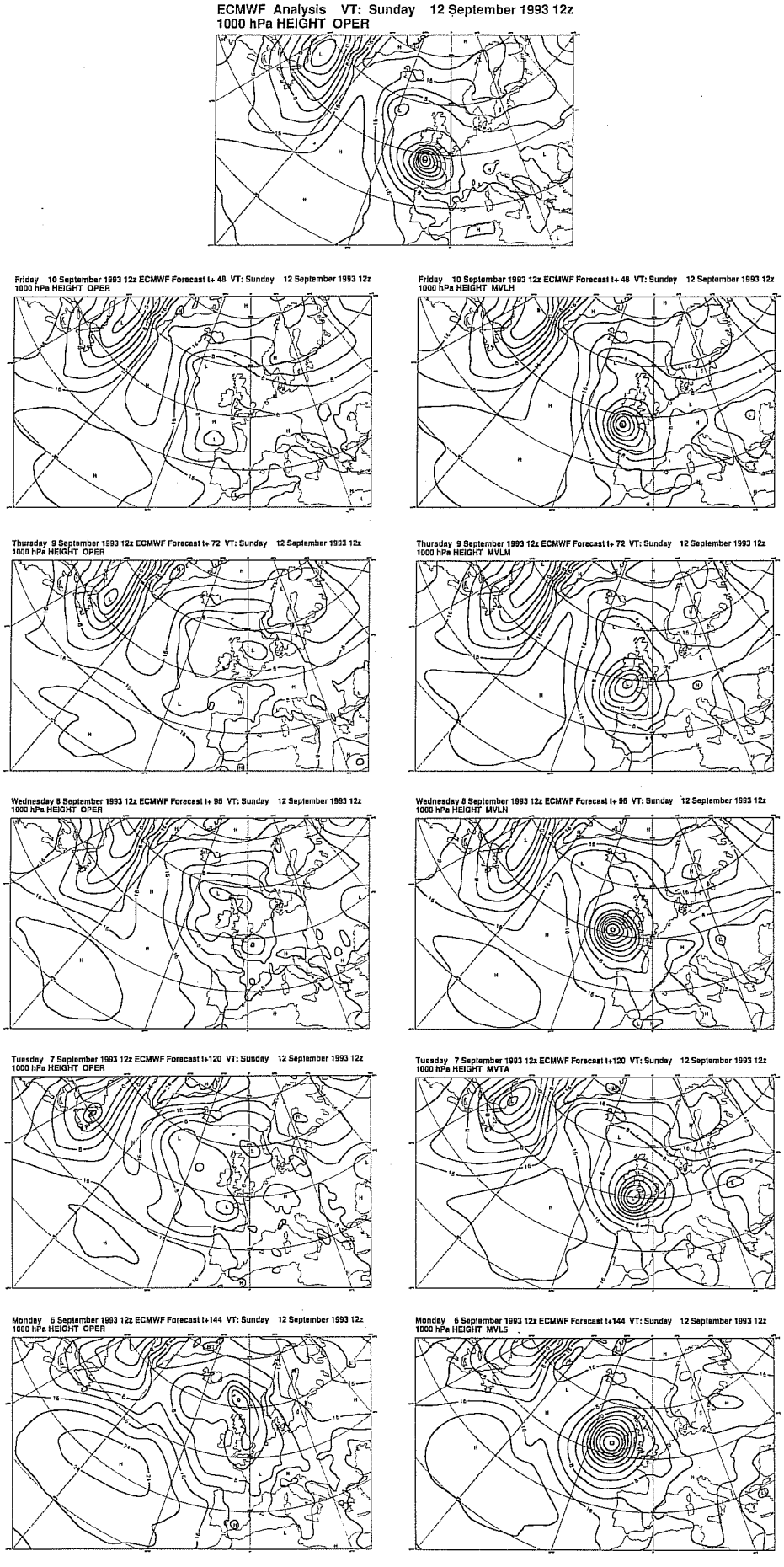


Fig 11 Maps of geopotential height at 500 hPa for 12 September 1993, analysis (top panel), operational forecast (left column), sensitivity forecast (right column). Forecast ranges 48,72,96,120,144 hours. Units: dam.

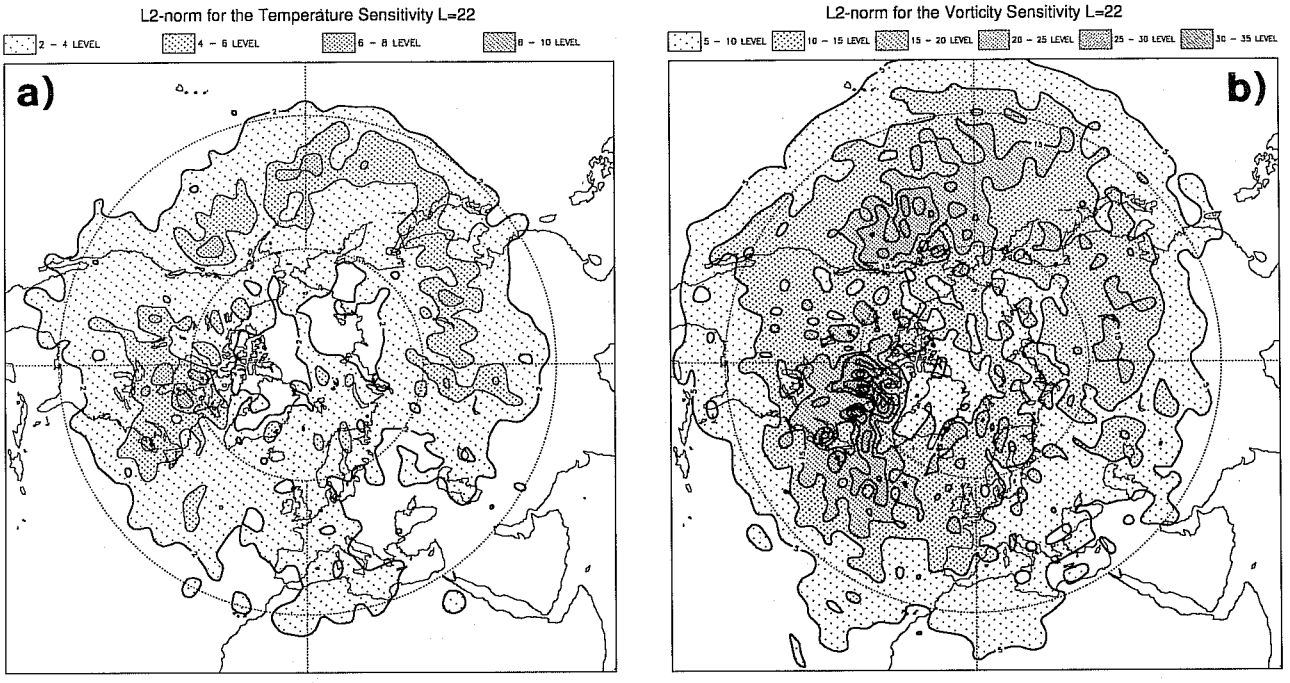


Fig 12 Root mean squares of sensitivity patterns in April 1994 at model level 22 (~ 700 hPa) for temperature (a) and vorticity (b).

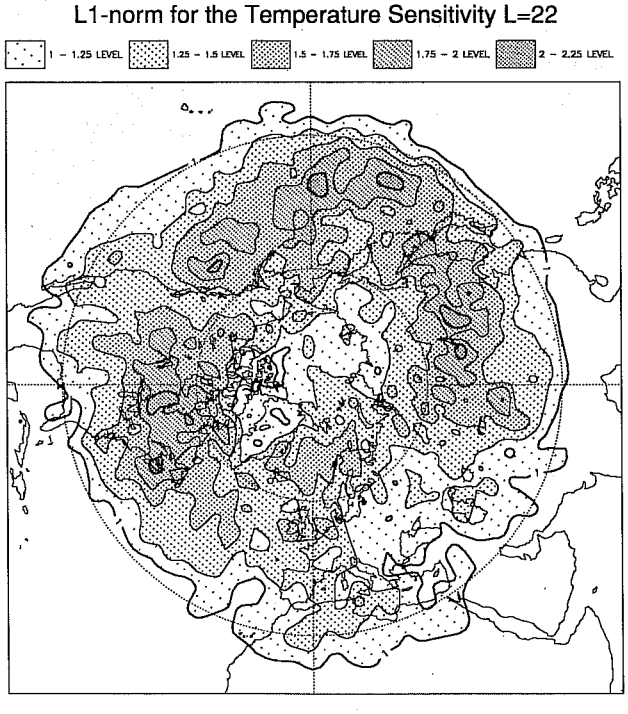


Fig 13 L1- norm of sensitivity patterns over April 1994 for temperature at model level 22 (~ 700 hPa).

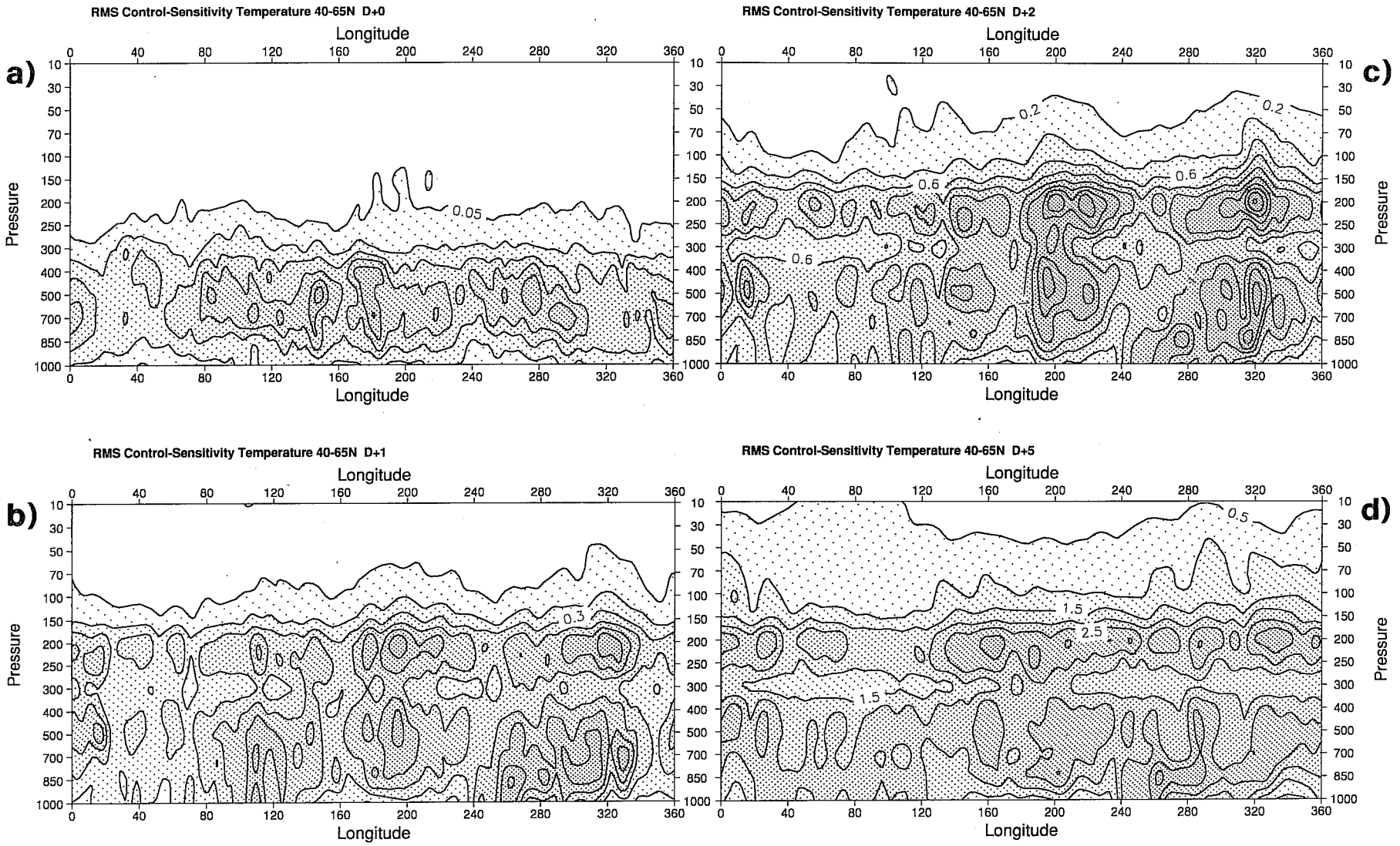


Fig 14 Vertical east-west cross-sections over the latitude band 40-65N of root mean squares of sensitivity perturbations in April 1994 for temperature at the initial time (a), after one day (b), two days (c) and five days (d). Units: K.

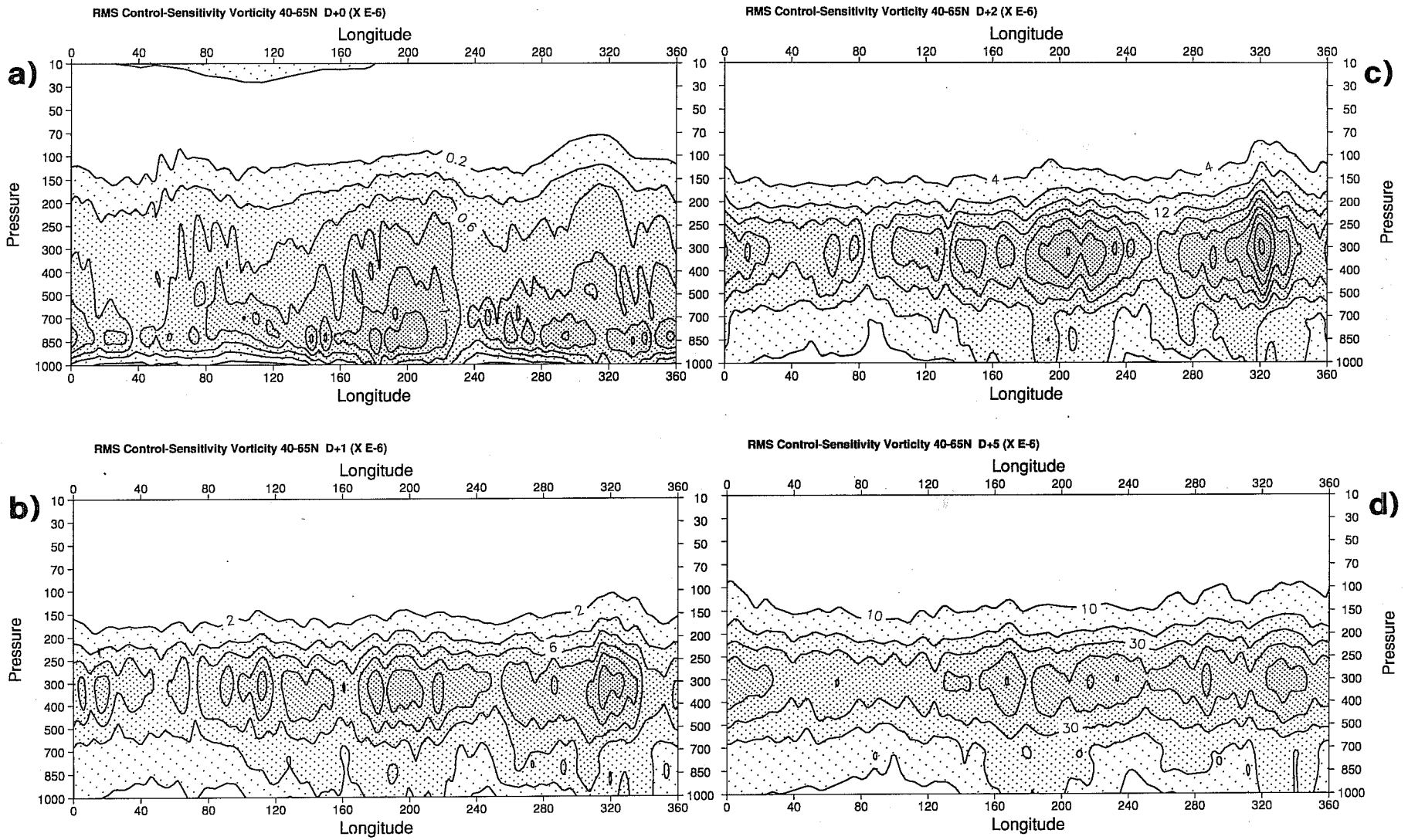


Fig 15 Vertical east-west cross-sections over the latitude band 40-65N of root mean squares of sensitivity perturbations in April 1994 for vorticity at the initial time (a), after one day (b), two days (c) and five days (d). Units: $10^{6}s^{-1}$.

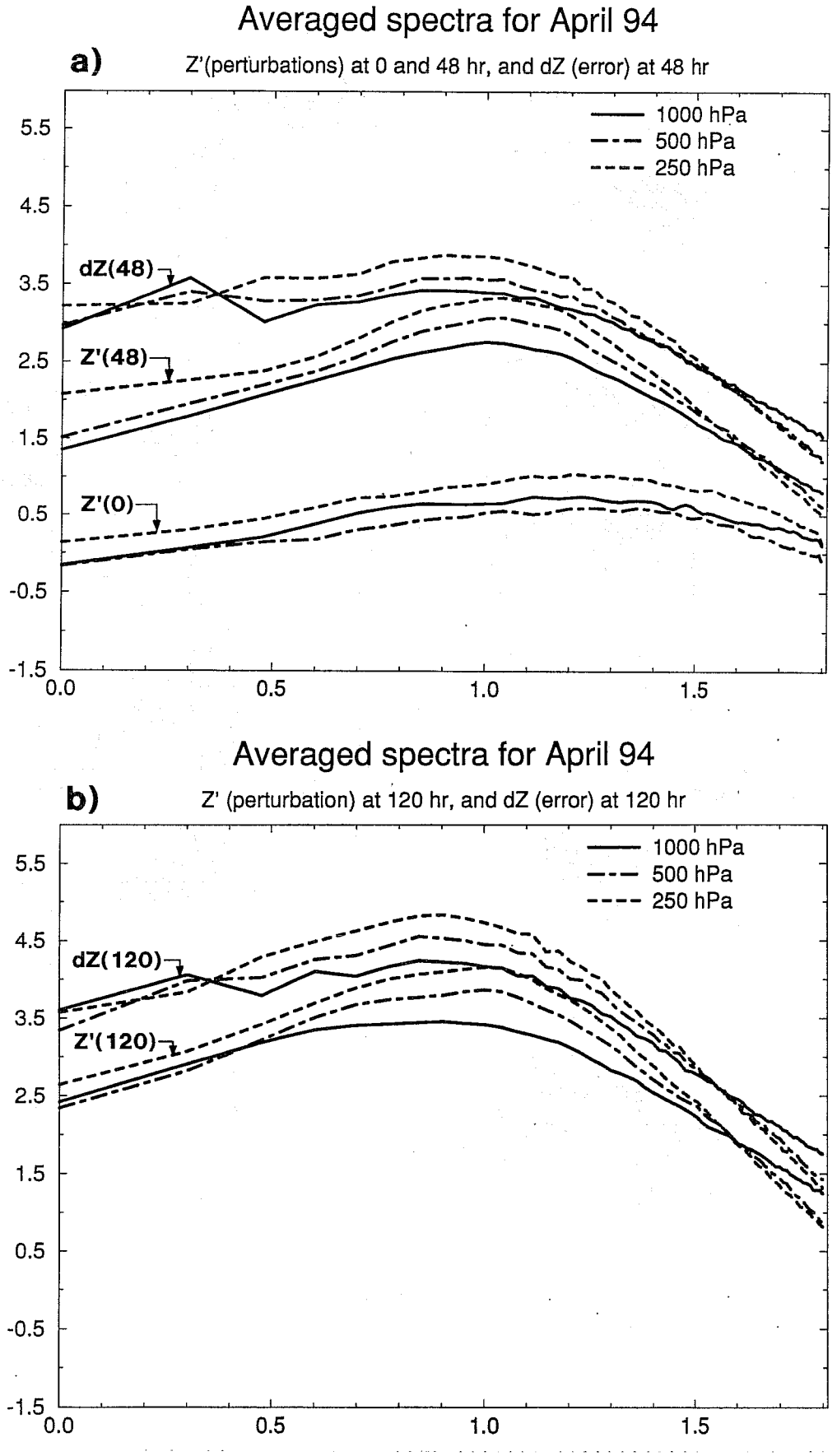


Fig 16 Spectra of geopotential perturbations averaged over April 1994 for various levels in log-log. The abscissa represents log(n) with n the horizontal wavenumber from 1 to 63. In panel a the bottom curves represent the initial perturbations, the middle ones represent the perturbations after 48 hours evolution and the top ones represent the 48-hour forecast error (control forecast minus truncated operational analysis). In panel b the bottom curves represent the perturbations after 5-day evolution and the top ones represent the 5-day forecast error (control forecast minus truncated operational analysis).

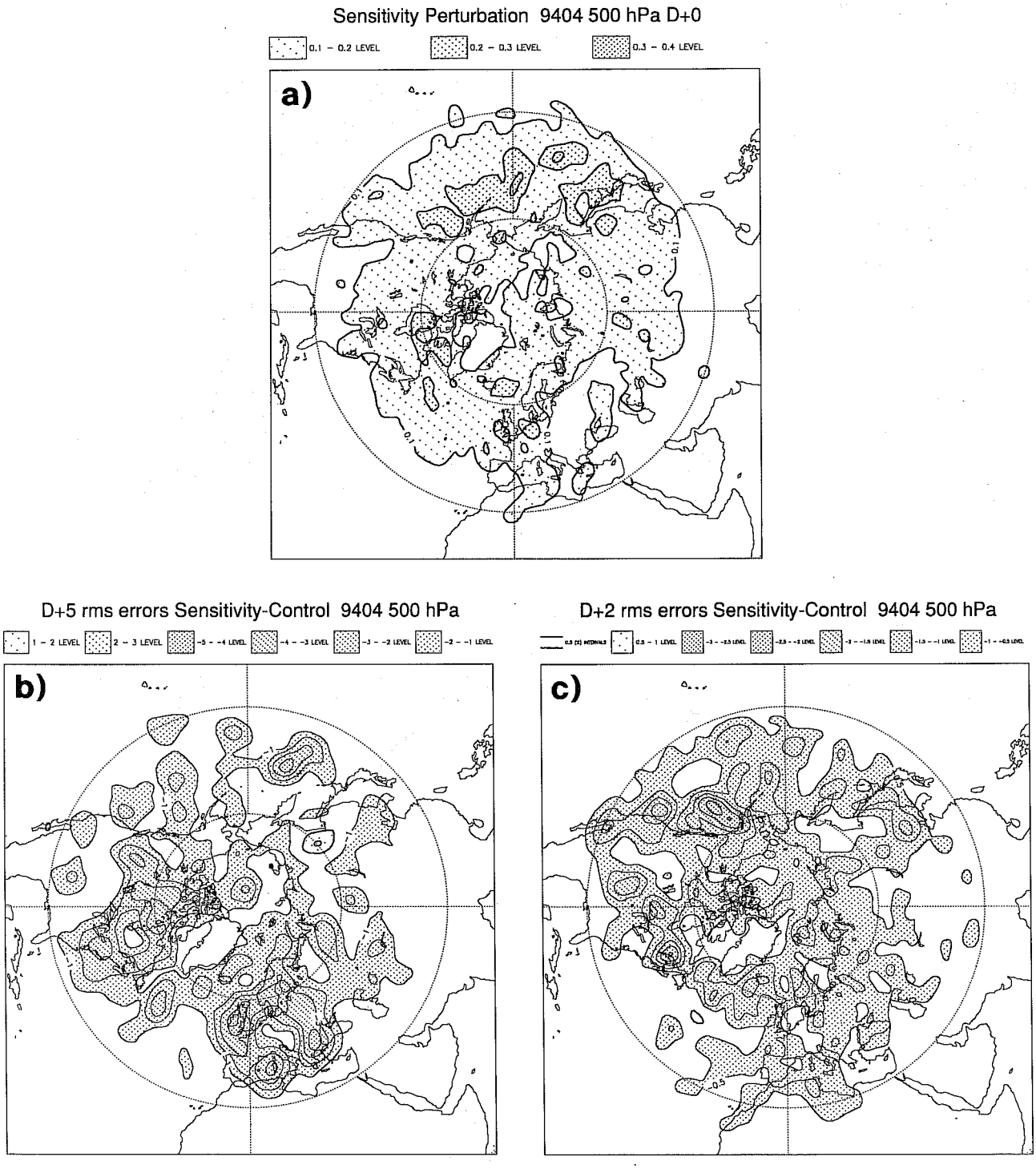


Fig 17 Root mean squares of the height sensitivity perturbations at the initial time (a) and the difference between the rms error of the sensitivity forecasts and the rms error of the control forecasts in April 1994 for day 2 (b) and for day 5 (c) at 500 hPa. Units: dam.

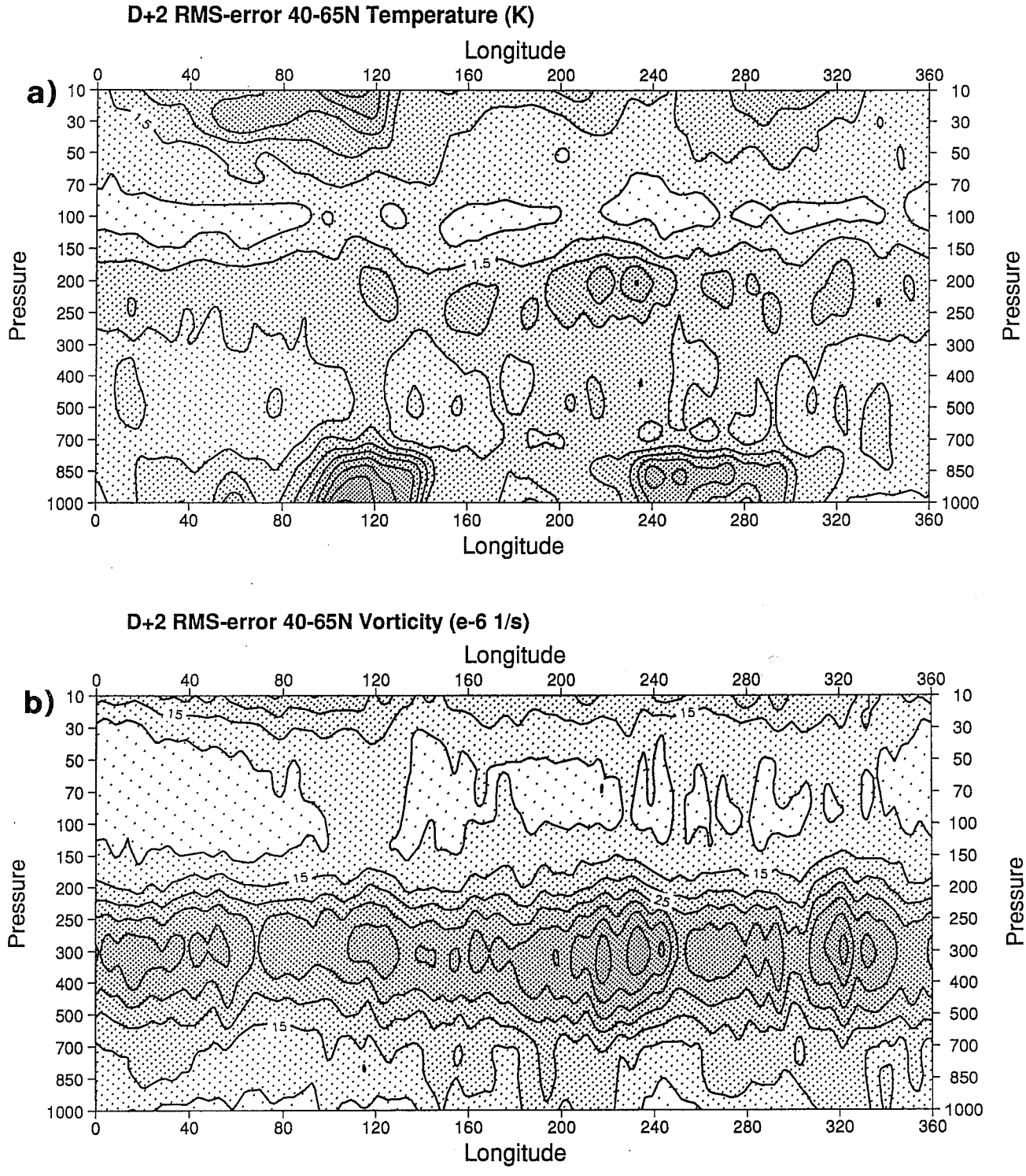
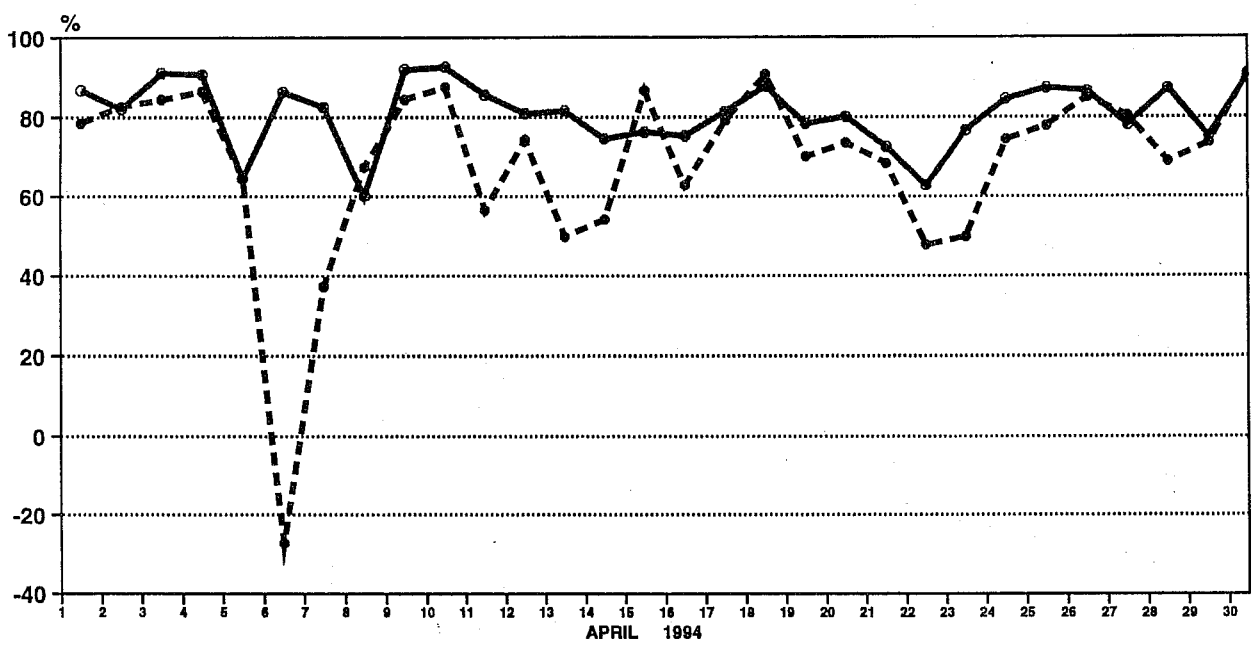


Fig 18 Two-day root mean square forecast errors over April 1994, for temperature (panel a), and vorticity (panel b).



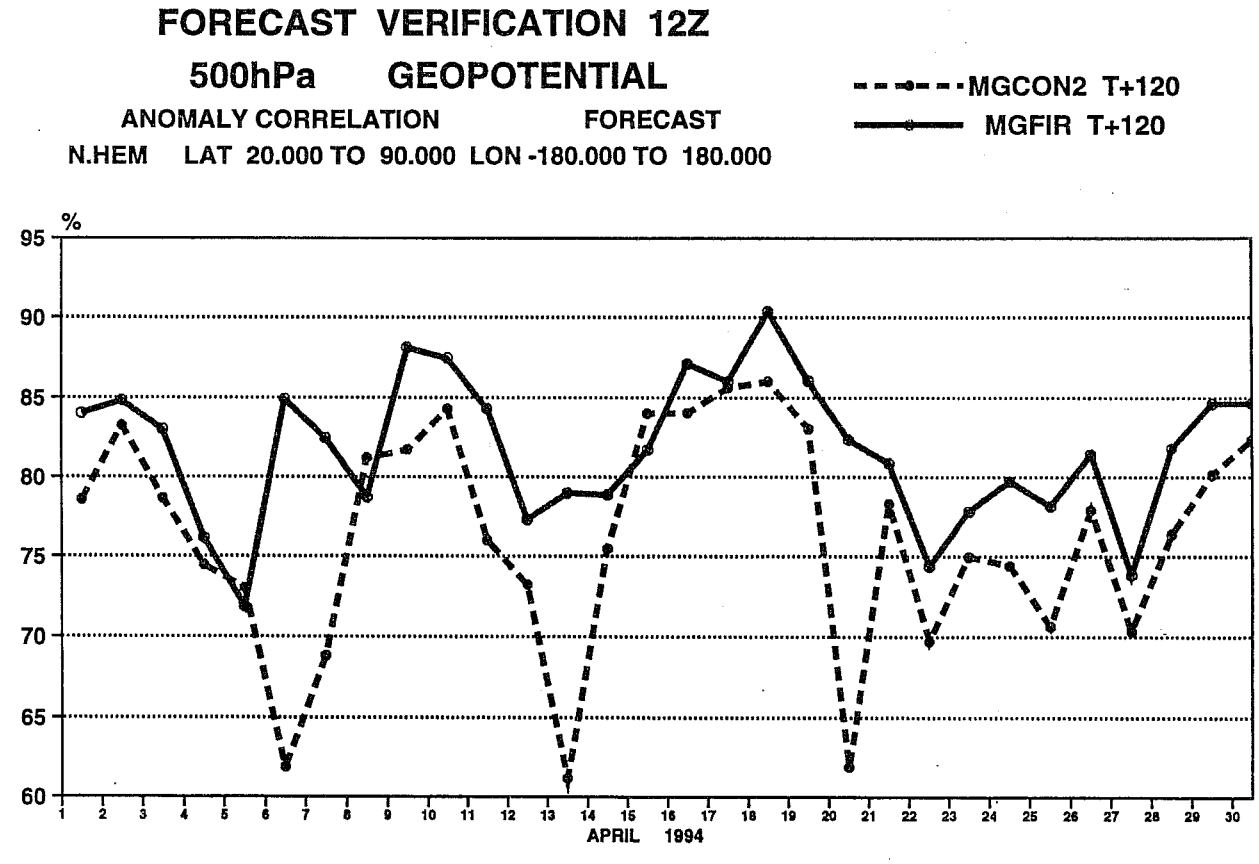
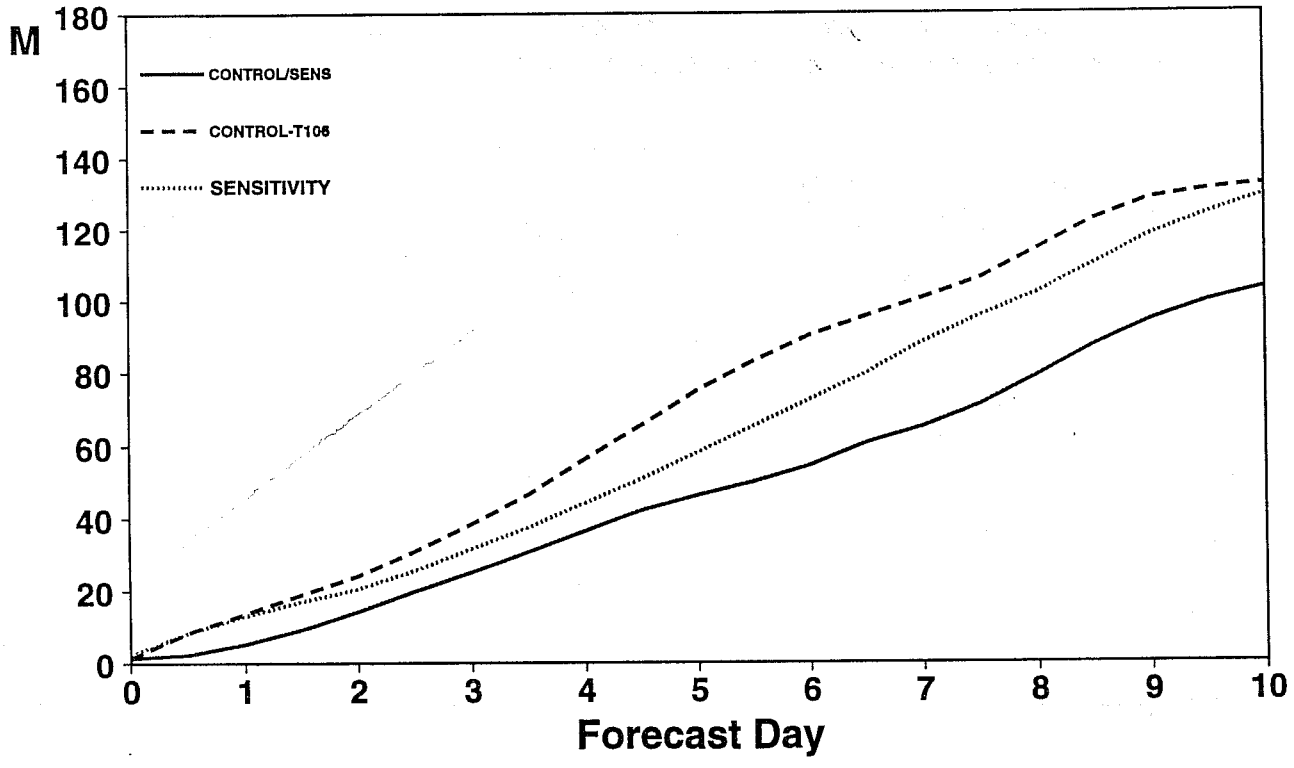


Fig 19 Time series of day-5 anomaly correlations for Europe (upper panel) and the Northern Hemisphere (lower panel). Control experiments: dashed line, sensitivity experiments: solid line.



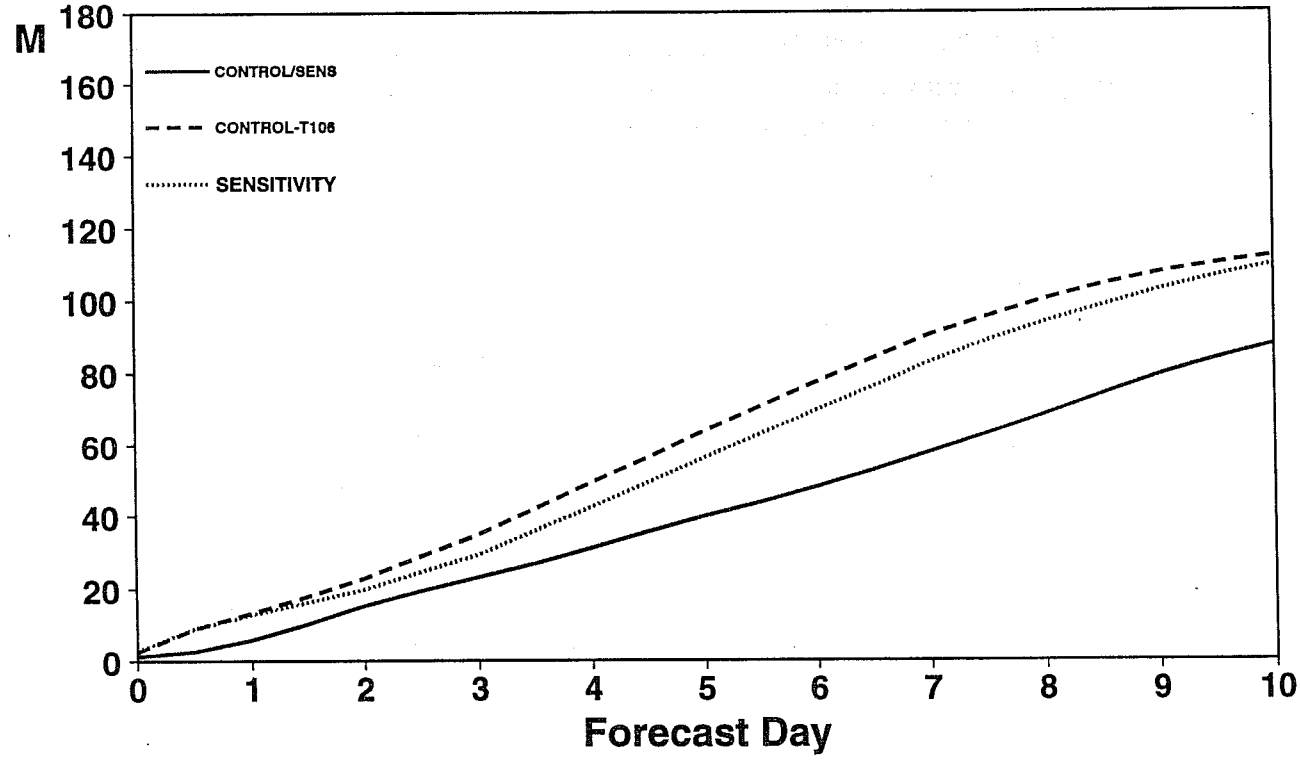


Fig 20 Monthly root mean square forecast errors and forecast differences of temperature at model level 18 (~ 500 hPa) for Europe (upper panel) and the Northern Hemisphere (lower panel). Control experiments: dashed line, sensitivity experiments: dotted line, difference between the control and sensitivity experiments (solid line). Units: m.