Q. J. R. Meteorol. Soc. (1997), **123**, pp. 1035–1053

Using forecast sensitivity patterns to improve future forecast skill

By ZHAO-XIA PU*, EUGENIA KALNAY, JOHN C. DERBER and JOSEPH G. SELA
National Centers for Environmental Prediction, USA

(Received 10 June 1996; revised 10 September 1996)

SUMMARY

A simple, relatively inexpensive technique has been developed for using past forecast errors to improve the future forecast skill. The method uses the forecast model and its adjoint and can be considered as a simplified 4-dimensional variational (4-D VAR) system. One- or two-day forecast errors are used to calculate a small perturbation (sensitivity perturbation) to the analyses that minimizes the forecast error. The longer forecasts started from the corrected initial conditions, although better than the original forecasts, are still significantly worse than the shorter forecasts started from the latest analysis, even though they both had access to information covering the same period.

As a much less expensive alternative to 4-D VAR, the adjusted initial conditions from one or two days ago are used as a starting point for a second iteration of the regular NCEP analysis and forecast cycle until the present time ($t = 0$) analysis is reached. Forecast experiments indicate that the new analyses result in improvements to medium-range forecast skill, and suggest that the technique can be used in operations, since it increases the cost of the regular analysis cycle by a maximum factor of about 4 to 8, depending on the length of the analysis cycle that is repeated. Several possible operational configurations are also tested.

The model used in these experiments is the NCEP's operational global spectral model with 62 waves triangular truncation and 28 σ-vertical levels. An adiabatic version of the adjoint was modified to make it more consistent with the complete forecast model, including only a few simple physical parametrizations (horizontal diffusion and vertical mixing). This adjoint model was used to compute the gradient of the forecast error with respect to initial conditions.

KEYWORDS: Adjoint model Data assimilation Numerical weather prediction Forecast errors Sensitivity 4-D VAR

1. INTRODUCTION

The accurate specification of initial conditions is vital for numerical weather prediction (NWP). Many poor forecasts can be traced to errors in the initial conditions. Recently, considerable attention has been focused on the 3-dimensional (3-D) and 4-dimensional (4-D) variational assimilation methods (Talagrand and Courtier 1987; Derber *et al.* 1991; Navon *et al.* 1992; Andersson *et al.* 1994, 1996) to improve the initial conditions. While the 3-D techniques have become feasible (Derber *et al.* 1991; Andersson *et al.* 1996), the 4-D variational techniques remain expensive, and still on the edge of feasibility for global operational implementation. In this paper we describe a simple, less expensive, alternative to the 4-D variational system that can result in improvements to the data assimilation and forecast system. It attempts to identify and partially remove from the background field some of the fast growing 'errors of the day', a problem identified as high priority in data assimilation.[†]

* Corresponding author: National Centers for Environmental Prediction, W/NCEP2, Room 207, 5200 Auth Road, Camp Springs, MD 20746, USA. At present on leave from Lanzhou University, Lanzhou, P. R. China.

† In the Panel Discussion of the 2nd WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography, Tokyo, March 1995, Philippe Courtier stated 'It is essential for data assimilation in a system which has instabilities to have 'errors of the day' that know about current instabilities (where the error grew and will grow). Rabier *et al.* (1996) demonstrated the practical importance of a careful treatment of the unstable manifold. As shown in Courtier *et al.* (1994) and Courtier (1995), it is possible to embed a variational algorithm (3-D VAR or 4-D VAR) within a simplified Kalman filter. This allows one to implement flow dependent structures which account for the current instabilities. Very similar ideas are expressed in Cohn and Todling (1996). As in Kalnay and Toth (1994) the background term is modified in a low-dimension space, but this dimension can go in practice up to a hundred (or even a thousand).' Other panelists agreed with this priority.

Lacarra and Talagrand (1988) and Vukicevic (1991) investigated the short-range evolution of small perturbations, and showed that the evolution of small errors in a nonlinear model can be approximated with a linear-tangent model for about two days. At the US National Centers for Environmental Prediction (NCEP), formerly the National Meteorological Center, there has been some experimentation to remove the short-term systematic errors and initial growing errors. Thiebaux and Morone (1990) investigated the spatial and temporal statistical coherencies in the departures of short-range forecasts from observed and analysed geopotential fields, and found that they can be used diagnostically as the basis for adjustment of forecast products. Kalnay and Toth (1994) used the growing modes obtained in the NCEP 'breeding cycle' to partially remove the growing errors in the analysis cycle through adjustment of the first guess, and their results indicated that the procedure has the potential of improving the first guess and subsequent forecasts. In this study, we attempt to use directly short-range-forecast errors to adjust the initial condition through a 4-D minimization process using the adjoint method. The experiments will show how we could use the adjusted initial condition to improve the *future* forecast skill.

In the last decade, adjoint models have been applied in 4-D data assimilation (Le Dimet and Talagrand 1986; Derber 1987, 1989; Zupanski 1993; Zupanski and Mesinger 1995), in optimal parameter estimation (Zou *et al.* 1992), in the determination of singular vectors (Buizza 1994), as well as in forecast sensitivity studies (Errico and Vukicevic 1992; Courtier *et al.* 1993; Zupanski 1995; Rabier *et al.* 1996, hereafter RKCH). One of the main advantages of the adjoint method is that it combines the dynamics of the forecast model and observed (or analysed) information in one system, so that the gradient of cost function with respect to all initial parameters (initial field and model input parameters) can be calculated with a single integration of the adjoint model. If the cost function is defined as the distance between a forecast and verifying analysis, then the gradient with respect to the initial conditions is denoted 'forecast sensitivity'. If the initial conditions are adjusted with a perturbation proportional to the forecast sensitivity, the difference between the adjusted and the original solutions will give an estimate of the distribution of the initial error explaining part of the forecast error. This idea was developed and applied by RKCH using the European Centre for Medium-Range Weather Forecasts (ECMWF) model. They calculated the perturbation to the two-day-old initial conditions that would minimize the two-day forecast error over the northern hemisphere, and compared the resulting forecast with the original two-day-old operational forecast. Their results showed a very substantial improvement of several particularly poor forecasts, although in almost all cases the updated forecasts from two days earlier were not better than those launched from the latest initial conditions. Zupanski (1995) found that the forecast sensitivity improved when several iterations were performed.

In this paper we present a simple extension of the RKCH technique, in which we use the known error from past forecasts to improve future forecasts. The initial conditions of forecasts 1 or 2 days old are altered based on their observed forecast error. The altered fields are then used as a starting point for a second iteration of the regular NCEP analysis cycle in order to improve *future* forecasts. The experiments are performed with the NCEP model with a triangular truncation of 62 waves and 28 vertical levels. The mathematical formulation is described in section 2. The adjoint model (section 3) contains the complete dynamics of the model, but only a few simple physical parametrizations (horizontal diffusion and vertical mixing), as in Buizza (1993). In section 4 we present forecast-sensitivity experiments. The extension of the method to improve future forecasts by performing a new assimilation cycle, and the corresponding results, are presented in section 5. Section 6 includes a summary and conclusions.

2. Mathematical formulation

At any forecast time t, for a forecast variable $\mathbf{X}$, the perceived forecast error, $\delta\mathbf{X}$, is given by the difference between forecast ($\mathbf{X}_t$) and analysis ($\mathbf{X}_t^{\mathrm{ref}}$)

$$\delta\mathbf{X} = \mathbf{X}_t - \mathbf{X}_t^{\mathrm{ref}} \tag{1}$$

where $\mathbf{X}_t = \mathrm{M}(\mathbf{X}(t_0))$; M is a forecast model and $\mathbf{X}(t_0)$ is the initial condition corresponding to $\mathbf{X}_0$. A measure of the error at forecast time t is based on a total energy norm, given by

$$\begin{aligned} F_t &= \|\delta\mathbf{X}\|^2 = \tfrac{1}{2}\delta\mathbf{X}^{\mathrm{T}}\mathrm{W}\delta\mathbf{X} \\ &= \tfrac{1}{2}\int_0^1 \int \int_\Gamma (\nabla\Delta^{-1}\zeta\,\nabla\Delta^{-1}\zeta \\ &\quad + \nabla\Delta^{-1}D\,\nabla\Delta^{-1}D + R_{\mathrm{a}}T_{\mathrm{r}}(\Pi)^2 + \frac{C_p}{T_{\mathrm{r}}}T^2)\,\mathrm{d}\Gamma\left(\frac{\partial P_{\mathrm{r}}}{\partial z}\right)\mathrm{d}\eta \end{aligned} \tag{2}$$

where $\delta\mathbf{X} = (\zeta, D, T, \Pi)$ denotes differences between the forecast and analyses for vorticity, divergence, temperature and natural logarithm of surface pressure, respectively, and $\delta\mathbf{X}^{\mathrm{T}}$ denotes its transpose. T_{r}, P_{r} are the reference temperature and pressure, respectively, R_{a} is the gas constant for dry air, and C_p is specific heat at constant pressure for dry air. Γ represents the integration domain. W is the matrix of weights defining the norm. The weights are a function of T_{r}, P_{r}, R_{a}, and C_p. η is the vertical coordinate. Thus, the cost function $\mathcal{J}$ over the forecast period is given by:

$$\mathcal{J} = \sum_{i=1}^{N} F_{t_i} \tag{3}$$

where N is the number of the analysis times used for calculating $\mathcal{J}$.

We would like to find an initial perturbation $\delta\mathbf{X}(t_0)$ that minimizes forecast error as measured by the objective function. The preconditioned gradient of the objective function with respect to the initial conditions is given by:

$$\nabla_{\mathbf{X}(t_0)}\mathcal{J} = \mathrm{W}^{-1}\sum_{i=1}^{N}\mathrm{L}^{\mathrm{T}}\mathrm{W}\delta\mathbf{X}(t) \tag{4}$$

where L is the linearization of the forecast model around the current solution, L^{T} denotes the adjoint of the operator L.

In the first iteration of most minimization algorithms, the correction to the initial conditions is equal to the gradient multiplied by a reasonable step size, ρ (Le Dimet and Talagrand 1986). From this we obtain the new initial conditions:

$$\mathbf{X}(t_0)\,|_{\mathrm{new}} = \mathbf{X}(t_0) - \rho\,\nabla_{\mathbf{X}(t_0)}\mathcal{J}. \tag{5}$$

In much of the following we take advantage of the fact that this single iteration of the minimization algorithm explains a substantial portion of the forecast error. The considerable cost of each iteration makes it impractical to perform the minimization with many iterations as recommended by Zupanski (1995). Examples on the impact of the use of several iterations will be shown in section 4.

The gradient in (4) is commonly called a forecast sensitivity pattern (Zupanski 1995; RKCH). We call the second term in (5) the sensitivity initial perturbation since it is related to such sensitivity patterns. RKCH and Buizza *et al.* (1995) showed that the sensitivity pattern can be approximated by projecting onto the subspace of growing singular vectors (which Buizza *et al.* (1995) denote unstable subspace), and that the projection is dominated by the leading singular vectors.

3. Modifications to the adjoint of the NCEP global spectral model

To ensure that we obtain a reasonably accurate gradient, it is necessary for the adjoint model to be as consistent as possible with the forecast model. However, the inclusion of all the model physics into the linear-tangent model and its adjoint is both difficult and computationally expensive, most applications so far have used simplified adjoint models (Errico *et al.* 1993; Buizza 1993, 1994; RKCH). At the NCEP an adiabatic version adjoint for the global spectral model was developed by Navon *et al.* (1992), including a simple surface-drag scheme and horizontal diffusion. It worked well for most of the atmosphere within a short period, but unrealistically larger growth rates were observed at low levels of the model (Szunyogh (1994), personal communication), a problem attributed to inadequacies in the model physics. Following Buizza (1993) and Thone and Mahrt (1986), we introduced, in addition to the surface drag and the horizontal diffusion, a linear vertical diffusion scheme. To assess the effect of the vertical diffusion parametrization on the spurious growth, the nonlinear NCEP global spectral model was integrated for 24 hours from 0000 UTC 16 May 1985 initial conditions, but using different subsets of physical parametrizations. The results were compared with those obtained from running the operational model with full physics. Figure 1 shows the kinetic energy norm of the different 24-hour forecasts at all model levels. In the run with no physics it is apparent that there is an unrealistically large increase of kinetic energy at low levels when compared with the run with full physics. The introduction of surface drag only improves the lowest model level, but the kinetic energy in levels 2–7 remains too large. The introduction of the new vertical diffusion and horizontal diffusion results in a solution much closer to the operational T62 model with full physics.

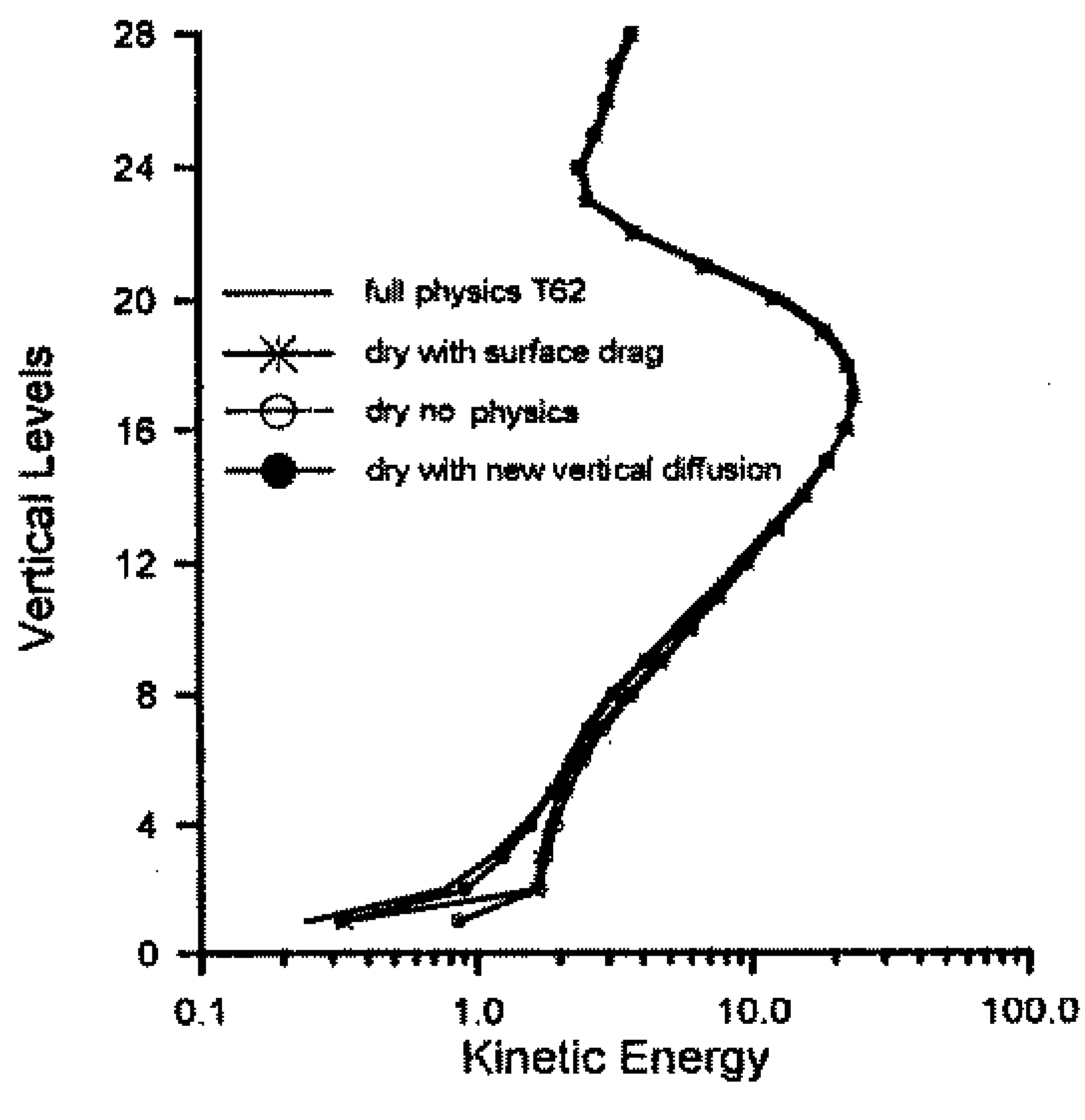

Figure 1. Vertical cross-section of kinetic energy norms after 24-hours integration of the T62 global spectral model with no physics, with surface drag, with surface drag and simple vertical diffusion, and with the full physics (operational model).

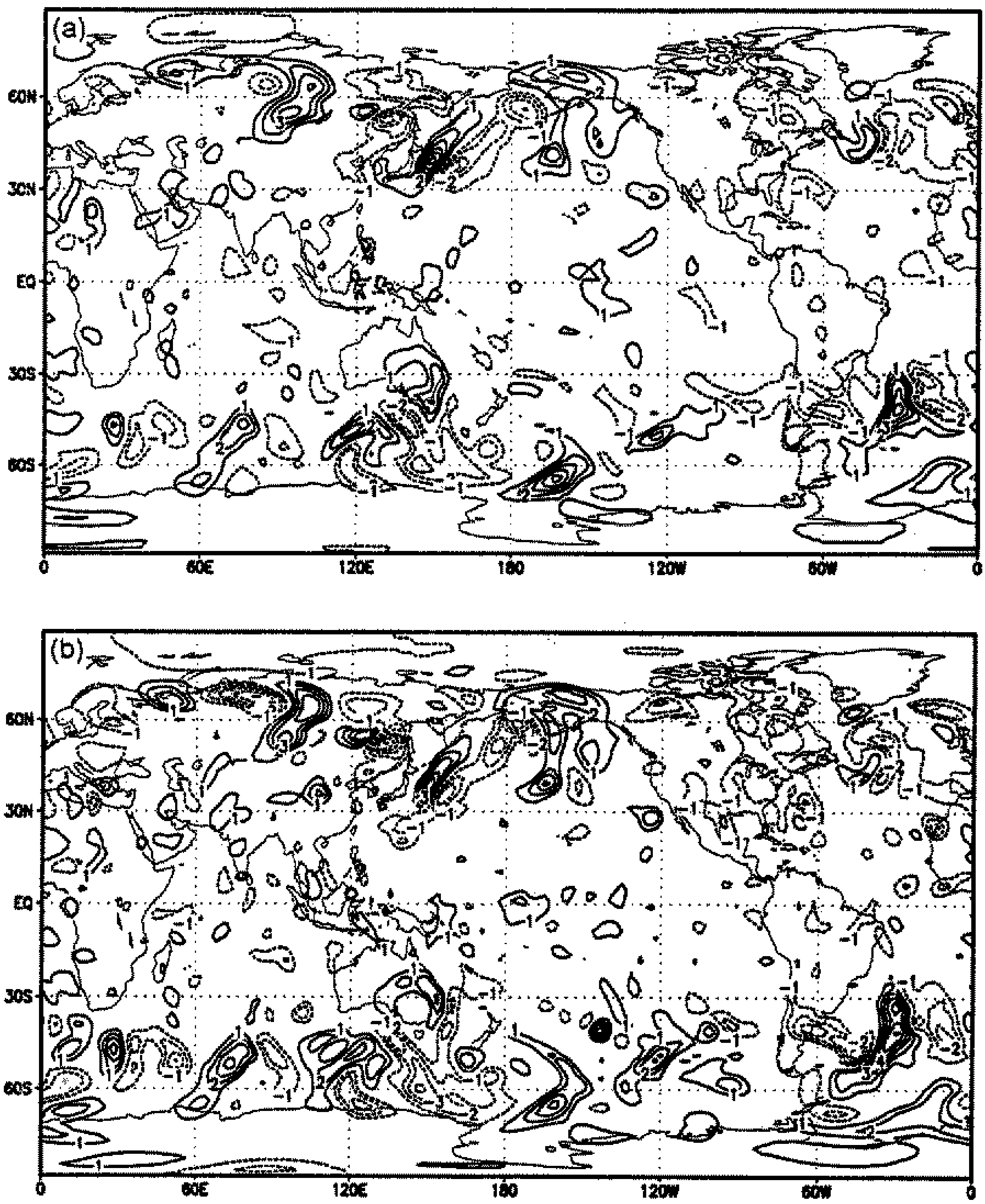

Figure 2. The evolution of a perturbation δX_0 after 24 hours at level 13 (sigma = 0.501); (a) linear and (b) nonlinear evolution. Contours show the temperature perturbation at 1 K intervals. Dashed lines represent negative values.

The new vertical diffusion scheme was included in the linear and adjoint models used in this paper. To examine further the approximation of this simplified version to the complete vertical diffusion parametrization, a small-amplitude initial perturbation field, $\delta \mathbf{X}_0$, was introduced into the nonlinear model integration with the full physics package, and the difference between the two nonlinear integrations (starting from perturbed and unperturbed initial condition at 0000 UTC 16 May 1985) was compared with a linear model integration which included the new linear vertical diffusion and the same perturbation ($\delta \mathbf{X}_0$) as initial condition. Figure 2 shows the comparison between the linear and nonlinear evolution for the temperature perturbation field at sigma level 13 (about 500 mb) after 24 hours of integration. The agreement between the perturbation fields obtained from the linear and the nonlinear integration is generally good, containing small differences in phase but reproducing all the main features. This suggests that, in agreement with previous investigators, the dry adiabatic linear model with surface friction and vertical diffusion can reproduce fairly well the nonlinear perturbations of the model with full physics for short-range (1–2 day) forecasts.

4. Sensitivity of short-range-forecast errors to perturbations in the initial conditions

(*a*) *Experiments*

In this section, sensitivity experiments similar to those performed at the ECMWF (RKCH) are performed using the NCEP adjoint system. Again, 0000 UTC 16 May 1985, available from the NCEP/NCAR* Re-analysis Project, was arbitrarily chosen as the initial condition. The current NCEP operational data-assimilation system (as of January 1995) was used to create the analysis on a 6-hour cycle. This system uses a 3-D variational analysis (3-D VAR) scheme developed by Parrish and Derber (1992) and Derber *et al.* (1991), implemented in 1991, and which has since then undergone further improvements (Derber *et al.* 1994; Parrish *et al.* 1996).

In order to test the sensitivity of the one- or two-day forecast errors to changes in the initial conditions, three sensitivity experiments are performed and compared. In experiment 1 we use a one-day forecast error, where the forecast error is defined as the difference between the forecast and the analysis at 0000 UTC 17 May 1985; in experiment 2 we use a two-day forecast error, similarly defined at 0000 UTC 18 May 1985; and in experiment 3 we also use a two-day forecast, but the forecast error in the cost function is defined as the sum of all the forecast differences with the analyses available every six hours within the period 0000 UTC 16 May 1985 through 0000 UTC 18 May 1985. A conjugate gradient method (Gill *et al.* 1981) was employed in our system, with an optimal step size defined as in Derber (1987).

Figure 3 illustrates the vertical distribution of the total energy norm of the forecast error before and after a single iteration in experiments 1 and 2, showing that the error decreases significantly at almost every sigma level. The effect of performing multiple iterations is shown in Fig. 4 for experiment 3. As in Zupanski (1995) we find that although the first iteration reduces substantially the cost function (by 18%), additional iterations continue to decrease it, with a total reduction of about 45% after ten iterations (Fig. 4(a)). Similarly, Fig. 4(b) shows that the forecast error throughout the two-day forecasts after one iteration is reduced by about 20%, and after another nine iterations by an additional 25%. Note that in the extended-period forecast error, we found that not only the 48 h forecast is improved, but the 3–5 day forecasts are also significantly better. There is a clear advantage of having more iterations, although one iteration is enough to improve the control forecast. Zupanski (1995), using a regional model and its adjoint, found that the sensitivity fields obtained after one and ten iterations were not very similar, with projections corresponding to angles of up to 50° between them. But we should also notice that the descent of the cost function in the first iteration is faster than for the other iterations.

Since the first iteration explains more forecast error than any of the following ones, and because of the large expense of performing many iterations of the descent algorithm, RKCH have chosen to perform a single iteration in their experiments, with a value of the step size $\rho = 1/100$, 1/150 or 1/200 (depending on the season) multiplied by the gradient, to generate the sensitivity perturbations. In this study, our main purpose is to use the improved initial conditions as the starting point for a second iteration of the regular NCEP analysis cycle (section 5). We are interested in the potential operational application of this method,

* National Center for Atmospheric Research.

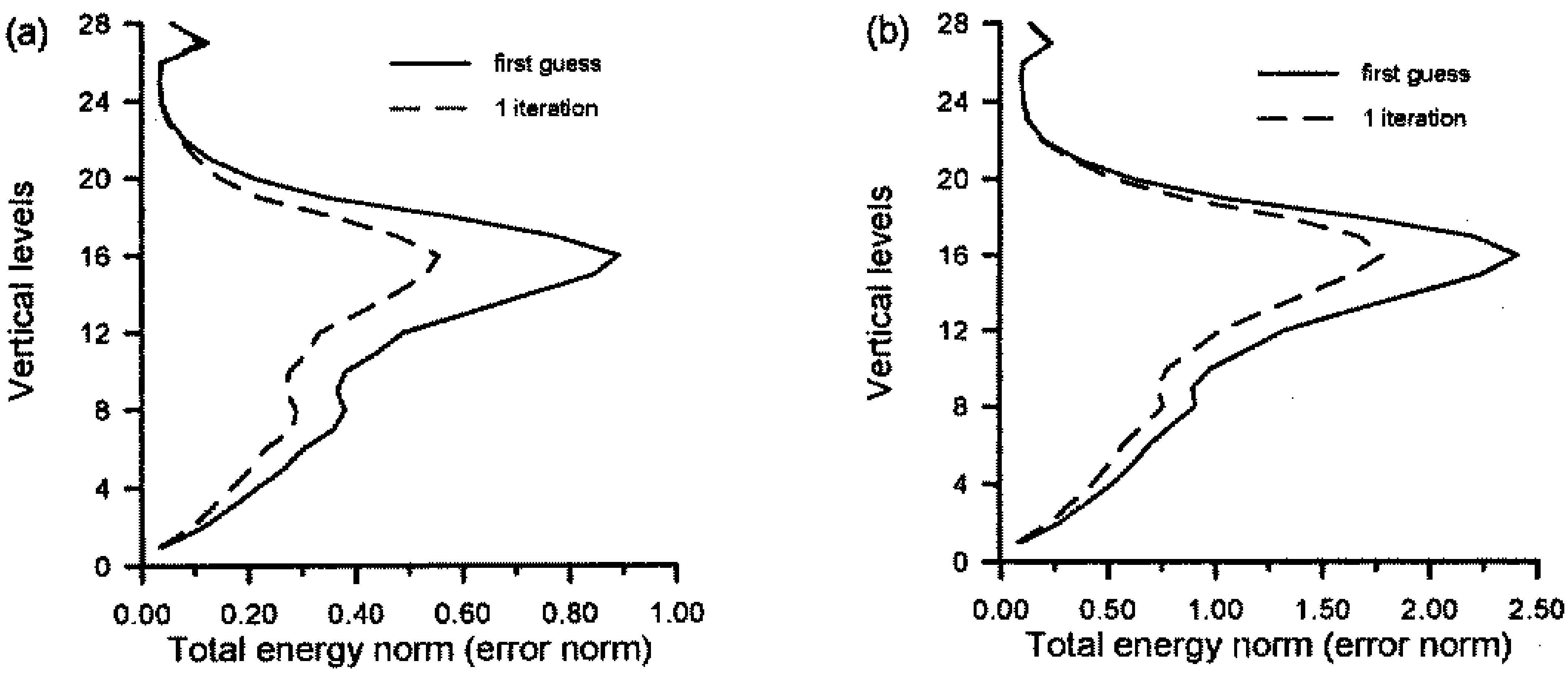

Figure 3. Vertical cross-section of the total energy norm of forecast error before (solid curve) and after (dashed curve) one iteration: (a) for experiment 1 and (b) for experiment 2.

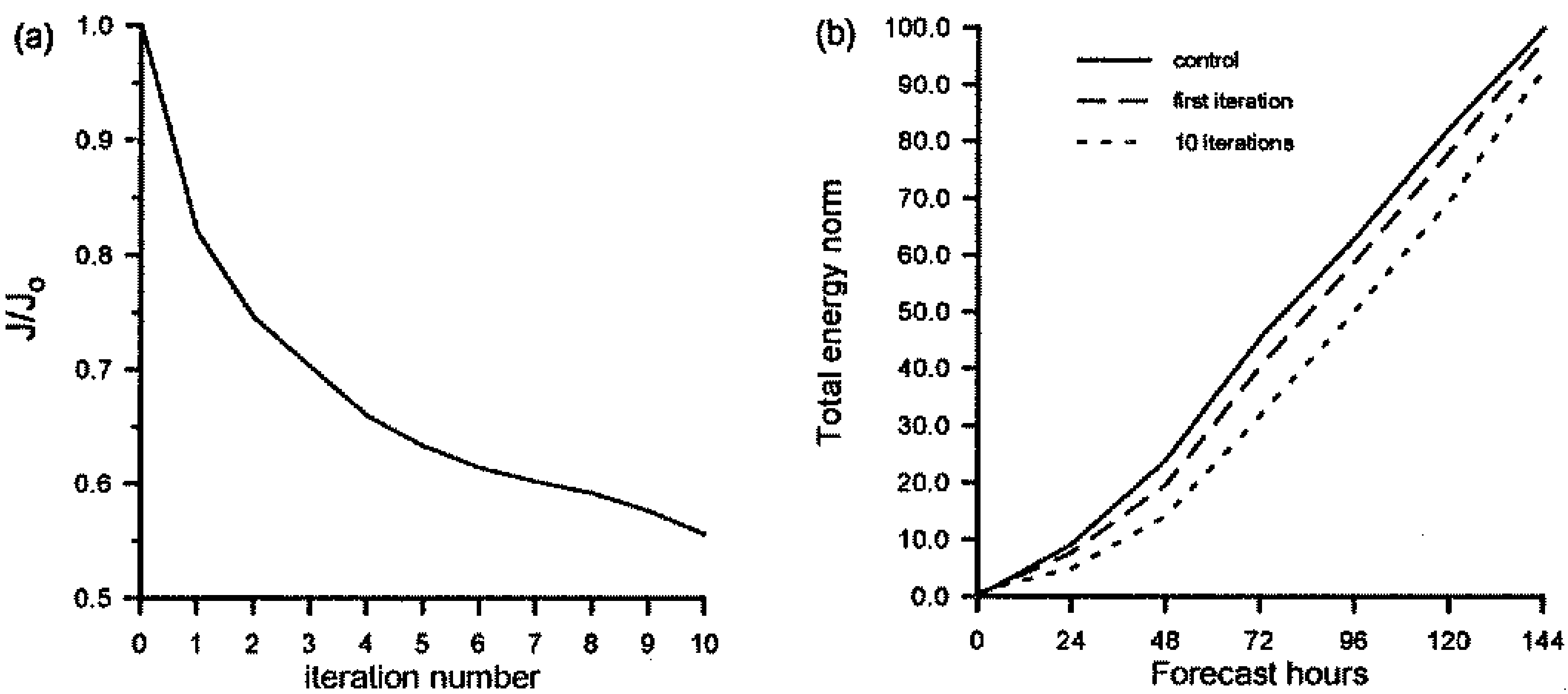

Figure 4. (a) Variation of the normalized cost function ($\mathcal{J}/\mathcal{J}_0$) (sum of total error norms) with the number of iterations. (b) The total energy norms at every 24 hours in two days and extended forecast period. Solid curve for the error norm of first guesses (control forecast), long-dashed line for an error norm after one iteration and short-dashed line for after ten iterations.

which limits the available computer resources, so that we have also chosen to perform only a single iteration assuming that such single iteration will result in a substantial reduction of the forecast error while remaining computationally efficient. Since the fixed step size may not be sufficiently accurate, we determine the step size using the technique described by Derber (1987), with 1/100 or 1/200 as the guess step size, ρ_g. The optimal step size ρ is given by:

$$\rho = \rho_g \frac{\sum_{i=1}^{N} \mathbf{W}\{(\mathbf{X}_i - \mathbf{X}_i^{\mathrm{ref}})\}^{\mathrm{T}} \mathbf{X}_g^e(t_i)}{\sum_{i=1}^{N} \{\mathbf{W}\mathbf{X}_g^e(t_i)\}^{\mathrm{T}} \mathbf{X}_g^e(t_i)} \tag{6}$$

where $\mathbf{X}_g^e(t_i)$ is the difference between the forecasts results from integrating the initial condition $\mathbf{X}(t_0)$ and $\mathbf{X}(t_0) + \rho_g \nabla_{\mathbf{X}(t_0)}\mathcal{J}$. To calculate the optimal step size for one iteration, the forecast model needs to be integrated one time from the initial condition $\mathbf{X}(t_0) + \rho_g \nabla_{\mathbf{X}(t_0)}\mathcal{J}$.

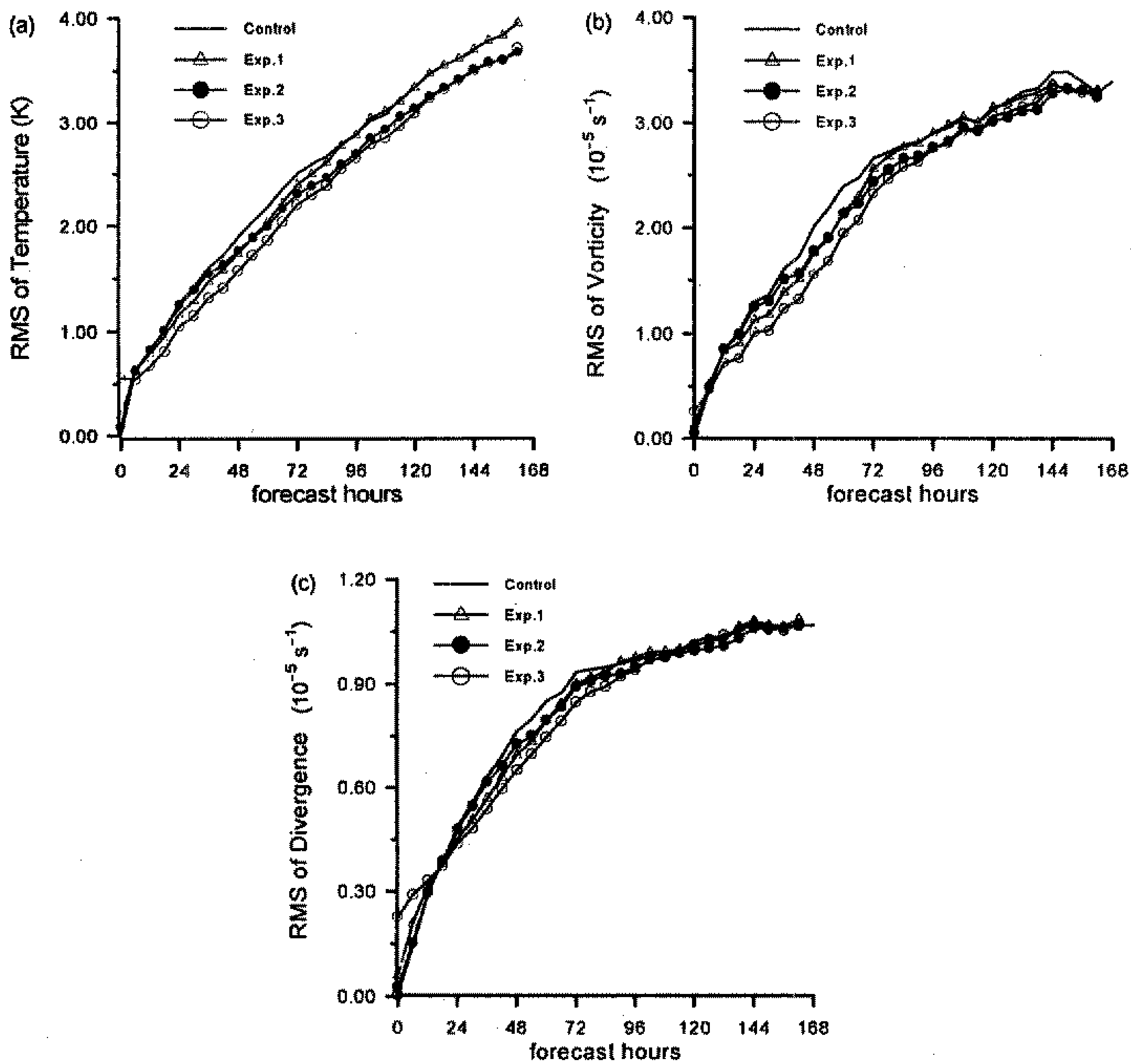

Figure 5. The root-mean-square forecast error of (a) the temperature, (b) the vorticity and (c) the divergence.

(*b*) *Impact of the sensitivity perturbations on medium-range forecasts*

In this subsection we compare the impacts on medium-range forecasts of the three forecast sensitivity experiments discussed in the previous subsection, using a single iteration, with the control forecast performed from the unperturbed analysis for 0000 UTC 16 May 1985.

Figures 5(a), (b), and (c) show the global root-mean-square (r.m.s.) forecast errors for temperature, vorticity and divergence, respectively, for the control and the three sensitivity experiments. The results indicate that the one-day sensitivity perturbation (experiment 1) improves not only the first day, but also that the benefit extends to days 2 through 4. The 2-day sensitivity-experiment (experiment 2 and 3) improvement extends throughout the whole forecast period (days 3 through 7). Comparing experiment 2 with experiment 3, in which the forecast error was defined throughout the two-day forecast and not just at the end of the two days, we observe that experiment 3 is clearly better during the first 4 days, but that at days 5 and 7 experiments 2 and 3 give similar errors. Note that experiment 3 results in considerable larger changes in the initial conditions than the other two experiments. Experiments indicated that the length of the period and the density of data insertion in the period affected the initial perturbation and subsequent forecasts. This effect will be further tested with the several possible operational configurations in subsection 5(b).

The results of these experiments agree with the RKCH experience that the benefit of the perturbed initial conditions extends beyond the period for which the forecast sensitivity perturbation was initially calculated.

(*c*) *Evolution of the sensitivity perturbation*

A second experiment on forecast sensitivity on short- and medium-range forecasts was performed with initial conditions corresponding to 0000 UTC 23 February 1995. The sensitivity was calculated from 2-day forecast errors at 0000 UTC 25 February 1995, as in experiment 2. Figure 6 shows the 500 mb geopotential-height difference between the sensitivity and the control forecast at 0, 24, 48, 72, 96, and 120 hours. We can see that the small initial differences (plotted with a 3 m contour interval) lead to very large differences after 120 hours (plotted with a 25 m contour interval). These large differences take place in both the northern and the southern hemispheres. Table 1 presents the anomaly correlation scores verified against operational analyses for the control and the sensitivity forecasts, showing large improvements in both hemispheres. An inspection of the distribution of the initial perturbation of geopotential height at 500 mb (corresponding to T+00 in Fig. 6) indicates that there may be two dominant error sources in the northern hemisphere, over Canada and the North Pacific areas. The 5-day forecast differences are also maximum over North America in this case. A check of the corresponding 5-day forecast of the NCEP operational model (at T126/28 level resolution) indicates that the anomaly correlation scores on 0000 UTC 28 February 1995 were also unusually low over North America, and within normal range for other areas of the northern hemisphere (P. Caplan (1995), personal communication). This confirms that the large error growth of the sensitivity perturbation is present even in the high-resolution model, and that the adjoint method is a useful tool to trace back growing initial perturbations associated with the forecast error.

TABLE 1. THE ANOMALY CORRELATION SCORES OF 1–5 DAY FORECASTS FROM 0000 UTC 23 FEBRUARY 1995 FOR 500 MB GEOPOTENTIAL HEIGHT

Hemisphere	Type	Day 1	Day 2	Day 3	Day 4	Day 5
Northern	Control	0.986	0.963	0.921	0.852	0.729
Northern	Sensitivity	0.989	0.978	0.969	0.934	0.877
Southern	Control	0.933	0.820	0.762	0.670	0.592
Southern	Sensitivity	0.943	0.833	0.767	0.725	0.666

5. USE OF PAST FORECAST ERRORS TO IMPROVE FUTURE FORECAST SKILL

(*a*) *Comparison of the sensitivity forecasts with regular forecasts from latest initial conditions*

So far, our results confirm the experience of RKCH, showing that the use of the short-range-forecast errors can improve substantially the original forecast, even beyond the period for which the error was computed. But the crucial question for an operational centre like the NCEP or the ECMWF is whether it can use this procedure to improve the skill of *future* forecasts. In other words, if the observed 2-day-forecast error is used to improve the initial conditions from two days ago, will the improved 5-day forecast from two days ago be also better than the 3-day forecast from today? Unfortunately, as will be seen, the answer to this question is that, in most cases, this is not the case. This lack of

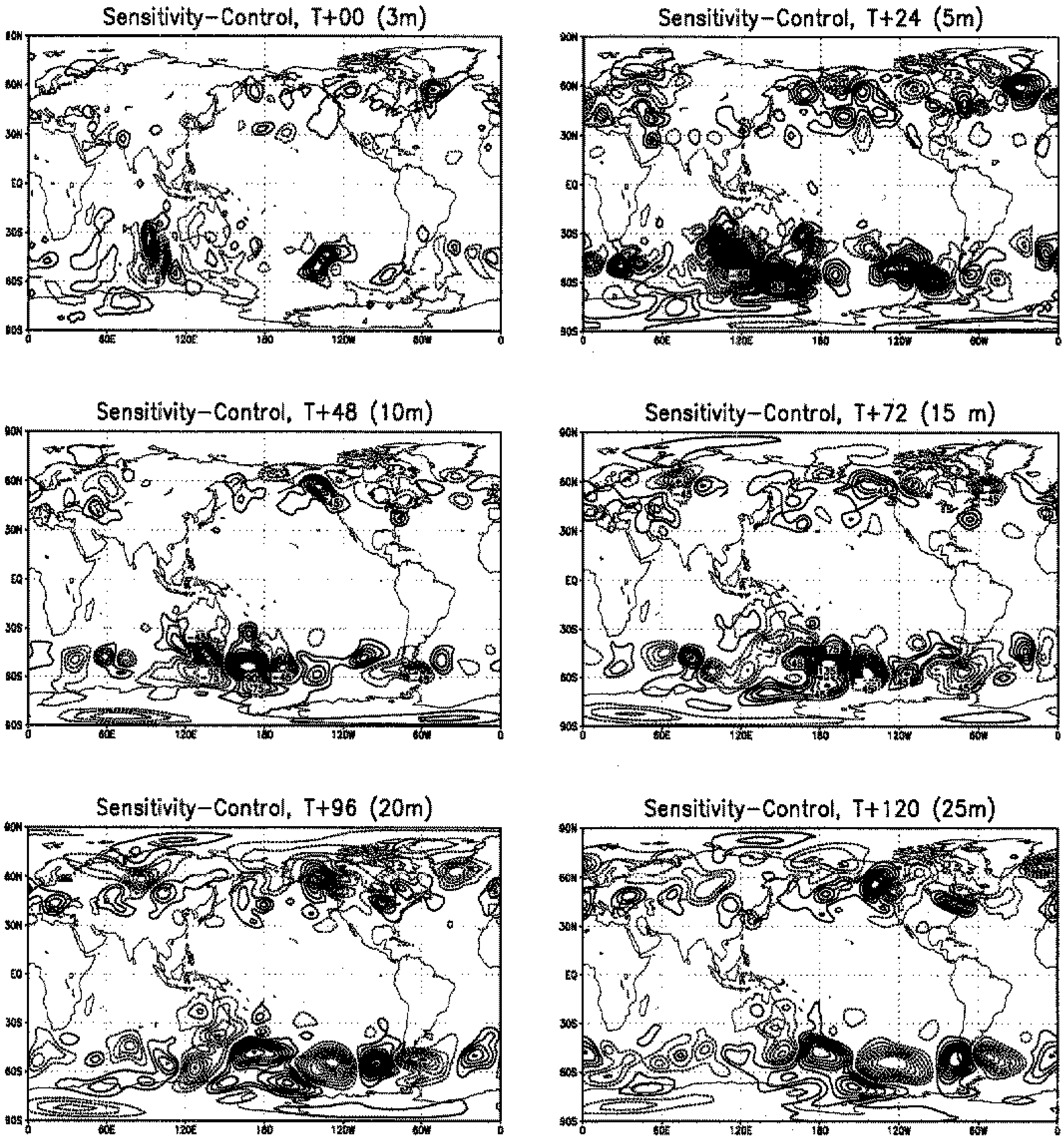

Figure 6. 500 mb height difference between the sensitivity and the control forecasts in steps from 0–120 hours from 0000 UTC 23 February 1995. Note the different contour intervals which are indicated in brackets above each panel.

improvement with respect to the latest forecast was also observed by RKCH. Therefore, taking advantage of our knowledge of the data available until the present time just to correct older initial conditions is not enough to make forecasts better than the regular forecasts made from the analysis cycle using all the data available until today. There may be at least three reasons for this result: (a) the use of the linear-tangent model as a strong constraint within the forecast sensitivity procedure; (b) the use of a single iteration to obtain the improved initial conditions; and (c) the fact that the latest analysis uses all the observations directly within a full data assimilation system, rather than just correcting a few degrees of freedom in the growing errors.

However, it seems natural that if we can improve the initial conditions from two days ago, it should be possible to perform a second iteration of the regular analysis cycle,

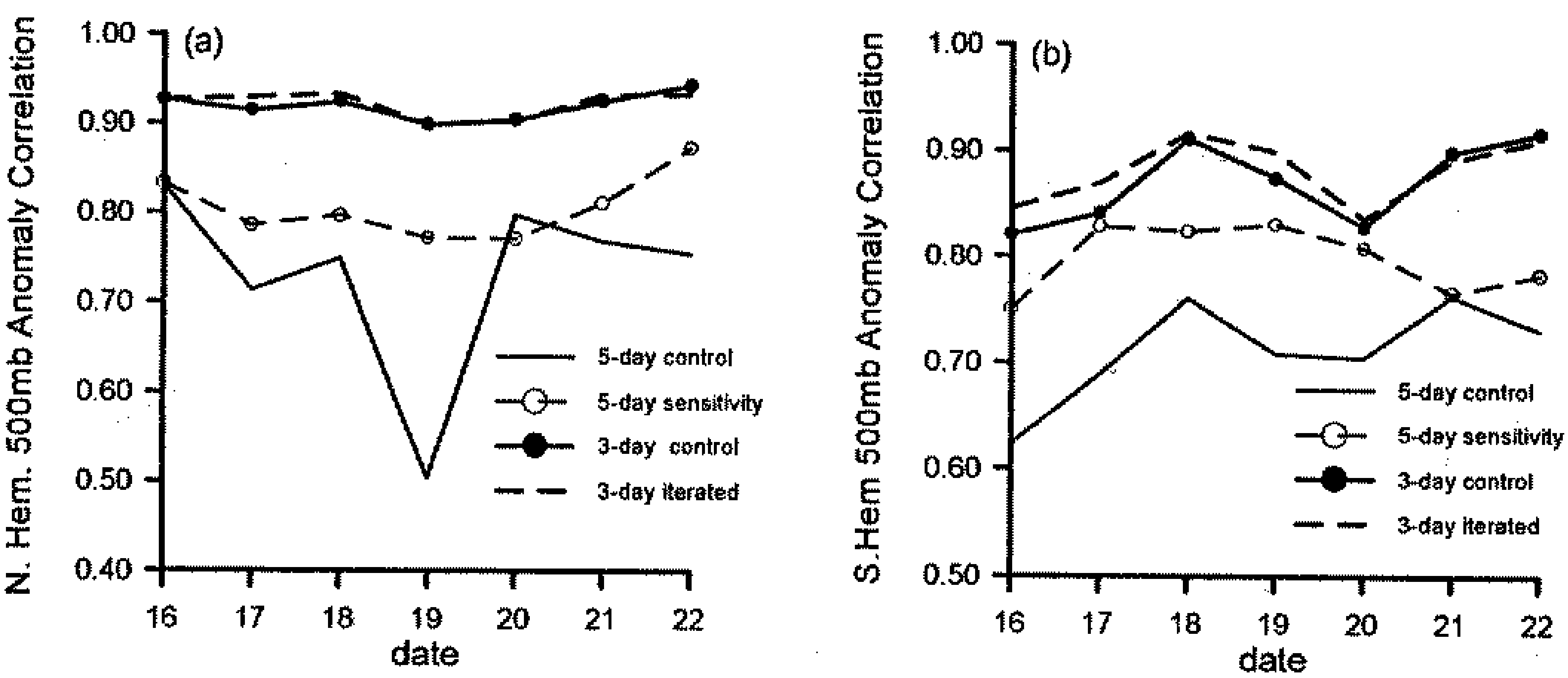

Figure 7. The anomaly correlation scores for the 500 mb geopotential heights, the 5-day sensitivity forecast, the corresponding 3-day control (operational) forecast, the 5-day control forecast, and the 3-day forecast from the new cycle (iterated), for (a) the northern hemisphere and (b) the southern hemisphere. Dates (March of 1995) on the horizontal axis denote the starting dates of forecasts.

starting from the improved initial conditions from two days ago, and that this iterated analysis could then lead to better forecasts from today.

We first tested this idea for seven consecutive cases from 16 to 22 March 1995. At 0000 UTC every day, we calculated the sensitivity perturbation from the 2-day forecast error, then compared the sensitivity 5-day forecasts with the original 5-day control forecasts. Figure 7 shows the comparison of the 5-day forecast anomaly-correlation scores verified against control analyses for 500 mb geopotential height. Not surprisingly, the 5-day sensitivity forecasts (which take advantage of the information available for two additional days in the computation of the 2-day forecast error) are in most cases substantially better than the original 5-day forecasts, in both hemispheres. However, the 3-day forecasts from the regular analysis cycle, also taking advantage of the data gathered over the last two days, show an even larger improvement (Fig. 7). A similar conclusion was reached using 1-day sensitivity forecasts: we found that the 5-day sensitivity forecasts were much better than the 5-day control forecasts, but did not exceed the skill of the corresponding 4-day forecasts (figure not shown).

As indicated above, we tried to solve the lack of improvement upon future forecasts by a simple extension of the sensitivity experiments, in which the sensitivity perturbation of short-range forecasts is used as a starting point for a second iteration in the analysis and forecast cycle for the last two days. We called this an iterated analysis–forecast cycle.

(*b*) *An iterated analysis–forecast cycle using the sensitivity perturbation*

The initial conditions for the operational medium-range global spectral model are provided by the NCEP spectral statistical interpolation (SSI) scheme, a 3-D variational analysis (Derber *et al.* 1991; Parrish and Derber 1992). In the NCEP 6-hour analysis/forecast cycle, the 6-hour forecast from the global spectral model is used as a first guess for the SSI, which finds the 3-D analysis fields that best fit both the first guess and all the observational data gathered for the next 6-hour period.

In the iterated analysis cycle, we introduce the 2-day sensitivity perturbation into the 2-day-old analysis. We then repeat the analysis cycle for the last two days using the improved initial analysis until we reach the end of the sensitivity period. The verifications are made against the control analysis. Figure 7 also shows the anomaly correlation of

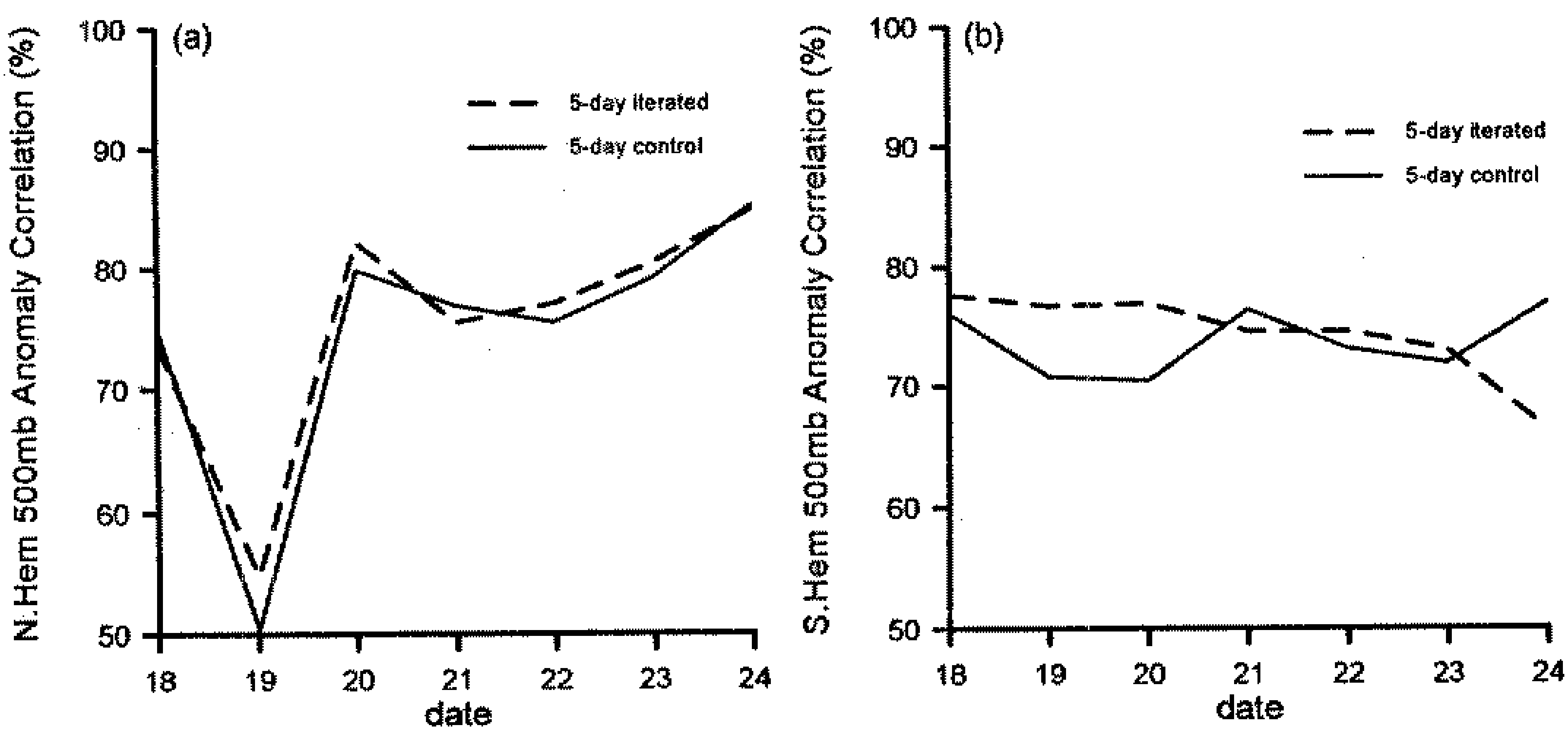

Figure 8. Same as Fig. 7, except 5-day control (operational) forecast and 5-day forecast from the new cycle (iterated).

the 3-day forecast that starts from the new analysis, indicating that, for most cases, the new 3-day forecast is better than the original forecast. Figure 8 shows the 5-day forecasts from the regular and iterated analysis cycle using the 2-day sensitivity perturbation. It is encouraging that, in five out of the seven cases, and in both hemispheres, the new 5-day forecast is equal or better than the regular (operational) forecast.

The conclusion of this experiment is that the iterated cycle has the potential of improving the forecast skill, because it improves the corresponding 3 to 5-day operational forecasts that used the same available information. Similar results were obtained using a 1-day sensitivity perturbation, and using it as starting point for a repeated 1-day analysis cycle (not shown).

(c) *Test of several possible operational configurations of the iterated analysis cycle*

In this final test, we chose a period from 18 March to 1 April 1995. All tests were done in a configuration that could be implemented operationally, and the same procedure was applied every day. The iterated cycle was performed in a continuous fashion as follows:

(i) Starting from 0000 UTC, the forecast error, i.e. the difference between the 0000 UTC analysis field and valid one- or two-day forecast, is taken as an initial condition for adjoint backwards integration to compute the sensitivity perturbation for the one or two-days previous analysis (at $t = -24$ h or $t = -48$ h).

(ii) At $t = -24$ h or $t = -48$ h, we modify the analysis field by the obtained sensitivity perturbation. Then, starting from this point to repeat the SSI analysis cycle until 0000 UTC is reached. The new current analysis (initial conditions) is obtained.

(iii) Starting from the new current initial condition at 0000 UTC, the medium-range forecasts are performed and compared with the forecasts from the original (not iterated) analysis cycle.

(iv) This new analysis is then used to calculate the next perturbation and start the next analysis cycle.

Four possible configurations were tested:

(1) A 24-hour forecast error was used in the one-day iterated assimilation period;

(2) A 48-hour forecast error was used in the two-day iterated assimilation period;

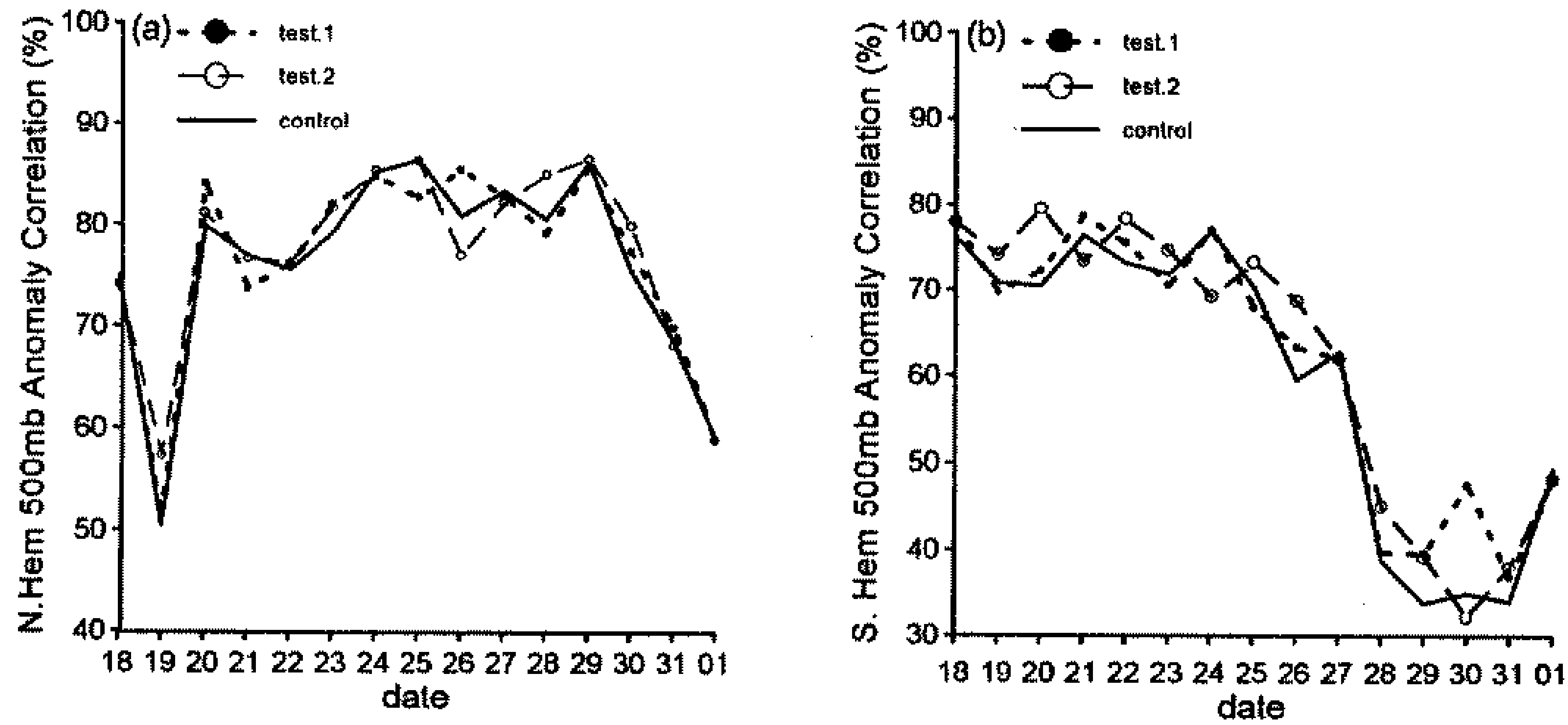

Figure 9. Comparison of 5-day forecast anomaly correlation scores verified against control analyses for 500 mb geopotential height, for the control forecast, and for test 1 and test 2 for (a) the northern hemisphere and (b) the southern hemisphere. Dates (March of 1995) on the horizontal axis denote the starting dates of forecasts.

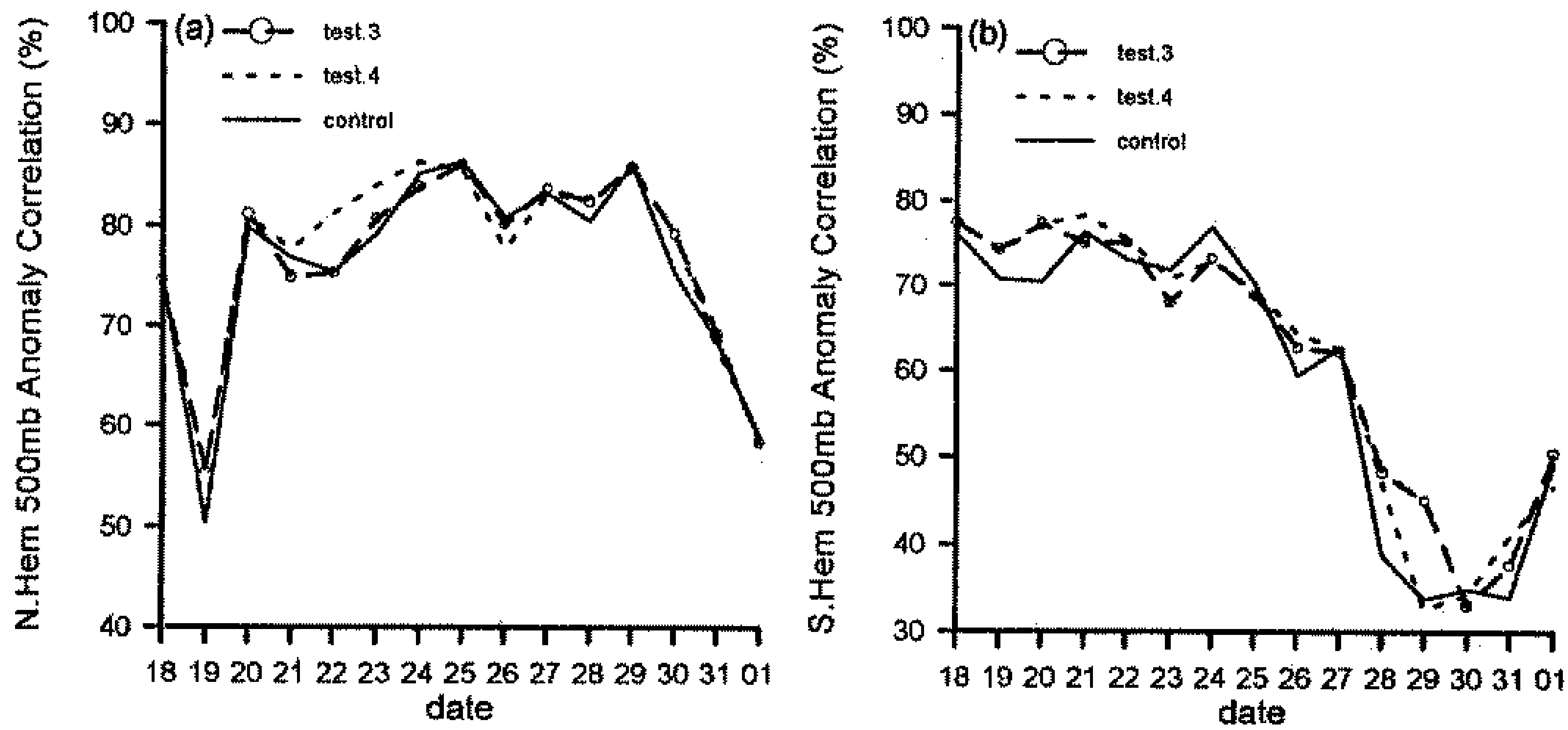

Figure 10. Same as Fig. 9, except for test 3 and test 4.

(3) As in (2), but 24, 36, and 48-hour forecast errors were used to determine the sensitivity perturbation; and

(4) As in (2), but 36 and 48-hour forecast errors were used to determine the sensitivity perturbation.

The individual forecast results corresponding to tests 1 and 2 are presented in Fig. 9, and those of tests 3 and 4 in Fig. 10. The average scores verified against the control analysis over the 15 days of experimental forecasts are presented in Table 2.

The iteration of the analysis cycle results in a small overall improvement with respect to the operational control, for both northern and southern hemispheres, even though it is not successful in every case. Overall, the use of the two-day sensitivity results in more improvements than the use of a one-day sensitivity. This is not surprising, since the errors are better defined after two days for two reasons: (a) the 2-day forecast errors are larger

TABLE 2. COMPARISON OF THE AVERAGE ANOMALY CORRELATION SCORES OF 1–5 DAY FORECASTS FOR GEOPOTENTIAL HEIGHT

	Northern hemisphere 1000 mb					Northern hemisphere 500 mb				
Day	Control	Test 1	Test 2	Test 3	Test 4	Control	Test 1	Test 2	Test 3	Test 4
1	0.972	0.971	0.974	0.973	0.973	0.989	0.989	0.990	0.990	0.990
2	0.937	0.934	0.942	0.938	0.938	0.965	0.963	0.966	0.966	0.966
3	0.884	0.878	0.888	0.886	0.883	0.924	0.920	0.926	0.926	0.925
4	0.788	0.787	0.797	0.796	0.793	0.852	0.849	0.859	0.858	0.859
5	0.685	0.688	0.689	0.691	0.691	0.760	0.763	0.770	0.767	0.774

	Southern hemisphere 1000mb					Southern hemisphere 500 mb				
Day	Control	Test 1	Test 2	Test 3	Test 4	Control	Test 1	Test 2	Test 3	Test 4
1	0.948	0.946	0.952	0.948	0.950	0.971	0.974	0.974	0.975	0.975
2	0.885	0.882	0.894	0.883	0.886	0.919	0.926	0.927	0.928	0.928
3	0.793	0.796	0.806	0.798	0.796	0.833	0.849	0.852	0.850	0.851
4	0.680	0.685	0.699	0.690	0.690	0.721	0.739	0.743	0.744	0.747
5	0.551	0.579	0.574	0.564	0.566	0.599	0.618	0.621	0.616	0.615

Starting dates of the forecasts ranged from 18 March 1995 to 1 April 1995.

and, therefore, the perceived 'forecast errors' are less affected by the unknown analysis errors; and (b) the growing errors are better organized after two days than after one day.

We also found that, at least over the northern hemisphere, the use of the 36-hour forecast errors in addition to the 48-hour errors within the cost function led to somewhat better results, presumably because it also improved the signal-to-noise ratio. In the southern hemisphere, where the analysis uncertainties are larger, the use of only 48-hour errors in the cost function was slightly better.

With respect to the use of just the 48-hour forecast errors or adding the 24- and 36-hour errors in determining the sensitivity patterns, the results seem to favour slightly the use of only the 36- and 48-hour forecast errors in the northern hemisphere. This would be in agreement with the two arguments presented above: by including the 36-hour forecast errors (at a time in which the growing errors are already organized) we increase the signal-to-noise ratios of the cost function. In the southern hemisphere, which is less well observed, and where the uncertainties in the analysis are larger, the results suggest that it may be better only to include the 48-hour forecast errors. However, given that the difference between the one and the two-day iterated assimilation impact is not large, it may still be desirable to perform this procedure operationally with a one-day sensitivity approach, which is half as costly as the two-day approach.

Figure 11 shows the impact of the iterated cycle for a typical 500 mb geopotential height 5-day forecast (0000 UTC 29 March 1995) over the US area. The 5-day control forecast misses a low-pressure cyclone over Utah, but the 5-day new forecast predicted this system much better, even though it is still slightly too far to the south and east.

6. SUMMARY AND DISCUSSION

There is a growing consensus among many leading experts in data assimilation that the single most important problem that has to be addressed in realistic data assimilation is the presence of fast growing (unstable) 'errors of the day' in the initial conditions. In agreement with this consensus, the purpose of this paper has been to identify and partially remove fast-growing errors from the first guess. We presented a simple, relatively inexpensive technique for using past forecast errors to improve the future forecast skill. The method is an extension of the forecast sensitivity studies of RKCH, and can be considered as a very simplified approach to 4-dimensional variational data assimilation.

We first tested the NCEP adjoint system and found that it was necessary to introduce a simplified vertical diffusion scheme into the linear and adjoint NCEP adiabatic models in order to avoid spurious increases of energy in the lower layers of the model. The sensitivity patterns are obtained with this system by minimizing a cost function defined by the one or two-day forecast error norm. Since we are interested in a cost-effective approach, we assume that the first iteration of the minimization algorithm will be enough to achieve a substantial reduction of the forecast error, but not necessarily a true indication of the analysis errors.

Not surprisingly, the sensitivity results obtained with this system are similar to those of RKCH: the 5-day forecasts from the 2-day-old analyses perturbed by the forecast sensitivity pattern (obtained from the 2-day forecast errors) were substantially better than the original 5-day forecasts. However, this important result, which proves that the use of the adjoint model is a powerful tool to estimate errors in the initial conditions, does not by itself lead to improved operational forecasts. When we compared the improved 5-day forecasts from 2-day old initial conditions with the 3-day forecasts from the latest available analyses, the latter were still considerably better.

In order to improve the future forecast skill, we developed an 'iterated cycle' technique, in which we use the perturbed one- and two-day-old analysis as an improved starting point for a repetition of the regular analysis cycle until the current initial conditions are reached. Forecasts from these iterated analyses are in general better than the original forecasts. Several possible operational configurations of this iterated cycle are also tested, indicating that the technique improves the future forecast skill. The technique requires only one iteration of the minimization algorithm per day, as well as an iteration of the analysis cycle for one or two days. Depending on whether one or two-day errors are considered in this procedure, it increases the cost of the analysis cycle by a factor of about 4 to 8, (compared with a much larger increase with full 4-D variational systems) and therefore becomes computationally feasible with present-day systems. It is possible that additional reduction in the cost of this system (and the 4-D variational system) can be made by including further approximations (e.g. lower resolution, etc.).

This study has proven that there is still scope for improvement in the forecast in medium-range forecasts by identifying and partially removing growing errors from the initial conditions. It is important to note that the sensitivity pattern projects very strongly on just a few singular vectors (Buizza *et al.* 1995; RKCH). They can be well approximated by projecting onto the subspace of growing singular vectors. Therefore, the improvement in the initial conditions only involves a few degrees of freedom. When the full 3-D VAR is performed starting from the improved initial conditions, only a few degrees of freedom (O(10)) have been modified, a number negligible compared with the total data available. Therefore, although introducing the sensitivity perturbation to the first guess and to the iterated 3-D VAR analysis introduces an influence of the observation errors, the sensitivity correction does not affect significantly the assumption made in the 3-D VAR about lack of correlation between model and observational errors. The modifications induced by the sensitivity approach are known to be very strongly baroclinic, while the SSI (3-D VAR) structure functions are essentially barotropic. The spherical harmonic energy spectra of the SSI and of the sensitivity patterns are also quite different, with the former being flatter. These arguments also indicate that the iterated cycle does not represent a significant violation of the assumption of uncorrelated model and observational errors in the SSI

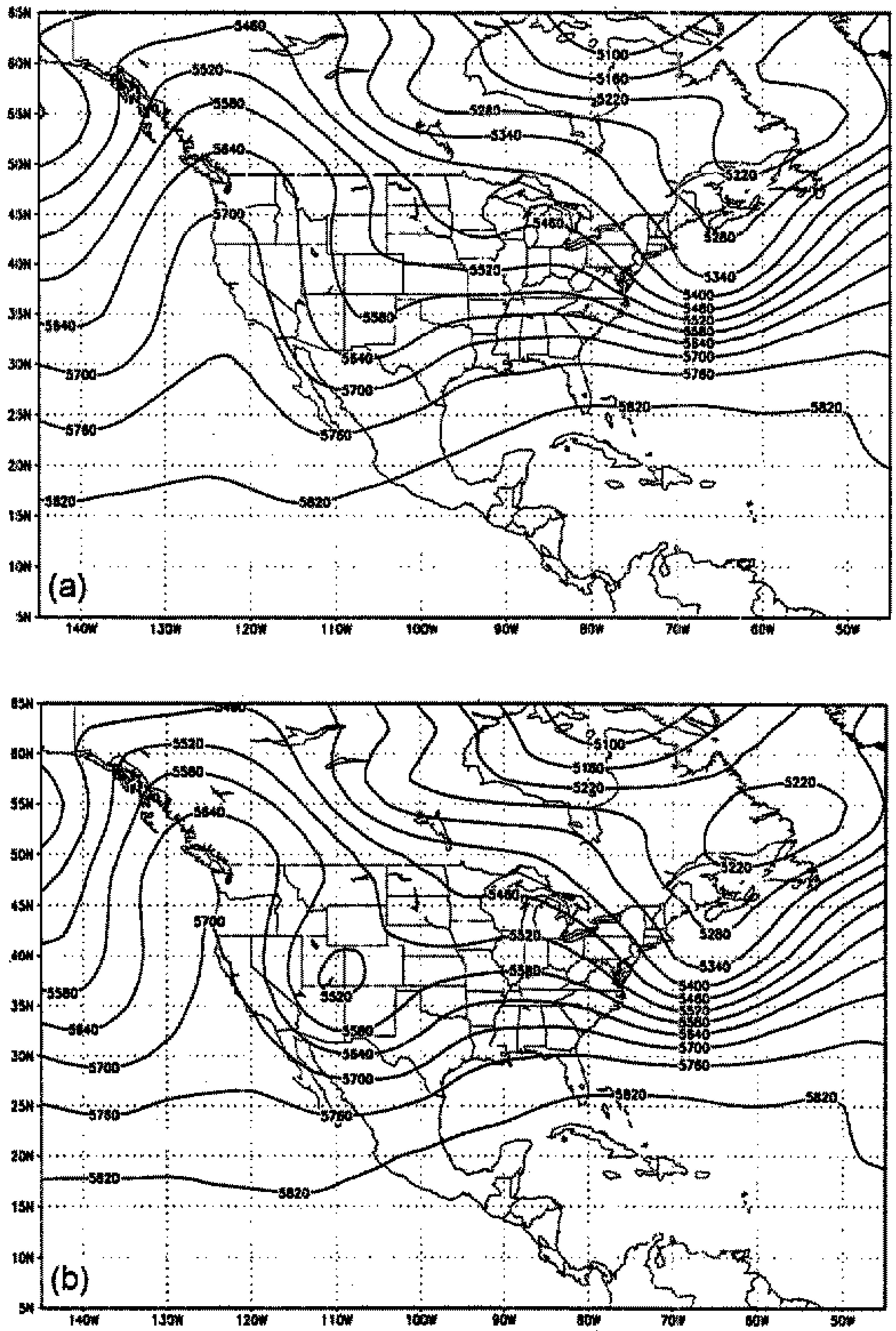

Figure 11. 500 mb geopotential-height field over the US area at 0000 UTC 29 March 1995: (a) 5-day control forecast, (b) 5-day forecast in iterated cycle, and (c) analysis.

system, which is also reinforced by the fact that the procedure leads to the improvement of the forecasts, and therefore to the analysis fields.*

There is increasing interest in methodologies that reduce the computational cost and thus bring the advanced 4-D VAR system in present operations (e.g. Zupanski and Zupanski 1996). We have recently made additional enhancements of the simplified 4-D VAR technique presented here (Pu 1996; Pu *et al.* 1996a; Pu *et al.* 1996b), using the quasi-inverse linear-tangent model, showing further promise for operational implementation at the NCEP.

* While this paper was under revision, we became aware that a similar technique was tested successfully by using a HIRLAM model and its adjoint in a 6 h data assimilation window with an optimum interpolation analysis scheme (Huang *et al.* 1996).

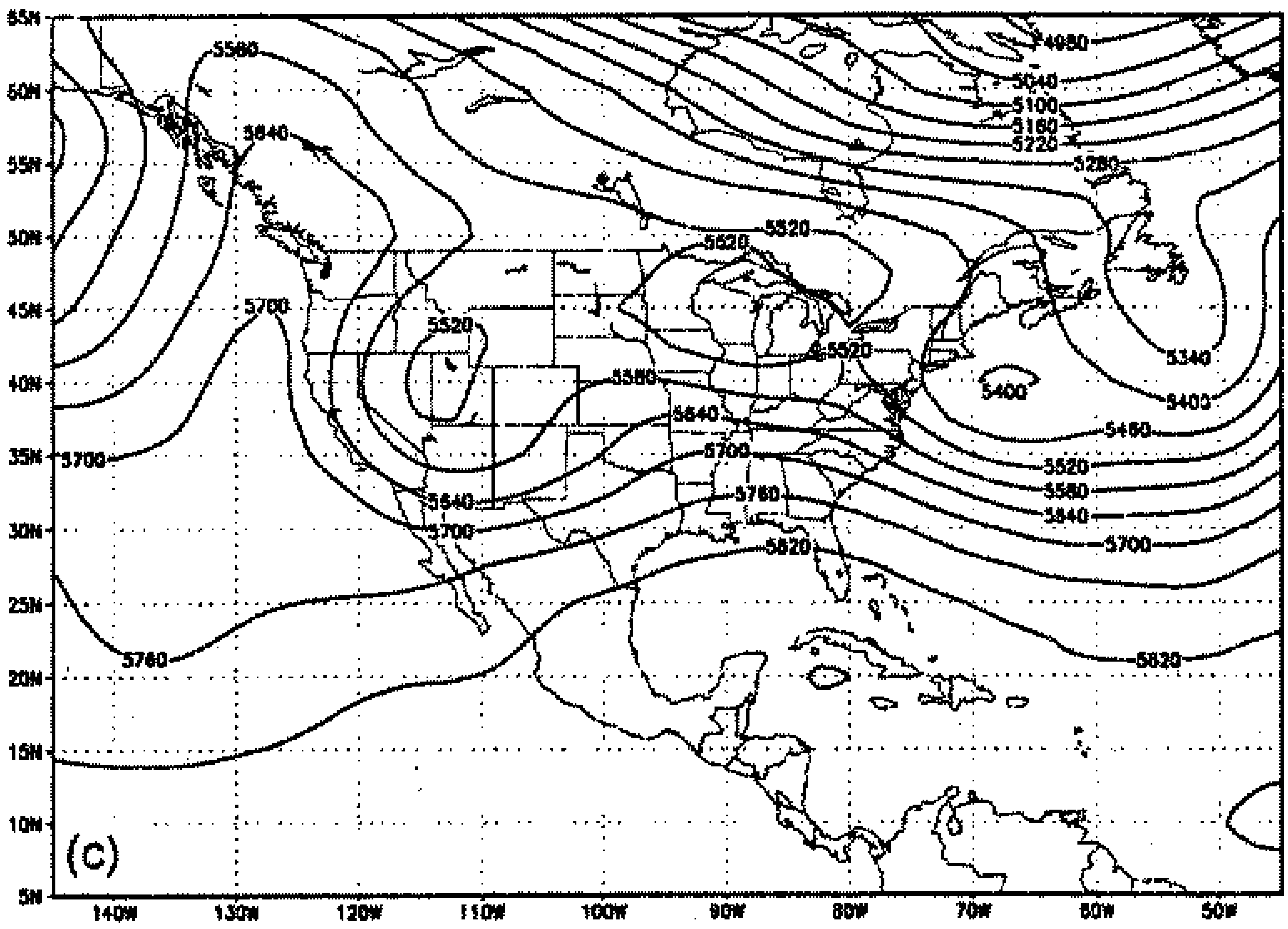

Figure 11. Continued.

Acknowledgements

We would like to thank Drs M. Zupanski, Z. Toth and S.-Y. Hong for very helpful discussions. Mr Yuejian Zhu provided all the data for this research. Drs Dusanka Zupanski, Zoltan Toth and Jean Thiebaux and three anonymous reviewers made helpful suggestions on the manuscript. The experiments presented in this paper constitute part of the first author's Ph.D. dissertation. The first author would like to thank Prof. Jifan Chou in Lanzhou University of P. R. China for his encouragement, and to the UCAR/NCEP Visiting Scientist Program, who supported her stay at the NCEP.

References

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

Andersson, E., Courtier, P., Gaffard, C., Haseler, J., Rabier, F., Undén, P. and Vasiljevic, D. 1996 '3-D VAR—the new operational analysis schemes'. Pp. 2–5 in ECMWF Newsletter No. 71

Buizza, R. 1993 'Impact of a simple vertical diffusion scheme and of the optimization time interval on optimal unstable structure'. ECMWF Technical Memorandum, No. 192

1994 Sensitivity of optimal unstable structures. *Q. J. R. Meteorol. Soc.*, **120**, 429–451

Buizza, R., Gelaro, R., Molteni, F. and Palmer, T. N. 1995 'Predictability studies with high resolution singular vectors'. ECMWF Technical Memorandum, No. 219

Cohn, S. E. and Todling, R. 1996 Approximate Kalman filters for stable and unstable dynamics. *J. Meteorol. Soc. Japan*, **74**, 63–75

Courtier, P. 1995 'Variational methods'. Pp. 25–30 in Proceedings of the 2nd WMO international symposium on assimilation of observation in meteorology and oceanography, Tokyo, March 1995. WMO

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

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994 A strategy for operational implementation of 4-D VAR using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120**, 1367–1387

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

1989 A variational continuous assimilation technique. *Mon. Weather Rev.*, **117,** 2437–2446

Derber, J. C., Parrish, D. F. and Lord, S. J. 1991 The new global operational analysis system at the National Meteorological Center. *Weather and Forecasting*, **6,** 538–547

Derber, J. C., Parrish, D. F., Wu, W.-S., Pu, Z.-X. and Rizvi, S. R. H. 1994 'Improvements to the operational SSI global analysis system'. Pp. 149–150 in Preprints of the 10th conference on numerical weather prediction. American Meteorological Society, USA

Errico, R. M. and Vukicevic, T. 1992 Sensitivity analysis using an adjoint of the PSU-NCAR mesoscale model. *Mon. Weather Rev.*, **117,** 2437–2446

Errico, R. M., Vukicevic, T. and Raeder, K. 1993 Examination of the accuracy of a tangent linear model. *Tellus*, **45A,** 462–477

Gill, P. E., Murray, W. and Wright, M. H. 1981 *Practical optimization.* Academic Press, London

Huang, X.-Y., Gustafsson, N. and Källén, E. 1996 'A poor man's 4-D VAR: The use of the adjoint model in an intermittent data assimilation system'. Pp. 132–134 in Preprints of the 11th conference on numerical weather prediction. American Meteorological Society, USA

Kalnay, E. and Toth, Z. 1994 'Removing growing errors in the analysis'. Pp. 212–215 in Preprints of the 11th conference on numerical weather prediction. American Meteorological Society, USA

Lacarra, J. F. and Talagrand, O. 1988 Short range evolution of small perturbation in a barotropic model. *Tellus*, **40A,** 81–95

Le Dimet, F. X. and Talagrand, O. 1986 Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A,** 91–110

Navon. I. M., Zou, X., Derber, J. C. and Sela, J. G. 1992 Variational data assimilation with an adiabatic version of the NCEP spectral model. *Mon. Weather Rev.*, **115,** 1479–1502

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

Parrish, D. F., Derber, J., Purser, J., Wu, W.-S. and Pu, Z.-X. 1996 The NCEP global analysis system: recent improvements and future plans. *J. Meteorol. Soc. Japan*, (in press)

Pu, Z.-X. 1996 'Sensitivity analysis and data assimilation using an adjoint system of the NCEP operational global spectral model'. Ph.D dissertation. Lanzhou University of P. R. China. (Available from NCEP, USA)

Pu, Z.-X., Kalnay, E., Sela, J. and Szunyogh, I. 1996a Sensitivity of forecast error to initial conditions with a quasi-inverse linear method. *Mon. Weather Rev.*, (in press)

Pu, Z.-X., Kalnay, E., Derber, J. and Sela, J. 1996b 'Using past forecast error to improve the future forecast skill'. Pp. 142–143 in Preprints of 11th conference on numerical weather prediction, American Meteorological Society, USA

Rabier, F., Klinker, E., Courtier, P. and Hollingsworth, A. 1996 Sensitivity of forecast errors to initial conditions. *Q. J. R. Meteorol. Soc.*, **122,** 121–150

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

Thiebaux, H. J. and Morone, L. L. 1990 Short-term systematic errors in global forecasts: their estimation and removal. *Tellus*, **42A,** 209–229

Thone, I. and Mahrt, L. 1986 A simple model of the atmospheric boundary layer: sensitivity to surface evaporation, *Boundary-Layer Meteorol.*, **37,** 129–148

Vukicevic, T. 1991 Nonlinear and linear evolution of initial forecast error. *Mon. Weather Rev.*, **119,** 1602–1611

Zou, X., Navon, I. M. and Le Dimet, F. X. 1992 An optimal nudging data assimilation scheme using parameter estimation. *Q. J. R. Meteorol. Soc.*, **118,** 1163–1186

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

1995 An iterative approximation to the sensitivity in calculus of variations. *Mon. Weather Rev.*, **123,** 3590–3604

Zupanski, D. and Mesinger, F. 1995 Four-dimensional variational assimilation of precipitation data. *Mon. Weather Rev.*, **123,** 1112–1127

Zupanski, M. and Zupanski, D. 1996 'A quasi-operational application of a regional four-dimensional variational data assimilation'. Pp. 94–95 in Preprints of the 11th conference on numerical weather prediction. American Meteorological Society, USA