FOUR DIMENSIONAL VARIATIONAL ASSIMILATION AT ECMWF

Florence Rabier, Jean-Noël Thépaut and Philippe Courtier
European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, UK

Summary : In this paper, some results of 4D-Var assimilations in cycling mode, over a few two-week assimilation periods are presented. 4D-Var is implemented in its incremental formulation, with a high resolution model with the full physical parametrization package to compare the atmospheric states with the observations, and a low resolution model with simplified physics to minimize the cost-function. The comparison of 4D-Var on several assimilation windows (6, 12 and 24 hours) and of 3D-Var (the equivalent of 4D-Var with no time-dimension) over a two-week period shows a clear benefit of using 4D-Var over a 6 or 12-hour window compared to the static 3D-Var scheme. It also exhibits some problems for the forecasts started from the 4D-Var on a 24-hour window. The poorer performance of 4D-Var on a relatively long assimilation window can be partly explained by the fact that, in these experiments, the tangent-linear and adjoint models used in the minimization are only approximations of the assimilating model (lower resolution, crude physics). The error these approximations introduce in the time evolution of a perturbation affects the convergence of the incremental 4D-Var, with larger discontinuities in the values of the cost-function when going from low to high resolution for longer assimilation windows. Additional experiments are performed comparing 4D-Var on a 6-hour window with the 3D-Var system. Two additional two-week periods show a consistent improvement in extratropical forecast scores with the 4D-Var system. The main 4D-Var improvements occur in areas where the 3D-Var errors were the largest. Local improvement can be as large as 35% for the root-mean square of the 5-day forecast error, averaged over a two-week period. A comparison of key analysis errors shows that, indeed, 4D-Var on a 6-hour window is able to reduce substantially the amplitude of its fast-growing error components. The overall fit to observations of analyses and short-range forecasts from 3D-Var and 4D-Var is comparable. In active baroclinic areas, the fit of the background to the data is considerably better for the 4D-Var system, resulting in smaller increments. It appears that in these areas (and in particular over West-Atlantic), 4D-Var is able to use better the information contained in the observations. The ability of 4D-Var to extrapolate some aircraft data in the vertical with a baroclinic tilt is illustrated. Some problems exist in the Tropics and mountainous areas due mainly to a lack of physics in the tangent-linear model. Possible improvements to the system (introduction of more physics, better behaviour of the incremental approach thanks to a line search at high resolution) are also discussed.

1. INTRODUCTION

Four-dimensional variational assimilation (4D-Var) is an advanced algorithm to solve the problem of estimating the atmospheric flow. Given the available information (observations, background, atmospheric model,..) the method consists in minimizing the distance between this

information and the atmospheric states over a period of time (the assimilation window). Most of the applications of 4D-Var up to now have made the underlying assumption of a perfect model, thus looking for a model trajectory which best fits the background and the observations over the assimilation window. The 4D-Var algorithm uses the adjoint equations for the computation of the gradient of the objective function measuring the distance between the model states and the information. 4D-Var was first applied to simple models (Le Dimet and Talagrand, 1986, Lewis and Derber, 1985, Talagrand and Courtier, 1987, Courtier and Talagrand, 1987, Courtier and Talagrand, 1990), before being tested in the context of primitive equations models (Thépaut and Courtier, 1991, Navon et al, 1992 and Zupanski, 1993). Lorenc (1986) and Ghil and Malanotte-Rizzoli (1991) have shown that under the perfect model assumption and when the dynamics are linear, 4D-Var produces the same result at the end of the assimilation period as a Kalman filter (and over the whole assimilation interval as a Kalman smoother). The advantage of 4D-Var over the full Kalman filter is its lower computational cost (one should note, however, that some promising possibilities exist to approximate the Kalman filter, eg Todling and Cohn, 1996). The link between 4D-Var and the Kalman filter implies that 4D-Var, although not computing them explicitly, implicitly uses flow-dependent structure functions within the assimilation window. These structure functions were illustrated by Bouttier (1993) and by Gauthier et al. (1993) for example in the Kalman filter context and by Thépaut et al (1993, 1996) in the 4D-Var context. Just as the Kalman filter can track atmospheric instabilities (eg Ghil and Todling, 1996), 4D-Var performs well in the presence of baroclinic instability (Rabier and Courtier, 1992). Its ability to "fill in" small scales from the observations of larger scales has been investigated in a baroclinic instability context (Rabier and Courtier, 1992) and in a case of fully-developed turbulent dynamics (Tanguay et al, 1996).

These promising results obtained in the framework of single analyses together with the incremental formulation of 4D-Var which makes its implementation computationally feasible on today's computers (Courtier et al, 1994), encouraged us to test the method in an assimilation cycling mode closer to that of numerical weather prediction centres.

In section 2, the 4D-Var set-up in its incremental formulation is presented. Some results on the comparison between three-dimensional variational analysis (3D-Var) and 4D-Var on different assimilation windows are described in section 3. Section 4 concentrates on the comparison between 3D-Var and 4D-Var on a 6-hour window. Developments on the convergence of the incremental 4D-Var method and the inclusion of more physics in the minimization are described in section 5. Finally, concluding remarks are given in section 6. The notations used in the paper follow Ide et al (1997).

2. INCREMENTAL FORMULATION OF 4D-VAR

In its incremental formulation (Courtier et al, 1994), 4D-Var attempts to minimize the objective function J

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i)^T\mathbf{R}_i^{-1}(\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i) \tag{1}$$

with i the time index, $\delta\mathbf{x}$ the increment to be added to the background $\mathbf{x}^b$ at the starting point i=0, $\delta\mathbf{x}(t_i) = \mathbf{M}(t_i, t_0)\delta\mathbf{x}$ the increment evolved according to the tangent linear model, $\mathbf{R}_i$ and $\mathbf{B}$ the covariance matrices of observation and background errors respectively, and $\mathbf{H}_i$ being a suitable linear approximation at time t_i of the observation operator H_i. In eq. 1, the innovation vector is given at each time step t_i by $\mathbf{d}_i = \mathbf{y}_i^o - H_i\mathbf{x}^b(t_i)$, $\mathbf{x}^b(t_i)$ being computed integrating the nonlinear model M from time t_0 to time t_i starting from $\mathbf{x}^b$.

The minimization algorithm used is provided by the Institut de Recherche en Informatique et Automatique (INRIA, France). It is a variable-storage quasi-Newton algorithm described in Gilbert and Lemaréchal (1989). The method uses the available in-core memory to update an approximation of the Hessian of the cost function. In practice, ten updates of this Hessian matrix are used.

The optimum $\delta\mathbf{x}^a$ is added to the background $\mathbf{x}^b$ in order to provide the analysis $\mathbf{x}^a$.

$$\mathbf{x}^a = \mathbf{x}^b + \delta\mathbf{x}^a \tag{2}$$

The nonlinear model M used to compute the innovation vector is the ECMWF operational model with the full set of physical parametrizations and run at relatively high resolution (triangular truncation T106 in the experiments described in this paper). The tangent linear model $\mathbf{M}$ used to propagate in time the increments is run at a lower resolution (T63) and with a very limited set of physical parametrizations (horizontal and vertical diffusion with a simple surface friction, as in Buizza, 1994). As explained in Courtier et al (1994), the physics is far more nonlinear than the dynamics, and the tangent linear approximation might not be valid for the full model for relatively large perturbations. A simple way to account in the final analysis for some non-linearities arising from the physical parametrizations is to define a sequence of minimization problems:

$$\begin{aligned} J(\delta\mathbf{x}^n) = &\frac{1}{2}(\delta\mathbf{x}^n + \mathbf{x}^{n-1} - \mathbf{x}^b)^T\mathbf{B}^{-1}(\delta\mathbf{x}^n + \mathbf{x}^{n-1} - \mathbf{x}^b) \\ &+\frac{1}{2}\sum_{i=0}^{n}(\mathbf{H}_i\delta\mathbf{x}^n(t_i) - \mathbf{d}_i^{n-1})^T\mathbf{R}_i^{-1}(\mathbf{H}_i\delta\mathbf{x}^n(t_i) - \mathbf{d}_i^{n-1}) \end{aligned} \tag{3}$$

with $\delta\mathbf{x}^n$ the increment to be added to the previous estimate $\mathbf{x}^{n-1}$ at the starting point i=0, $\delta\mathbf{x}^n(t_i) = \mathbf{M}(t_i, t_0)\delta\mathbf{x}^n$ the increment evolved according to the tangent linear model.

The innovation vector is given at each time step t_i by $\mathbf{d}_i^{n-1} = \mathbf{y}_i^o - H_i\mathbf{x}^{n-1}(t_i)$, $\mathbf{x}^{n-1}(t_i)$ being computed by integrating the nonlinear model M from time t_0 to time t_i starting from $\mathbf{x}^{n-1}$.

The optimum $\delta\mathbf{x}^n$ is added to the previous estimate $\mathbf{x}^{n-1}$ in order to provide the new estimate $\mathbf{x}^n$.

$$\mathbf{x}^n = \mathbf{x}^{n-1} + \delta\mathbf{x}^n \tag{4}$$

This algorithm thus consists of a pair of nested loops. The outer loop uses the complete model to compute the sequence of innovation vectors at high resolution. The inner loop uses the tangent linear and adjoint of a simpler and lower resolution model to minimize the cost function for the increments calculated with respect to this current estimate of analysis. (Remark: the actual trajectory needed for the inner loop is computed at low resolution with all the physical paprametrizations). For more details on these aspects, the reader is referred to Courtier et al (1994).

In the experiments described in this paper, 3 or 4 outer loops have been used, and for each minimization 15 to 25 iterations of the quasi-Newton algorithm, totalling 60 to 100 iterations in all. With the particular set-up of 4 outer loops with 15 iterations for each minimization, the CPU cost of a 4D-Var on a 6-hour window was found to be about 4 times the cost of a 3D-Var. The formulation of the different terms in the cost function, the structure functions used in $\mathbf{B}$ and the observations used are the same as in the operational 3D-Var system. For the description of the 3D-Var system, see Courtier et al (1997), Rabier et al (1997) and Andersson et al (1997).

The assimilation windows considered in the experiments are the following. For both 3D-Var and 4D-Var on a 6 hour period, observations are used with +/- 3 hours around the synoptic times. For 4D-Var on a 12 hour period, the assimilation is performed from 15 to 03Z and from 03 to 15Z. For 4D-Var on a 24 hour period, the assimilation is performed from 15 to 15Z. Every day a 10-day model run is started at 12Z from the analyses using observations up to 15Z. All the experimentation described in this paper was obtained with a proper cycling of the atmospheric analyses in which the result of the assimilation for one cycle is used as initial point for the next. However, as far as the analysis and background error variances for each cycle are concerned, their computation relies on the Optimal Interpolation scheme operational at ECMWF at the time these experiments were performed. Within one 4D-Var assimilation window, the OI is performed as many times as there are 6-hour periods in the assimilation window to perform the computation of analysis and background error variances at each cycle prior to the main 4D-Var analysis within this window. Compared to an optimal Kalman filter, the background error statistics are then not properly cycled taking into account the tangent linear model. They are taken equal to the statistics produced by the OI for a typical 6-hour forecast, whatever the assimilation window on which the cycle is performed. One should note that the OI is also used for the quality control of the observations in this set-up.

3. COMPARISON BETWEEN 3D-VAR AND 4D-VAR ON DIFFERENT ASSIMILATION WINDOWS: OCTOBER 1993 PERIOD

A first set of experiments was conducted to investigate the performance of 4D-Var compared to its sequential counterpart 3D-Var, when cycled over a two-week period. Four assimilation cycles were run in parallel: 3D-Var, 4D-Var on a 6-hour window, 4D-Var on a 12-hour window and 4D-Var on a 24-hour window. The configuration chosen for the 3D-Var analysis is the so-called First-Guess at the Appropriate Time (FGAT), which means that the innovation vector is computed by a model integration over the 6-hour assimilation window (exactly as in 4D-Var on a 6-hour period). Then, the analysis increments are evaluated at the central synoptic time and added to the background at this central time. The assimilation period extends from the 1st October 1993 to the 13th October 1993. The horizontal resolution of the model is T106/T63 as explained previously. The vertical resolution is 19 levels. The set-up of the incremental 4D-Var is the following: 25 iterations are performed for each minimization stage, and 3 outer loops (4 in the case of the 4D-Var on a 24-hour window) giving a total of 75 iterations (100 for 4D-Var on a 24-hour window). The 3D-Var is run with 70 iterations. The code version (cycle 12r1) used for the model and the analysis was operational in 1994. The formulation of the background term in the cost function is known to be defective in the Tropics (this was before the introduction of a univariate tropical formulation, as explained in Andersson et al, 1997).

The extratropical scores averaged over thirteen forecasts are given in Figure 1. All 4D-Var configurations are seen to perform reasonably. 4D-Var on a 6-hour and 12-hour windows are seen to perform the best, with a better performance than 3D-Var at all ranges in both hemispheres. In contrast, 4D-Var on a 24-hour window does not perform quite as well as any of the other assimilations, being in particular worse than the forecasts from the 3D-Var scheme up to day 4. On the one hand, the good performance of 4D-Var on 6-hour and 12-hour windows is encouraging; on the other hand, the poor performance of 4D-Var on a 24-hour window is disappointing as earlier idealized studies on single cases had shown that with the 24-hour window the dynamics had more effect than with a 6-hour window (see for instance Bouttier, 1993, Thépaut et al, 1996). Moreover, in Rabier and Courtier (1992) and Tanguay et al (1996), it was shown that it was beneficial to have a reasonably long window in order to reconstruct the small scales from the observations of the large scales. In cycling mode, it is also known that a full Kalman filter, performs better than a 3D-Var cycling (eg Todling and Cohn, 1994).

However, the way 4D-Var is implemented in these experiments involves three major approximations to an optimal implementation following the full Kalman filter. Firstly, the model error is not taken into account: the assimilating model is assumed to be perfect. Secondly, the tangent linear and adjoint models used in the minimization are only approximations of the assimilating model (lower resolution, very simplified physics). Thirdly, the statistics are not cycled according to the Kalman filter, from one assimilation period to the next. The background error covariance matrix at the starting time of the assimilation window is thus static.

The impact of this third approximation has not been investigated. A simplified Kalman

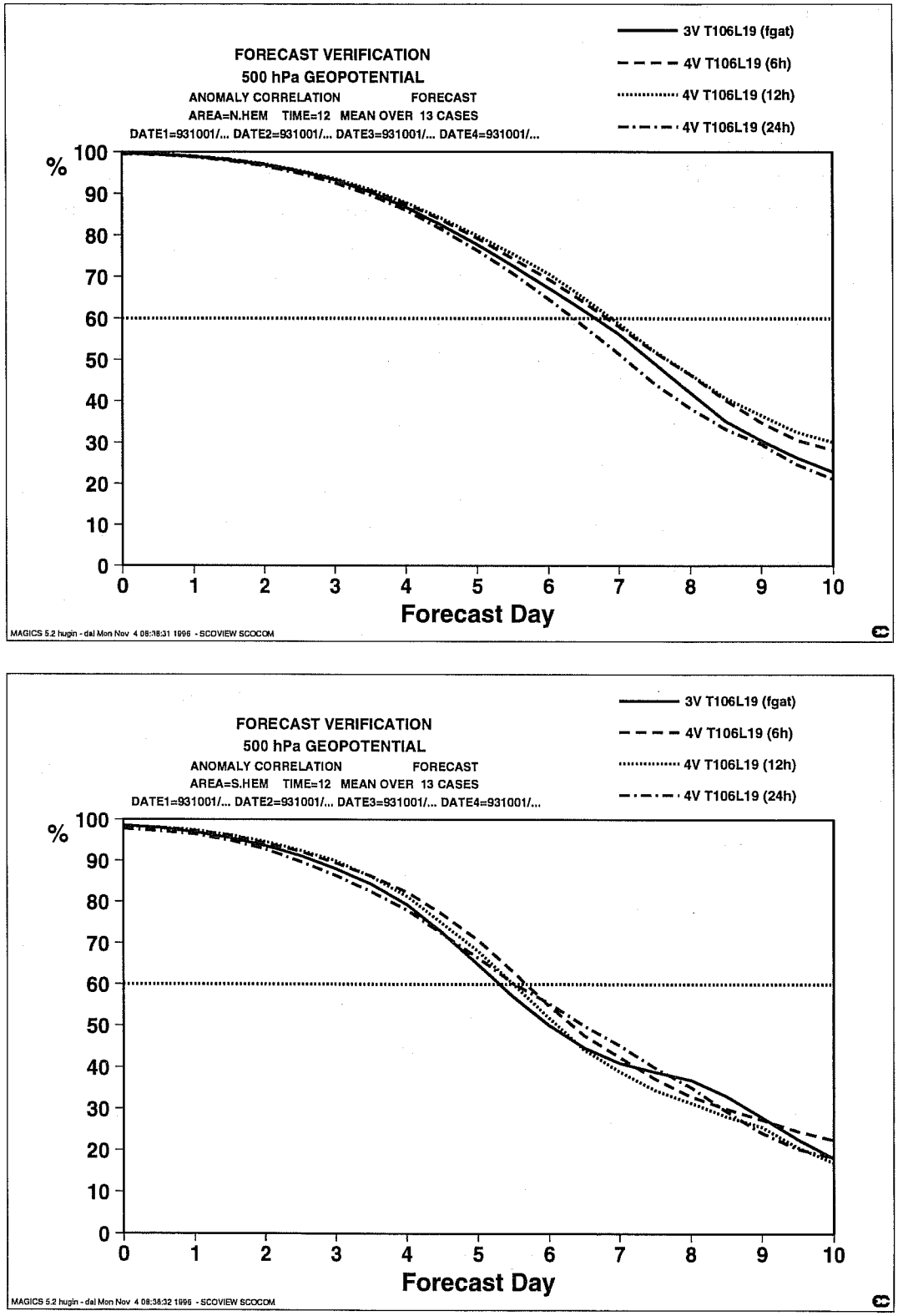

Figure 1: Anomaly correlation for the Geopotential height at 500 hPa, computed over 13 consecutive forecasts from 931001 for the Northern Hemisphere (top panel), and the Southern Hemisphere (bottom panel). 3D-Var is represented by a solid line, 4D-Var on a 6-hour window by a dashed line, 4D-Var on a 12-hour window by a dotted line and 4D-Var on a 24-hour window by a dash-dotted line

filter would be needed to perform more elaborate experiments. The second approximation is easier to document. In order to illustrate the impact of using a simplified model for the evolution of perturbations, an experiment was performed on a more recent case (25th August 1995). A perturbation typical of a key part of the analysis error was calculated by performing three iterations of the minimization of the short-range forecast error as in the sensitivity experiments by Klinker et al (1997). This perturbation is known to be smaller than the full analysis error, while able to explain a significant part of the two-day forecast error. The evolution of this perturbation in time was investigated for three different models: the T106L31 model with full physical parametrizations, the T63L31 model with full physical parametrizations,and the T63L31 model with simplified physics (as described in the previous section). The tangent-linear evolution has been approximated by finite-differences in all cases, substracting the result from a control forecast to the one obtained from a perturbed forecast. Figure 2 shows the evolution of this perturbation after 6 hours (top panels) and 24 hours (bottom panels), for T106L31 model with full physical parametrizations and the T63L31 model with simplified physics.

For both ranges, one can notice a great similarity between the perturbations evolved with the high-resolution model (left panels) and the simplified one (right panels). However, after 24 hours, the differences start to be visible. In particular, a pattern over Asia (45N, 90E) is too intense in the perturbation evolved with the simplified model (right-hand side panel). This particular deficiency of the T63L31 simplified model comes from a lack of physics and not from a lack of resolution (the T63L31 model with full physics ,not shown, did not display this deficiency to such an extent). The results might be slightly different if the T63L31 simplified model had used a trajectory computed with full physics, as is actually done in the 4D-Var set-up. In terms of global root-mean-squares (RMS) in the Northern hemisphere North of 30N, Figure 3 illustrates the RMS of the evolved perturbations for the three different models and the RMS of the error for the two low resolution models as a function of time. The error is defined as the difference between the perturbation evolved with the low resolution model (with full physics or simplified physics) and the perturbation evolved with the reference model (T106L31, full physics). The reference model exhibits a substantial growth of the perturbation, from an RMS value of 20m at a 6-hour range to a 50m value at a 24-hour range. The T63L31 model with full physics tends to under-estimate the perturbation growth, whereas the T63L31 model with simplified physics overdoes it. The RMS error for the T63L31 model with simplified physics is increasing from 5m (20%) at the 6-hour range to 15m (30%) at the 24-hour range. Thus, although the correlation between the fields is still high at the 24-hour range (more than 0.95), the relative error has increased compared with the 6-hour range. The values are slightly lower for the T63L31 model with full physics.

On this example, the approximated evolution of a typical fast-growing perturbation of the analysis has been illustrated and quantified, showing how this approximation is less accurate for longer time ranges. The question now is how this can affect the convergence of the incremental 4D-Var. In the incremental set-up, the discrepancy between the high-resolution model used for the computation of the innovation vectors in the outer loops and the low-resolution model used for the minimizations (inner loops) implies that there is a jump in the value of the cost-function

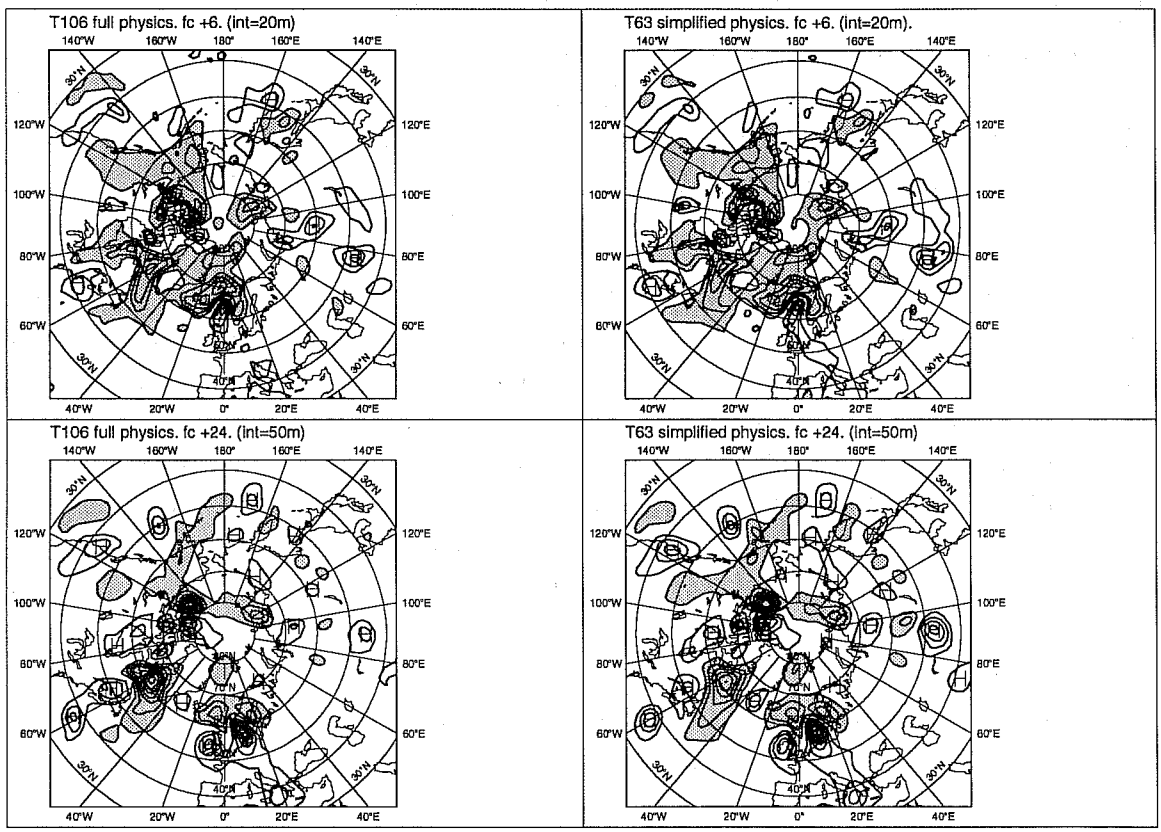

Figure 2: Evolution of a perturbation consisting of key analysis errors for the 25th August 1995, for the geopotential height at 500 hPa. The evolution after 6 hours is shown in the top panels, with a contour interval of 20m. The evolution after 24 hours is shown in the bottom panels, with a contour interval of 50m. The left-hand side panels are from the T106L31 model with full physics and the right-hand side panels are from the T63L31 model with simplified physics. Negative contours are shaded.

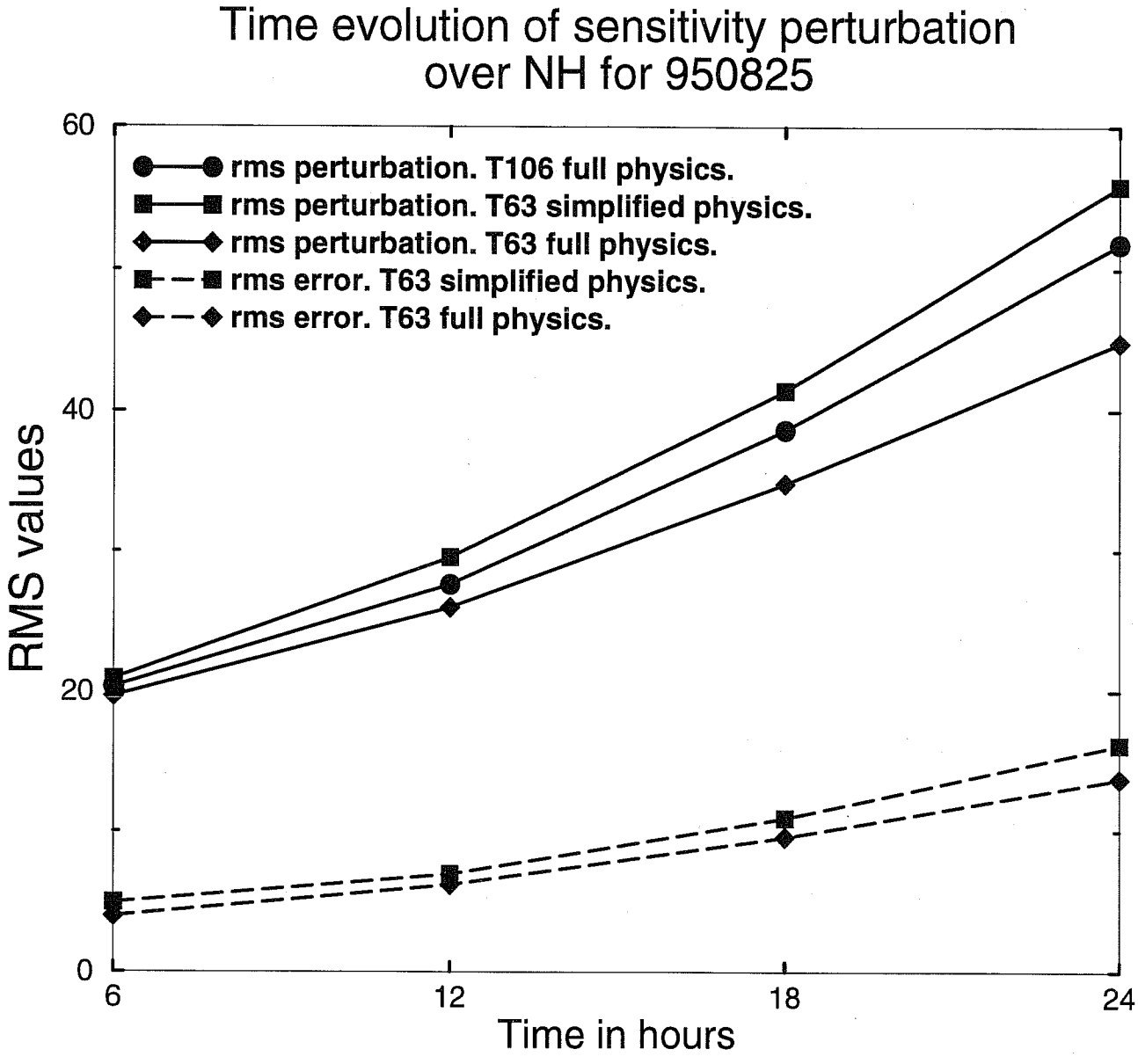

Figure 3: Time evolution of the sensitivity perturbation obtained by the minimization of the short-range forecast error for the 25th August 1995. The perturbation is evolved with three different models: the T106L31 model with full physics (shown as circles), the T63L31 model with simplified physics (shown as squares) and the T63L31 model with full physics (shown as diamonds). The RMS of the perturbations are shown as solid lines for the three models, and the RMS of the errors are shown as dashed lines (these are defined as the difference between the perturbation evolved with the low resolution models and the perturbation evolved with the reference model T106L31with full physics). The RMS are computed over the Northern hemisphere, defined as latitudes higher than 30N. The values are in geopotential meters for the 500hPa geopotential height field.

when going from low to high resolution due to the change of resolution and the mismatch in physical parametrizations. For each outer loop, the value of the cost-function at low resolution is given by eq. 3 which we repeat here

$$J_l(\delta\mathbf{x}^n) = \frac{1}{2}(\delta\mathbf{x}^n + \mathbf{x}^{n-1} - \mathbf{x}^b)^T\mathbf{B}^{-1}(\delta\mathbf{x}^n + \mathbf{x}^{n-1} - \mathbf{x}^b) + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{H}_i\delta\mathbf{x}^n(t_i) - \mathbf{d}_i^{n-1})^T\mathbf{R}_i^{-1}(\mathbf{H}_i\delta\mathbf{x}^n(t_i) - \mathbf{d}_i^{n-1}) \quad (5)$$

whereas its value at high resolution after minimization is

$$J_h(\delta\mathbf{x}^n) = \frac{1}{2}(\delta\mathbf{x}^n + \mathbf{x}^{n-1} - \mathbf{x}^b)^T\mathbf{B}^{-1}(\delta\mathbf{x}^n + \mathbf{x}^{n-1} - \mathbf{x}^b) + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{d}_i^n)^T\mathbf{R}_i^{-1}(\mathbf{d}_i^n) \quad (6)$$

with $\mathbf{d}_i^{n-1} = \mathbf{y}_i^o - H_i(\mathbf{x}^{n-1}(t_i))$, and $\mathbf{d}_i^n = \mathbf{y}_i^o - H_i(\mathbf{x}^{n-1}(t_i) + \delta\mathbf{x}^n(t_i))$. The first part of these equations (fit to the background) is actually computed only at low resolution and the same value is used at high resolution.

If similar models and similar resolutions were used in inner and outer loops these cost-functions would be identical (in the limit of validity of the tangent-linear hypothesis for the observation operator). But, with different models, there is a mismatch which is illustrated in Figure 4. The ratio $\frac{J_h - J_l}{J_l}$ is shown as a function of minimization number (or equivalently outer loop index) for the four assimilation systems, averaged over the 13 days of assimilation.

In the three-dimensional case, when the change of resolution affects the computation of the observation departures with no model integration, there is 16% jump after the whole minimization (70 iterations). In the 4D-Var cases, the jumps are computed after partial minimizations (25 iterations each). These ratios are a decreasing function of the minimization number, which illustrates the convergence of the incremental approach. One can also notice that, the shorter the 4D-Var window, the smaller the jumps. This can be seen as a consequence of a better approximation for shorter time ranges. In particular, 4D-Var on a 24-hour window is clearly performing worse than the other systems. Even after an additional outer loop (minimization number 4), the jump is still around 5%, whereas it is already less than 4% for 4D-Var on a 6 and a 12 hour windows after the third minimization. This worse convergence of 4D-Var on a longer period of 24 hours, due to approximations in the tangent-linear model is certainly an indication that these approximations are reaching their limit of validity in this set-up. It is not believed that performing another minimization would have helped in this case, as other indicators showed that the cost-function was not decreasing anymore, although is is likely that more iterations would be needed for a relatively long assimilation window. In order to get the

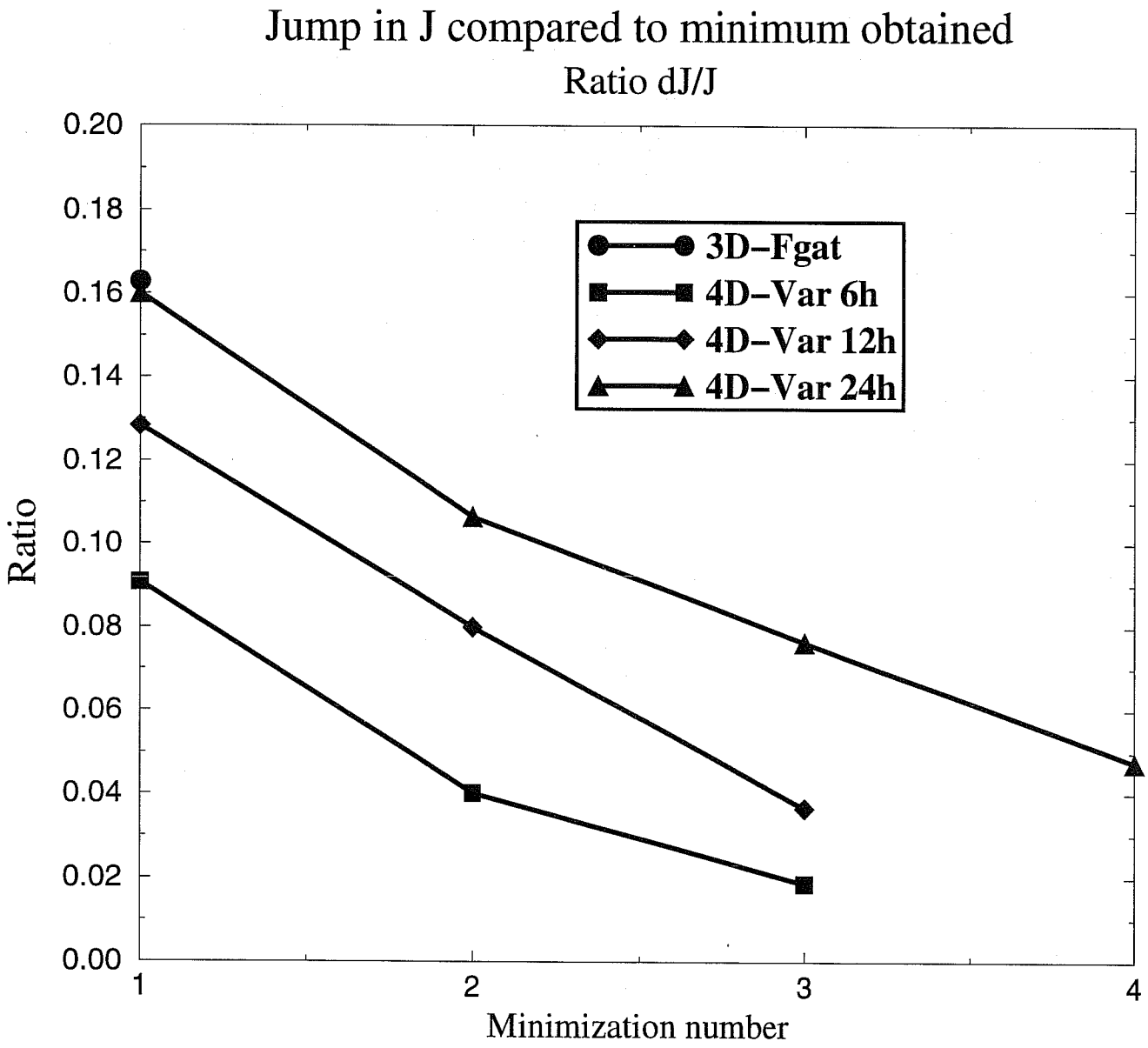

Figure 4: Ratio $\frac{J_h - J_l}{J_l}$ as a function of the number of minimization steps for the four assimilation systems, averaged over the 13 days of assimilation. 3D-Fgat is shown in circles, 4D-Var on a 6-hour period in squares, 4D-Var on a 12-hour period in diamonds and 4D-Var on a 24-hour period in triangles.

benefit of performing more iterations, the current set-up should be changed to perform more iterations within each minimization for instance. This has not been investigated.

Finally, an approximation more difficult to illustrate is the perfect model assumption. It is known that the model is inaccurate, some of its deficiencies are documented but the model error is far from being known precisely. Some figures on the systematic part of the error indicate that this part grows in time by a factor of around two between a 6-hour and a 24-hour range (Ernst Klinker, pers comm), but this is only a fraction of the total error computed by taking the long-term average of forecast errors. The impact of model error on advanced assimilation schemes has been studied by a number of authors (Wergen, 1992, Miller et al, 1994, Cohn et al, 1994). In practice, results on 4D-Var on a period of 12 hours have benefitted from the introduction of a model error term (Zupanski, 1993, Zupanski, 1996), while 4D-Var on short periods (2 to 3 hours) with the perfect model assumption has been quite successful (Zou et al, 1995, Kuo et al, 1995). In the present experiments, there is an indication that 4D-Var on a 24-hour period does not fit the data within the observational error, which shows that the assimilation is defective. Furthermore, the fit to the data within an assimilation window tends to be worse at the beginning and at the end of the period. For instance, over the last 24-hour assimilation including the synoptic times: 931012, 18Z; 931013, 00Z; 931013, 06Z and 931013, 12Z, the ratio of the RMS fit to synop surface pressure data between the 4D-Var on a 24-hour analysis and the corresponding 4D-Var on a 6-hour period analyses are 1.32, 1.23, 1.36 and 1.45. The same type of behaviour had already been observed in previous 4D-Var experiments (Courtier and Talagrand, 1990). In theory, in a dynamical system with both stable and unstable modes, the error in a perfect model 4D-Var analysis should be minimum around the middle of the period, where there is more information, as explained in Pires et al (1996). Thus, the fit to the data, which is basically the observational error minus the analysis error should be worse around the middle of the period in this perfect model context, exhibiting a concave shape. In the case with model error, a recent paper by Ménard and Daley (1996) also illustrates such a concave shape in the cost-function in the case of an optimal analysis taking into account the model error, whereas the 4D-Var making the perfect model assumption exhibits a convex cost-function. It then seems that an optimal system should present a concave shape in the cost function, which is not the case in our 4D-Var in a 24-hour period. More investigations should be performed to confirm the link between the curvature of the cost-function and the non-validity of the perfect model assumption.

Two of the approximations made in this set of experiments will be addressed in the near-future. The proper cycling of the background error statistics will be tackled in a simplified Kalman filter context (Fisher and Courtier, 1995). The approximations made in the tangent-linear and adjoint models will be partly solved by introducing more physics (Mahfouf et al, 1996) , as explained in section 5. It will then be appropriate to test again the quality of 4D-Var on a relatively long period (24 hours), and see if the model error approximation is still sufficiently detrimental to outweigh the benefits of a longer assimilation window. Possibilities to introduce model error in 4D-Var exist and will be investigated (Courtier, 1996). In the shorter term, the scientific question which will be investigated is the better performance of the

4D-Var on a 6 hour window over a 3D-Var scheme.

4. COMPARISON BETWEEN 3D-VAR AND 4D-VAR ON A 6-HOUR WINDOW OVER TWO TWO-WEEK PERIODS

4.1 Extratropical scores

Two extended experiments were performed comparing 4D-Var on a 6 hour window and 3D-Var: two weeks in August 1995 (24th August to 7th September) and two weeks in January 1996 (16th to 29th). In these experiments, the vertical resolution is 31 levels, the number of outer loops is 4 and the number of iterations for each minimization is 15 leading to a total of 60 iterations for each cycle of the 4D-Var algorithm. (This particular minimization set-up was found to give good convergence in these cases). The 3D-Var minimization still uses 70 iterations as in the previous experiments. The version of the code used for the model and the analysis is different from the previous set of experiments (it is now cycle 14r3, which was used operationally in 1996).

The extratropical forecasts scores averaged over the two periods are presented in Figure 5. The forecasts started from 4D-Var on a 6-hour window are always performing better than the forecasts from the 3D-Var analyses. In the Northern hemisphere, both periods are positive (the January period being the better of the two, not shown), giving an overall advantage to 4D-Var. In the Southern hemisphere, the August period is positive for 4D-Var, whereas the January period is more neutral, combining into a small positive impact of 4D-Var. It thus seems that in the winter hemisphere 4D-Var performs much better than 3D-Var, with a more neutral impact in the summer hemisphere. However, it should be noted that 4D-Var did not produce worse scores than 3D-Var (when averaged over two weeks) for either hemisphere and either of the periods investigated. In general, the scatter between the two experiments is reasonably small, showing mostly better forecasts for 4D-Var and very few slightly negative. The day-to-day consistency between forecasts produced by the 4D-Var analyses is also improved compared to that of the 3D-Var (not shown). The geographical distribution of the improvement brought by 4D-Var in the Northern hemisphere in the medium range (day 5) is shown in Figure 6 for the geopotential height at 500 hPa. The top panels are relative to the August period, the bottom ones to the January period. The left-hand side panels represent the RMS error for the 5-day forecasts started from the 3D-Var analyses. The right-hand side panels represent the difference between the RMS error for the 5-day forecasts started from the 3D-Var analyses and the RMS error for the 5-day forecasts started from the 4D-Var analyses (positive values indicating where 4D-Var performs better than 3D-Var are shaded). For both periods, there is quite a good correspondence between the areas of improvement brought by 4D-Var and the areas where the forecasts issued from 3D-Var had large errors. In particular, for the summer period (top panels), 4D-Var brings an improvement of more than 30m in Northern Europe, where the errors from the 3D-Var forecasts were around 140m (local improvement of around 25%). Similarly, for the winter period (bottom panels), 4D-Var brings an improvement of around 60m or more in

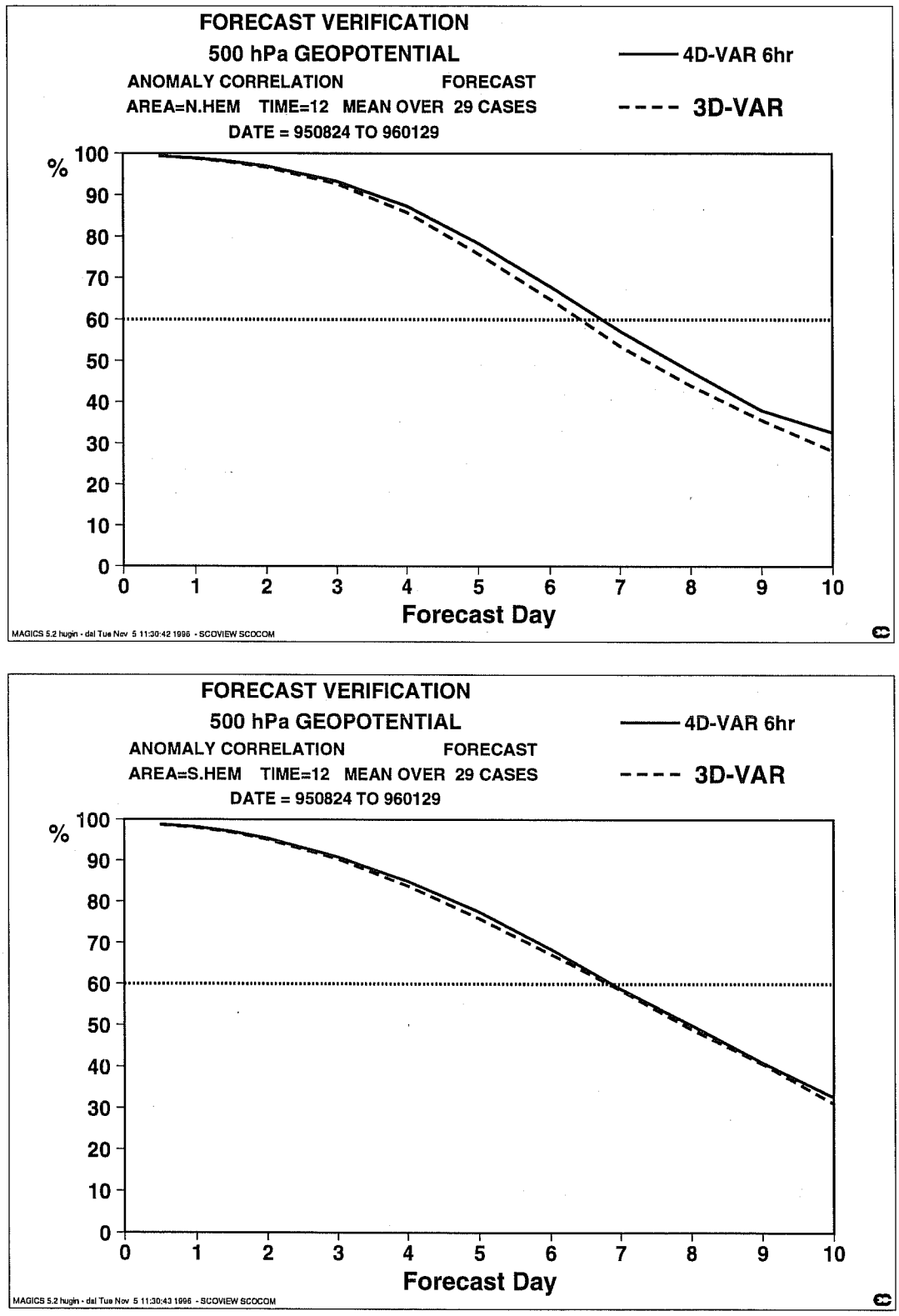

Figure 5: Anomaly correlation for the Geopotential height at 500 hPa, computed over 29 forecasts (14 consecutive days in August 1995 and 15 consecutive days in January 1996) for the Northern Hemisphere (top panel), and the Southern Hemisphere (bottom panel). 3D-Var is represented by a dashed line, 4D-Var on a 6-hour window by a solid line

the Pacific area where the errors from the 3D-Var forecasts reached 190m (local improvement of around 35%). The forecast performance brought by 4D-Var though not large on averaged scores, is always positive and is largest in the areas of large errors; the improvement reaches 30% locally.

The fact that 4D-Var reduces the largest errors implies that the analyses are mainly different in components which have a relatively high growth rate. We have compared the analysis changes made by 4D-Var and the key analysis errors computed operationally (Rabier et al, 1996, Klinker et al, 1997). The correspondence is quite good, in particular for the January period in the Central Pacific, Hudson Bay, and Labrador sea areas. Comparisons in a case study are illustrated in the following section.

4.2 Comparison with key analysis errors

Key analysis errors are computed for the 25th August 1995 based on the two-day forecast error. These represent the part of the analysis error which grows substantially in the following forecast, as explained in Klinker et al (1997). Besides the control forecast started from the control analysis, a sensitivity integration is also run from the analysis from which the key analysis errors are subtracted. The medium-range forecast error is noticeably reduced, particularly over Europe in the sensitivity integration. The question one might ask is whether 4D-Var on a 6-hour period is able to cancel some of this fast growing component of the error. Both the 3D-Var and the 4D-Var systems are cycled for 24 hours, from the 25th August 1995 at 12Z to the 26th August 1995 at 12Z. Then, two forecasts are run from the resulting analyses. The divergence of the two pairs of forecasts (4D-Var versus 3D-Var and sensitivity versus control) is remarkably similar for the 31st August 1995, as shown in Figure 7 over Europe. Both the 4D-Var and the sensitivity bring changes of around 150m at 60N, 30E, thus reducing the error by respectively 30% and 40% from an initial error of about -300m for both the control experiment and the 3D-Var forecast. This error can be traced back to a comparable deepening of the low over the Labrador sea on the 26th August 1995: 9.5m for 4D-Var and 14m for the sensitivity perturbation in the area 57N, 55W for the geopotential at 500 hPa. The 3D-Var and 4D-Var analyses are shown in Figure 8 together with the radiosonde observations at the 500 hPa level. One can notice that the 4D-Var analysis (dashed lines) is slightly deeper around 60W than the 3D-Var analysis (solid lines). This change, although very small, is compatible with the available observations, and is located in a very sensitive baroclinic area. In this comparison, it was not expected to find a very good correspondence between 4D-Var and the sensitivity, as the two systems have used very different information, and in particular the sensitivity integration used the analysis one day ahead of the observations used in both 3D-Var and 4D-Var. However, it is known that this sensitivity integration captures a key part of the analysis errors and this comparison illustrates that 4D-Var can also capture a fraction of it.

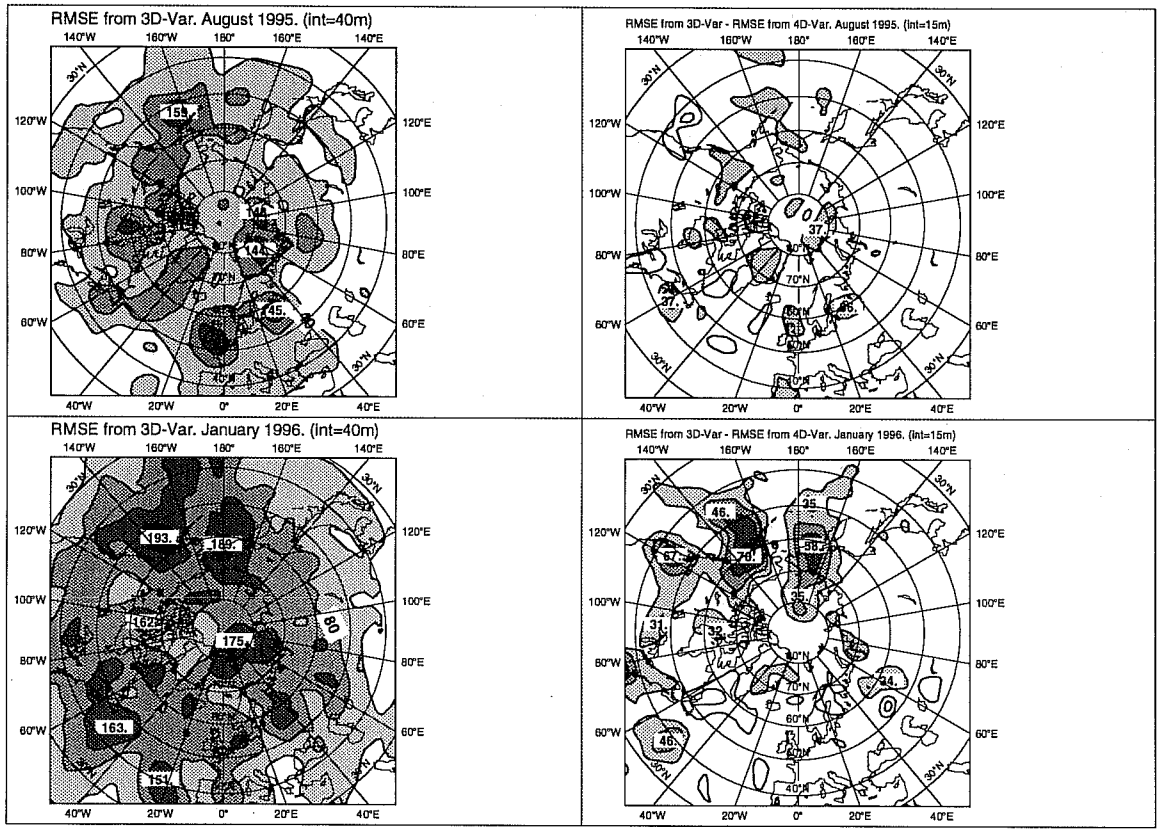

Figure 6: Geographical distribution of the forecasts errors in the medium range (day 5) for the geopotential height at 500 hPa. The top panels are for the August 1995 period, the bottom ones for the January 1996 period. The left-hand side panels represent the RMS error for the 5-day forecasts started from the 3D-Var analyses (contour interval is 40m, shading starts at 40m). The right-hand side panels represent the difference between the RMS error for the 5-day forecasts started from the 3D-Var analyses and the RMS error for the 5-day forecasts started from the 4D-V ar analyses (contour interval is 15m, positive values indicating where 4D-Var performs better than 3D-Var are shaded for values above 15m).

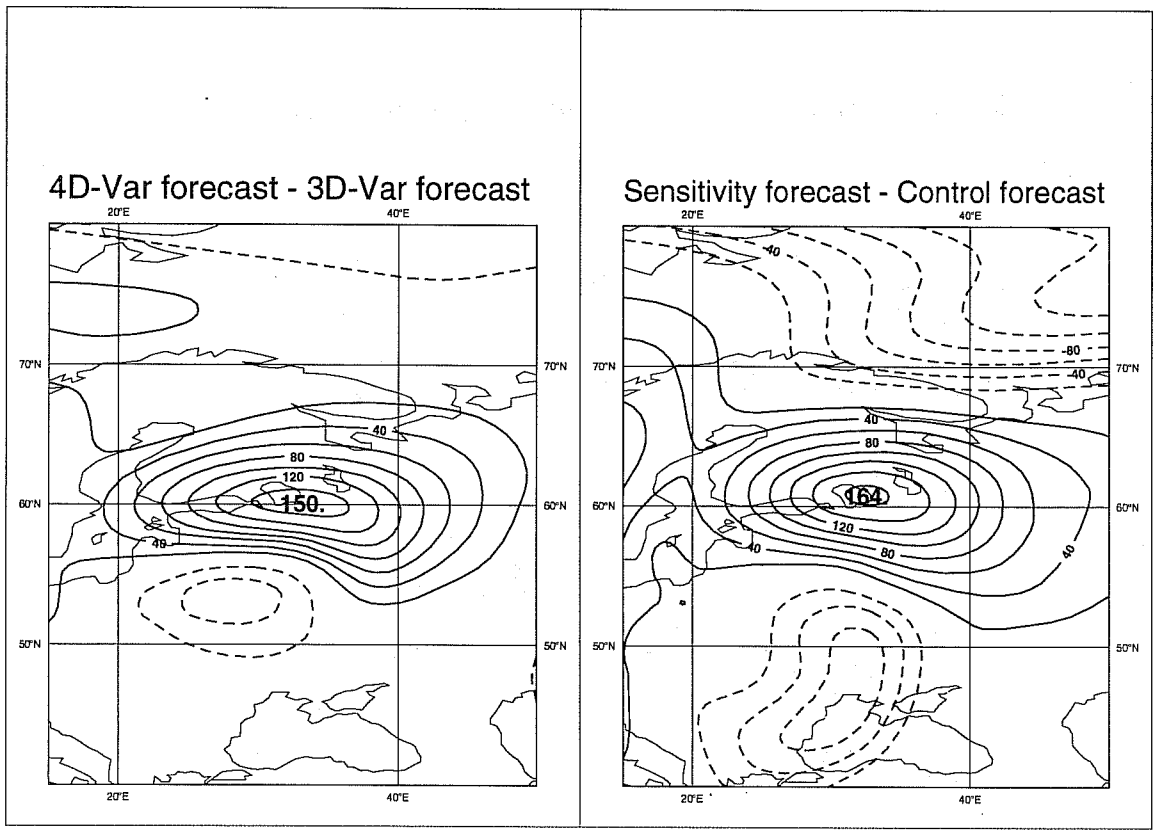

Figure 7: Divergence of forecasts for the geopotential height at 500 hPa valid for the 31st August 1995. The left-hand side panel represents the difference between the forecasts started from 4D-Var and from 3D-Var on the 26th August 1995. The right-hand side panels represents the difference between the sensitivity integration and the control forecast started on the 25th August 1995. Contour interval is 20m.

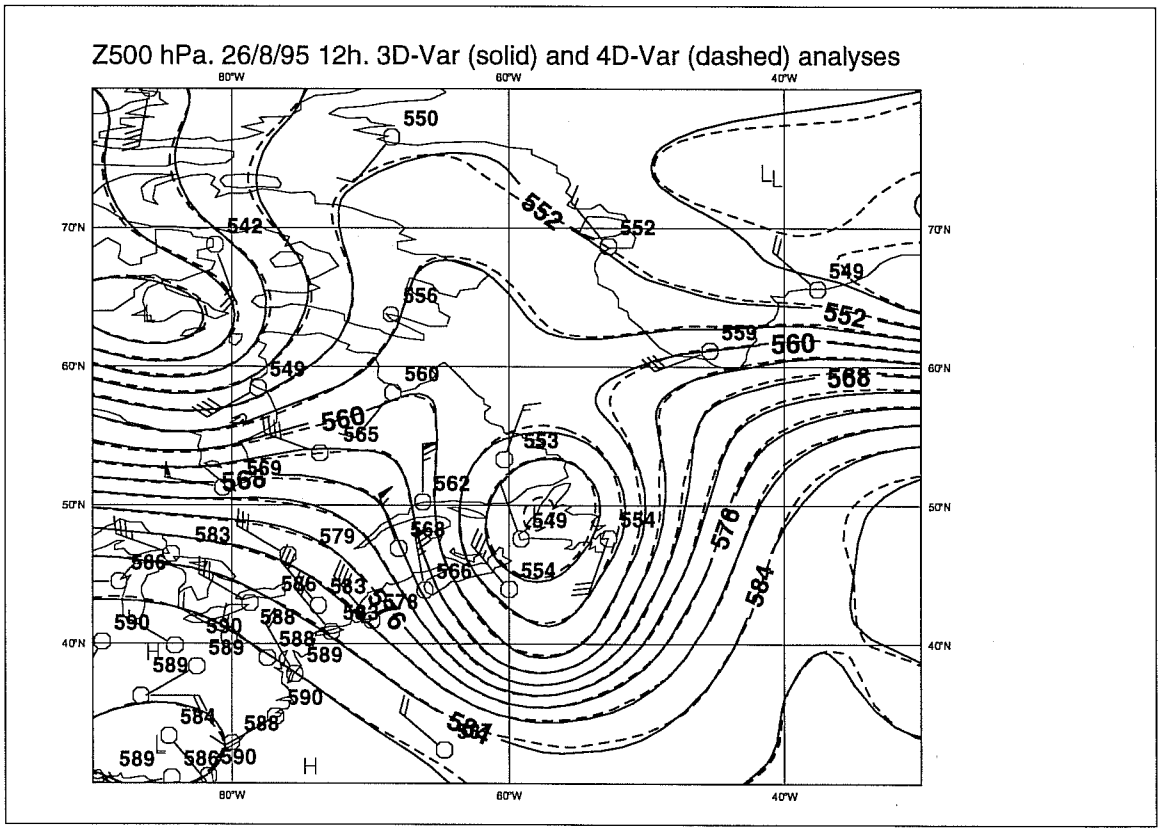

Figure 8: Analyses for the geopotential height at 500 hPa for the 26st August 1995. 4D-Var is shown as dashed lines, 3D-Var as solid lines. Contour interval is 4m. Radiosonde observations are also plotted.

4.3 Increments and fits to observations

The fit of the analyses and the background fields (6-hour forecasts) to the available data is investigated for the August 1995 period (for the first 8 days of assimilation). The root-mean-squares and averages of the differences between model equivalent of the observations and the observations, averaged over the period are plotted for various observation types as a function of height. The results, averaged over large areas (Northern hemisphere, Tropics, Southern hemisphere) are quite similar for both systems. The fit of the background fields to the observations are computed using a quadratic interpolation for the background at ranges 3, 6 and 9 hour for both systems. The fit of the analyses to the observations, on the other hand, are computed along the 4D-Var model integration for the 4D-Var system, and in a static way for the 3D-Var system. This implies that the comparison of the fit of the analyses to the observations will be fair only for data reporting in the middle of the 6-hour assimilation window for which the 3D-Var analysis is valid (basically radiosonde data). Results for the standard-deviations of the difference between analyses (respectively backgrounds) and observations are presented in Table 1(respectively 2) for the radiosonde data (TEMPs) at three levels and aircraft observations (AIREPs) at 250 hPa. The significance of the results is computed using a Fisher's test at the 90% confidence level. Whenever the standard-deviations for the two systems are significantly different, it is indicated in the Tables. For the radiosonde data, only results for which the two systems are significantly different for at least two out of three levels are presented. From Table 1, one can notice that the 4D-Var analyses fit better the radiosonde mass data than the 3D-Var analyses everywhere and fit worse the radiosonde wind data in the Northern hemisphere. However, the worse or better fit of the analyses to the data is not in itself a performance criterion. The relevant comment which can be made from this table is that both systems fit the data reasonably. The fit of the background fields to the data is more relevant to judge the quality of the assimilation system as the short-range forecasts and observations are not correlated. In Table 2, the fit of the 4D-Var background fields to the radiosonde mass observations is better than the fit of the 3D-Var backgrounds to the observations in the Northern hemisphere. It is also better for the fit to aircraft wind data in both hemispheres.

These small differences can hide some larger local differences, and the averaged difference between the RMS of the 4D-Var increments and RMS of the 3D-Var increments for the geopotential height at 200 hPa show that, indeed, the size of the increments is not uniformally identical in the Northern hemisphere (Figure 9).

4D-Var increments tend to be smaller around 50N, especially over the oceans. The maximum difference occurs over the West Atlantic area, with 4D-Var increments up to 6 meters smaller than 3-Var increments. To clarify the reasons for this difference, the RMS fits to the data produced over this West Atlantic area (30 to 60N, 80 to 40W) are presented in figure 10. The solid lines represent the RMS fits of the background fields to the observations, the dashed ones the RMS fits of the analyses to the observations. There are around 300 observations in that area over the period. The 4D-Var analyses and forecasts (shown as squares) have a better fit to the data, especially the background (16.5m compared to 20m for the 3D-Var background at

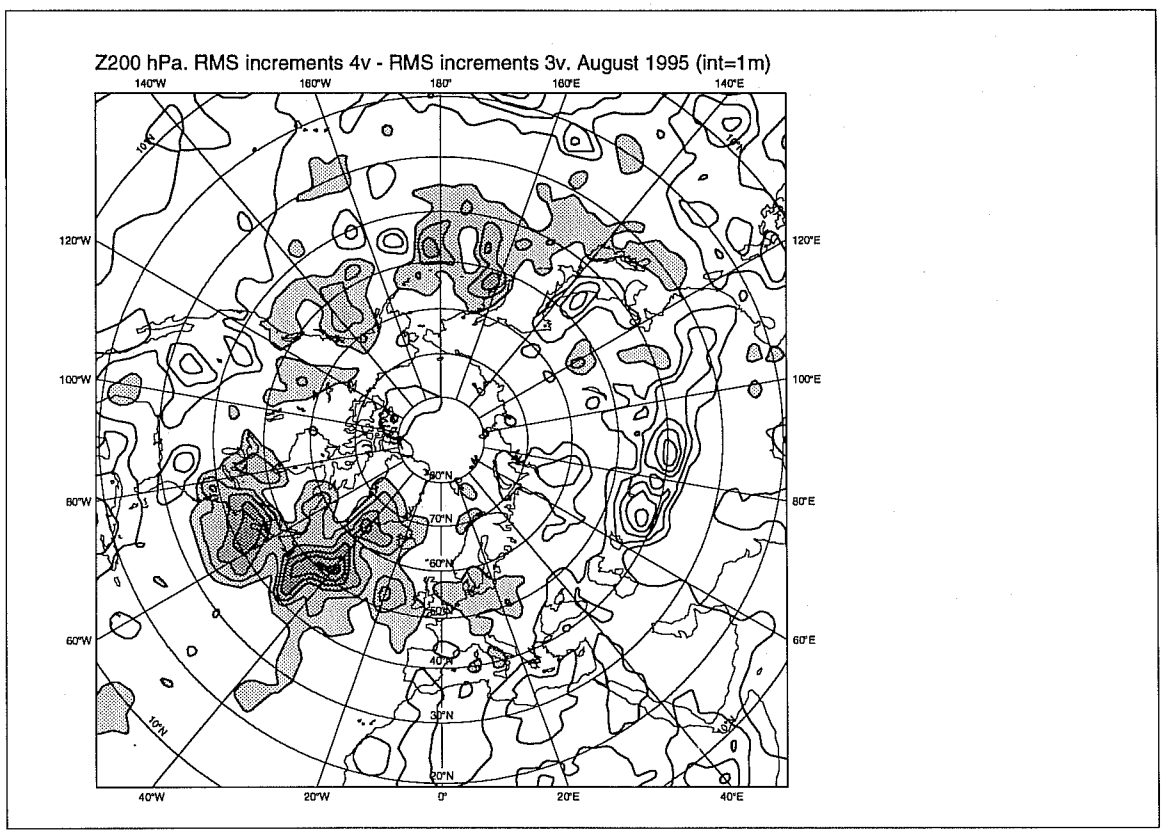

Figure 9: Averaged difference between the RMS of the 4D-Var increments and the RMS of the 3D-Var increments for the geopotential height at 200 hPa for the August 1995 period. Contour interval is 1m. Negative values are shaded.

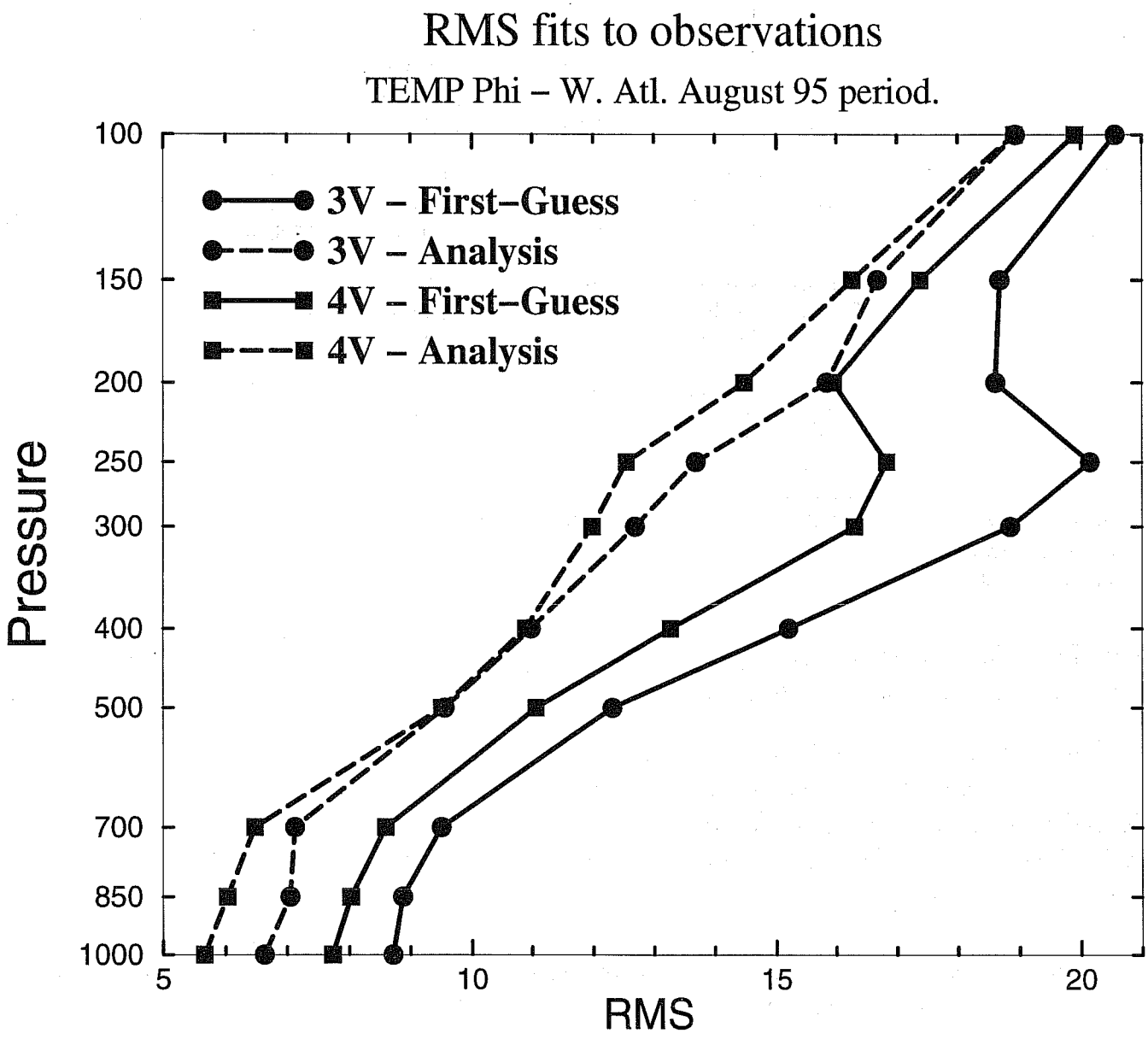

Figure 10: RMS fits to the radiosonde geopotential data produced over the West Atlantic area (30 to 60N, 80 to 40W). The solid lines represent the RMS fits of the backgrounds to the observations, the dashed ones the RMS fits of the analyses to the observations. 4D-Var is shown as squares, 3D-Var as circles. The abscissa is the RMS in geopotential meters. The ordinate is the pressure in hPa.

TABLE 1. STANDARD DEVIATIONS OF DEPARTURES OF ANALYSES FROM THE OBSERVATIONS FOR THE AUGUST 1995 PERIOD.

Type of obs	Variable	Area	Level	Number of obs	3V(AN) std	4V(AN) std	Significant
TEMP	U	NH	850 hPa	6567	2.16 m/s	2.22 m/s	yes
TEMP	U	NH	500 hPa	7152	2.18 m/s	2.25 m/s	yes
TEMP	U	NH	250 hPa	6931	2.85 m/s	2.87 m/s	no
TEMP	V	NH	850 hPa	6567	2.12 m/s	2.17 m/s	yes
TEMP	V	NH	500 hPa	7152	2.15 m/s	2.21 m/s	yes
TEMP	V	NH	250 hPa	6931	2.88 m/s	2.87 m/s	no
TEMP	Z	NH	850 hPa	6986	6.93 m	6.12 m	yes
TEMP	Z	NH	500 hPa	7216	9.69 m	9.63 m	no
TEMP	Z	NH	250 hPa	6866	17.13 m	16.60 m	yes
TEMP	Z	Tropics	850 hPa	925	8.98 m	6.97 m	yes
TEMP	Z	Tropics	500 hPa	923	11.98 m	11.31 m	yes
TEMP	Z	Tropics	250 hPa	894	19.84 m	18.08 m	yes
TEMP	Z	SH	850 hPa	822	6.78 m	5.06 m	yes
TEMP	Z	SH	500 hPa	824	8.09 m	7.41 m	yes
TEMP	Z	SH	250 hPa	800	14.24 m	12.80 m	yes

Three levels (850, 500 and 250 hPa) are indicated for radiosonde data (TEMPs) whenever two at least show significant differences in the fit to observations between 3D-Var and 4D-Var. The significance test is carried out with a Fisher's test at 90 % confidence.

250 hPa). The smaller increments for the 4D-Var system are then a consequence of a much better background field from which to start the minimization.

This area is very active from a baroclinic point of view. Furthermore, there are a lot of single-level aircraft data, which are better assimilated in a 4D-Var system due to the flow-dependent structure functions. An example is given in Figures 11 to 13. For the very first cycle of the assimilation period (24/08/95, 0Z), a cross-section of the vorticity background field along 45N is presented in Figure 11 just off the American coast. One can notice a westward tilt with height typical of baroclinic systems. The 4D-Var increments (Figure 12) are more tilted than the rather barotropic 3D-Var increments (Figure 13). 3D-Var extrapolates data with the specified barotropic structure functions, whereas 4D-Var knows about the dynamical evolution over the assimilation period. Although 4D-Var is implemented here with a 6-hour assimilation window, the interaction between the dynamics and the observations is sufficiently strong in active baroclinic areas to provide some useful information leading to a better analysis and short-range forecast in these areas.

Figure 9 also exhibits some areas where 4D-Var increments are substantially larger than the 3D-Var ones. In particular, there are some differences of up to 5m over Central Asia. The RMS fits to the data produced over this area (35 to 45N, 60 to 100E) for geopotential data indicate that the 4D-Var background is worse by about 2m in the upper troposphere. The worse quality of the 4D-Var system can be explained by the importance of orographic processes in this area

TABLE 2. STANDARD DEVIATIONS OF DEPARTURES OF BACKGROUNDS FROM THE OBSERVATIONS FOR THE AUGUST 1995 PERIOD.

Type of obs	Variable	Area	Level	Number of obs	3V(FG) std	4V(FG) std	Significant
TEMP	Z	NH	850 hPa	6986	9.38 m	8.94 m	yes
TEMP	Z	NH	500 hPa	7216	13.44 m	13.02 m	yes
TEMP	Z	NH	250 hPa	6866	21.35 m	21.27 m	no
AIREP	U	NH	250 hPa	27327	4.04 m/s	3.88 m/s	yes
AIREP	U	Tropics	250 hPa	2616	4.86 m/s	4.91 m/s	no
AIREP	U	SH	250 hPa	3426	3.70 m/s	3.57 m/s	yes
AIREP	V	NH	250 hPa	27327	4.26 m/s	4.05 m/s	yes
AIREP	V	Tropics	250 hPa	2616	4.75 m/s	4.63 m/s	no
AIREP	V	SH	250 hPa	3426	3.85 m/s	3.59 m/s	yes

Three levels (850, 500 and 250 hPa) are indicated for radiosonde data (TEMPs) whenever two at least show significant differences in the fit to observations between 3D-Var and 4D-Var. Aircraft observations (AIREPs) are represented at the level 250 hPa for which there are more data. The same computation of First-Guess At the appropriate Time came into the computation of the two systems. The significance test is carried out with a Fisher's test at 90 % confidence.

which are not included in the tangent-linear and adjoint models used in 4D-Var. In section 3, the comparison between the evolution of a perturbation by the model with full physics and the simplified model used in the minimization already hinted at possible problems there.

4.4 Tropical performance

Turning to the Tropics, there are at least two indications that the present 4D-Var system does not perform as well as 3D-Var there. The first indication is that the short-range tropical wind scores (verified against its own analysis or against observations) are slightly worse, in particular at 200 hPa. The second indication is a worse balance in the hydrological budget in the first day of the forecasts. The forecasts started from the 3D-Var analyses have a slight imbalance between global precipitation and evaporation rates (less than 5%) in our experiments, but this imbalance reaches 10% in the 4D-Var case. 4D-Var forecasts have greater precipitation than evaporation rates at the very beginning of the forecast.

The 4D-Var analyses are moister than the 3D-Var ones (see Figure 14 for the mean analyses difference for the August 1995 period) by up to 0.3 g/kg zonally averaged along the Equator. The 4D-Var analysis is moister from the very first analysis cycle, and the averaged increments clearly add more humidity to the Tropical atmosphere than the 3D-Var ones. Figures 15 and 16 show the mean zonally averaged increments for 3D-Var and 4D-Var respectively. Both analyses, on average, add humidity in the tropical band, but 4D-Var increments are about twice as large. When looking at the geographical distribution of the 3D-Var and 4D-Var increments, they seem to occur at the same locations, but with a different amplitude.

The fact that 4D-Var produces larger increments could mean that the effective background

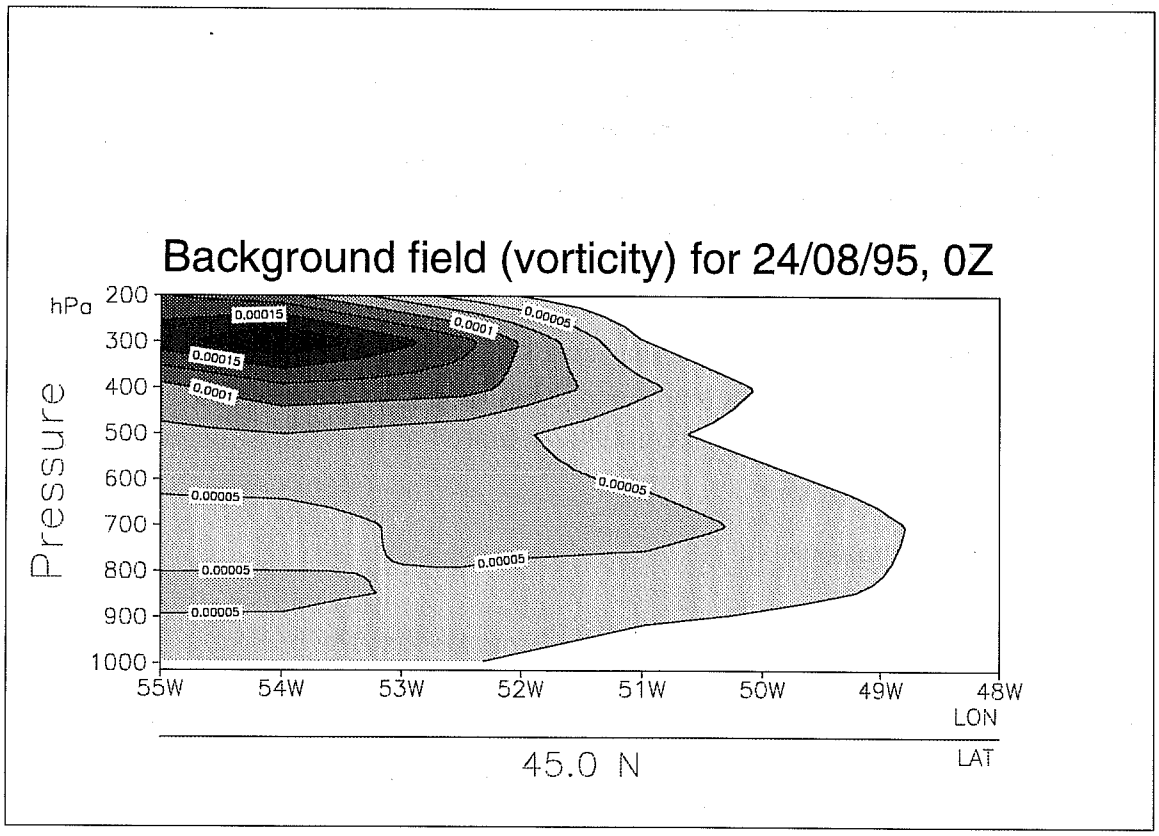

Figure 11: Cross-section longitude/height along 45N, showing the background vorticity field for the analysis on the 24/08/95 at 0Z. Contour interval is 0.000025 s^{-1}.

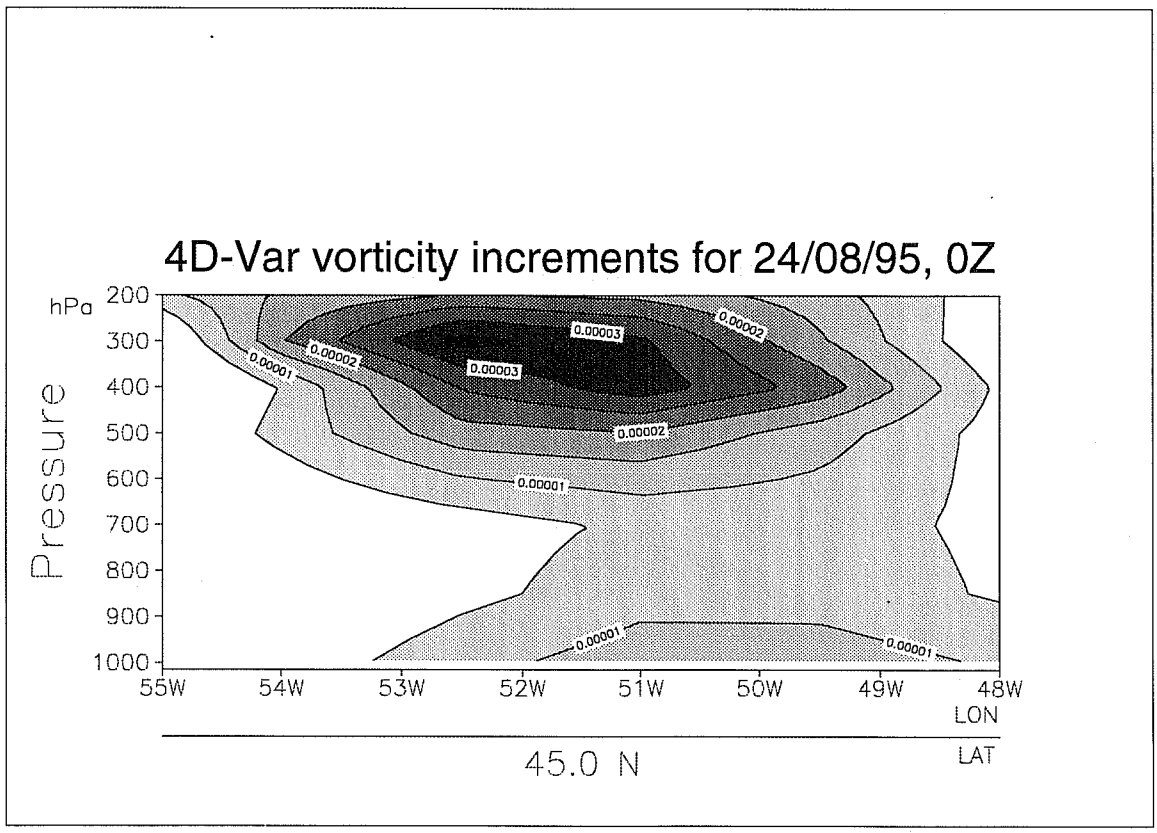

Figure 12: Cross-section longitude/height along 45N, showing the 4D-Var increments for the analysis on the 24/08/95 at 0Z. Contour interval is 0.000005 s^{-1}.

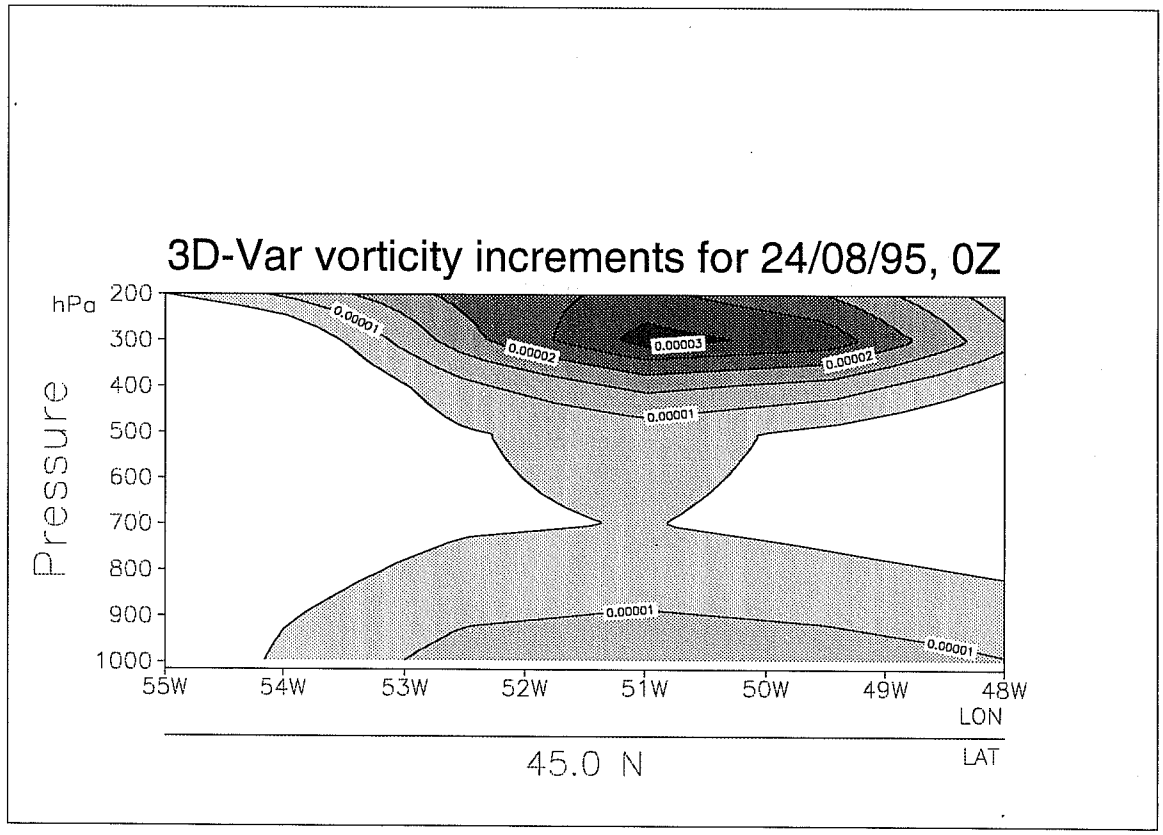

Figure 13: Cross-section longitude/height along 45N, showing the 3D-Var increments for the analysis on the 24/08/95 at 0Z. Contour interval is 0.000005 s^{-1}.

error standard-deviations are larger than for 3D-Var. In 4D-Var, the implied background error standard-deviations are those which would be produced by a Kalman filter implemented over the assimilation window. This means that they evolve according to the dynamics over the 6-hour period. Although this is one of the main advantages of the method if the dynamics and simplified physics used in the tangent-linear and adjoint models are accurate, it could be a drawback if this simplified model is not a good approximation of the evolution of the atmosphere. Budget diagnostics show a large dynamical tendency of humidity in the mid to low troposphere in the tropical area. In the model with full physics, this is almost balanced by parametrized physical tendencies of opposite sign. But in the adiabatic tangent-linear model, the humidity advection tendencies which are driven by the dynamics are larger than they would be if relevant parametrized physical processes such as convection were included. 4D-Var "sees", during the minimization, a larger net change in humidity than what it should, which implies a larger uncertainty in the background field, thus a larger background error. The absence of relevant physics is thus one factor which might explain the larger humidity increments in this area.

Other factors were also seen to be important, such as the version of the atmospheric model used in the assimilation. The performance of 4D-Var in the tropical area will certainly improve when more physical processes are included in the tangent-linear and adjoint models (see next section). However, as explained by Wergen (1992), the Tropics might still remain a potential problem area for 4D-Var because of large model errors there and because there is less tendency information and multivariate information to be extracted from the data.

5. FUTURE DEVELOPMENTS

5.1 Development of the tangent-linear and adjoint of the physics

The tangent-linear and adjoint of the main physical parametrizations are under development (Mahfouf et al, 1996). These include vertical diffusion, sub grid scale orography, large scale condensation, diagnostic cloud scheme and a simplified solar radiation scheme for surface fluxes. As far as the convection is concerned, an approximation of the Jacobian is being used to investigate the scientific problems of developing a suitable formulation.

We have performed a very simple sensitivity experiment to evaluate the impact of introducing more physics in the Tropics. To simulate the effect of convection, the tangent-linear model dynamics are modified by multiplying the vertical velocity perturbation by a factor less than one in the vertical advection and energy conversion terms. This factor is a smooth function going from 0 at the Equator to 1 at 30 degrees North and South. The change is intended to simulate (in an over-simplified way) the fact that the dynamics are balanced by the convective processes in the Tropical area. The adjoint model was modified accordingly. Three cycles of assimilation from the 24th August 1995, 0Z to the 24th August 1995, 12Z were re-run with 4D-Var on a 6-hour window with this change included in the minimization stages. Comparison

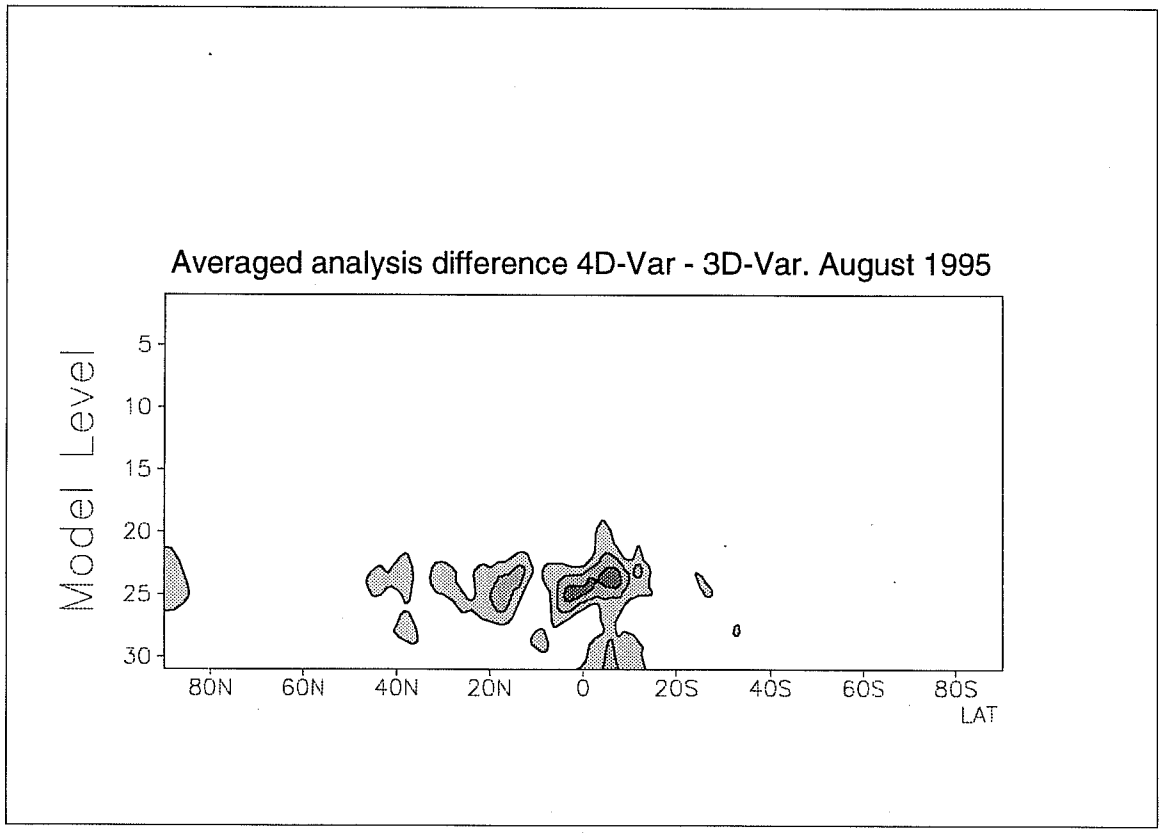

Figure 14: Cross-section latitude/height, showing the difference between the 4D-Var and the 3D-Var specific humidity analyses, averaged over the August 1995 period, and over longitudes. Contour interval is 0.1 g/kg.

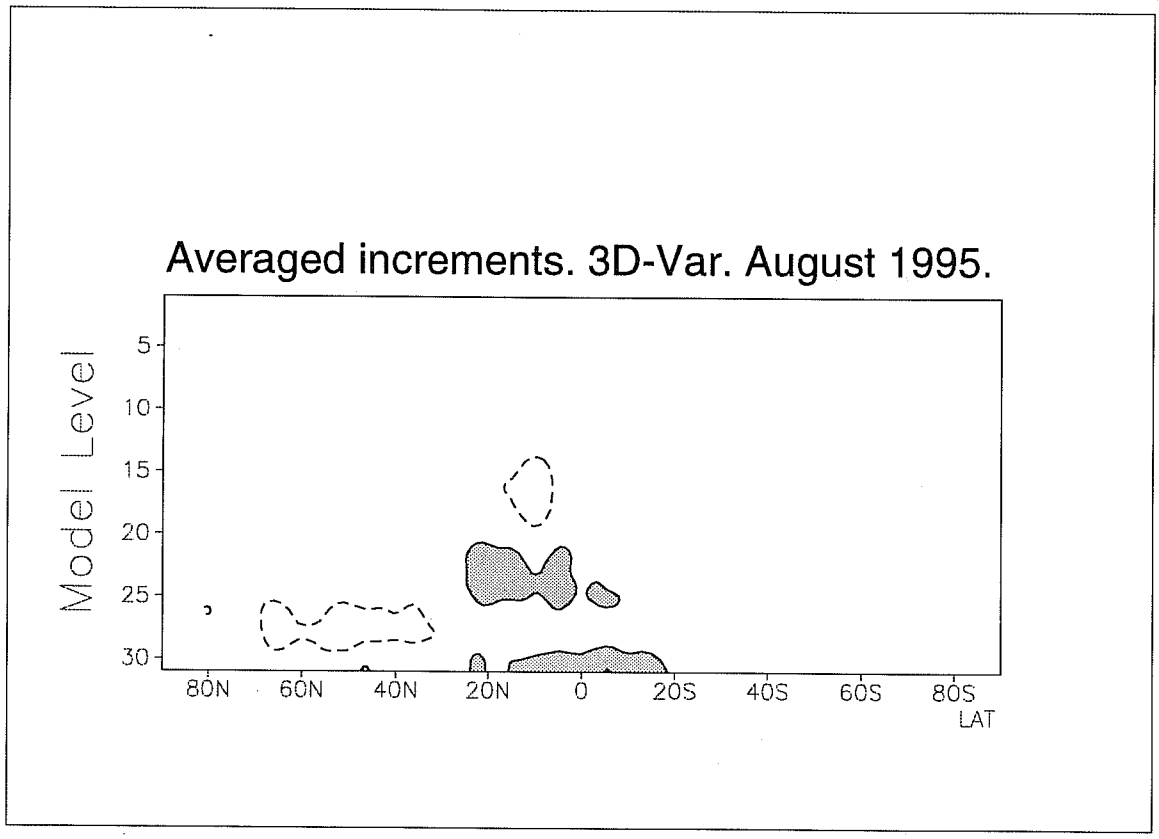

Figure 15: Cross-section latitude/height, showing the 3D-Var specific humidity increments, averaged over the August 1995 period, and over longitudes. Contour interval is 0.05 g/kg.

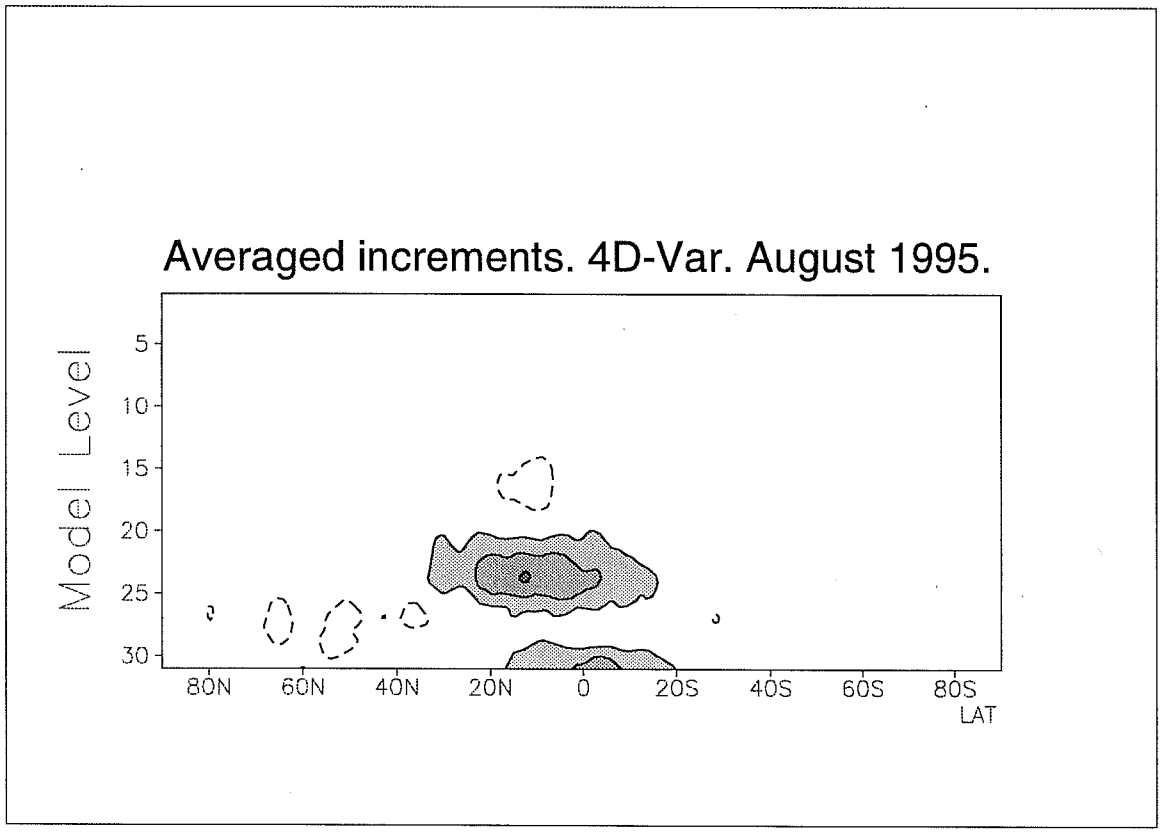

Figure 16: Cross-section latitude/height, showing the 4D-Var specific humidity increments, averaged over the August 1995 period, and over longitudes. Contour interval is 0.05 g/kg.

with the reference 4D-Var analysis for 12Z shows that the specific humidity has indeed been decreased by about 0.2 g/kg in a narrow latitude band at the Equator (not shown), leading to values much more similar to the corresponding 3D-Var values. As a consequence, the forecast run from this analysis exhibits a much smaller difference between initial precipitation and evaporation rates (0.02 mm/day) than the control 4D-Var (0.08 mm/day), thus reducing considerably the precipitation spin-down. This simple experiment is an indication that accounting for parametrized physical processes, and particularly convection, in the minimization phase of the assimilation will be beneficial for the tropical analysis.

5.2 Line-search for outer loops of 4D-Var incremental

The present formulation of 4D-Var incremental does not guarantee a decrease in the cost-function at high resolution after each update. Although the minimization achieves a decrease in the cost-function at low resolution during the inner loops, this decrease might be compensated by a similar or even larger increase when going to high resolution, when the fit to observations is re-computed with the full model in the outer loop. An example of this behaviour is shown in Figure 17 for the standard 4D-Var with a 6-hour window on one particular analysis (solid line). One can see that the decrease in the cost-function is quite large during the first two minimizations, and that it starts getting smaller for the subsequent two minimizations, as one comes closer to the minimum. The jumps in the cost-function occurring for each outer-loop also tend to become smaller, which indicates that the method converges. For instance, the last jump occuring at iteration 60 is only a fraction of the first one at iteration 15. However, it appears that this last increase almost completely cancels the previous decrease in the last minimization. The values of the cost-function at iteration 45 (at the beginning of the last minimization) is 55860, at iteration 60 (at the end of the last minimization) 54920, and at the last outer loop (iteration 60) 55829. The final cost-function is only marginally smaller than it was at the previous outer-loop, which is not satisfactory.

A possible solution to eliminate this problem is to perform a line-search at high resolution (in the outer loop), in the direction given by the minimization at low resolution. Instead of just injecting the increments found at low resolution into the high resolution field, one would then find an appropriate scaling to optimize the decrease of the cost-function at high resolution along the direction given by these increments. The way this is achieved is simple: once the increments at low resolution are computed, one performs several computations of the observational cost-function at high resolution and of the background cost-function (only available for technical reasons at low resolution) from different scalings of the increments. Then, one fits a parabola to the total cost-function evaluations, and associates the optimal scaling factor to the minimum of the parabola. This idea is tried on three cycles of assimilation from the 24th August 1995, 0Z to the 24th August 1995, 12. It is found unnecessary to compute the scaling factors for the first two outer-loops (because they are so close to 1), but the optimal scaling was computed for the two last outer-loops (factors as low as 0.4 are occasionally found). The performance of the minimization for the first analysis is shown as dashed line in Figure 17. The behaviour

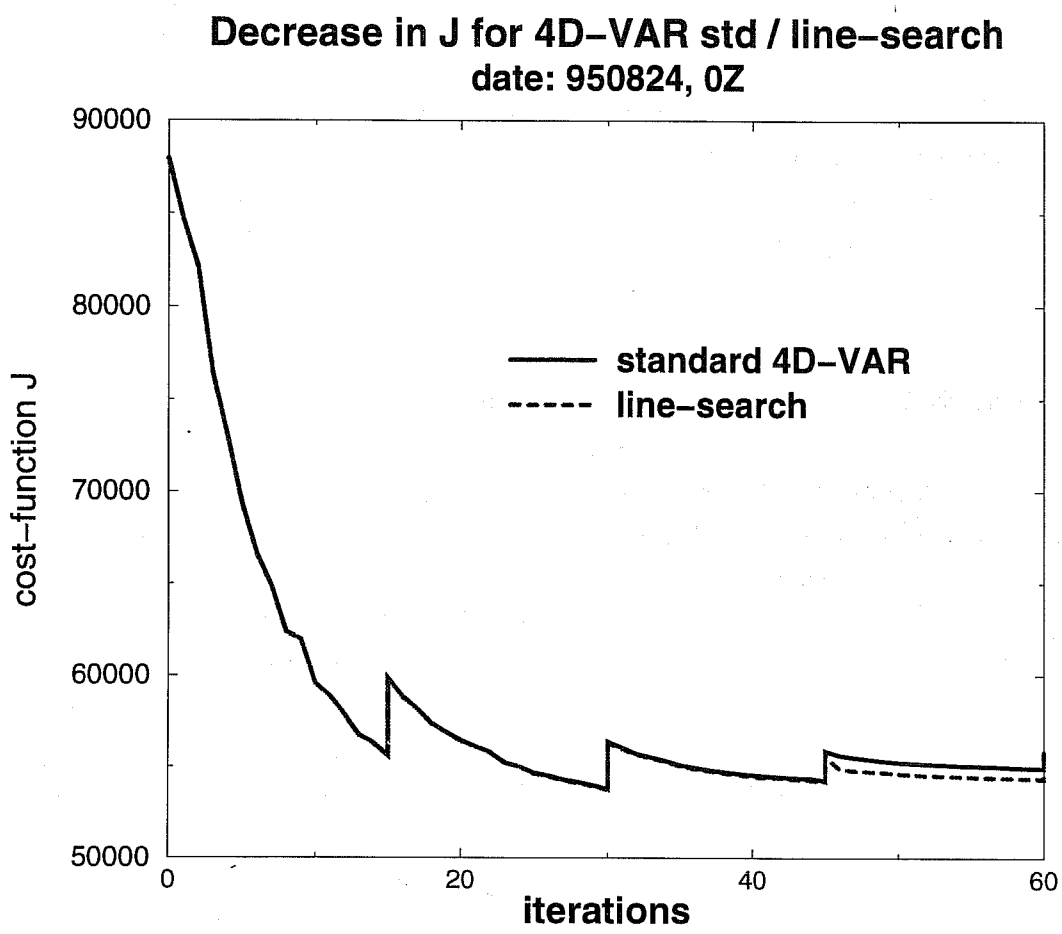

Figure 17: Decrease of the cost-function as a number of iteration number for 4D-Var on a 6-hour period, for the 24th August 1995, 00Z analysis. Standard 4D-Var is shown as solid line, 4D-Var with line search for the outer loop is shown as dashed line.

of the minimization, although quite similar to that of the standard 4D-Var, is better at the last outer-loop. The values of the cost-function at iteration 45 (at the beginning of the last minimization) is 55574, at iteration 60 (at the end of the last minimization) 54346, and at the last outer loop (iteration 60) 55410. The final cost-function is now considerably smaller than it was at the penultimate outer-loop, which is precisely what we wanted to achieve with this method.

6. CONCLUDING REMARKS

In this paper, results of extended 4D-Var assimilations in cycling mode, over several two-week assimilation periods have been presented. 4D-Var is implemented in its incremental formulation (Courtier et al, 1994) with a high resolution (T106) model with the full physical parametrization package to compare the atmospheric states with the observations and a low resolution (T63)

model with simplified physics to minimize the cost-function. The comparison of 4D-Var on several assimilation windows (6, 12 and 24 hours) with 3D-Var (the equivalent of 4D-Var with no time-dimension) over a two-week period showed a clear benefit of using 4D-Var over a 6 or 12-hour window over the static 3D-Var scheme. It also exhibited some problems for the forecasts started from the 4D-Var on a 24-hour window. This poor performance of the 4D-Var on a 24-hour system is disappointing as the influence of the dynamics had been found to be beneficial in earlier studies on such an assimilation window (Rabier and Courtier, 1992, Thépaut et al, 1996). However, there are at least two possible explanations. The first one is that, in the set-up of these experiments, the tangent-linear and adjoint models used in the minimization are only approximations of the assimilating model (lower resolution, very simplified physics). An example of the error these approximations introduce in the time evolution of a perturbation shows an increase from about 20% at the 6-hour range to 30% at the 24-hour range. This error affects the convergence of the incremental 4D-Var, with larger discontinuities in the values of the cost-function when going from low to high resolution for longer assimilation windows. The second main reason is the fact that a perfect model is assumed in this 4D-Var context. This is certainly not ideal (as shown for instance by Wergen, 1992), and probably more detrimental at the 24-hour range than at the 6-hour range. Experiments performed with this hypothesis have been successful for short 2 to 3-hour windows (Kuo et al, 1995, Zou et al, 1995), but other experiments indicate a benefit of relaxing this constraint on a longer 12-hour assimilation window (Zupanski, 1993, Zupanski, 1996). Some possibilities have been tried to introduce some model error terms in 4D-Var (Derber, 1989, Wergen, 1992, Zupanski, 1996). A full weak-constraint approach is equivalent to the representers method (Bennet, 1992, Egbert et al, 1994) as shown by Amodei (1995) and Courtier (1996). However, before changing our 4D-Var formulation, more diagnostics are needed on the fit to observations (Dee, 1995a, Ménard and Daley, 1996) and more work should be carried out on evaluating model error statistics (Daley, 1992, Dee, 1995b, Wahba et al, 1995).

In the mean-time, we have performed more experiments comparing 4D-Var on a 6-hour window with the 3D-Var system. Two additional two-week periods show a consistent improvement in extratropical forecast scores when using the 4D-Var system. The main improvements occur in areas where the 3D-Var errors are the largest. Local improvement can be as large as 35% for the root-mean square of the 5-day forecast error, averaged over a two-week period. A comparison with key analysis error, as computed in Klinker et al (1997) shows that, indeed, 4D-Var on a 6-hour window is able to reduce substantially the error in its fast-growing components. The overall fit of analyses and short-range forecasts to observations is comparable or slightly better in 4D-Var. In active baroclinic areas, the fit of the background to the data is considerably better for the 4D-Var system, resulting in smaller increments. It appears that in these areas (and in particular over West-Atlantic), 4D-Var is able to use better the information contained in the observations. The way 4D-Var can extrapolate some aircraft data in the vertical with a baroclinic tilt is illustrated in the paper.

Some problems exist in other areas (Tropics and mountainous areas) mainly due to a lack of parametrized physics. The tangent-linear and adjoint of the main physical parametriza-

tion schemes are under development (Mahfouf et al, 1996). Once they are available, more experiments will be performed for 4D-Var on a 6-hour window and for longer periods as well. Experiments will also be performed with higher resolution models. However, care should be taken not to exceed the range of validity of the tangent-linear hypothesis, as it is known to break down first for the small scales (Tanguay et al, 1996).

More potential advantages of using 4D-Var in an operational context have not been investigated in this study. For instance, the same amount of data is currently used in a 4D-Var on a 6-hour window analysis as in a 3D-Var analysis, with a synoptic screening of data. A screening on one-hour time-slots, more adapted to 4D-Var is being developed (Järvinen, pers comm). 4D-Var is known to use consistently asynoptic data such as radiances (Andersson et al, 1994). More types of observations have already been tested in the four-dimensional framework, such as ozone (Fisher and Lary, 1995, Riishøjgaard, 1996), precipitable water (Filiberti et al, 1996, Kuo et al, 1995), radio refractivity (Zou et al, 1995),... 4D-Var will continue to be tested in a quasi-operational framework, with more physical processes, and more observations. Although this algorithm is expensive, some scenarii exist to reduce its cost such as a multi-incremental strategy (Veersé, pers comm) or quasi-continuous variational assimilation (Järvinen et al, 1996) which should keep it within the possibilities of operational testing and allow for a full examination of the method.

Acknowledgments

We would like to thank Ernst Klinker for his help with the diagnostic part of this study, and in particular the comparison with key analysis errors. We are particularly indebted to Mats Hamrud, Jan Haseler and Drasko Vasiljevic for the management of the code, the scripts and the observations processing. All the people developing the 3D-Var system have indirectly contributed to this work: Per Undén, Erik Andersson, Mike Fisher, François Bouttier and Heikki Järvinen. It is a pleasure to thank other colleagues at ECMWF, and in particular Dave Burridge, Anthony Hollingsworth, Jean-François Mahfouf, Martin Miller, Adrian Simmons, Clive Temperton and Pedro Viterbo for their valuable comments. This work benefitted from numerous discussions with Olivier Talagrand.

References

Amodei, L.,1995: Solution approchée pour un problème d'assimilation de données météorologiques avec prise en compte de l'erreur de modèle. *Comptes Rendus de l'Académie des Sciences*, **321,** série II a, 1087–1094.

Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J., McNally, A. P., Kelly, G., 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., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljevic, D., Brankovic, C.,

Cardinali, C., Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier, F., Simmons, A., Strauss, B., Thépaut, J.-N., and Viterbo, P., 1997: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part III: Experimental results. *Q. J. R. Meteorol. Soc.*, submitted.

Bennett A.F.,1992: *Inverse methods in physical oceanography.*, Cambridge monographs on mechanics and applied mathematics, Cambridge University Press.

Bouttier, F., 1993: The dynamics of error covariances in a barotropic model. *Tellus*, **45A,** 408–423.

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

Cohn, S.E., Sivakumaran, N.S. and Todling, 1994: A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.*, **122,** 2838–2867.

Courtier, P., and Talagrand, O., 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., and Talagrand, O., 1990 : Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus*, **42A,** 531–549.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A., 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120,** 1367–1388.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hollingsworth, A., Rabier, F., and Fisher, M., 1997: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Q. J. R. Meteorol. Soc.*, submitted.

Courtier, P., 1996: Introduction to variational assimilation. In Proceedings of the ECMWF seminar on Data Assimilation, Shinfield Park, Reading, RG2 9AX, 2–6 September 1996.

Daley, R., 1992: Estimating model-error covariances for application to atmospheric data assimilation. *Mon. Wea. Rev.*, **120,** 1735–1746.

Dee, D.P., 1995a: Testing the perfect-model assumption in variational data assimilation. *Proceedings of the second international symposium on assimilation of observations in Meteorology and Oceanography*, **Volume I,** 49–54.

Dee, D.P., 1995b: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.*, **123,** 1128–1145.

Derber, J., 1989: A variational continuous assimilation technique. *Mon. Wea. Rev.*, **117,** 2437–2446.

Egbert, G.D., Bennett, A.F. and Foreman, M.G.G., 1994: Topex/Poseidon tides estimated using a global inverse model. *J. Geophys. Res.*, **99,** 24,821–24,852.

Filiberti, M.-A., Rabier, F., Thépaut, J.-N., Eymard, L. and Courtier, P, 1996: Four-dimensional variational asimilation of SSM/I water vapour data. *Q. J. R. Meteorol. Soc.*, submitted.

Fisher, M. and Lary, D. J., 1995: Lagrangian four-dimensional variational data assimilation of

chemical species. *Q. J. R. Meteorol. Soc.*, **121,** 1681–1704.

Fisher, M. and Courtier, P., 1995: *Estimating the covariance matrices of analysis and forecast error in variational data assimilation*, ECMWF Tech. Mem. 220 (Available from ECMWF, Reading, UK).

Gauthier, P., Courtier, P., and Moll, P., 1993: Assimilation of simulated wind lidar data with a Kalman filter. *Mon. Wea. Rev.*, **121,** 1803–1820.

Gilbert, J-C., and Lemaréchal, C., 1989: Some numerical experiments with variable-storage quasi-Newton algorithms. *Mathematical Programming.*, **B25,** 407–435.

Ghil, M. and Malanotte-Rizzoli, P., 1991: Data assimilation in meteorology and oceanography. *Advances in Geophysics*, **33,** 141–266.

Ghil, M. and Todling, R., 1996: Tracking atmospheric instabilities with the Kalman filter. Part II: Two-layer results. *Mon. Wea. Rev.*, **124,** 2340–2352.

Ide, K., Bennett, A.F., Courtier, P., Ghil, M. and Lorenc, A.C., 1997: Unified notation for data assimilation: operational, sequential and variational. *J. Met. Soc. Japan*, In press.

Järvinen, H., Thépaut, J. N., and Courtier, P., 1996: Quasi-continuous variational data assimilation. *Q. J. R. Meteorol. Soc.*, **122,** 515–534.

Klinker, E., Rabier, F., and Gelaro, R., 1997: Estimation of key analysis errors using the adjoint technique. *Q. J. R. Meteorol. Soc.*, submitted.

Kuo, Y. H., Zou, X., and Guo, Y. R., 1995: Variational assimilation of precipitable water using a nonhydrostatic mesoscale adjoint model. Part I: moisture retrieval and sensitivity experiments. *Mon. Wea. Rev.*, **124,** 122–147.

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

Lewis, J. and Derber, J., 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus.*, **37,** 309–327.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112,** 1177–1194.

Mahfouf, J-F., Buizza, R., and Errico, R. M., 1996: Strategy for including physical processes in the ECMWF variational data assimilation system. In Proceedings of the ECMWF workshop on non-linear aspects of data assimilation, Shinfield Park, Reading, RG2 9AX, 9–11 September 1996.

Ménard, R., and Daley, R., 1996: The application of Kalman smoother theory to the estimation of 4DVAR error statistics. *Tellus*, **48A,** 221–237.

Miller, R.-N., Ghil, M. and Gauthiez, F., 1994: Advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sci.*, **51,** 1037–1056.

Navon, I.-M., Zou, X., Derber, J.C. and Sela, J., 1992: Variational data assimilation with an adiabatic version of the NMC spectral model. *Mon. Wea. Rev.*, **120,** 1433–1446.

Pires, C., Vautard, R., and Talagrand, O., 1996: On extending the limits of variational as-

similation in nonlinear chaotic systems. *Tellus*, **48A,** 96–121.

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

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

Rabier, F., Mc Nally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A. and Bouttier, F., 1997: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part II: Structure functions. *Q. J. R. Meteorol. Soc.*, submitted.

Riishøjgaard, L. P., 1996: On four-dimensional variational assimilation of ozone data in weather-prediction models. *Q. J. R. Meteorol. Soc.*, **122,** 1545–1571.

Talagrand, O. and Courtier, P., 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I. Theory. *Q. J. R. Meteorol. Soc.*, **113,** 1321–1328.

Tanguay, M., Bartello, P., and Gauthier, P., 1996: Four-dimensional data assimilation with a wide range of scales. *Tellus*, **47A,** 974–997.

Thépaut, J.-N. and Courtier, P., 1991: Four-dimensional variational assimilation using the adjoint of a multilevel primitive-equation model. *Q. J. R. Meteorol. Soc.*, **117,** 1225–1254.

Thépaut, J.-N., Hoffman, R.N., and Courtier, P., 1993: Interactions of dynamics and observations in a four-dimensional variational assimilation. *Mon. Wea. Rev.*, **121,** 3393–3414.

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

Todling, R., and Cohn, S., 1994: Suboptimal schemes for atmospheric data assimilation based on the Kalman filter. *Mon. Wea. Rev.*, **122,** 2530–2557.

Todling, R., and Cohn, S., 1996: Some strategies for Kalman filtering and smoothing. In Proceedings of the ECMWF seminar on Data Assimilation, Shinfield Park, Reading, RG2 9AX, 2–6 September 1996.

Wahba, G., Johnson, D. R., Gao, F., and Gong, J., 1995: Adaptive tuning of numerical weather prediction models: randomized GCV in three- and Four-dimensional data assimilation. *Mon. Wea. Rev.*, **123,** 3358–3369.

Wergen, W., 1992: The effect of model errors in variational assimilation. *Tellus*, **44A,** 297–313.

Zou, X. Kuo, Y. H., and Guo, Y. R., 1995: Assimilation of atmospheric radio refractivity using a nonhydrostatic adjoint model. *Mon. Wea. Rev.*, **124,** 2229–2249.

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

Zupanski, D., 1996: A general weak constraint applicable to operational 4DVAR data assimilation systems. *Mon. Wea. Rev.*, submitted.