TECHNICAL MEMORANDUM

271

The ECMWF operational implementation of four dimensional variational assimilation. Part I: Experimental results with simplified physics

F. Rabier, H. Järvinen, E. Klinker,
J F. Mahfouf and A. Simmons

Research Department

February 1999

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

The ECMWF operational implementation of four dimensional variational assimilation. Part I: Experimental results with simplified physics.

By **F. Rabier, H. Järvinen, E. Klinker, J-F. Mahfouf, A. Simmons**

Abstract

This paper presents results of a comparison between four-dimensional variational assimilation (4D-Var) using a 6-hour assimilation window and simplified physics during the minimisation, and three-dimensional variational assimilation (3D-Var). Results have been obtained at "operational" resolution T213L31/T63L31. The sensitivity of the 4D-Var performance to different set-ups is investigated. In particular, the performance of 4D-Var in the Tropics revealed some sensitivity to the way the adiabatic non-linear normal mode initialization of the increments was performed. Going from four outer loops to only one (as in 3D-Var) helped to improve the 4D-Var performance, together with a change to the 1997 formulation of the background constraint and an initialization of only the small scales. Tropical scores then became only marginally worse for 4D-Var than for 3D-Var. Twelve weeks of experimentation with the one outer-loop 4D-Var and the 1997 background formulation have been studied. The averaged scores show a small but consistent improvement in both hemispheres at all ranges. In the short range, each two to three-week period has been found to be slightly positive throughout the troposphere. The better short-range performance of the 4D-Var system is also shown by the fits of the background fields to the data. More results are presented over the Atlantic Ocean area during the FASTEX period, during which 4D-Var is found to perform better. In individual synoptic cases corresponding to interesting Intensive Observing Periods, 4D-Var has a clear advantage over 3D-Var during rapid cyclogeneses. The very short-range forecasts used as backgrounds are remarkably closer to the data over the Atlantic for 4D-Var than for 3D-Var. The 4D-Var analyses also display more day-to-day variability. Some structure functions are illustrated in the 4D-Var case for a height observation inserted at the beginning of the assimilation window, in the middle or at the end. The dynamical processes seem to be relevant, even on a short 6-hour assimilation period, which explains the overall better performance of the 4D-Var system.

1. INTRODUCTION

Four-dimensional variational assimilation (4D-Var) minimizes a cost-function measuring the distance between a model trajectory and the available information (background, observations) over an assimilation interval or window. The 4D-Var system is the temporal extension of the three-dimensional variational analysis operational at ECMWF since January 1996 (Courtier et al., 1998; Rabier et al., 1998a; Andersson et al., 1998). The 4D-Var algorithm uses the adjoint equations for the computation of the gradient of the cost-function. 4D-Var was first applied to simple models (Lewis and Derber, 1985; Le Dimet and Talagrand, 1986; Courtier and Talagrand, 1987; Talagrand and Courtier, 1987), before being tested in the context of primitive equations models (Thépaut and Courtier, 1991; Rabier and Courtier, 1992; Navon et al., 1992; Zupanski, 1993). It is being developed at ECMWF in its incremental formulation (Courtier et al., 1994) which comprises running a high-resolution model with the full physical parametrization package to compare the atmospheric states with the observations as part of the evaluation of the cost-function and a low resolution model with simplified physics to minimize the cost-function. Results obtained at resolution T106L31/T63L31 with 3 to 4 outer-loops and 15 to 25 inner-loops, and very simplified physics (horizontal and vertical diffusion and a surface drag) are described in Rabier et al. (1998b). In summary, it was found that 4D-Var using a 6 or 12-hour window performed better than 3D-Var over a 2-week assimilation period, whereas 4D-Var using a 24-hour window did not. The poorer performance of 4D-Var with a relatively long assimilation window could be partly explained by the fact that, in these experiments, the tangent-linear and adjoint models used in the minimization were only approximations of the assimilating model (lower resolution, crude physics). The error these approximations introduced in the time evolution of a perturbation affected 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. While the tangent-linear and adjoint of a more accurate physical parametrization package was being developed, the strategy was to concentrate on the 4D-Var with a 6-hour window with a view to operational

implementation, before revisiting the 4D-Var with longer windows. Two additional two-week periods were then run, which showed a consistent improvement of 4D-Var over 3D-Var in the extratropics. Some problems emerged however in the tropical area.

This article is the first of three papers describing the steps leading to the operational implementation of 4D-Var with a 6-hour window at ECMWF on 25th November 1997. As this is such an important achievement (a first worldwide, the development of which involved many people and resources and with much potential for the exploitation of existing and forthcoming sources of observations), it was decided to describe thoroughly the stages leading to the implementation. This paper concentrates on the experimental results produced with 4D-Var using very simplified physics in the minimisation (as in Buizza 1994); the second paper will describe a set of more elaborate linear physics, their validation and impact on 4D-Var performance; the third paper will deal more precisely with the diagnostics of the operational system chosen as the result of the main experimentation described in papers I and II.

After the work reported by Rabier et al. (1998b), the handling of observations in 4D-Var underwent some modifications. The observation screening and quality control which used to be done by the Optimum Interpolation are now performed within the variational assimilation (Järvinen and Undén, 1997; Andersson, 1996), and the observation operators are activated in their tangent-linear versions within the minimization (a finite-difference approximation was previously used). Furthermore, there is now the possibility of using more observations from frequently reporting stations in the 4D-Var scheme, and of performing the experiments at higher resolution T213L31/T63L31. Further updates of 4D-Var include using a Lanczos algorithm for the evaluation of the analysis and background errors (Fisher and Courtier, 1995, Bouttier et al., 1997).

The validation of these changes was performed by re-running the January 1996 period previously run at T106L31/T63L31 resolution with the old set-up. As discussed in section 2 which describes the sensitivity to various 4D-Var set-ups, the improvement of 4D-Var over 3D-Var seen earlier is reproduced using the same amount of data in both schemes. The use of extra off-time observations is not found to be beneficial in the current set-up. Tropical wind scores are substantially poorer in 4D- than in 3D-Var in these new experiments, as they had been previously. Subsequent developments involving reduction of the number of outer loops from four to one, use of a new background term introduced in operational 3D-Var in 1997 (Bouttier et al., 1997), and removal of initialization of large scales, resulted in better tropical behaviour as also explained in section 2. Section 3 presents the results of the baseline 4D-Var chosen on the basis of these experiments (one outer loop, 1997 background term) on a total of twelve-weeks and discusses in detail results obtained over the Atlantic area during part of the FASTEX experiment (Joly et al., 1997). Section 4 illustrates the structure functions used implicitly in 4D-Var. Discussion of results is given in section 5.

2. Sensitivity of results to different 4D-Var set-ups

2.1 Validation of changes in resolution, observation processing and forecast error specification

One of the periods (16 to 29 January 1996) previously tested at resolution T106L31 was repeated at T213L31 on a new distributed memory computer, with some technical and scientific changes. As mentioned in the introduction, the observation screening and quality control are now performed within the variational assimilation (Järvinen and Undén, 1997, Andersson, 1996), the evaluation of the analysis and background errors involves a Lanczos algorithm (Fisher and Courtier, 1995, Bouttier et al., 1997), and the observation operators are activated in their tangent-linear versions within the minimization. The version of the analysis which is used is with the 1996 formulation of the background term (Courtier et al., 1998) and with 4 outer loops and 20 iterations within each minimization. 4D-Var's configuration uses the same amount of data as 3D-Var. Results are presented in Fig. 1. 3D-Var at low-resolution is shown as a solid line, 3D-Var at high-resolution as a dashed line, 4D-Var at low-resolution as a dotted line and at high-resolution as a dash-dotted line. One can notice the improvement of both 3D-Var and 4D-Var when going to the high-resolution version. The important point for

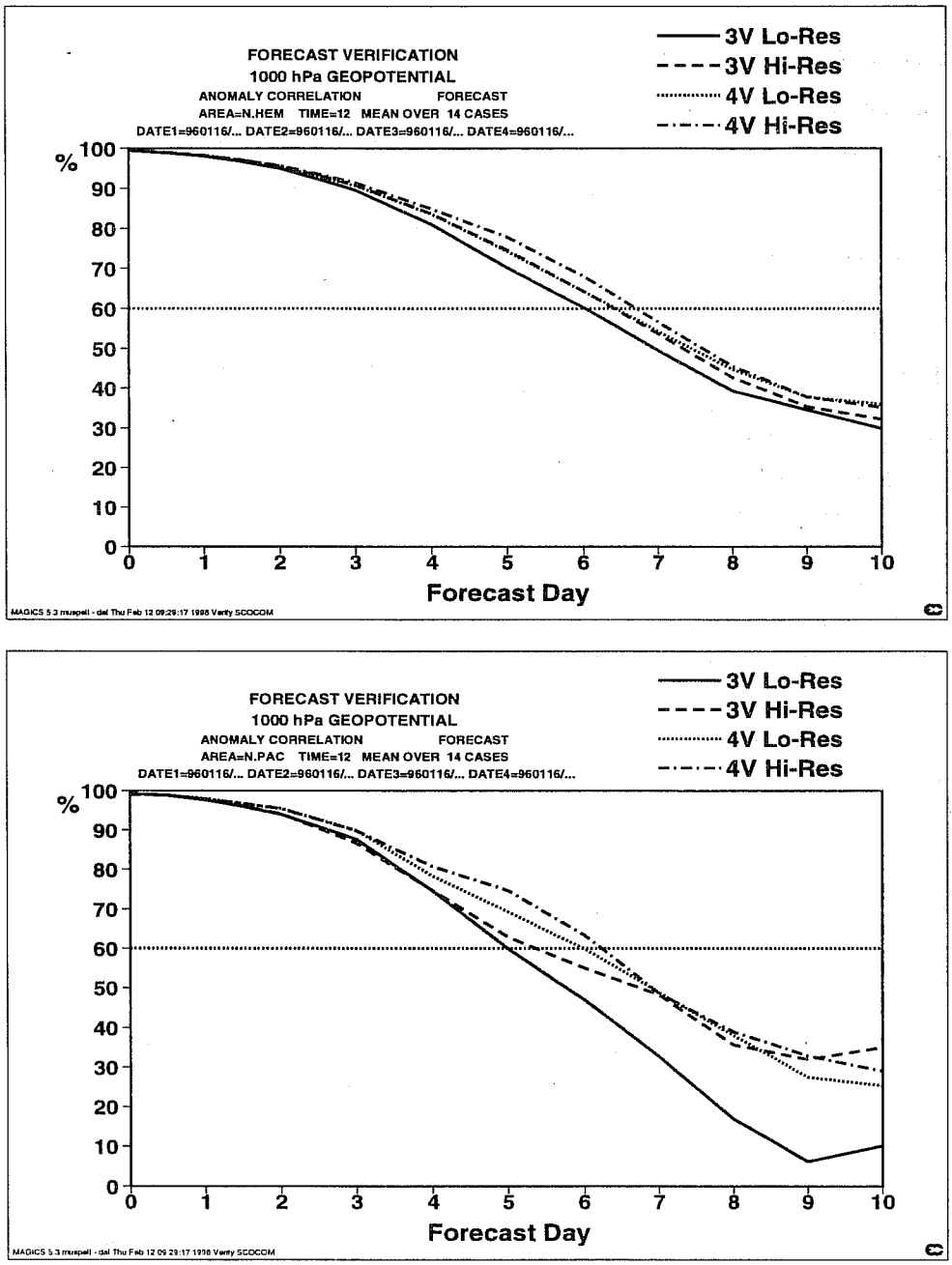

Figure 1 . Anomaly correlation scores for 3D-Var and 4D-Var at T106L31 and at T213L31, using the same amount of observations, for the Northern Hemisphere (top panel) and North Pacific (bottom panel) for two weeks in January 1996.

our validation exercise is that the improvement when going from 3D- to 4D-Var is retained in the later version. This is true for the Northern Hemisphere as a whole (top panel), and also for the dramatic improvement in the North Pacific area at least up to day six (bottom panel). The tropical scores are also consistent with the previous experiments, with scores

markedly worse for 4D- than for 3D-Var. As an example, the averaged high-resolution tropical wind scores verified against their own analysis at 850 hPa are 4.1 m/s for 3D-Var and 4.8m/s for 4D-Var at day 3.

2.2 Influence of extra off-time data

Unlike a static assimilation scheme, 4D-Var assimilates the observations along the trajectory that extends over the assimilation window. There are two related benefits. First, the observations are used at the appropriate time. Second, many observations from frequently reporting stations can be used within one assimilation period. These extra observations are a resource that has not been fully utilized with earlier static assimilation schemes. The observations are selected for the assimilation during the initial high resolution trajectory integration. At this stage, all the necessary information is available for the quality assessment of the observations. From the subset of good quality observations, all redundant information is rejected so that a unique set of observations is left for the assimilation. In 3D-Var, the preference is given to the observations that are close to the middle of the assimilation period. In 4D-Var, the observations are organized into one hour time slots, and the comparison of the trajectory with observations is accordingly done once per hour. Within a time slot, the preference is given to observations that are close to the middle of the time slot. The IFS (Integrated Forecasting system) is however designed so that either an hourly or a six-hourly observation screening can be performed for 4D-Var. The effect of this choice on the number of observations used in the assimilation is largest for observation types that report frequently, like SYNOPs and DRIBUs. An illustrative example is given in Fig. 2 which displays the number of available observations in different time slots and the number actually used. The number of SYNOP surface pressure observations used in a two-week period in the 4D-Var assimilation is roughly twice the number in a corresponding 3D-Var assimilation for the same period. The difference arises from the observations made at times other than the main observing times.

The impact of these off-time observations was tested for the January 1996 period in addition to the previous experiments: 3D-Var and 4D-Var without the off-time observations. As shown in Fig. 3, both 4D-Var systems produce equally good forecast scores in the Northern Hemisphere. In the medium-range the scores are 6 to 12 hours ahead of the 3D-Var scores both at 1000hPa and 500hPa (not shown). The Southern Hemisphere scores however are better for the 4D-Var system without the off-time observations (dotted line). There the inclusion of the off-time observations deteriorated the 4D-Var scores (dashed line) to the level of, or even below, the 3D-Var scores (solid line).

Investigation of the observation statistics revealed certain stations with significant biases against the background for all time slots. Some of these stations got an increased weight in the analysis as there was an observation contributing in each time slot and therefore large analysis increments were produced in the vicinity of those stations. For isolated observations, particularly in the Southern Hemisphere, we have currently no mechanism to prevent these unrealistic increments from appearing and developing into forecast errors. In the Northern Hemisphere, in contrast, there are many more observations to constrain the analysis and therefore the forecast scores there are equally good for the two 4D-Var systems. It is clear now that time-correlated observation errors have to be taken into account in 4D-Var in order to make proper use of the off-time observations. This involves changes to the way the observation term of the cost function is calculated for these observations. It is also necessary to perform the variational quality control simultaneously for all the observations from the same station. The actual form of the time-correlation of the observation errors has to be studied and modelled. Work is in progress on this subject and recent results will be discussed in the summary.

2.3 Sensitivity of the tropical performance to the 4D-Var set-up

In order to understand better the behaviour of the 4D-Var system, some analysis experiments were performed in which only a few simulated isolated observations were inserted. In particular a geopotential datum at 850 hPa was simulated for a given date (5 December 1996, 0UTC) at 20S, 80W, with a height departure from the background of 10m. The resulting mass and wind increments at 850hPa are illustrated for both 4D-Var (Fig. 4.a) and 3D-Var (Fig. 4.c) with the use of the 1997 background constraint (Bouttier et al., 1997). The comparison of 4D- and 3D-Var geopotential increments shows a

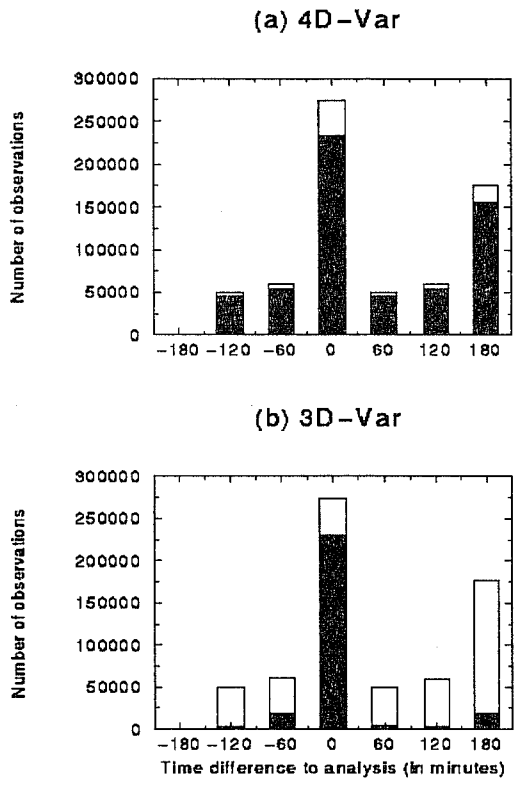

Figure 2 . Time distribution of SYNOP surface pressure observations during a 2-week period. The column height gives the number of all the available observations while the shaded part displays those actually used by the analysis in a) 4D-Var and in b) 3D-Var.

larger amplitude in the 4D-Var case (6.6m versus 5.4m). The 4D-Var wind increments also appear to be larger. The interpretation of this result requires us to go back to the way the incremental variational assimilation is implemented.

Each minimization problem can be written

$$J(\delta\mathbf{x}_l^n) = \frac{1}{2}(\delta\mathbf{x}_l^n + \mathbf{x}_l^{n-1} - \mathbf{x}_l^b)^T\mathbf{B}^{-1}(\delta\mathbf{x}_l^n + \mathbf{x}_l^{n-1} - \mathbf{x}_l^b) \tag{1}$$

$$+\frac{1}{2}\sum_i(\mathbf{H}_i\delta\mathbf{x}_l^n(t_i) - \mathbf{d}_i^{n-1})^T\mathbf{R}^{-1}(\mathbf{H}_i\delta\mathbf{x}_l^n(t_i) - \mathbf{d}_i^{n-1})$$

with subscript l indicating that fields are at low resolution, subscript i the time index, superscript n the minimization index. $\mathbf{x}_l^b$ is the background field truncated at low resolution and $\mathbf{x}_l^{n-1}$ the current estimate of the atmospheric flow (it is equal to the background for the first minimization). $\delta\mathbf{x}_l^n$ is the increment at low resolution at initial time, and $\delta\mathbf{x}_l^n(t_i)$ the increment evolved according to the tangent linear model from the initial time to time index i. $\mathbf{R}_i$ and $\mathbf{B}$ are the covariance matrices of observation and background errors respectively. $\mathbf{H}_i$ is a suitable linear approximation at time index i of the observation operator H_i. The innovation vector is given at each time step by $\mathbf{d}_i^{n-1} = \mathbf{y}_i^o - H_i\mathbf{x}^{n-1}(t_i)$, where $\mathbf{y}_i^o$ is the observation vector at time index i. This innovation vector is computed by integrating the model at high resolution from our current n-1 estimate. The way the increment $\delta\mathbf{x}_l^n$ is then added to the current estimate can be written

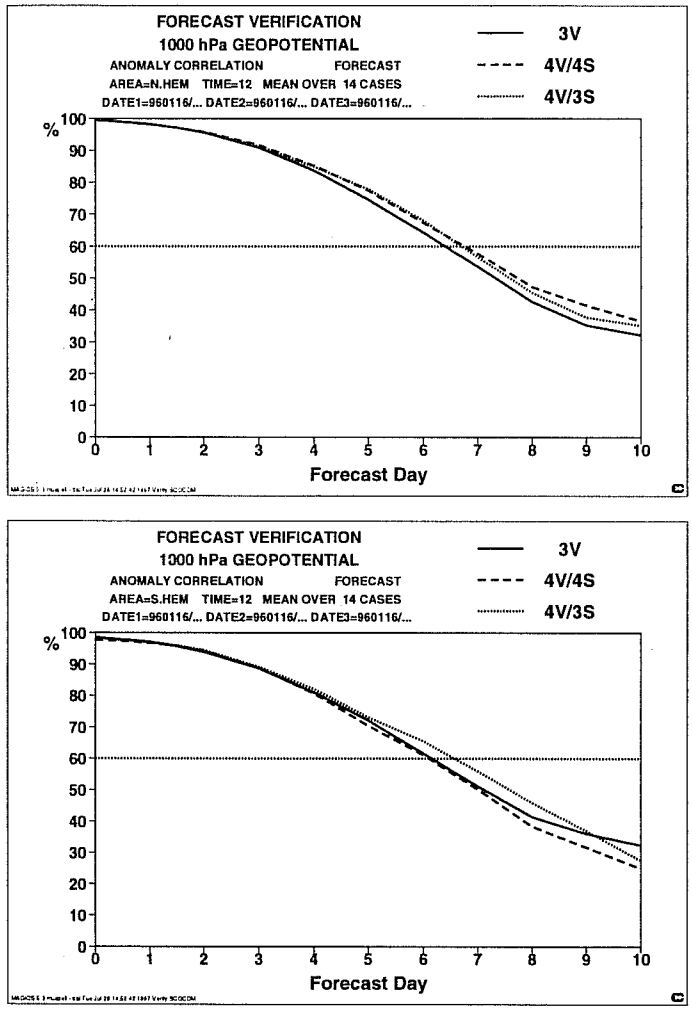

Figure 3 . The anomaly correlation in the Northern Hemisphere (top) and the Southern Hemisphere (bottom) at 1000hPa for two weeks in January 1996. 3V stands for 3D-Var, 4V/3S for 4D-Var with 3D-screening and 4V/4S for 4D-Var with 4D-screening

$$\mathbf{x}^n = \mathbf{x}^{n-1} + NMI(\mathbf{x}^{n-1} + \delta\mathbf{x}^n) - NMI(\mathbf{x}^{n-1}) \qquad (2)$$

where *NMI* stands for adiabatic non-linear normal mode initialization. The original purpose of this use of initialization was to ensure that the analysis was adjusted to the high-resolution orography of the forecast model. As adjustment was likely to be needed predominantly on smaller scales, the initialization was restricted to total wavenumbers 20 and above. During the pre-operational development of 3D-Var it became desirable to initialize all scales of motion in the final incremental initialization step, and this form of initialization was used for the first operational version of 3D-Var. It was also used for each outer loop (or, equivalently, each update of the trajectory) in 4D-Var. In Fig. 4.a, it has thus been used 4 times in 4D-Var. If one uses for 4D-Var a set-up similar to 3D-Var, i.e. one outer loop and only one initialization of the increments, results come much closer, as can be seen by comparing panels b) and c) in Fig. 4. The increment in Fig. 4.b is now only 5.5m (versus 6.6m in Fig. 4.a), and the wind increments are only marginally larger than in 3D-Var. The impact of this initialization of the increments can be seen by comparing panel b) (4D-Var with one update and NMI) with panel d) (4D-Var with one update and no NMI). Initialization reduces the amplitude of the geopotential increment and creates

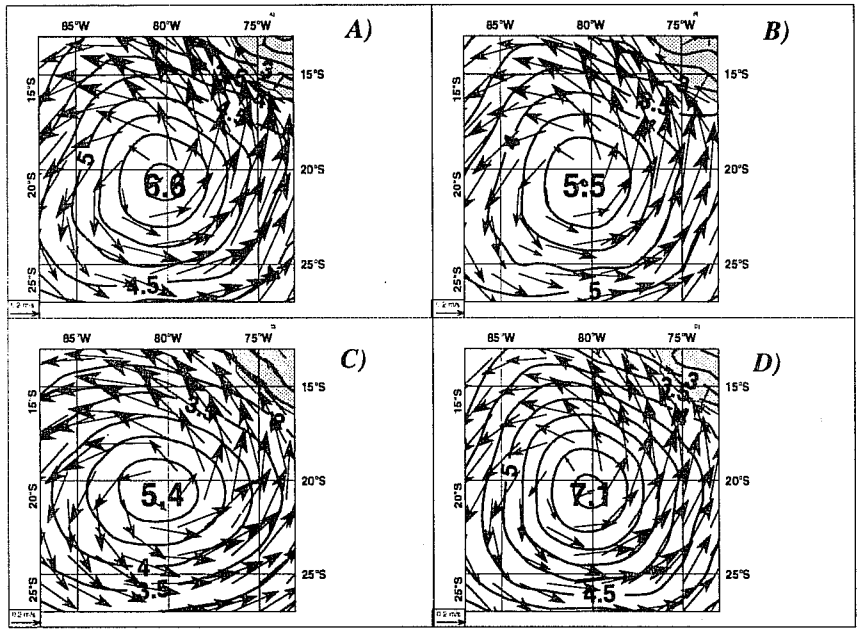

Figure 4 . Mass and wind increments from an isolated mass observation at 850 hPa. Panel a) is 4D-Var with 4 outer-loops and initialization, panel c) is 3D-Var with initialization, panel b) is 4D-Var with 1 outer-loop and initialization, and panel d) is 4D-Var with 1 outer-loop and without initialization. Contour intervals are every 0.5 geopotential meters.

more rotational wind increments associated with mass observations. The actual increment at low resolution created by the first minimisation (not shown) is quite large: 7.3m in the case of the 4D-Var experiments and 7.6m in the 3D-Var case. Not using NMI when going to high resolution alters this value slightly to 7.1m in the 4D-Var case (the resulting increment is displayed in Fig. 4.d). The impact of initialisation on this low resolution increment is dramatic. It reduces the value to 5.5m in the 4D-Var case (Fig. 4.b) and 5.4m in the 3D-Var case (Fig. 4.c). When several outer-loops are used in 4D-Var, the effect of imposing initialization several times and minimizing several times is that the 4D-Var algorithm minimizes as much in terms of mass as when no initialization is performed, and creates rotational winds associated with these mass increments. From an increment value of 5.5m after the first outer-loop, the second minimisation increases the increment to 6.7m. Then, the second initialisation reduces it to 6.5m, before reaching a stable 6.6m value after 4 outer-loops. The cost-function is better minimized by performing several outer loops in that way, but the enforced balance constraint creates large wind increments from mass observations. A test was run performing several outer loops without applying any initialization. The results were then very similar to when only one outer loop was performed without initialization. The main impact was thus confirmed to come from the NMI. The conclusion of these simulated observation experiments is that, with the initialization of the increments performed at each outer loop of the 4D-Var algorithm, the dynamical balance implied by adiabatic non-linear normal-mode initialization is enforced quite strongly in the tropical area. To evaluate whether this is beneficial or detrimental in 4D-Var, it was decided to compare the behaviour of 4D-Var with 4 outer-loops and with one outer-loop on a two-week period (1 to 14 February 1997). The experiments are using the 1997 background constraint. The tropical wind scores are shown in Fig. 5. These scores clearly show that performing 4 outer loops is detrimental for the 4D-Var tropical wind scores. 4D-Var with only one outer loop is competitive with the 3D-Var system in the Tropics. The hydrological budgets are presented for the tropical band (30N to 30S) for the three systems in Fig. 6. 4D-Var has a smaller evaporation spin-up than 3D-Var, but a larger precipitation spin-down. Going to one outer loop

slightly reduces the spin-down of precipitation, but it is still larger than that for 3D-Var. We will see in Part II that this can be remedied by using more physical processes in the 4D-Var minimization.

Besides the number of outer loops performed in 4D-Var, two other factors were found to affect the tropical performance of both 3D- and 4D-Var. Firstly, the change to the 1997 background formulation dramatically improved the tropical scores (Bouttier et al., 1997) in both 3D-Var and 4D-Var. The balance provided by the 1997 background constraint appeared to be sufficiently good to allow the initialization of total wavenumbers below 20 to be abandoned (Simmons and Rabier, 1997). We shall re-examine the impact of the number of outer loops in 4D-Var later. For the time being, experimentation was resumed with 4D-Var with only one outer loop. This baseline configuration has the advantage that it is relatively cheap, and ensures a reasonable tropical behaviour.

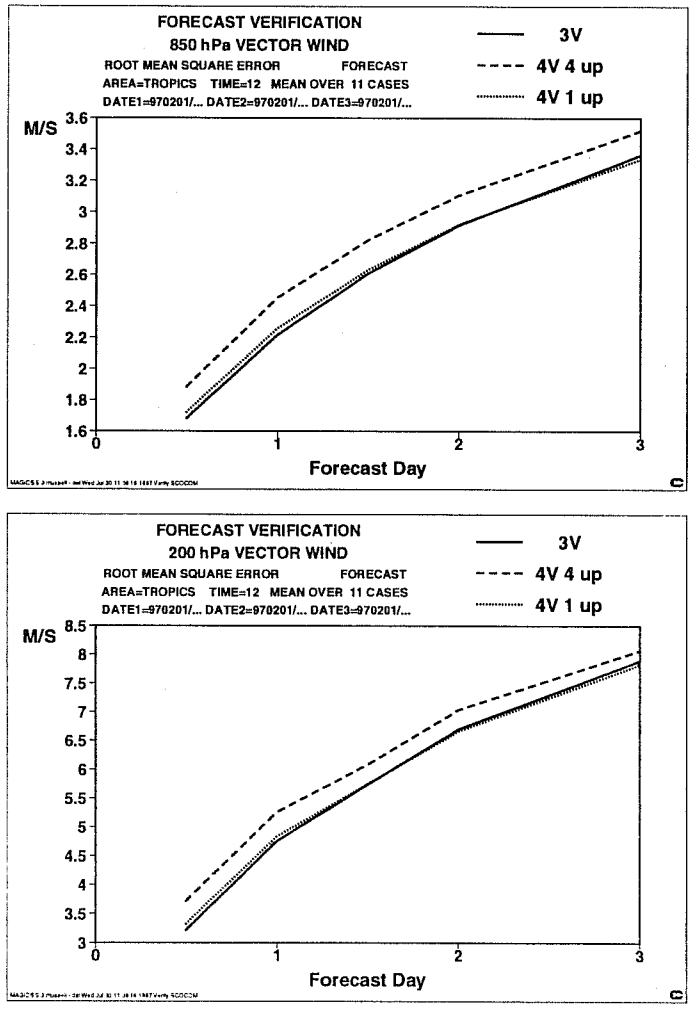

Figure 5 . Tropical wind scores verified against own analysis at 850 and 200 hPa. 3D-Var is shown as a solid line, 4D-Var with 4 outer loops as a dashed line and 4D-Var with one outer loop as a dotted line

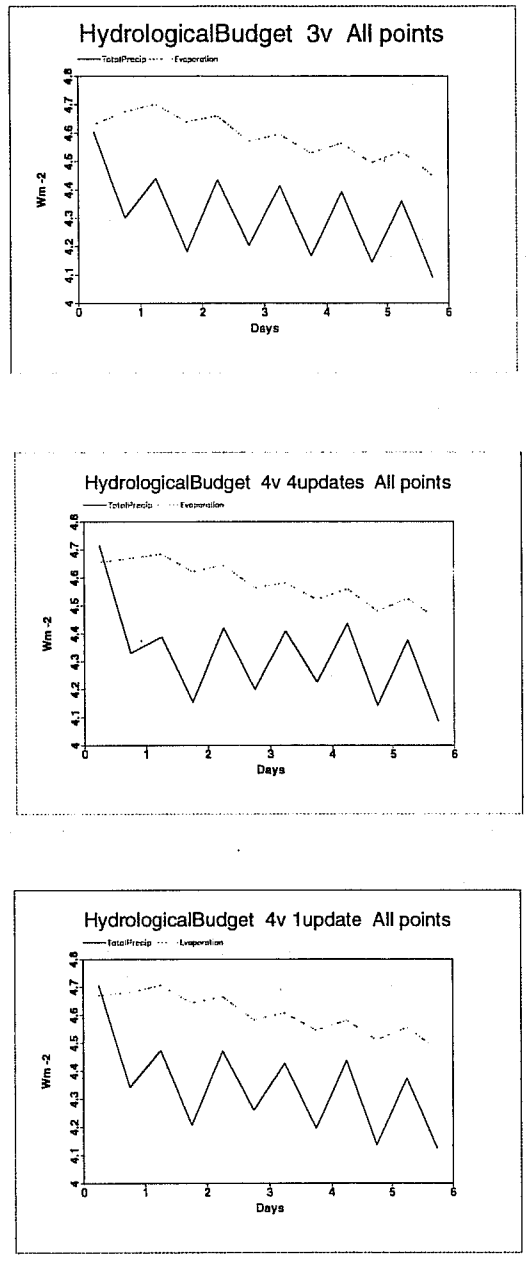

Figure 6 . Hydrological budgets for the tropical band (30N to 30S) averaged over 14 days for 3D-Var (top panel), 4D-Var with 4 outer loops (middle panel) and 4D-Var with one outer loop (bottom panel). Total precipitation is shown as a solid line, and evaporation as a dotted line.

3. BASELINE EXPERIMENTATION WITH 4D-VAR WITH ONE OUTER LOOP

3.1 Results over 12 weeks

Several two to three-week periods were run with 4D-Var with one outer loop and with the 1997 formulation of the background term. The list of periods is : 24 August to 6 September 1995, 25 June to 15 July 1996, 15 to 28 January 1997, 1 to 21 February 1997 and 27 June to 10 July 1997.

In each experiment period, the 3D-Var control and the experimental 4D-Var system were based on the same model version (mainly IFS cycle 16r3), the same background error term and the same data usage. Averaged scores verified against operations and averaged over the 12 weeks are presented in Fig. 7. Impact in the medium range varies from neutral to significantly positive from one period to another, producing a slightly better overall performance from 4D-Var. In the Southern Hemisphere, scores are clearly positive up to day 5. In Europe, scores are positive in the medium-range. The scatter in the medium-range indicates that 4D-Var generally performs better in the cases of relatively bad forecast performance, while having at least as many very good forecasts as 3D-Var (not shown). In the Northern Hemisphere, for the anomaly correlation of geopotential height at 500 hPa at day5, the number of very good forecasts scoring better than 85% is 17 for 3D-Var versus 20 for 4D-Var, while the number of very bad forecasts scoring worse than 65% is 6 for 3D-Var and only 2 for 4D-Var. Similarly for the same score at day7, the number of very good forecasts scoring better than 70% is 9 for 3D-Var versus 12 for 4D-Var, while the number of bad forecasts scoring worse than 40% is 15 for 3D-var versus 12 for 4D-Var. A selection of scores computed with respect to radiosonde observations in the Northern Hemisphere is presented in Fig. 8. One can notice a better performance of the forecast at all ranges for parameters very relevant for the bench forecaster (temperature and wind at low levels for instance).

Since short-range forecast scores show less scatter than medium-range scores, the short-range performance differences represent a larger statistical significance than those for the medium-range. For each individual 2-to-3 week period, 4D-Var was found to behave better in the short-range (up to day 3) than 3D-Var. The difference in performance at day 1, averaged over the total of 12 weeks, is illustrated in Fig. 9. Cross-sections of differences between root-mean-square (RMS) error at day 1 between 4D-Var and 3D-Var are presented for both hemispheres. They are computed for the mid-latitude band 40 to 70 degrees. The values are predominantly negative throughout the atmospheric depth, which means that 4D-Var has a smaller error than 3D-Var almost everywhere. The contour interval is double in the Southern Hemisphere cross-section where the impact of 4D-Var is larger. In the Northern Hemisphere, the largest improvement is over the oceans (160E to 240E for the North Pacific and 280E to 340E for the North Atlantic). The RMS errors are also reduced over Europe (350E to 30E). The picture is not so clear over Asia with a large positive impact of 4D-Var in the mid-troposphere at around 100E, but a negative one at 50E and 150E.

3.2 Results over the February 1997 period

The very short range forecasts can be investigated by computing the root-mean-square fit of the background field to the observations for both systems. Figure 10 shows the averaged fit to the radiosonde data over the first two weeks of February 1997. The fit of background to observations is very relevant to judge the quality of the assimilation system as the short-range forecast errors and the observation errors are generally uncorrelated. For these (solid lines) 4D-Var clearly outperforms 3D-Var for both mass and wind (for wind mainly at the jet level in the Northern hemisphere) in all areas. A statistical test found a significantly better fit of the 4D-Var background fields to the observations for the tropospheric height in all areas, and for the wind at upper levels in the Northern hemisphere (the statistical test was also carried out for other periods, including summer periods, and the significance of a better fit for geopotential height data was confirmed in all areas). Differences between 3D-Var and 4D-Var in the fit of the analyses to the data used in the analyses are not of themselves a performance criterion. The relevant conclusion which can be made from the analysis fits (dashed lines) is that both systems fit the data reasonably well. 4D-Var analysis increments are generally smaller than 3D-Var ones which is consistent with an improved short-range forecast.

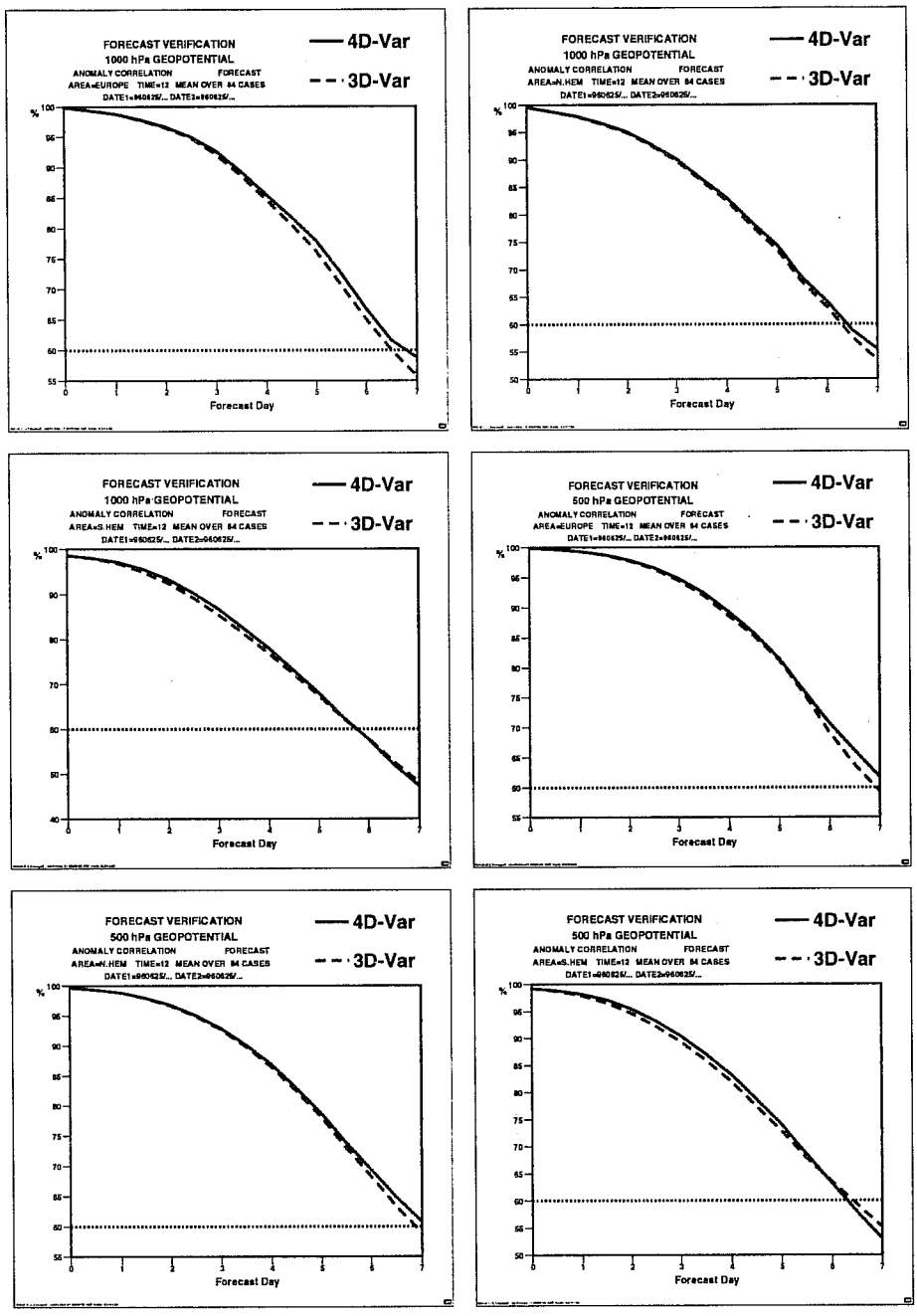

Figure 7 . Anomaly correlation scores for forecasts from 4D-Var (solid line) and from 3D-Var (dashed line), averaged over 12 weeks. Scores are shown for geopotential height at 1000 hPa and 500 hPa, for Europe, the Northern Hemisphere and the Southern Hemisphere.

This period of February 1997 is particularly interesting because of the FASTEX experiment (Joly et al., 1997) which took place in the Atlantic Ocean storm-track. In the remainder of this section, we will concentrate on the discussion of the performance of 4D-Var in the Atlantic area over this period. Both 3D-Var and 4D-Var have used the same observations: on top of the standard set of observations, the extra radiosonde measurements have been taken into account. As far as

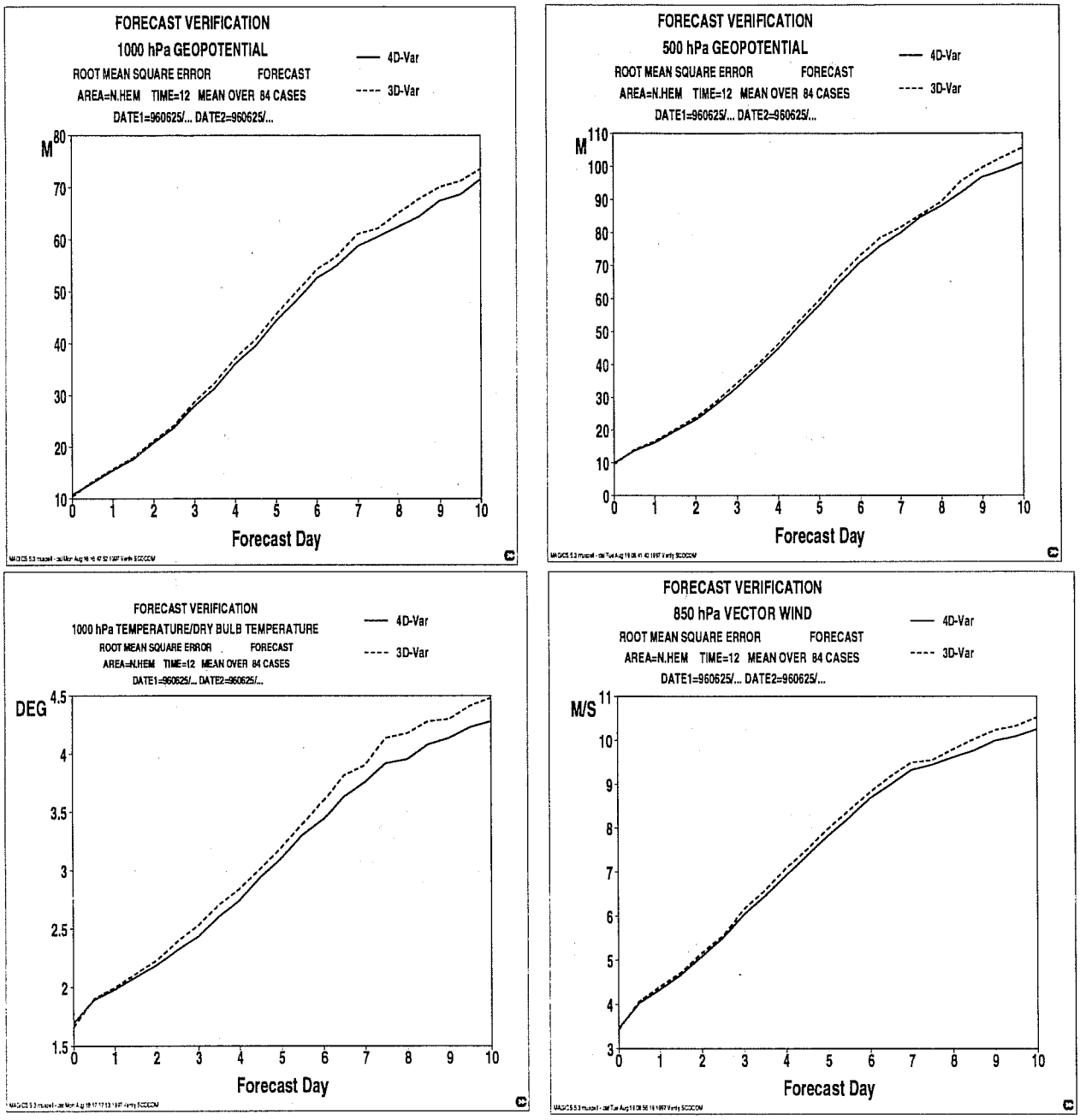

Figure 8 . Root-mean-square of the difference between forecasts from 4D-Var and observations (solid lines) and forecasts from 3D-Var and observations (dashed line), averaged over 12 weeks in the Northern Hemisphere for various parameters (geopotential height at 1000 hPa and 500 hPa , temperature at 1000 hPa, vector wind at 850 hPa).

synoptic cases are concerned, two examples are shown in Figs. 11 and 12. They correspond to interesting dates during the FASTEX experiment (Joly et al., 1997). The first synoptic case is the Intensive Observing Period IOP 12, during which a cyclone evolved in the vicinity of Iceland with a deepening of 19 hPa in 24 hours seen in the manual UKMO analysis of surface pressure: the surface pressure fell from 966 hPa the 9 February 1997 at 12UTC to 947 hPa the 10 February 1997 at 12UTC. The deepening in the 24-hour forecast from the 4D-Var analysis is very similar to the analysed one, pressure

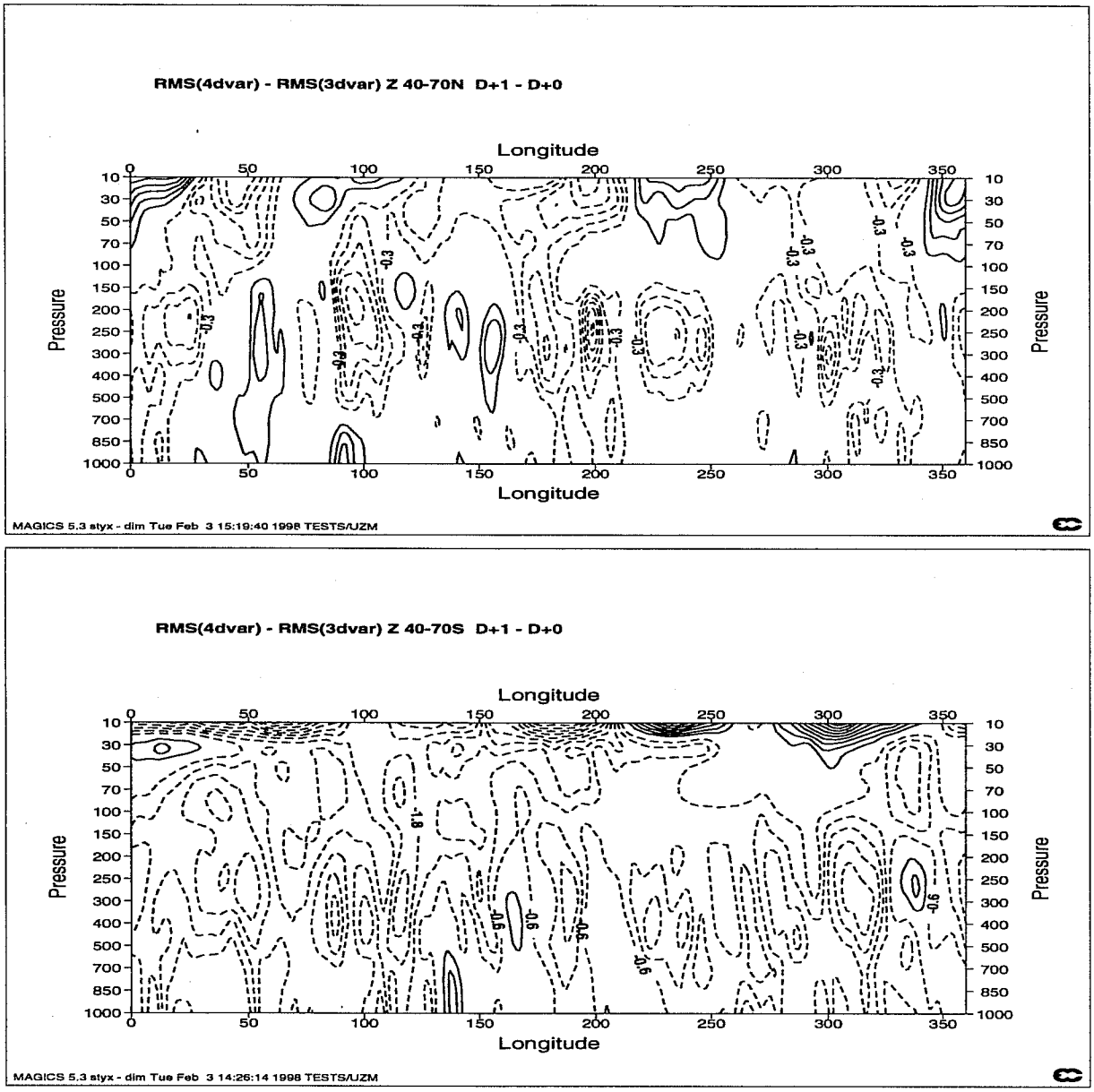

Figure 9 . Cross-sections of the difference between the root-mean-square error (RMSE) of forecasts from 4D-Var and from 3D-Var at day1, averaged over 12 weeks and over the latitude band 40 to 70 degrees. Contour interval is 0.3 m for the top panel representing the Northern Hemisphere and 0.6 m for the bottom panel representing the Southern Hemisphere. Negative contours mean that 4D-Var has a smaller RMSE.

dropping from 970 hPa to 951 hPa in the same period of time. In contrast, cyclogenesis in the corresponding forecast from 3D-Var is not intense enough, with a deepening of only 10 hPa, from 969 hPa to 959 hPa. The analyses and 24-hour forecasts for both 3D-Var and 4D-Var are shown in Fig. 11, together with the verifying analyses. The second synoptic case (IOP 17) is a depression, with a manually analysed central pressure of 937 hPa on 20 February 1997 at 0UTC. In the medium-range, the forecast from 12 February 1997 at 12UTC is significantly better for 4D-Var, as shown in Fig. 12 comparing the forecasts from 3D-Var and 4D-Var with their verifying analyses. Of course, this only corresponds to one forecast, and this dramatic improvement was not found systematically for all forecasts verifying on this date. To judge the overall quality of the 4D-Var forecasts for this particular event, we computed the averaged root-mean-square error of the intensity of the low, averaged over all the forecasts from the 11th to the 19th verifying on 20 February 1997 at 0UTC is

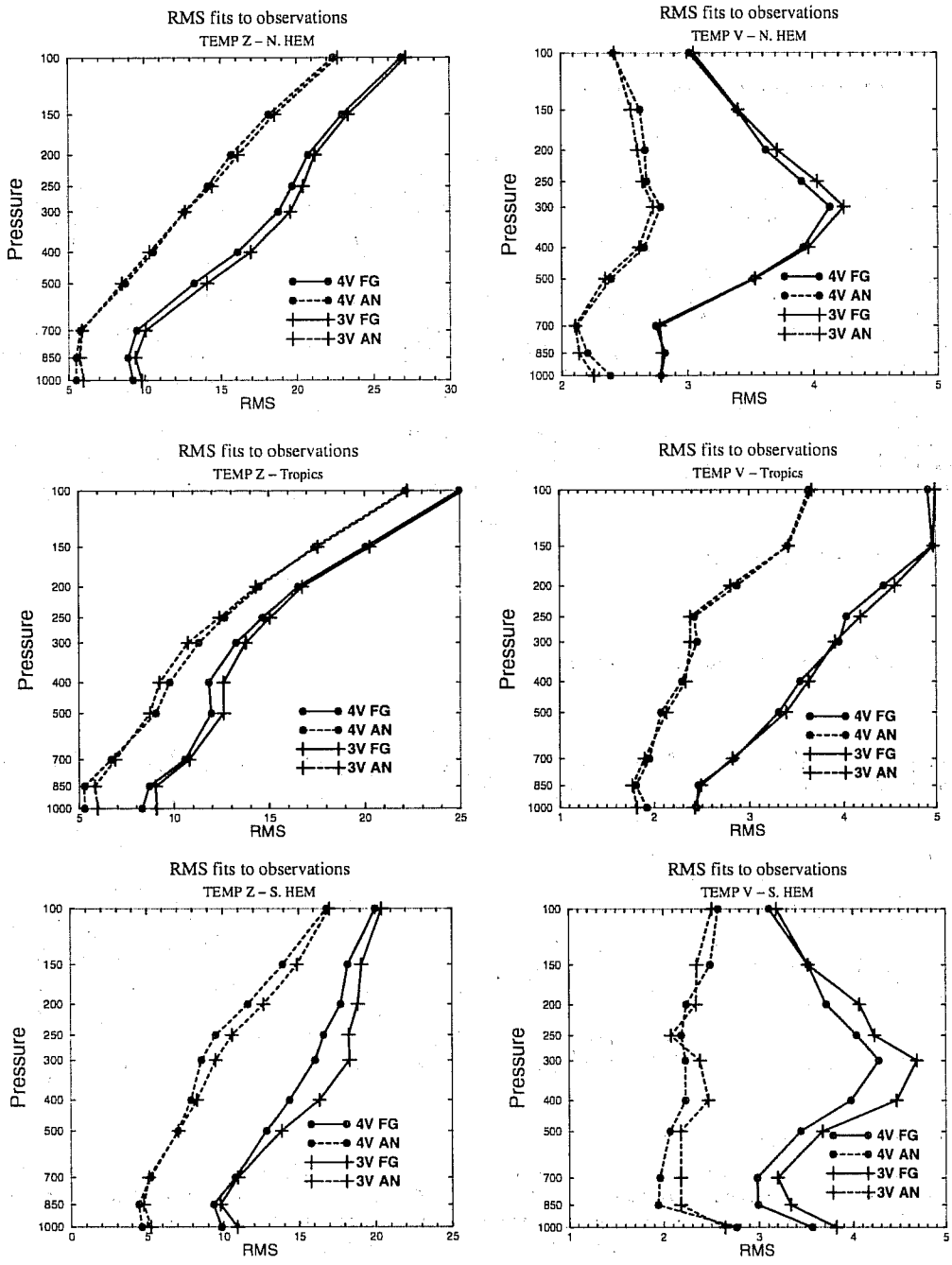

Figure 10 . Root-mean-square (RMS) fits to the radiosonde height and meridional component of the wind data produced over the Northern Hemisphere, the Tropics and the Southern Hemisphere, averaged over 1 to 14 February 1997. The solid lines represent the RMS fits of the backgrounds to the observations, the dashed ones the RMS fits of the analyses to the observations. 4D-Var is shown as circles, 3D-Var as plusses. The abscissa is the RMS in geopotential units and m/s. The ordinate is the pressure in hPa.

15 hPa for 4D-Var versus 19 hPa for 3D-Var, which corresponds to a 20% improvement. One can also notice a better analysed surface pressure low with 4D-Var (940 hPa versus 943 hPa).

Over the whole period for which both 3D-Var and 4D-Var were run (1 to 21 February), 4D-Var performed better on average in the Atlantic area. Fig. 13 represents the cross-sections of differences of root-mean-square error against its own

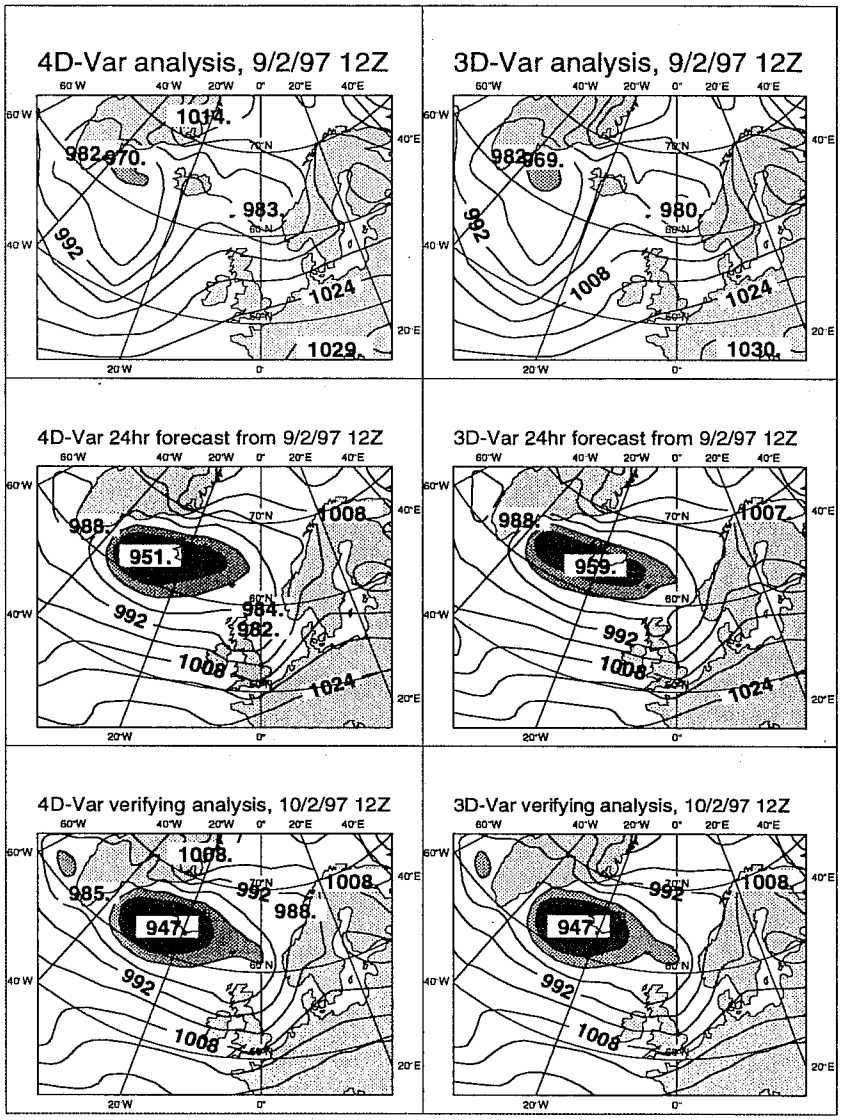

Figure 11 . Maps of mean-sea-level pressure for the synoptic case corresponding to IOP 12 of FASTEX. The top panels show the analyses for the 9 February 1997 at 12UTC, the middle panels the 24-hour forecasts from 9 February 1997 at 12UTC, and the bottom panels the verifying analyses. 4D-Var charts are on the left, 3D-Var on the right.

analysis at day 1 between 4D-Var and 3D-Var for the mid-latitude band 40 to 70 degrees over the Atlantic area (310E to 350E). The values are predominantly negative throughout the atmospheric depth for both height (top panel) and zonal component of the wind (bottom panel), which means that 4D-Var has a smaller error than 3D-Var. The largest impact can be seen in the higher troposphere and lower stratosphere, with a maximum around 300 hPa (this is also where the largest improvement was found in terms of temperature and meridional component of the wind, not shown). In terms of values

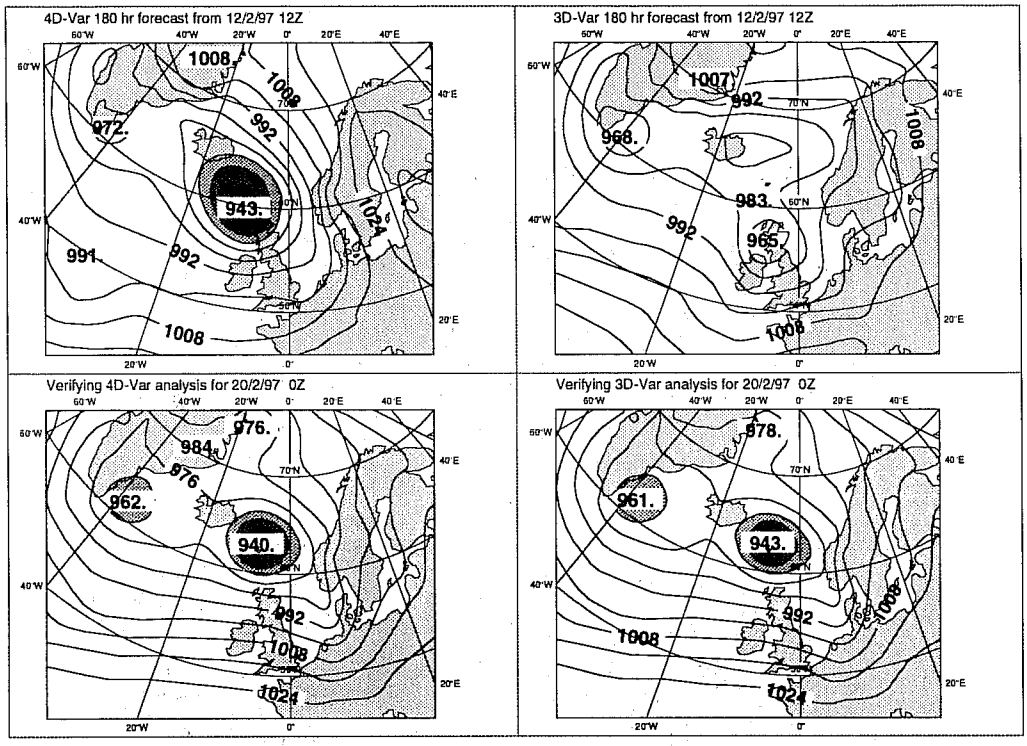

Figure 12 . Maps of mean-sea-level pressure for the synoptic case corresponding to IOP 17 of FASTEX. The top panels show the forecasts from 12 February 1997 at 12UTC verifying 20 February 1997 at 0UTC, and the bottom panels the verifying analyses. 4D-Var charts are on the left, 3D-Var on the right.

of 500 hPa geopotential scores over the Atlantic area, at the 12-hour range the root-mean-square error (RMSE) are 10.8m and 8.8m respectively for 3D-Var and 4D-Var, while at the 24-hour range they are 13.2m and 11.9m. The same scores computed over the European area for slightly longer ranges are consequently improved by 4D-Var: at the 24-hour range, the RMSE are 14.3m and 13.5m respectively for 3D-Var and 4D-Var, while at the 48-hour range they are 27.3m and 25.6m. The better results for 4D-Var are also confirmed when comparing forecasts with available observations (not shown).

The Atlantic ocean is a particularly sensitive area over this period as indicated by the operational "key analysis errors". As explained in Klinker et al. (1998), these are obtained by finding some increments to the analysis which would significantly reduce the two-day error of the ensuing forecast. They usually highlight the areas which are both dynamically unstable and not so well resolved by observations. In this case, the "key analysis errors" (not shown) have a relatively large amplitude of more than 2m South-West of Iceland (around 20W to 40W, and 50N to 70N). One can then compute the difference between the 3D-Var and 4D-Var analyses in this area: the RMS of the differences between the analyses is the largest at around 20 to 30W and 50 to 60N (not shown). The analyses are mostly different at the same longitudes as the ones highlighted by the "key analysis errors" (20 to 40W), although the maximum is shifted towards lower latitudes. The maximum differences between 3D-Var and 4D-Var analyses are around 15m in RMS terms, which is significantly larger than the corresponding 2 to 3 m values for the "key analysis errors". This is perfectly understandable, as the difference between the two assimilation systems gives an estimate of the total uncertainty within the analysis, whereas the "key analysis errors" only describe the fast-growing part of this uncertainty. In any case, the two sets of analyses appear to be significantly different over an area known to be sensitive for the quality of the resulting forecasts.

For a given meteorological situation (9 February 1997 already studied above), 3D-Var and 4D-Var are compared to "sensitivity analysis". This "sensitivity analysis" is obtained by substracting the "key analysis errors" to the 4D-Var

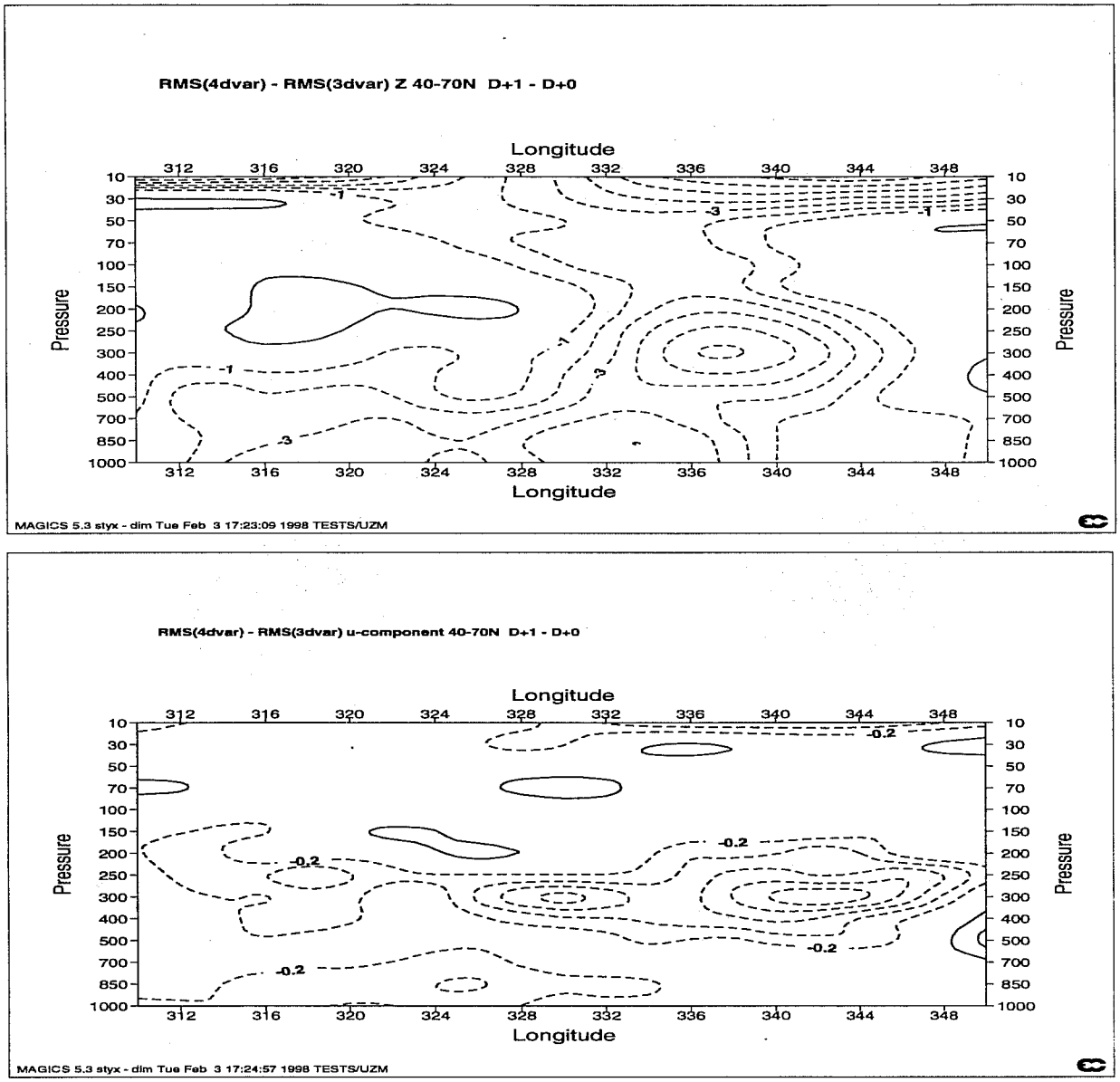

Figure 13 . Cross-sections of the difference between the root-mean-square error (RMSE) forecasts from 4D-Var and from 3D-Var at day1, averaged over 3 weeks and over the latitude band 40 to 70 degrees over North Atlantic. Contour interval is 3 m for the top panel representing the geopotential height and 0.2 m/s for the bottom panel representing zonal component of the wind. Negative contours mean that 4D-Var has a smaller RMSE.

analysis, and is the best estimate of the truth at our disposal (it was shown in Klinker et al., 1998 that the "sensitivity analysis" can both provide a better fit to the data and a better ensuing forecast). The difference between 3D-Var and the sensitivity analysis on the one hand and between 4D-Var and the sensitivity analysis on the other hand are then considered as 3D-Var and 4D-Var errors. These errors are projected on the first singular vectors describing the unstable manifold (Buizza, 1994). The first three singular vectors contributing to the Atlantic area (80W to 20W; 30N to 90N) are chosen . The energy norm of the 4D-Var and 3D-Var errors, projected onto this unstable manifold are respectively 2.43 kgs^{-2} and 3.22 kgs^{-2}. The 4D-Var error is then 25% smaller than the 3D-Var error on the unstable manifold defined by the first three singular vectors over the Atlantic area for this particular case. Of course, this computation is performed on one meteorological situation only, and the choice of the "truth" is not unique. However, the results indicate that 4D-Var is

behaving better on the unstable manifold than 3D-Var. Another indication will be given in the third paper of the series (Klinker et al., 1999) in which key analysis errors are shown to be smaller than for 3D-Var on a 40-day parallel experimentation.

When decomposed in mean and standard-deviation components, the large values of the RMS of the difference between 3D-Var and 4D-Var analyses are seen to come mainly from the standard-deviation part (not shown). In the mean difference, 4D-Var analyses exhibit slightly higher geopotential values, and a slight reduction of the Eady index estimating the instability of the flow (not shown). These differences do not seem to be very well structured and significant. More importantly, one can compare the standard-deviations of analysis for both systems. These are shown in the bottom panels of Fig. 14. The standard-deviation of the 4D-Var analysis (right panel) is quite noticeably larger than the one for 3D-Var (left panel). Differences between the two can reach up to 6m around 50N, 30W, where a typical value of standard-deviation of analysis is 115m. It thus corresponds to a 5% increase in the variability of the analysis.

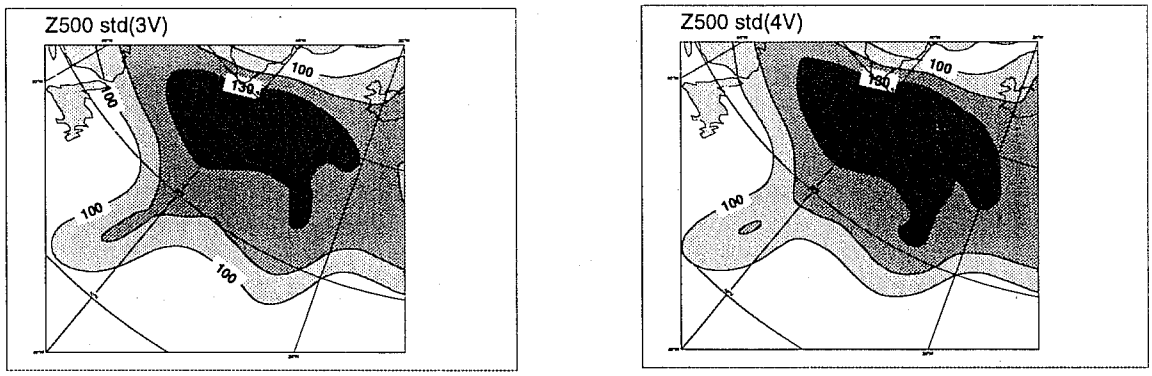

Figure 14 . Maps of quantities related to the geopotential height at 500 hPa averaged over the period 1 to 21 February 1997 over the Atlantic Ocean. The left panel shows the standard-deviation of the 3D-Var analysis computed over all the 12 UTC analyses of the period and the right panel shows the standard-deviation of the 4D-Var analysis computed over all 12 UTC analyses over the period. The units are the geopotential metre. Contour intervals are 15 m starting at 120 m.

The larger variability does not appear to be caused by larger increments at each analysis cycle. Similarly to what was shown in the previous paper by Rabier et al. (1998b), in the Atlantic area 4D-Var actually produces smaller increments than 3D-Var (not shown). To clarify the use of observations in the area, the fit to the data was computed explicitly for the area 40N to 70N and 50W to 10W. The number of radiosondes in the area is slightly higher than normal for the period due to the extra observations taken during the FASTEX experiment. In total, there are around 300 data at each pressure level entering the root-mean-square calculations illustrated in Fig. 15. Even though this is a limited sample, it is obvious from Fig. 15 than the background fields used in the 4D-Var assimilation fit the data much more closely than the background fields used in the 3D-Var assimilation. In particular, at the 300 hPa level, the RMS fit to the height data is 17.6 m for 4D-Var versus 21.6 for 3D-Var, and the equivalent values for the meridional component of the wind are 4.3m/s versus 5.4 m/s respectively. The very-short range forecasts from 4D-Var agree much better with the synoptic observations than the equivalent forecasts from 3D-Var. Even though the 4D-Var analysis produces smaller increments than 3D-Var, it can produce an analysis of more variance, and presumably better quality. Comparing the background fields of both systems with asynoptic observations is more problematic as 4D-Var backgrounds are compared with the observations at their valid time within the assimilation window whereas 3D-Var backgrounds are compared with the observations gathered at the central analysis time only. This is illustrated for aircraft observations over the Atlantic area in Fig. 16. This figure represents the fit of the backgrounds to the data, in RMS terms, averaged over all levels between 150hPa and 400hPa, by

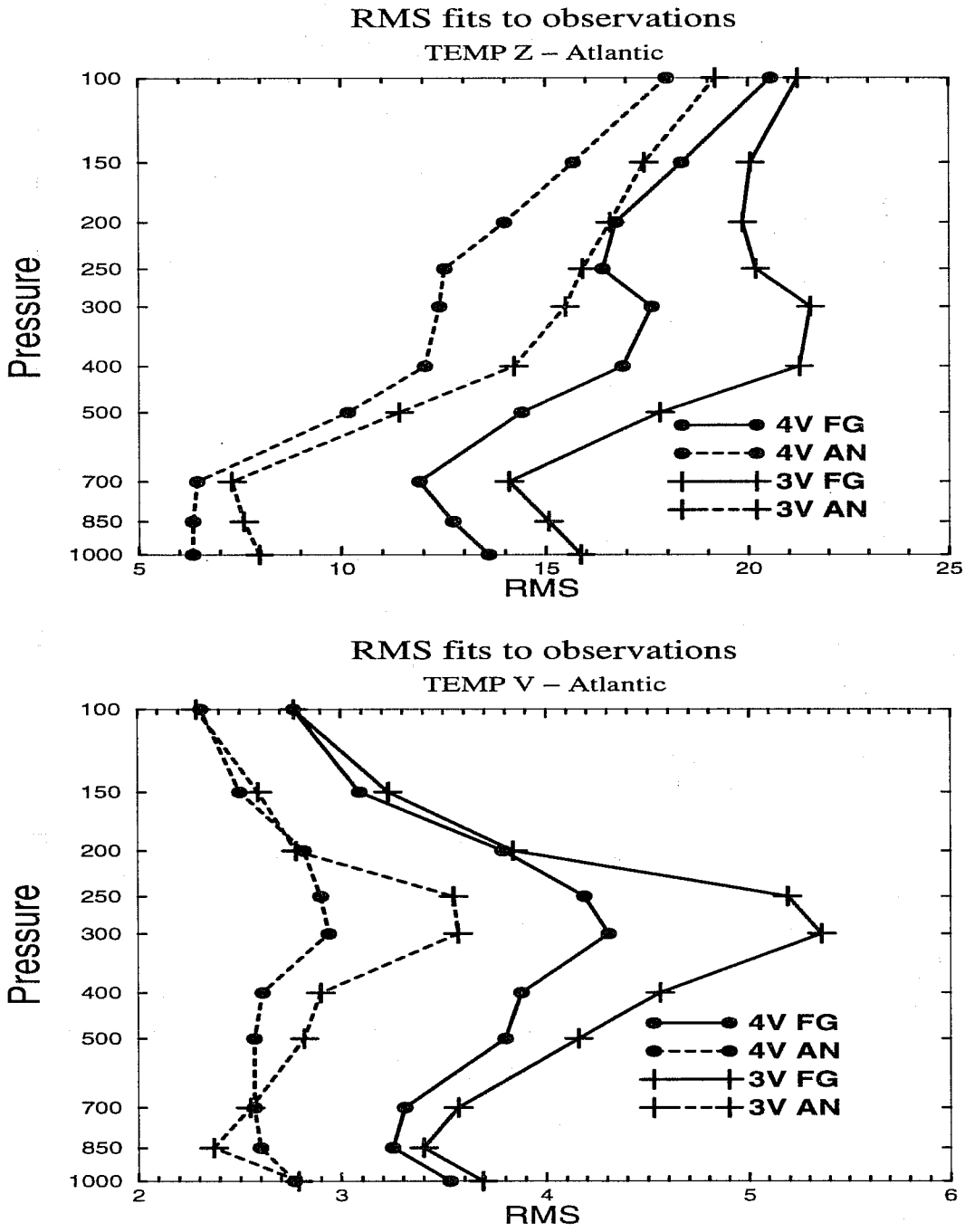

Figure 15 . Root-mean-square (RMS) fits to the radiosonde height and meridional component of the wind data produced over the Atlantic area, averaged over 1 to 21 February 1997. The solid lines represent the RMS fits of the backgrounds to the observations, the dashed ones the RMS fits of the analyses to the observations. 4D-Var is shown as circles, 3D-Var as plusses. The abscissa is the RMS in geopotential units and m/s. The ordinate is the pressure in hPa.

timeslots. There are 7 timeslots: the first and seventh are half-hour slots, while the remaining 5 are one-hour slots. Timeslot number 4 is then centered around the main synoptic time. Data are accumulated every 12 hours, over the entire 21 day assimilation period. The average number of observations is 1000 per timeslot. The RMS-fit to the data is shown as a solid line for 3D-Var and as a dashed line for 4D-Var. For 4D-Var, the fit is rather constant with respect to timeslot number,

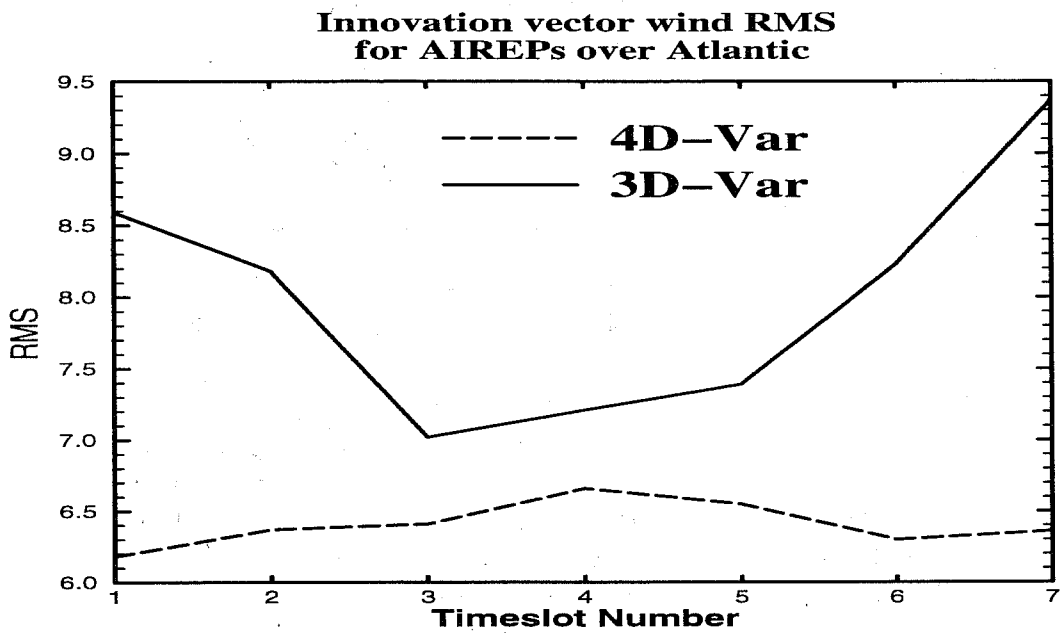

Figure 16 . Root-mean-square (RMS) of the innovation vector wind for aircraft data produced over the Atlantic area, averaged over 1 to 21 February 1997, every 12 hours. 4D-Var is shown as dashed line, 3D-Var as solid line. The abscissa is the timeslot number. The ordinate is the RMS in m/s.

around 6 to 6.5 m/s in vector wind. For 3D-Var, there is a convex curve, showing a better fit around the middle of the assimilation window (around 7 to 7.5 m/s) and a much degraded fit towards the sides of the assimilation window, reaching values around 9 m/s. At the central time, when the comparison between 3D-Var and 4D-Var is fair, one can note that the 4D-Var background is closer to the observations than the 3D-Var one, which is consistent with what has been observed with the radiosonde data. Obviously, the sample is not large enough to obtain smooth curves, but the figure describes quite clearly how the innovation vector can be degraded if the observations are not used at their appropriate time. Although this difference in the innovation vectors is quite spectacular, it is mainly presented to illustrate the importance of dynamical processes on a 6-hour window. It is not believed that the main improvement brought is only due to a better computation of the innovation vector, but primarly to using the tangent-linear dynamics in evolving the increments. However, this is indeed a particular area where we expect an improvement from 4D-Var: the proper use of asynoptic data.

4. Structure functions in 4D-Var

The results presented above point to a better short-range forecast from 4D-Var than from 3D-Var in the Atlantic area during February 1997, probably implying a better analysis. In order to understand the better behaviour of 4D-Var in cyclogenetic situations, it is useful to illustrate the impact of the dynamics on the increments caused by one observation in this sort of meteorological situation. Structure functions can easily be illustrated by single observation experiments, as in Thépaut et al. (1996). For a particular date (5 December 1996, 0UTC), a baroclinic area was chosen in the West Atlantic. The background for this date, which corresponds to a 6 hour forecast from the last operational analysis exhibits fields which are tilted in the vertical. This is illustrated in Fig. 17 in a cross-section of the zonal wind component. One observation of geopotential height was inserted at location 60W, 40N, at 850 hPa in 4D-Var, either at time 21UTC (3 hours before the main synoptic time 0UTC), or at 0UTC, or at 03UTC (3 hours after the main synoptic time 0UTC). Each time, the initial departure from the background is equal to 10m. The structure functions can then be illustrated for three different scenarios:

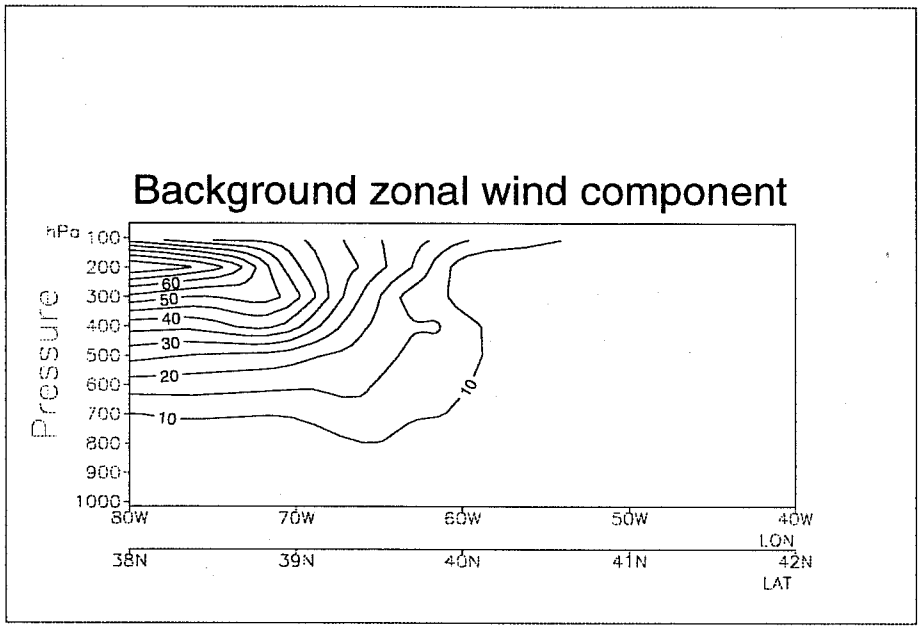

Figure 17 . Cross-section for the zonal wind component of the background for 5 December 1996, 0UTC. Unit is m/s.

an observation at the beginning of the assimilation window (21UTC), in the middle (0UTC) and at the end (03UTC). These are shown, at the time of the observation, in Fig. 18 (which shows cross-sections along the same line as Fig. 17). This particular meteorological situation is not rapidly developing which allows comparison of the increments even if their validity time can be up to 6 hours different. The top panel in Fig. 18 corresponds to the increments at 21UTC created from an observation at 21UTC. These are similar to the increments which would have been created by the 3D-Var system. They are barotropic, the value decreasing with height and horizontal distance from the observation location. When the observation is located 3 or 6 hours after the initial time of the assimilation window, some influence of the dynamics is noticeable. The increments in the middle and bottom panels of Fig. 18 are tilted in the vertical, in a way consistent with the meteorological situation. The longer the time gap between the initial time and the observation time, the tighter the structure functions get at the surface and the more the impact of the observation spreads vertically (following the baroclinic tilt). One can also illustrate the increments produced by the 4D-Var system at the main synoptic time 0UTC from these three configurations. These are shown in Fig. 19. Although Fig. 18 illustrates how different the structure functions can be, the actual increments at a given time (here, the main synoptic time, corresponding to the middle of the assimilation window) are much more similar. This can be regarded as reassuring, as it implies that there is no real conflict in the 4D-Var system between observations inserted at different times within the window. However, the structure functions illustrated in Figure 18 still appear to be constrained to a large extent by the initial static covariance matrix illustrated in the top panel. This indicates that 4D-Var will still benefit from a better specification of the initial covariance matrix by a Kalman filter for example.

5. SUMMARY AND DISCUSSION

This article is the first of three papers describing the steps leading to the operational implementation of incremental 4D-Var at ECMWF on 25th November 1997. This first paper concentrates on results produced with 4D-Var on a 6-hour window at "operational" resolution T213L31/T63L31, using very simplified physics during the minimisation. Firstly, the sensitivity to different 4D-Var set-ups is investigated. It is found that using additional asynoptic data is not useful at the present stage. Later results using a full four-dimensional variational quality control allowing for time-correlation of

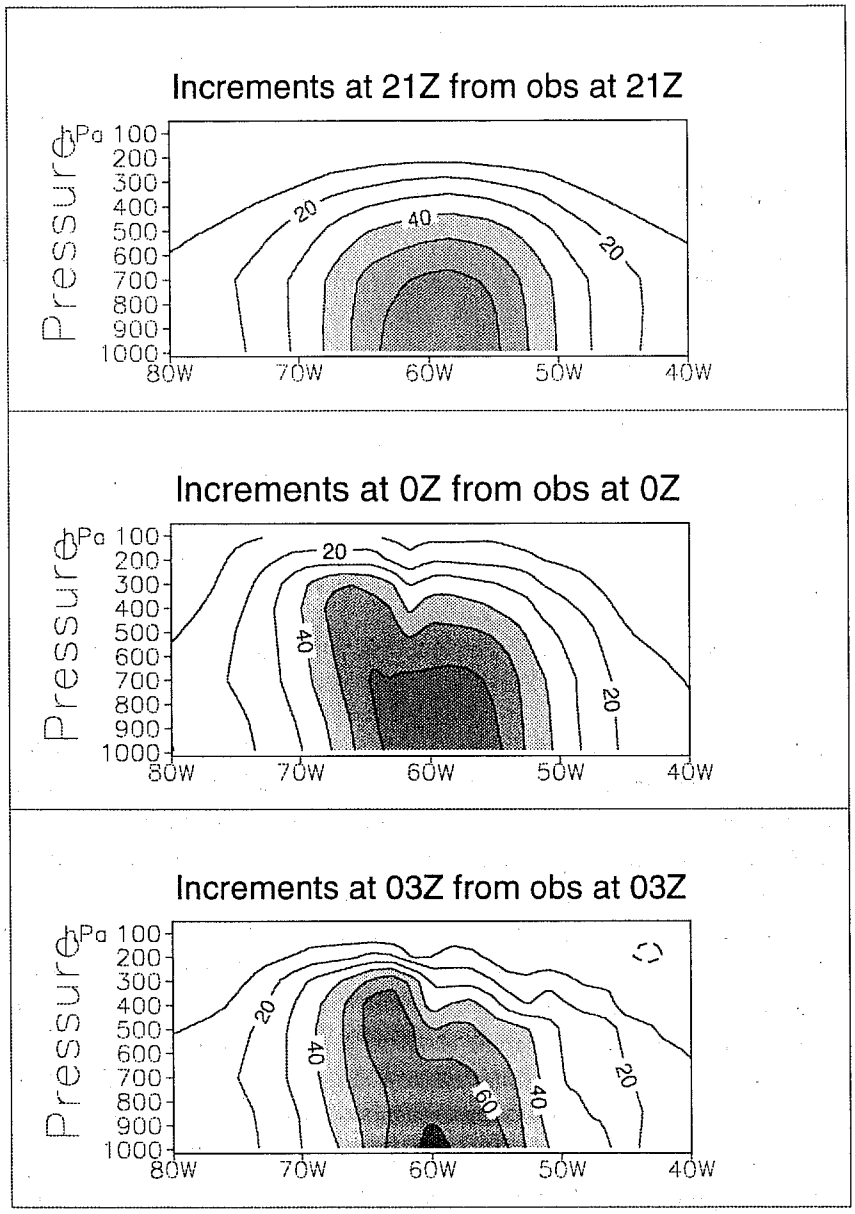

Figure 18 . Structure functions for a height observation at 850 hPa, 40N, 60W. Isolines show the resulting increment, in geopotential unit. The top panel corresponds to an observation at 21UTC, the middle panel to an observation at 0UTC, and the bottom panel to an observation at 03UTC.

observation errors have been found to improve 4D-Var performance, but these developments were not available for operational implementation and will be part of later improvements to the system. Investigation of the poor performance of 4D-Var in the Tropics revealed some sensitivity to the way the adiabatic non-linear normal mode initialization of the increments was performed. Going from four outer loops to only one (as in 3D-Var) helped to reduce the problem, together

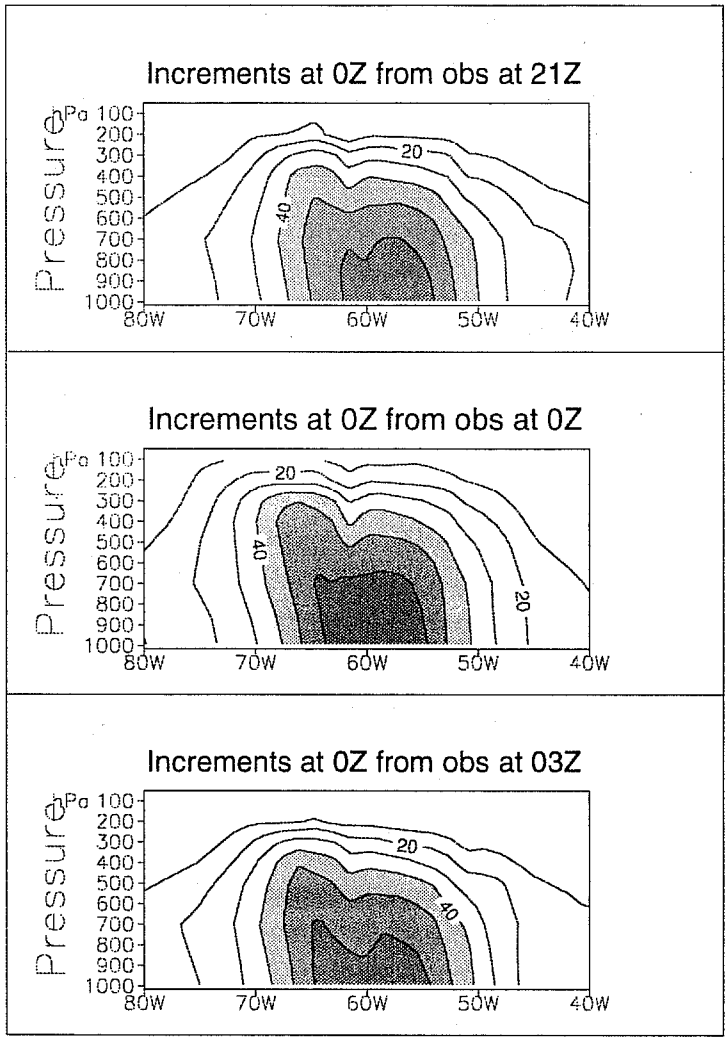

Figure 19 . Increments at 0UTC for a height observation at 850 hPa, 40N, 60W. Isolines show the resulting increment, in geopotential unit. The top panel corresponds to an observation at 21UTC, the middle panel to an observation at 0UTC, and the bottom panel to an observation at 03UTC.

with a change to the 1997 background formulation and an initialization of only the small scales. Tropical scores then became only marginally worse for 4D-Var than for 3D-Var.

Twelve weeks of experimentation with one outer-loop in 4D-Var and with the 1997 background formulation have been studied. These include 7 weeks of summer and 5 weeks of winter. In the medium range, each two to three-week period has been found to be either neutral or positive, resulting in slightly positive averaged scores computed either with respect to a verifying analysis or with respect to observations. In the short range, each two to three-week period has been found to be slightly positive throughout the troposphere. The largest improvement is seen to come from the Southern Hemisphere as a whole and from the Northern Hemispheric mid-latitude oceanic areas. The better short-range performance of the 4D-Var system is also shown by the fits of the background fields to the data. Detailed results were presented over the Atlantic Ocean area during the FASTEX period, during which 4D-Var is found to perform better. In individual synoptic cases corresponding to interesting IOPs during the FASTEX period, 4D-Var has a clear advantage over 3D-Var during rapid

cyclogeneses. This advantage does not appear to stem from differences in mean analysis fields, but rather from relatively large day-to-day differences between the 4D-Var and 3D-Var analyses in a sensitive area. Results on one IOP indicate that 4D-Var is behaving better on the unstable manifold than 3D-Var. 4D-Var analyses display more day-to-day variability than their 3D-Var counterparts. This is believed to bear improvement as in general, the 4D-Var system does not produce larger analysis increments than the 3D-Var system. The very short-range forecasts used as backgrounds are remarkably closer to the data over the Atlantic area for 4D-Var than for 3D-Var, which explains why smaller increments can lead to more variable analyses. The small adjustments made by 4D-Var at each analysis cycle are sufficient to depict accurately the rapid evolution of the flow in this storm-track area.

Structure functions were illustrated in the 4D-Var case for a height observation inserted either at the beginning of the assimilation window, in the middle or at the end. As anticipated from previous results, the dynamical processes seem to be relevant, even on a short 6-hour assimilation period. More influence of the dynamics could be taken into account by a proper cycling of 4D-Var using a simplified Kalman filter which is currently being developed (Fisher, 1998) . To be cycled in a cost-effective way, the simplified Kalman filter would need longer 4D-Var windows, which is an option to be studied in the near-future.

As far as cost issues are concerned, the main analysis job of 4D-Var with only one outer loop and no physics (our baseline 4D-Var) costs about twice as much as the main analysis job of 3D-Var. The additional costs of 4D-Var are mainly due to the extra model integrations which run efficiently. Comparing the relative costs of a whole day of assimilation plus one 10-day forecast, the baseline 4D-Var costs 40% more than 3D-Var, in line with expectations.

Although this paper has emphasized the role of the dynamics in 4D-Var results, there is scope for improvement in both the extratropics and the Tropics by including more elaborate physical processes within the 4D-Var minimisation. Paper II (Mahfouf and Rabier, 1999) will describe a set of physical processes developed, tested, and included in the 4D-Var experimentation. The resulting 4D-Var system was then tested over 80 days of experimentation with positive results. An additional 6-week parallel suite prior to implementation clearly showed the benefit of 4D-Var, with significant improvements with respect to 3D-Var. They will be described extensively in Paper III (Klinker et al., 1999).

6. ACKNOWLEDGEMENTS

It is a pleasure to acknowledge all the colleagues at ECMWF who have contributed to this 4D-Var experimentation. This includes all the staff of the data assimilation section (P. Undén, Erik Andersson, F. Bouttier, M. Fisher, J. Haseler and D. Vasiljevic), of the migration project (M. Hamrud, L. Isaksen and S. Saarinen) as well as of the satellite section (R. Saunders, G. Kelly, B. Harris, D. Lemeur, T. McNally). The success of the 4D-Var project relies heavily on its history and in particular on the contributions by P. Courtier and J.-N.Thépaut. A. Hollingsworth has been extremely supportive, providing us with all the resources and incentive to achieve this work. C. Temperton is acknowledged for his coordination of the project.

The minimisation package used is provided by the Institut National de Recherche en Informatique et Automatique (INRIA, France). It is described in Gilbert and Lemaréchal, 1989.

7. REFERENCES

Andersson, E., 1996: Implementation of variational quality control. In Proceedings of the ECMWF Workshop on non-linear aspects of data assimilation, Shinfield Park, Reading, RG2 9AX, 9-11 September 1996.

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., 1998: The ECMWF implementation of three dimensional variational assimilation (3D–Var). Part III: Experimental results. *Q. J. R. Meteorol. Soc., 124, 1831-1860.*

Bouttier, F., Derber, J., and Fisher, M., 1997: The 1997 revision of the J_b term in 3D/4D-Var. ECMWF Research Department Technical Memorandum No. 238.

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

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. , 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., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M., 1998: The ECMWF implementation of three dimensional variational assimilation (3D–Var). Part I: Formulation. *Q. J. R. Meteorol. Soc.*, **124**, 1783-1807

Fisher, M., and Courtier, P., 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. ECMWF Research Department Technical Memorandum No. 220.

Fisher, M., 1998: Development of a simplified Kalman filter. ECMWF Research Department Technical Memorandum No. 260.

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

Järvinen, H, and Undén, P., 1997: Observation screening and first-guess quality control in the ECMWF 3D-Var data assimilation system. ECMWF Research Department Technical Memorandum No. 236.

Joly, A., Jorgensen, D., Shapiro, M., A., Thorpe, A., Bessemoulin, P., Browning, K., A., Cammas, J-P., Chalon, J-P., Clough, S.,A., Emanuel, K.,A., Eymard, L., Gall, R., Hildebrand, P., H., Langland, R.,H., Lemaître, Y., Lynch, P., Moore, J., A., Persson, P. , O., G., Snyder, C., Wakimoto, R., M., 1997: Definition of the Fronts and Atlantic Storm-Track Experiment (FASTEX). *Bull. Amer. Merorol. Soc.*, **78**, 1917-1940.

Klinker, E., F. Rabier and R. Gelaro, 1998: Estimation of key analysis errors using the adjoint technique. *Q. J. R. Meteorol. Soc.,* **124**, 1909-1934.

Klinker, E., F. Rabier, G. Kelly and J.-F. Mahfouf, 1999: The ECMWF operational implementation of four dimensional variational assimilation. Part III: Experimental results and diagnostics with operational configuration. *Q. J. R. Meteorol. Soc.* (submitted).

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.

Mahfouf, J.-F. and F. Rabier, 1999: The ECMWF operational implementation of four dimensional variational assimilation. Part II : Experimental results with improved physics. *Q.J.R. Meteorol. Soc.* (submitted).

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.

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., McNally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A. and Bouttier, F., 1998a: The ECMWF implementation of three dimensional variational assimilation (3D–Var). Part II: Structure funtions. *Q. J. R. Meteorol. Soc.* **124**, 1809-1829

Rabier, F., Thépaut, J-N., and Courtier, P. , 1998b: Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Q. J. R. Meteorol. Soc.*, **124**, 1861-1887

Simmons, A., and Rabier, F., 1997: Removal of the incremental initialization of medium and large scales in the ECMWF analysis. ECMWF Memorandum Research Department, May 1997.

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

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

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