TECHNICAL MEMORANDUM

273

The ECMWF operational implementation of four dimensional variational assimilation. Part III: Experimental results and diagnostics with operational configuration

E. Klinker F. Rabier, G. Kelly
and J-F. Mahfouf

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 III: Experimental results and diagnostics with operational configuration.

E. Klinker[1], F. Rabier[2], G. Kelly[1], J-F. Mahfouf[1]

[1] *European Centre for Medium-Range Weather Forecasts, UK*

[2]*Meteo France, Toulouse, France*

Abstract

The first two papers of this series describe the development of the operational 4D-Var configuration implemented at ECMWF. The basic features are a 6-hour incremental 4D-Var setup with two minimization steps, using very simplified physics in the first minimization and a more complete set of linear physics in the second one. This paper describes the validation of this configuration. Prior to implementation, 12 weeks of experimentation showed a consistent improvement relative to 3D-Var. After an additional 6 weeks of encouraging parallel operational suite, 4D-Var with physics was introduced in operations at ECMWF in November 1997. The difference in scores is statistically significant and the fast-growing components of the 4D-Var analysis errors are shown to be smaller than its 3D-Var counterpart. The performance of this new operational assimilation system is studied for the month of January 1998 for which the 4D-Var analyses exhibit more realistic baroclinic waves than 3D-Var, especially in the Pacific area. A case study illustrates the improvement one can expect in forecast terms in the mid-latitudes. The 4D-Var system improved the forecast skill in the tropics in general. Observing System Experiments show that the current 4D-Var operational system benefits from the assimilation of both satellite data and conventional observations.

1. Introduction

Atmospheric data assimilation is the process of determining a consistent four-dimensional atmospheric structure from information on the observational network, a dynamical model and other physical constraints. The time dimension can be accounted for in different ways. Four-dimensional variational assimilation (4D-Var) is the natural temporal generalisation of the three-dimensional variational analysis operational since January 1996 at ECMWF (Courtier et al., 1998, Rabier et al., 1998a, Andersson et al., 1998). It minimizes a cost-function measuring the distance between a model trajectory and the available information (observations, background field coming from a previous short-range forecast) over an assimilation interval or window. Under the assumptions of linear dynamics and of a perfect model, it gives the same result at the end of the assimilation interval as the full Kalman filter and at a smaller cost (but does not provide explicitly the analysis and forecast error stastistics).The potential of the 4D-Var algorithm was first shown in simple models in the mid-80s (Lewis and Derber, 1985, Le Dimet and Talagrand, 1986, Courtier and Talagrand, 1987, Talagrand and Courtier, 1987) before being applied to primitive equation models at a relatively low resolution (e.g. Thepaut and Courtier, 1991, Rabier and Courtier, 1992, Navon et al., 1992, Zupanski, 1993). The incremental formulation of 4D-Var (Courtier et al, 1994), which comprises running a high-resolution model with the full physical parameterization package to compare the atmospheric states with the observations to evaluate the cost-function and a low resolution model with simplified physics to minimize the cost-function, made its implementation feasible for operational models at high resolution. After extensive evaluation (Rabier et al., 1998b), it became operational on a 6-hour assimilation time-window at ECMWF on the 25th November 1997. The first paper (Rabier et al, 1999) presented the results of the baseline 4D-Var (one outer loop using simplified physics in the minimization) on a total of twelve weeks from different seasons. It illustrated in particular the structure functions used implicitly in 4D-Var and showed the impact of the dynamics on those. A comprehensive linearized physics and its adjoint, presented in the second paper (Mahfouf and Rabier, 1999) has been implemented in 4D-Var in an efficient configuration. Results of this first operational 4D-Var version with more complete physics (referred to as simply 4D-Var in the following) are presented in this paper.

ECMWF's main duty is to provide accurate medium range forecasts, for which it needs accurate analyses. The aim of this paper is to provide a broad view of the impact of analysis changes related to the introduction of 4D-Var. Besides standard diagnostics, our evaluation of analysis differences includes the estimation of errors in the initial conditions, the so called "key analysis errors", which can be obtained by additional adjoint integrations (Klinker et al, 1998). Model output parameters such as diabatic heating in the short range, which are less constrained by observations, can also help in the validation of the assimilation. However, forecast skill remains one of the main indicators for evaluating the quality of analyses.

The first operational configuration of 4D-Var has undergone several months of stringent tests including a parallel suite with data assimilation and medium range forecasts. Results of the pre-operational experimentation are discussed in section 2 followed by a discussion of further experiments after the operational implementation in section 3. Tropical performance is discussed in section 4. The first operational 4D-Var system has also been used to estimate the contributions of different observing systems to forecast skill as presented in section 5, concluding remarks are in section 6.

2. Pre-operational experimentation

2.1 Results from comparing 4D-Var versus 3D-Var over 80 days spread over several seasons

The comparison of 4D-Var with and without linear physics in the adjoint model (Mahfouf et al,1999) led to the decision to continue the 4D-Var experimentation with 2 outer loops (i.e. two minimizations and two updates of high resolution comparisons with observations), using simplified physics in 50 iterations of the first minimization and linearized physics in 20 iterations of the second minimization. Variational quality control of observations (Andersson and Järvinen,1999) is switched on after 40 iterations in the first minimization and throughout the second minimization. Additional periods were investigated and a total of 80 days were available to compare 3D-Var with this experimental 4D-Var, before the decision to start a parallel suite was taken. The list of periods is: 24 August to 6 September 1995, 15 January 1997 to 14 February 1997, 27 June 1997 to 10 July 1997, and 1 to 21 September 1997. As in Rabier et al's (1999) 12 week comparison between the baseline 4D-Var and 3D-Var, the comparison of 4D-Var with physics and 3D-Var shows a very consistent improvement in forecast performance (not shown). Scores up to day 7 with respect to observations in the Northern Hemisphere are presented in Figure 1. The comparison of forecasts to observations confirms the better scores for 4D-Var for all parameters (geopotential, temperature, wind). The extra-tropical scores are also generally better from 4D-Var with physics with respect to its own verifying analysis than those from 3D-Var in the early part of the forecast range (not shown), except for tropical wind at 200 hPa and tropical stratospheric temperature scores, which are slightly worse in the first day or so. In conclusion, 4D-Var with physics over a 6-hour assimilation window performs better than 3D-Var in a quite systematic way. The results motivated the decision to move to a final pre-operational test.

2.2 Pre-operational parallel suites

Prior to operational implementation of 4D-Var in November 1997, 4D-Var and 3D-Var were run in parallel suites for 6 weeks, from 9 October 1997 to 24 November 1997. As the analyses of the two systems can have significant differences, forecast verification statistics have been calculated for each run with respect to its own analysis. Figure 2 shows rms-errors averaged over 40 days (9 October 1997 to 17 November 1997). One clearly sees a positive impact of 4D-Var on forecast performance. Statistical tests have been applied to assign significance to the differences in rms-errors. In Figure 2, the 90% level confidence intervals calculated from their estimated standard-deviation at each range have been plotted on top of the mean of the score-differences. (For the t-test considered here the error bars are +/-1.645 times the standard-deviation). 4D-Var is found to be significantly better than 3D-Var at all ranges in both hemispheres (except for day 7 in the Southern hemisphere where the significance is slightly lower. Usually the significance in the short-range is higher than in the medium-range in all areas.

The better performance in terms of forecast scores is almost certainly coming from a better analysis. To judge the quality of the analyses, the "key analysis errors" were computed for both systems as 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 areas which are both dynamically unstable and not well covered by observations. However, "key analysis errors" represent only the fast growing part of the total forecast errors. Additional error evolution can be assigned to stable growth from errors in the initial condition and to model errors. Therefore we find normally total errors to be 3 times as large as evolved "key analysis errors" at a forecast range of one day.

The root-mean-square of the "key analysis errors" over the parallel suite period for the geopotential at 500 hPa and latitudes poleward of 30N are smaller for 4D-Var (1.6m) than for 3D-Var (1.7m). The time evolution of these "key analysis errors" can be computed by differencing a model integration starting from a "corrected" analysis and a control model integration starting from an uncorrected analysis. The rms values of the "key analysis errors" evolved for 24 hours have been calculated for the geopotential at 500 hPa over the Northern Hemisphere. They show a reduction by 0.3 m from 5.0m for 3D-Var to 4.7m for 4D-Var. The geographical distribution of the difference between the two "key analysis errors" are shown in Figure 3. Shaded areas indicate negative values, where 4D-Var has smaller evolved "key analysis errors" than 3D-Var. They are widespread, with a large extent, in particular, from the Eastern Pacific to Europe. In order to estimate to what extent differences in "key analysis errors" describe different forecast error evolution in 4D-Var and 3D-Var, the difference in "key analysis errors" is compared to the difference in the full 24-hour forecast errors, shown in Figure 4 (the Northern Hemisphere rms value of the 24-hour forecast error for 3D-Var is 14.0m and 13.4m for 4D-Var). There is quite a good correspondence between the two figures 3 and 4. For both measures, "key analysis errors" and 24-hour forecast errors, the rms values for 4D-Var forecasts are considerably smaller than for 3D-Var, and the maximum impact of the different assimilation system is seen in similar areas extending from the Eastern Pacific to Europe. In summary, in the short-range, the 4D-Var forecasts have smaller errors than 3D-Var over widespread areas; the Southern Hemisphere is almost everywhere favourable to 4D-Var (not shown). In the Northern Hemisphere, these smaller errors are consistent with smaller "key analysis errors" of the 4D-Var system. We can therefore link reduced short range forecast errors directly to improved initial conditions using the 4D-Var system in the analysis. A larger error reduction in absolute terms for the 24-hour forecasts (0.6 m) compared to the reduction in evolved "key analysis errors" (0.3 m), however, suggests that other aspects of the analysis have been improved as well.

In the data assimilation system the standard evaluation of the fit of the six hour forecast (or the so called background) to observations represents a convenient tool to estimate the forecast performance in the very short range (6 hours). Figure 5 shows for the two systems the averaged fit to the radiosonde data over the first 4 weeks of the parallel suites (971009 to 971105) in the Northern and Southern Hemispheres for the geopotential height. As explained in the first paper of this series Rabier et al (1999), 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 un-correlated (the root-mean-square plots then represent the square-root of the sum of both mean-square errors). For these background fits, 4D-Var clearly outperforms 3D-Var for both hemispheres at all levels. For other parameters, such as wind, better results are also found, similar to results presented in Rabier et al (1999). The better fit was found to be statistically significant at the 95% confidence level (around 20000 individual data entered the computation at every level in the Northern Hemisphere and 2500 in the Southern Hemisphere).

4D-Var analysis increments are generally smaller than 3D-Var ones which is consistent with an improved short-range forecast. The norms of the increments have been computed over the Northern Hemisphere for latitudes north of 30N at 12Z on a daily basis and plotted as a time series in figure 6. The norm is the energy norm, as used in the sensitivity computations by Rabier et al (1996), Klinker et al (1998), and in the singular vector computations by Buizza et al., 1998. It takes into account all components of the increments and all variables (mass and wind). The 3D-Var norms of the increments (dashed line) are systematically larger than the 4D-Var norms (solid line), by up to 10%. The increments can also be projected onto the first singular vectors which span part of the unstable sub-space. Each day, the 3D-Var and 4D-Var increments were projected over the first 30 singular vectors and the norms of the increments in this unstable sub-space were computed. Finally, the ratio of the norms of the increments in the unstable sub-space over the total norm was computed. The mean of this ratio over the 40 days is equal to 0.09546 for 3D-Var and 0.09638 for 4D-Var. Both systems have roughly 10% of the increments projecting on the unstable sub-space described by the first 30 singular vectors. Although the difference is not large, it is worth noting that 4D-Var has a marginally larger proportion of its increment on the unstable sub-space, despite having 10% smaller total increments

3. Performance in January 1998

3.1 Global diagnostics

In terms of scores, the ECMWF system performed particularly well over the first winter after operational implementation. In order to document whether 4D-Var was indeed one of the contributing factors of this good performance, 3D-Var was also run in parallel. The comparison of scores (not shown) shows a similar improvement of 4D-Var upon 3D-Var as in the parallel experimentation for the previous season as described above.

For January 1998 the differences between 4D-Var and 3D-var rms forecast errors is shown on maps for various ranges (day-1 in figure 7, day-3 in figure 8 and day-5 in figure 9). For negative values (dark shading) 4D-Var has smaller errors than 3D-Var. The smallest differences between the two systems are found over the Asian continent. All other areas in the Northern Hemisphere show predominantly smaller errors for 4D-Var in the short-range and in the medium-range. In the short-range (day-1 in Figure 7), the impact is mostly seen in the data-sparse oceanic areas. This is not surprising, as previous experience with the 4D-Var system shows that the 4D-Var and 3D-Var analyses are most different over the mid-latitude data-sparse areas. There, 4D-Var can exploit the dynamical processes to interpolate more appropriately the information coming from the rather sparse data. The fact that the 4D-Var assimilation system is able to create more baroclinic structures can be measured by computing the northward eddy heat-flux averaged over January 1998, which gives an indication of vertical tilts of baroclinic systems. Figures 10 to 12 show respectively the difference in eddy heat-flux at 500 hPa between the 4D-Var and the 3D-Var analyses, between the 4D-Var and the 3D-Var first-guesses, and the eddy heat-flux for the 4D-Var analysis. From Figure 12, the eddy heat-flux, averaged over all 4D-Var analyses for the month of January 1998, is mainly positive, consistent with a normal westward tilt of geopotential with height. Largest eddy heat-flux values are found over the oceanic storm-tracks. Figure 10, illustrating the difference in eddy heat-flux between 3D-Var and 4D-Var exhibits mainly negative values which are generally of the order of 10% of the analysis fluxes but can locally be close to 100% as in the central Pacific.This means that the vertical structures of the 3D-Var analyses are generally less baroclinic than the 4D-Var ones. This is also true, to a lesser extent for the first-guesses (6-hour forecasts) as shown in Figure 11. It is clear from these results that 4D-Var produces a more baroclinic analysis over the dynamically unstable mid-latitude oceanic areas by producing more baroclinic increments.

Dynamically influenced structure functions have already been illustrated in a number of papers on 4D-Var (e.g. Thepaut et al, 1996, Rabier et al, 1998b). Examples of tilted geopotential increments caused by real observations have also been shown in Rabier et al, 1998a. Figure 13 illustrates geopotential increments caused solely by the presence of surface observations. The geographical locations are chosen in the Southern ocean for an experiment using no satellite information such that there were no upper-air observations in the vicinity. (Note that in those areas the standard-deviations of background error is homogeneous enough not to contribute in creating such structures). The cross-sections are chosen along latitude rows for simplicity. The tilt in the increment can reach 15 to 20 degrees in longitude over the depth of the troposphere. These increments illustrate the ability of 4D-Var to create flow-dependent increments.

3.2 A synoptic example of differences between forecasts from 3D-Var and 4D-Var

Given that 4D-Var forecasts show better verification statistics than 3D-Var forecasts one may ask if the forecast differences are significant in synoptic terms. To answer this question we have selected a case from the January 1998 period discussed above. For this period 4D-Var performed better than 3D-Var at all ranges particularly over the Atlantic and European areas. The case discussed here is not the best, but is one of the noticeably better 4D-Var forecasts. On 15 January 1998, a low pressure area was located on the South-East coast of the United-States. From there, several lows propagated down stream the Atlantic Ocean. “Key analysis errors” on that date show quite a large sensitivity of forecast errors to the atmospheric flow around 30 to 40N, just off the American coast (not shown). 3D-Var and 4D-Var analyses are indeed quite different in this area, although it is impossible to know which is better by looking at the fit to the available data. It is always difficult to establish the precise location and the reason for a better analysis, when testing a new data assimilation system. In a long parallel run, improvements will arise initially from a better analysis system resulting in better short range forecasts that serve as a background field for later analyses. Strictly speaking, the effect of the analysis system can only be identified precisely in single cycle experiments run from identical background fields. One can, however, assess the quality of the initial conditions from the evolution of the forecasts run with identical models from the different analyses.

Figures 14 and 15 show the D+2 forecasts and D+5 forecasts (respectively) from 3D-Var and 4D-Var on 15 January 1998, together with the 4D-Var verifying analyses. As early as in the two-day range, one can see that the 4D-Var forecast is of superior quality than the 3D-Var forecast: the low pressure area in the Western Atlantic is too deep in the 3D-Var forecast and not elongated enough. At day 5, the system has evolved and moved to the mid-Atlantic Ocean. The intense low (960hPa) is slightly underestimated by the 4D-Var forecast (at 975 hPa), but is not present at all in the 3D-Var forecast. 4D-Var produces clearly superior synoptic short range and medium range forecasts from this date.

4. Tropical performance

The 4D-Var system implemented operationally in November 1997 improved the forecast in the tropics as well as in extra-tropical latitudes. The results for the parallel suite (period as discussed in section 2) are characterized by an improved fit of the background to observations from using a 4D-Var analysis. Figure 16 shows the rms-fits to observations for the meridional wind component and the height field. In particular the background for height field is significantly improved at all levels, largely a result of a better fit of surface pressure and mid-level thicknesses (not shown). For the wind field (v-component here shown) the fit of the background to observations is better at low levels and upper levels of the troposphere.

Just as in the mid-latitudes of both hemispheres the 4D-Var system also leads to improved tropical forecasts at almost all forecast ranges as shown by root mean square error of the vector wind at 850 hPa (Figure 17). Marginally larger errors after 12 hours have to seen in the light of slightly different forecast length. The initial conditions in the 4D-var system are in fact a 3 hour forecast, during which model errors, which represent a comparatively large part of the total errors in the tropics compared to mid-latitudes, can already evolve to noticeable amplitudes. The performance of the tropical analysis has been tested for the TOGA COARE period as well, when the coverage of atmospheric soundings was exceptionally good. The analysis and forecast statistics for the West-Pacific were again in favour of 4D-Var.

5. Observing-system experiments

The operational observing network, which uses both conventional and satellite measurements, influences how accurately the initial atmospheric state can be prescribed and therefore to a large extent the resulting forecast accuracy. In order to evaluate the contribution made by the main ground-based and satellite-based observing systems to forecast quality a series of Observing System Experiments (OSE's) was performed. Similar experiments based on the 3D-VAR system operational in early 1997 are described in a previous study (Kelly et al., 1997). The 3D-Var system has been shown to benefit from the assimilation of both satellite data and conventional observations. The inclusion of each data type almost always improved the forecast which was not invariably the case in previous studies which were based on an analysis system using optimum interpolation (Kelly et al., 1993). The assimilation of the Tiros Operational Vertical Sounder (TOVS) radiances directly in the 3D-Var analysis contributed largely to the improved benefit from satellite data.

A series of OSE experiments have been run using the operational 4D-Var system for 34 days in May and December 1997. For each experiment, the following observing systems have been removed from the full operational system:

(a) Satellite clear radiance data from the TOVS satellites (NOTOVS)

(b) Geostationary Atmospheric Motion Winds (AMW's) from cloud and water vapour (NOAMW)

(c) Radiosonde wind, temperature and humidity data (NORAOB)

(d) Aircraft winds and temperatures (NOAIREP)

(e) The combined removal of both a. and b. (NOSAT).

All these experiments have been compared to the full operational system (CONTROL). All verification statistics use the operational ECMWF analysis which had used all data, as it is considered to be the best for verification purposes. The influence of the data on forecast performance will be illustrated for the wind at 200 hPa where the impact is among the largest.

The results are grouped into two sections, first the experiments on the satellite measurements and second the experiments on the conventional upper-air measurements.

5.1 Experiments withdrawing satellite data

To evaluate the impact of satellite data in the 4D-Var framework we compared forecasts from the following four configurations:

(a) CONTROL which assimilates all data

(b) NOTOVS which excludes radiance data from the TOVS satellites

(c) NOAMW which excludes geostationary atmospheric motion wind data

(d) NOSAT which excludes both radiance data and geostationary atmospheric motion wind data

Comparing errors of the CONTROL forecasts (solid line) and the NOSAT forecasts (dashed line) in Figure 18 shows that the overall impact of satellite data around the 3-day forecast range varies from 1/3 of a day in the Northern Hemisphere, to 1 day in both the Tropics and the Southern Hemisphere. Both TOVS radiances and AMWs Geostationary wind data have a positive impact in all areas. However, the combination of the two data sources produces the best forecasts. AMWs have most value in the Tropics. TOVS have a significant impact in the Tropics and large impact in the Southern Hemisphere. With the exception of AMWs in the Northern Hemisphere, all combinations of observational data show a positive influence on forecast performance throughout the forecast range up to 5 days.

5.2 Experiments withdrawing conventional upper air data

For the evaluation of the impact of conventional upper air measurements we have made assimilations for two more configurations excluding radiosondes or aircraft measurements and we compare forecast performance of the following experiments

(a) CONTROL which uses all data

(b) NORAOB which excludes both radiosonde winds, temperatures and humidity data

(c) NOAIREP which excludes both aircraft wind and temperature observations.

For the purposes of comparison, the NOSAT experiment will also be considered in this group in order to assess the relative importance of the AIREP and SAT wind observing systems.

The impact on the 200 hPa winds is shown in Fig. 20. In the Northern Hemisphere the radiosondes have the largest effect, as their exclusion (NORAOB) reduces the forecast accuracy by 1/2 of a day over most forecast ranges. The satellite data and AIREPs have a lesser effect and both have about equal weight, each degrading the forecast by 1/3 of a day in the short range but the impact of the AIREP's decreases with forecast time.

In the Tropics and the Southern Hemisphere, satellite data has the largest impact of about 1 day at a forecast range of 3 days. Removing radiosondes has more impact in the Tropics (about 1/2 a day) than in the Southern Hemisphere (1/3 a day) in the short range. AIREPs have negligible effect in the Southern Hemisphere and a small impact in the Tropics of about 1/3 of a day. Both radiosonde and satellite data bring an improvement in forecast skill that is almost constant as a function of forecast range.

The results obtained in this first set of 4D-Var observing system experiments are encouraging. The current operational 4D-VAR system benefits from the assimilation of both satellite data and conventional observations. Its performance in each of the Northern Hemisphere, the Tropics and the Southern Hemisphere is broadly satisfactory. The results are very similar to those obtained for the 3D-Var system (Kelly et al., 1997). A difference might be that in the Tropics and Southern Hemisphere the 4D-Var forecasts keep the benefit of AMWs and radiosonde data further into the forecast range than 3D-Var. The two week periods of data assimilation that form the basis of the OSEs are most certainly too short to estimate any long term drift in the analysis quality arising from excluded observations. Therefore the impact of different observing systems is likely to represent an underestimation.

6. Summary and discussion

Rabier et al (1999) presented a baseline 4D-Var system, using very simplified physics in the minimization. Twelve weeks of experimentation resulted in better performance compared with 3D-Var, on average and in particular during rapid cyclogenesis. The dynamical processes seem to be relevant, even on a short 6-hour assimilation period. Mahfouf and Rabier (1999) described the development of a first comprehensive set of linear physics for 4D-Var applications. The inclusion of this physics package in a 4D-Var "2-outer-loop" configuration has a positive impact on the performance of the analysis in the tropics, with a reduction of the spin-down of precipitation in the subsequent forecasts, and improved wind scores. The extra-tropical scores averaged over 8 weeks show a slight improvement brought by the physics. In view of these results, it was decided to continue the experimentation with 4D-Var with physics and two outer loops.

In this article, the forecasts starting from 4D-Var with physics as implemented in operations on the 25th November 1997 have been discussed. They are shown to be significantly better than the forecasts starting from 3D-Var in both hemispheres, in particular in the short-range.

In general, the background fields in 4D-Var are closer to observations than in 3D-Var and are therefore more accurate since the background errors and random observational errors are un-correlated. Because 4D-Var background fields are more accurate, 4D-Var increments are smaller than 3D-Var increments. Significant differences were found on the vertical structure of increments in mid-latitude oceanic areas. More baroclinic increments in 4D-Var result in a more baroclinic analysis, in particular over the Pacific ocean. Despite the more intense baroclinic waves in the analysis, "key analysis errors" which arise in dynamically unstable areas are actually smaller in the 4D-Var system. This is an indication that the improvement brought by the 4D-Var system in these key areas must have been particularly large. 4D-Var has been able to produce more realistically tilted structures, in agreement with observations. Consequently, short-range forecasts from 4D-Var are better than from 3D-Var, in particular over the oceans. This positive signal then spreads to other areas in the medium-range.

To facilitate validation, the first operational version of 4D-Var which described these papers has been tightly constrained to be as close as practicable to 3D-Var. The two systems use the same 6-hour assimilation window, the same amount of data and the same quality control of observations. There are also no differences in the number of iterations and the spectral resolution used during the minimization. There is considerable scope for further development of 4D-Var. Important issues are the use of more a-synoptic observations, a higher resolution in the inner loop and a longer assimilation time window. Further future work on 4D-Var will concentrate on developing the physics and improving the convergence of the incremental approach. Short-term issues include studying the sensitivity to initialization parameters (both in the initialization of the background at low resolution and of the increments), re-tuning the background and observation errors in the light of the experience so far. The use of more a-synoptic data already has been studied with encouraging results (Järvinen et al, 1999).

In the long term the 4D-Var system will be extended to handle 12 or 24 hour assimilation windows. Optimal use of longer time windows in the 4D-Var assimilation will be made possible by a proper cycling of background error covariances provided by a simplified Kalman filter, which is currently being developed (Fisher, 1998).

For the main analysis job, the operational configuration of 4D-Var costs three times as much as 3D-Var. Comparing the costs of a whole day of assimilation plus one 10-day forecast, the 4D-Var costs 65% more than 3D-Var. Viewed in the context of the overall cost of the operational suite, including the Ensemble Prediction System for instance, this increased cost is an acceptable fraction, and is in line with expectations.

With further optimization of the analysis and model code it has been possible to use T106 instead of T63 in the inner loop, which defines the maximum resolution of the increments and T_L319 (linear grid for spectral resolution 319) instead of T213 in the outer loop. The first results show an encouraging improvement of the quality of the analysis in particular in the tropics as measured from the analysis and background departures from observations.

The 4D-Var system makes good use of different types of observations as demonstrated in a series of observing system experiments. On the basis of these results one can expect further improvement of analyses by assimilating additional future data sources. In particular 4D-Var is a natural framework to exploit fully observations which are not directly linked to large scale atmospheric fields, but represent the result of physical processes such as precipitation (Kuo et al, 1995).

7. Acknowledgements

The authors would like to thank all colleagues at ECMWF who contributed to the 4D-Var implementation. O.Pesonen carried out the technical implementation of 4D-Var with great enthusiasm and efficiency. The staff from the Meteorological Operations and Applications Sections are acknowledged.

Mike Fisher provided the significance test to evaluate the scores. Thanks to A. Hollingsworth for his useful comments which helped to improve the original manuscript. The minimisation package (Gilbert and Lemaréchal, 1989) has been provided by the Institut National de Recherche en Informatique et Automatique (INRIA, France).Some plots have been generated using the UPDATE package of R. Krishna (Meteorological Office, Bracknell)

References

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.*

Andersson, E. and Järvinen, H, 1999: Variational quality control. Accepted for publication in Q. J. R. Meteorol. Soc.

Buizza, R., Petroliagis, T., Palmer, T., N., Barkmeijer, J., Hamrud, M., Hollingsworth, A., Simmons, A., and Wedi, N., 1998: Impact of model resolution and ensemble size on the performance of an ensemble prediction system. Submitted to Q. J. R. Meteorol. Soc.

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., 1998: Development of a simplified Kalman filter. ECMWF Technical Memorandum 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.

Järvinen, H, Andersson, E and Bouttier, F.: 1999:Variational assimilation of time sequences of surface observations with serially correlated errors. ECMWF Research Department Technical Memorandum

Kelly, G.A., 1997: Influence of Observations on the Operational ECMWF System. WMO. Bulletin Vol 46 No. 4 Oct. 1997. pp 336-342

Klinker, E., Rabier, F. and Gelaro, R., 1998: Estimation of key analysis errors using the adjoint technique. Q. J. R. Meteorol. Soc., 124,1909-1933.

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.

Lin X. and Johnson, R.H., 1996: Heating, moistening and rainfall over the western Pacific warm pool during TOGA COARE. J. Atmos. Sci., 53, 3368-3383

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., 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., 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 functions. *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.

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 multi-level 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 assimilation: 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.

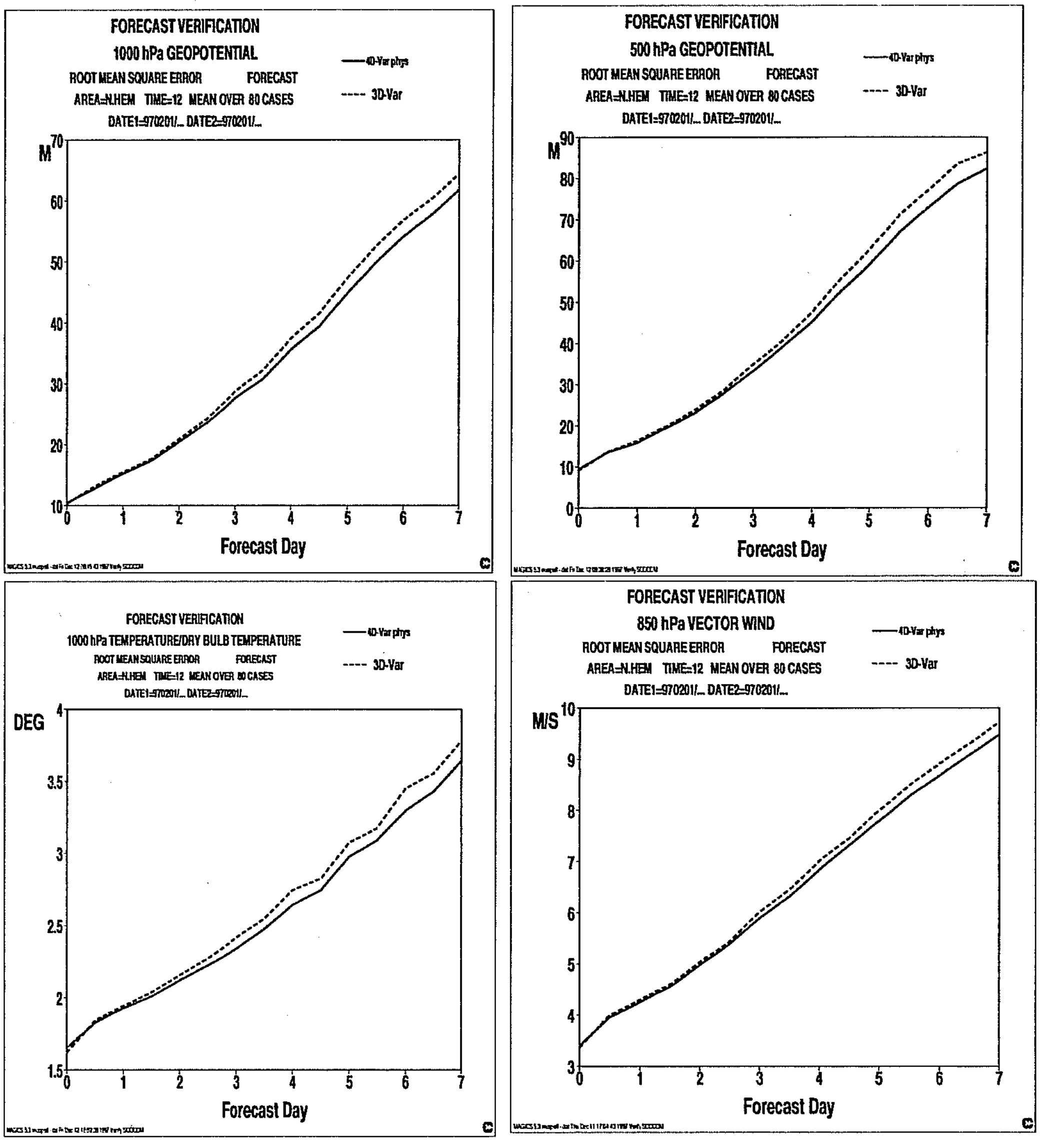


Figure 1 : Root mean square of the difference between forecasts from 4D-Var and observations (solid lines) and forecasts from 3D-Var and observations (dashed lines), averaged over 80 cases in the Northern Hemisphere, for various parameters (geopotential height at 1000 hPa and 500 hPa, temperature at 1000 hPa, vector wind at 850 hPa).

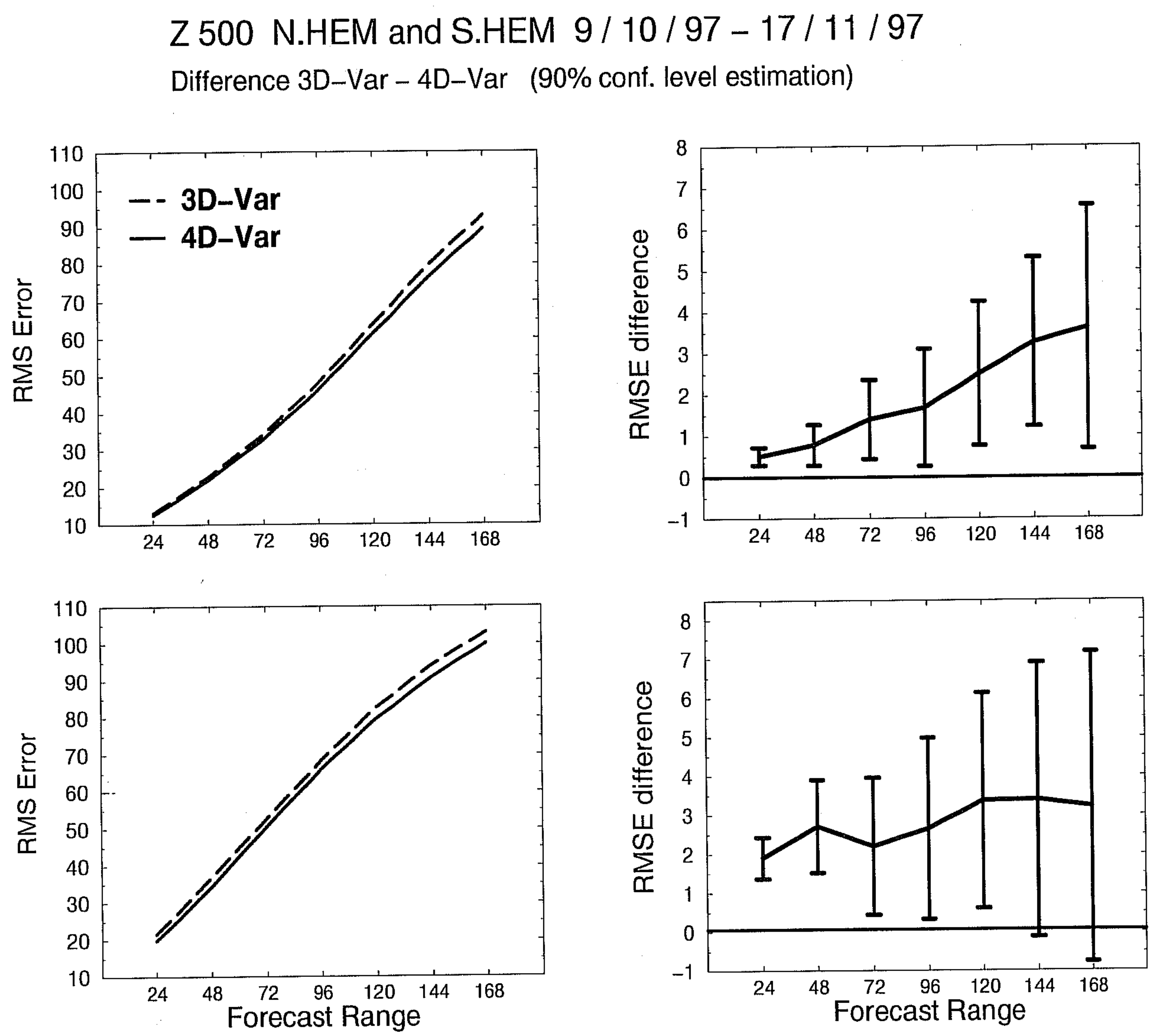


Figure 2: Root-mean-square of forecast errors in the Northern Hemisphere (top-left panel) and the Southern Hemisphere (bottom-left panel) for 40 days of the parallel operational suites. 4D-Var is represented by a solid line, and 3D-Var by a dashed line. Difference in root-mean-square of forecast errors in the Northern Hemisphere (top-right panel) and the Southern Hemisphere (bottom-right panel) with error bars at 90% confidence level.

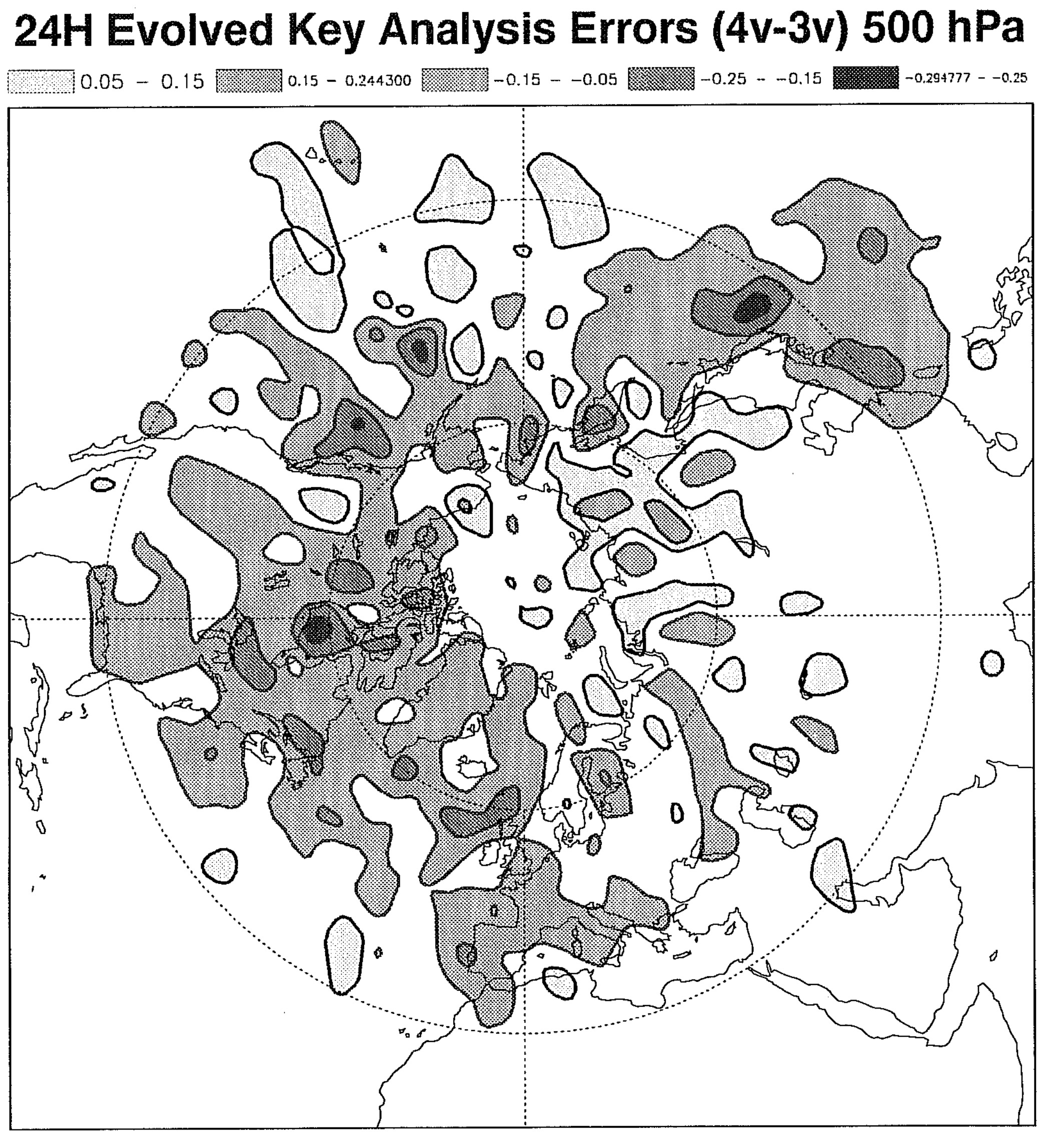


Figure 3: Difference between the 4D-Var and 3D-Var root-mean-square of the “key analysis errors” evolved for 24 hours for the geopotential height at 500hPa. The averaging period is the parallel suite period (971009 to 9711121). Dark shaded areas indicate negative values, where 4D-Var has smaller evolved “key analysis errors” than 3D-Var.

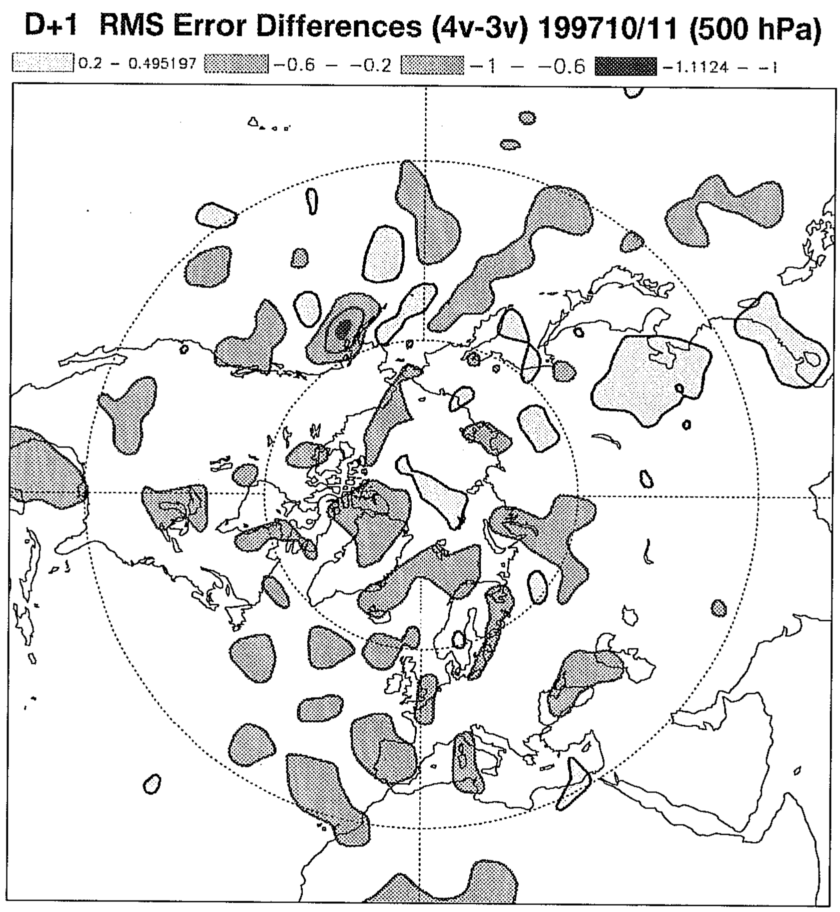
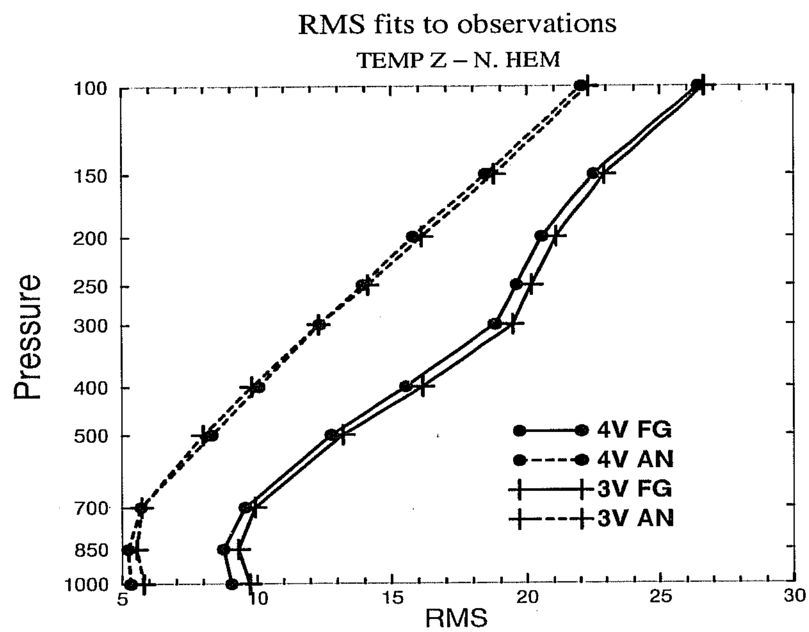


Figure 4: Differences between 4D-Var RMS errors and 3D-Var RMS errors for day-1 forecasts at 500 hPa. Dark shaded areas indicate negative values, where 4D-Var has smaller RMS errors than 3D-Var. Units: dam.The averaging



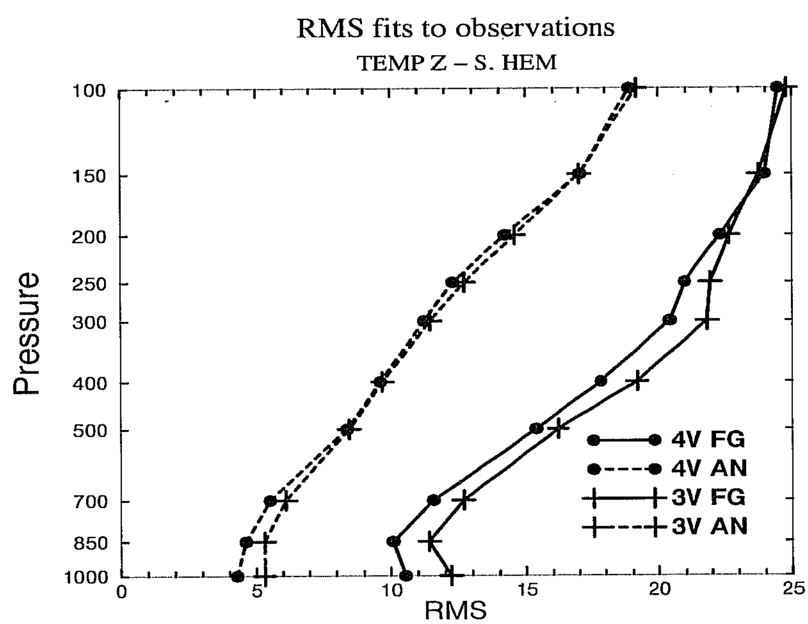


Figure 5: Root-mean-square (RMS) fits of geopotential height to radiosonde measurements produced over the Northern Hemisphere (North of 20N) and the Southern Hemisphere (South of 20S), averaged over 4 weeks of the parallel suites (October-November 1997). The solid lines represent the RMS fits of the backgrounds to the observations, the dashed lines the RMS fits of the analyses to the observations. 4D-Var is shown as circles, 3D-Var as plus signs. The abscissa is the RMS in geopotential. The ordinate is the pressure in hPa.

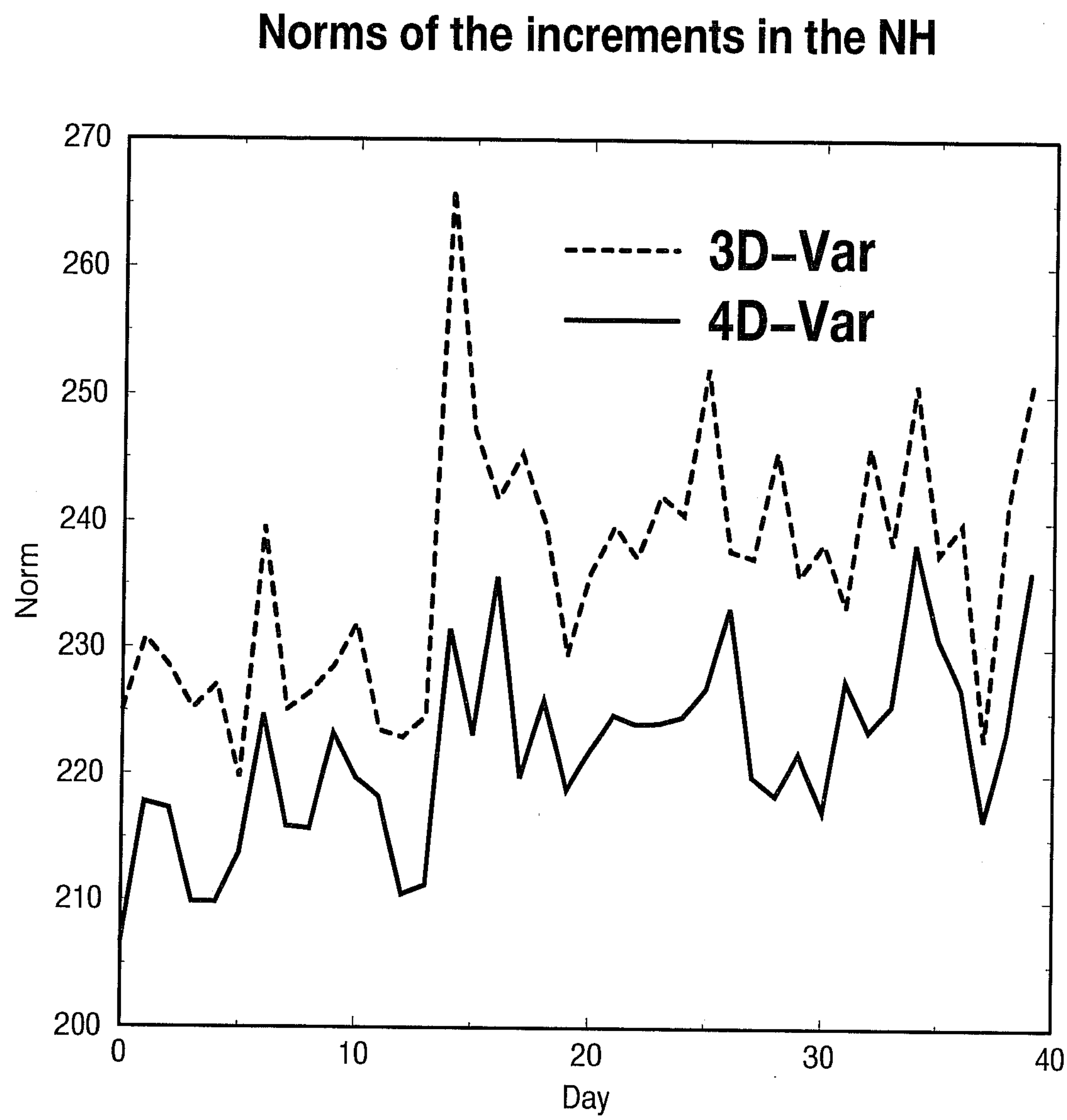


Figure 6: Norms of the increments produced by the 3D-Var analysis (dashed line) and 4D-Var analysis (solid line) at 12Z for 40 consecutive days during the parallel suite. The norm is the energy norm and the area over which it is computed is the Northern Hemisphere, for latitudes north of 30N.

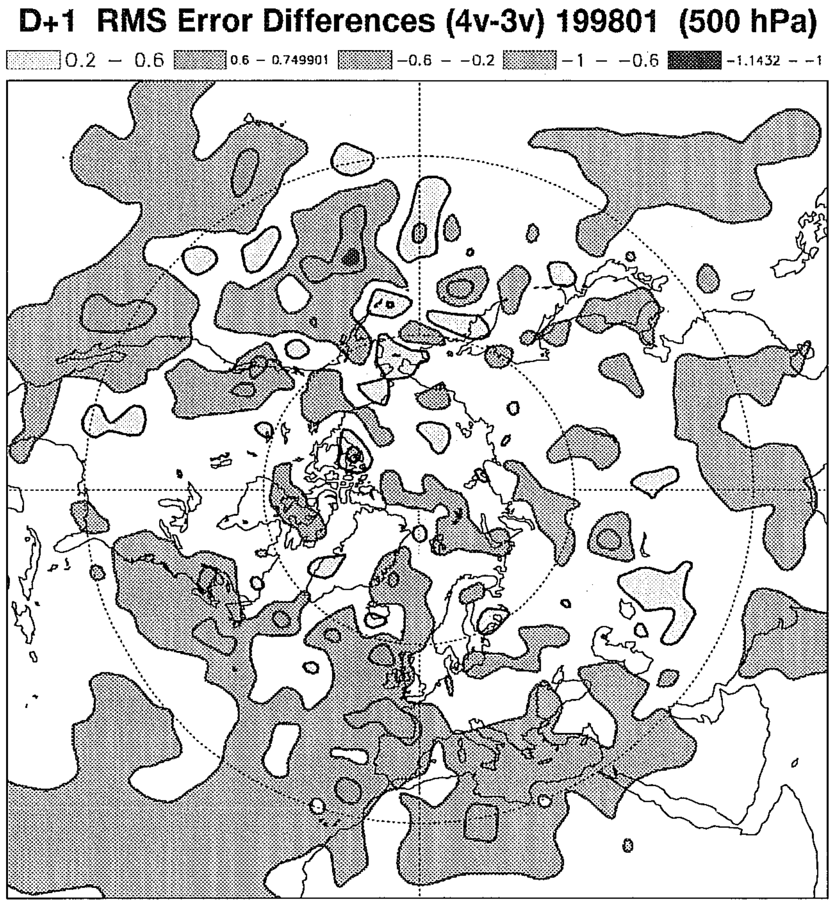


Figure 7: Differences between 4D-Var RMS errors and 3D-Var RMS errors for day-1 forecasts at 500 hPa, averaged over January 1998. Dark shaded areas indicate negative values, where 4D-Var has smaller RMS errors than 3D-Var. Units: dam.

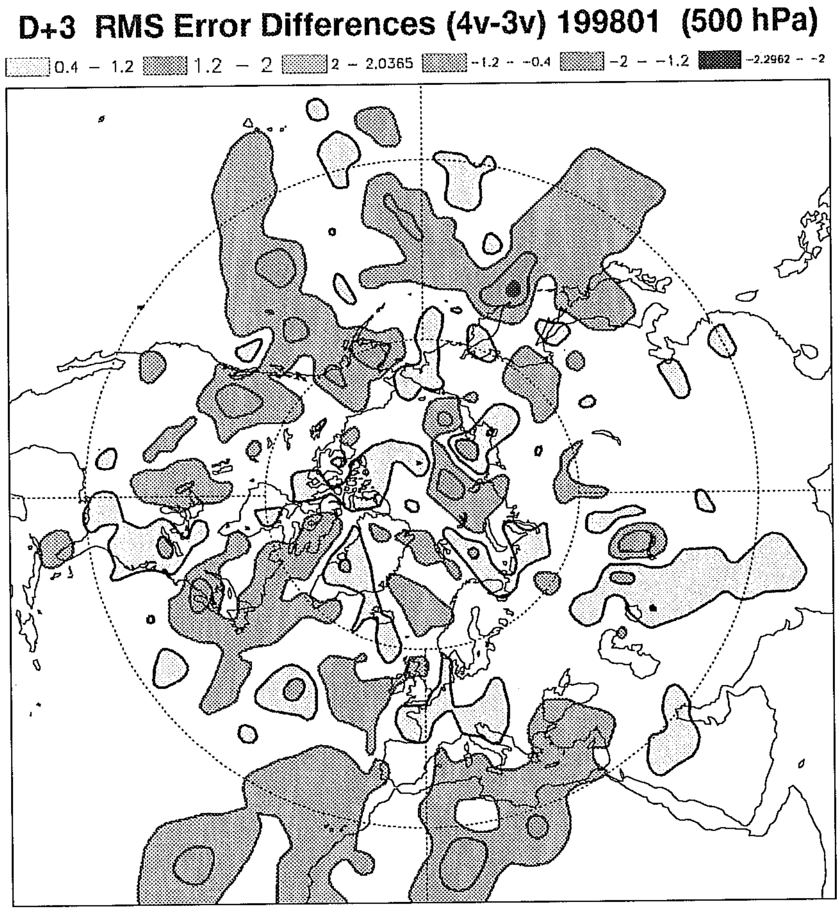


Figure 8: Differences between 4D-Var RMS errors and 3D-Var RMS errors for day-3 forecasts at 500 hPa, averaged over January 1998. Dark shaded areas indicate negative values, where 4D-Var has smaller RMS errors than 3D-Var. Units: dam.

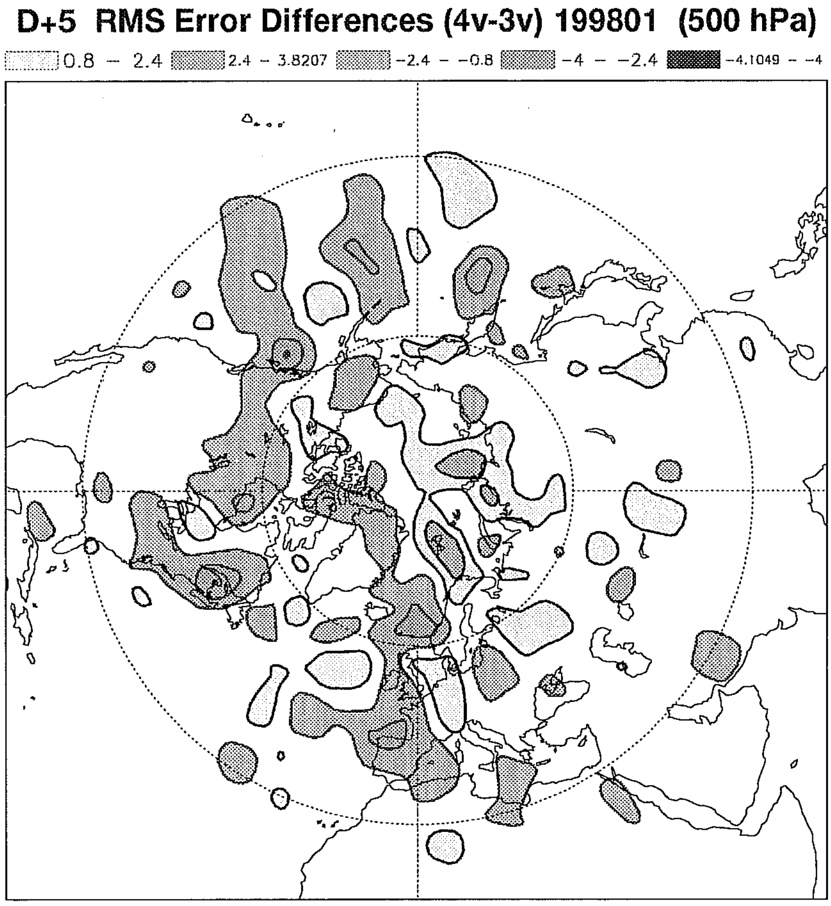


Figure 9: Differences between 4D-Var RMS errors and 3D-Var RMS errors for day-5 forecasts at 500 hPa, averaged over January 1998. Dark shaded areas indicate negative values, where 4D-Var has smaller RMS errors than 3D-Var. Units: dam.

Eddy Heat Flux [v*t*] 3v-4v 9801 Analysis 500 hPa

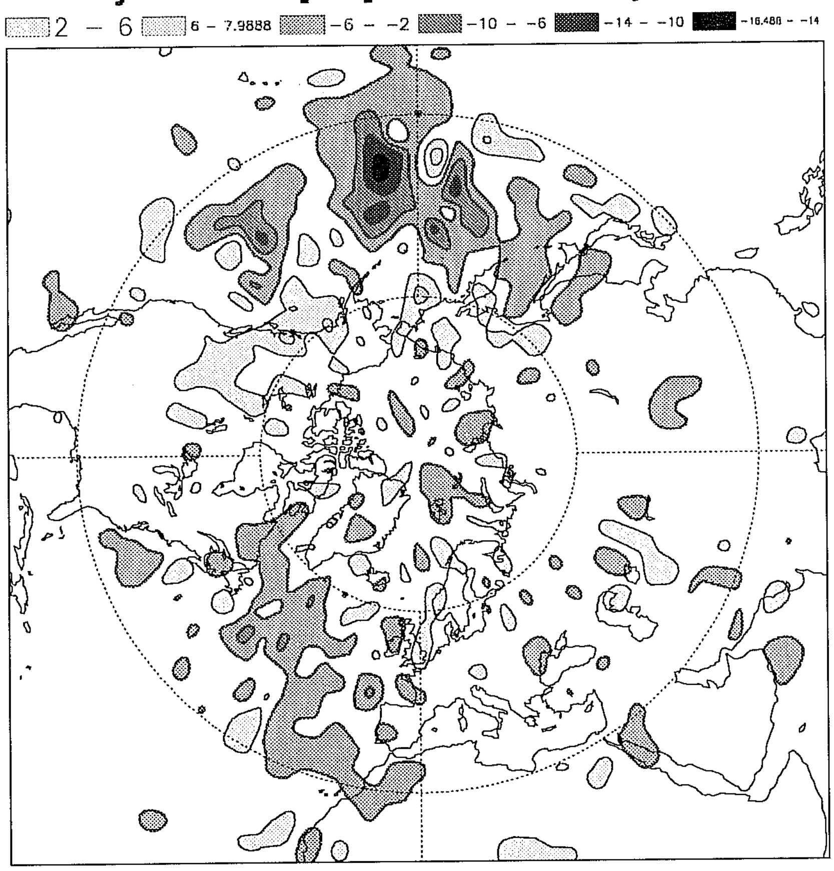


Figure 10: Difference between 3D-Var and 4D-Var eddy heat flux calculated from the analysis of January 1998 at 500 hPa. Dark shaded areas indicate negative values, where 3D-Var has a smaller northward eddy heat flux than 4D-Var. Units: m/s K

Figure 12: Mean eddy heat-flux for the 4D-Var analysis for January 1998 at 500 hPa. Light shaded areas indicate positive values, where eddy heat flux is poleward. Units: m/s K

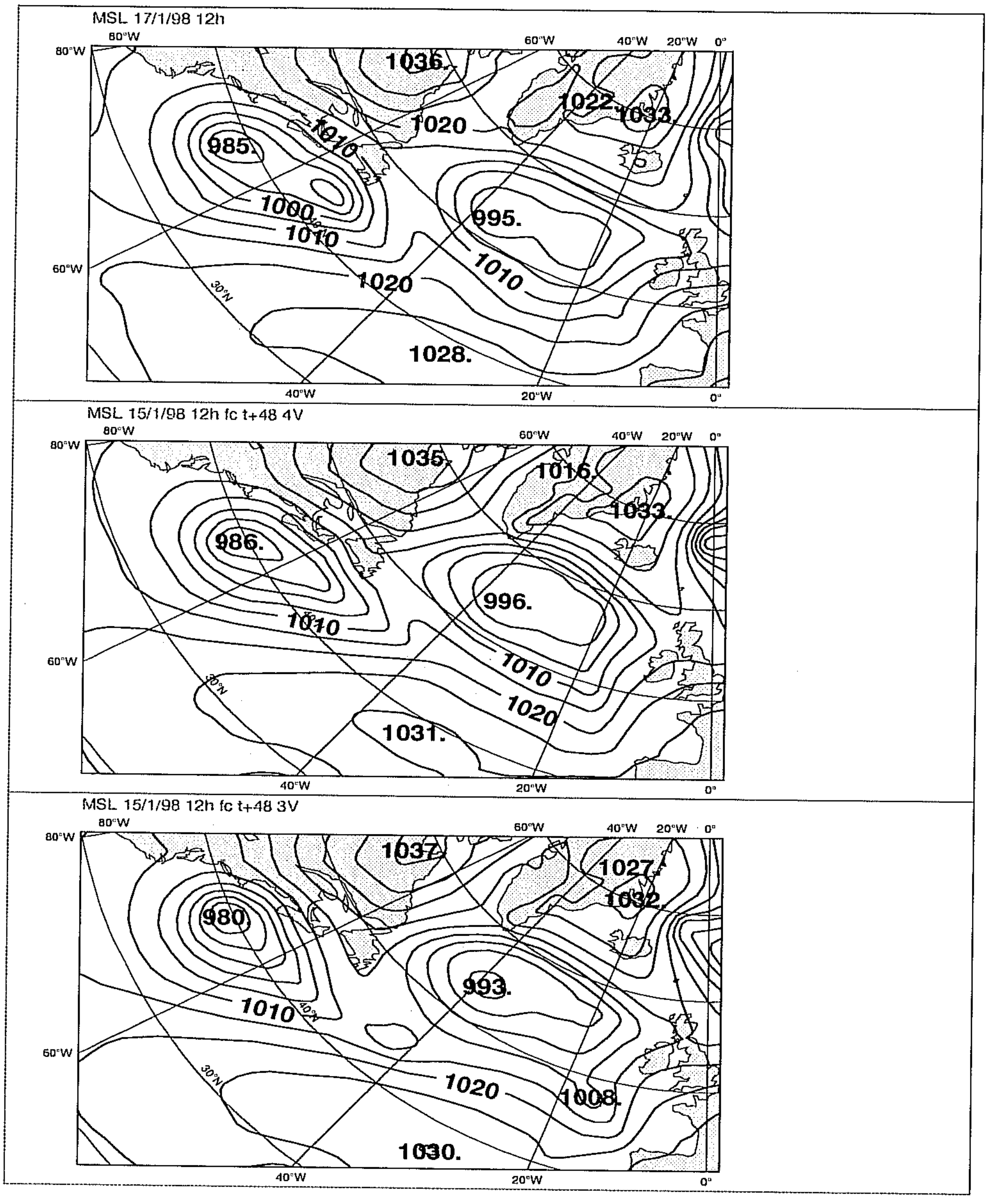


Figure 14: Mean sea-level pressure fields valid for 17/01/98 at 12Z. Top panel: verifying 4D-Var analysis. Middle panel: 4D-Var two-day forecast from 15/1/98. Bottom panel: 3D-Var two-day forecast from 15/1/98. Isolines are every 5hPa.

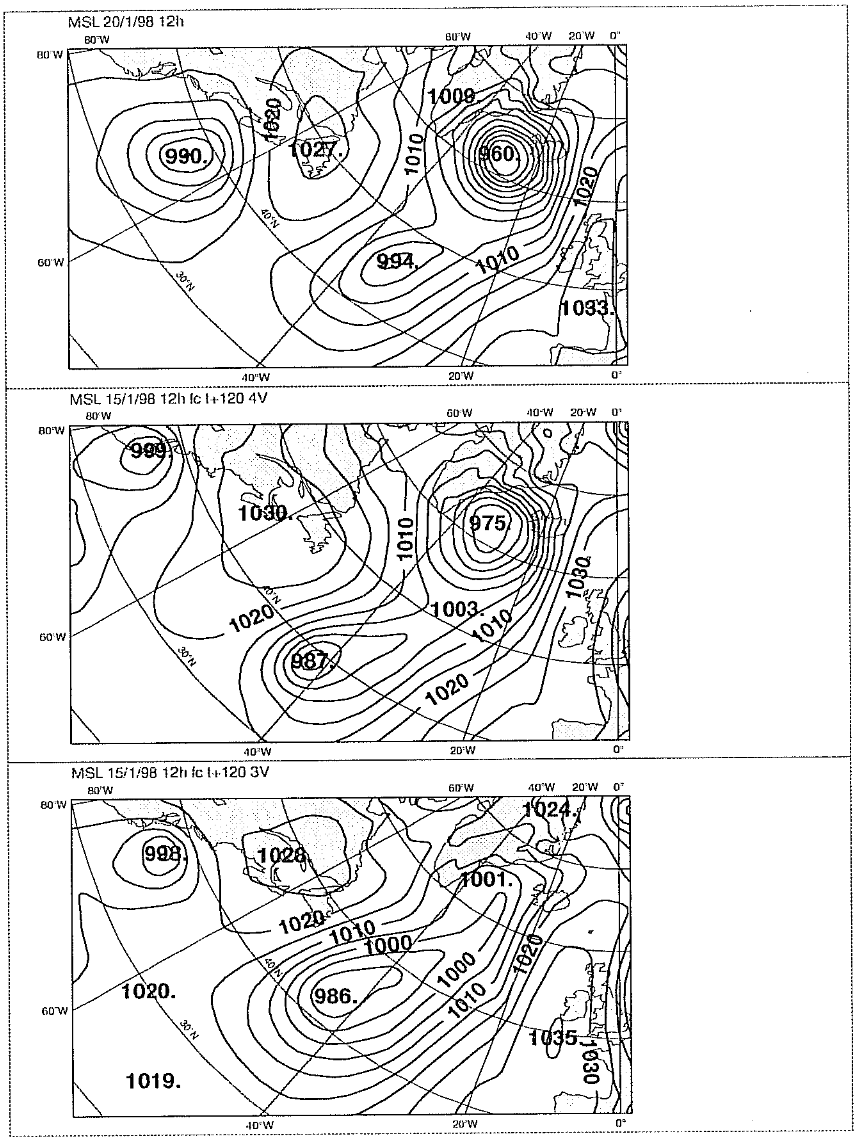
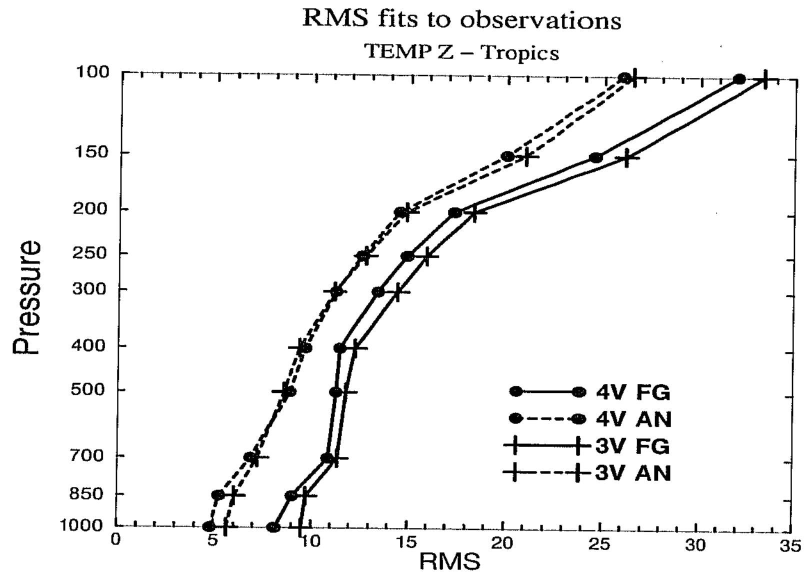


Figure 15: Mean sea-level pressure fields valid for 20/01/98 at 12Z. Top panel: verifying 4D-Var analysis. Middle panel: 4D-Var five-day forecast from 15/1/98. Bottom panel: 3D-Var five-day forecast from 15/1/98. Isolines are every 5hPa.



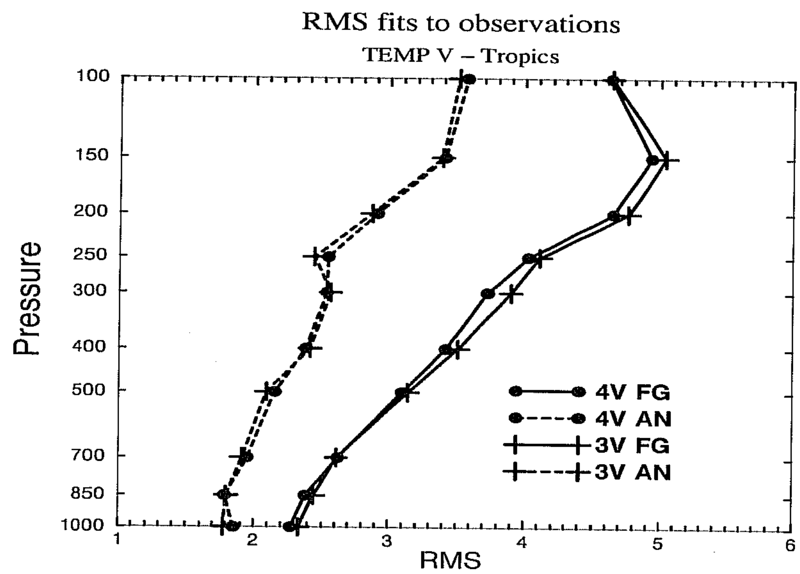


Figure 16: Root-mean-square (RMS) fits of geopotential height (upper panel) and meridional wind (lower panel) to radiosonde measurements for the tropics (20N-20S), averaged over 4 weeks of the parallel suites (October-November 1997). The solid lines represent the RMS fits of the backgrounds to observations, the dashed lines the RMS fits of the analyses to observations. 4D-Var is shown as circles, 3D-Var as plus signs. The abscissa is the RMS in geopotential meters for height and m/s for wind. The ordinate is the pressure in hPa.

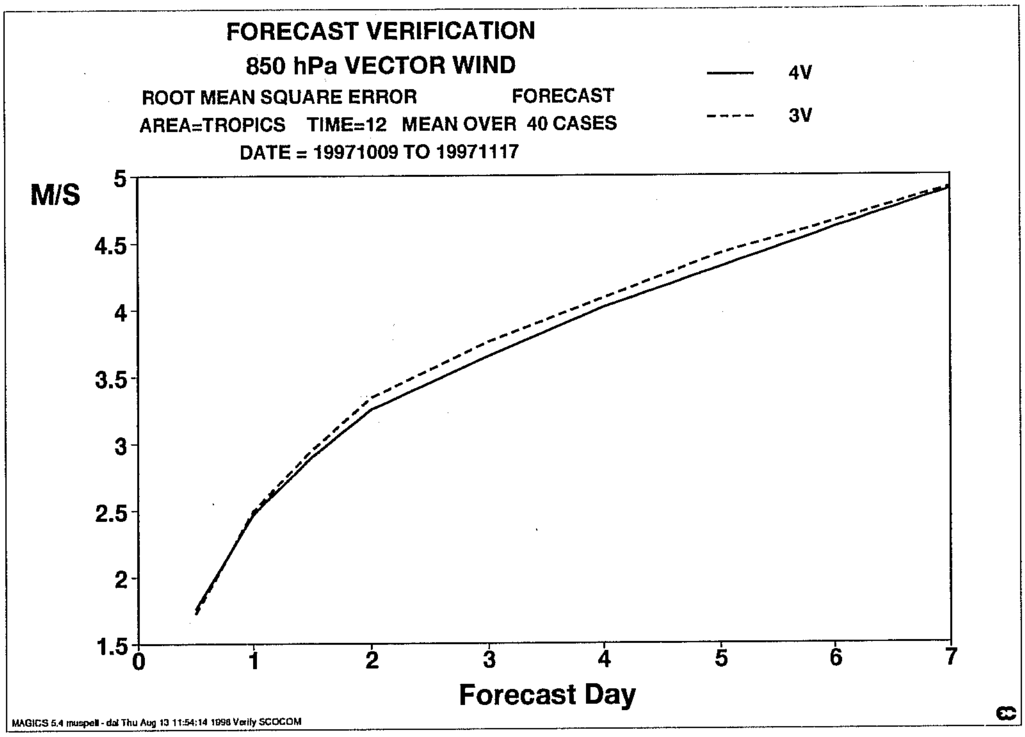


Figure 17:Figure 2: Root-mean-square of forecast errors for the vector wind in the Tropics (20N-20S) for 40 days of the parallel operational suites. 4D-Var is represented by a solid line, and 3D-Var by a dashed line. The abscissa is the RMS in m/s and the ordinate is the pressure in hPa.

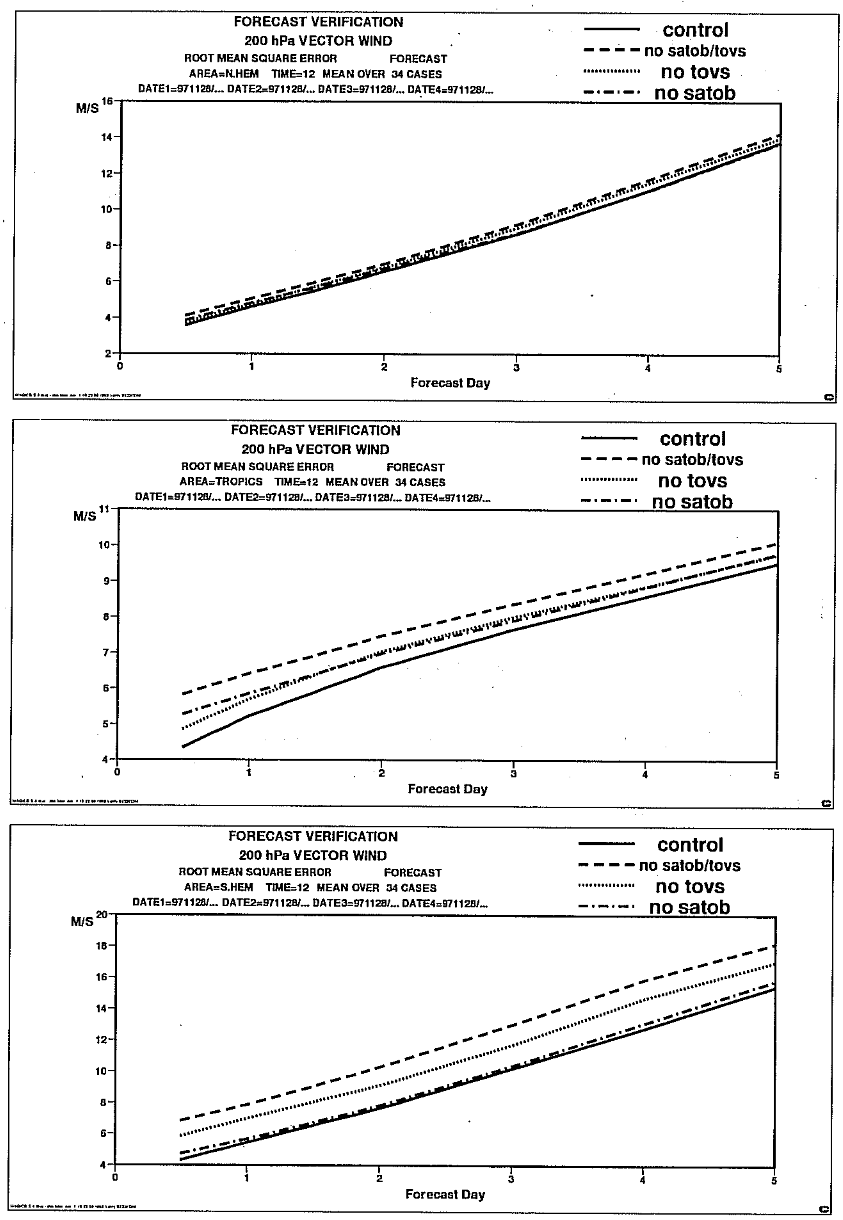


Figure 18: 200hPa vector wind root-mean-square forecast error for the Northern Hemisphere (top panel), the Tropics (middle panel) and the Southern Hemisphere (bottom panel) for the satellite OSEs.

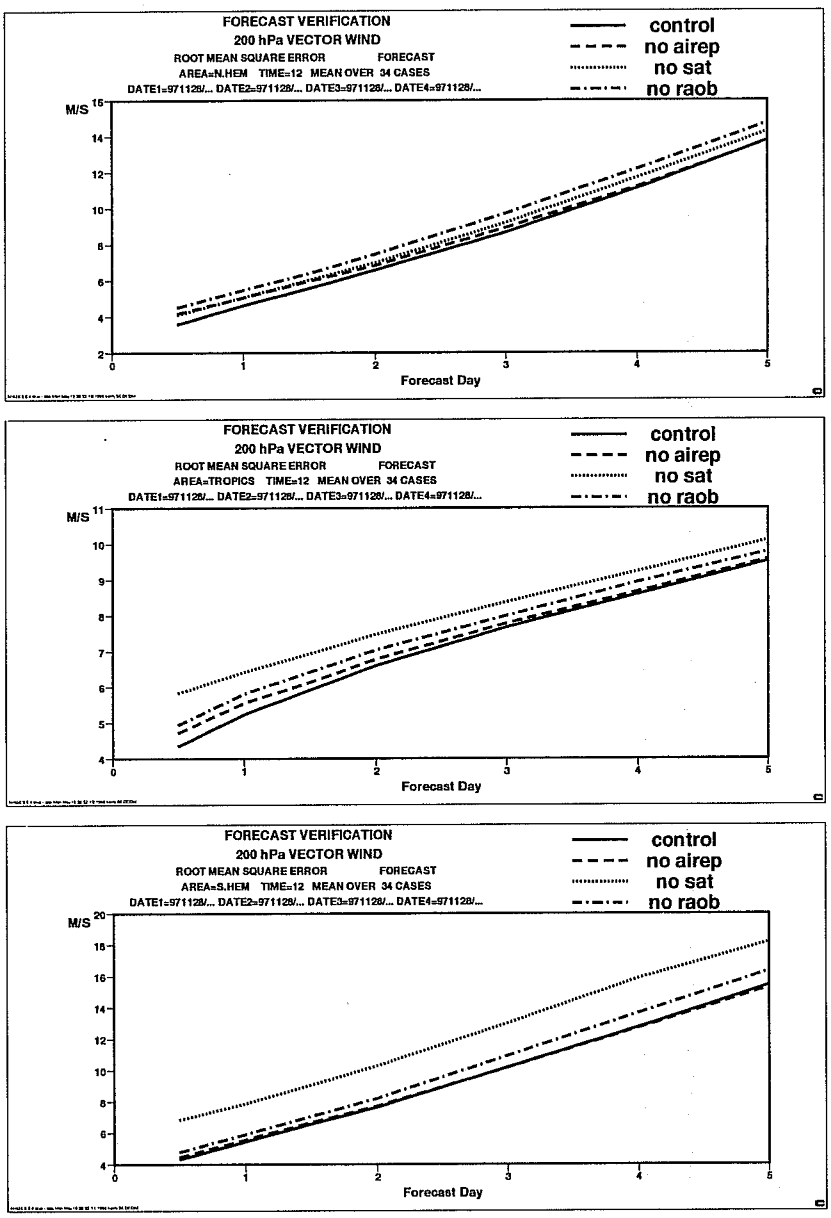


Figure 19: 200hPa vector wind root-mean-square forecast error for the Northern Hemisphere (top panel), the Tropics (middle panel) and the Southern Hemisphere (bottom panel) for the conventional upper air OSEs.