3D-Var - the new operational analysis scheme

Reprinted from ECMWF Newsletter Number 71

Introduction

A new analysis scheme, based on variational methods, became operational at the end of January, 1996. It replaces the current OI (Optimum Interpolation) scheme, which has been operational at ECMWF for more than 15 years, in fact from the beginning of operational forecasting in 1979. The development of the variational scheme started in 1987, in parallel with continued development of OI. The new scheme gives similar forecast accuracy and comparable analysis quality to the OI in the Northern Hemisphere, and it gives an improvement in the Southern Hemisphere.

New features

One of the advantages of the new scheme is that it is more flexible in its use of observations, and more easily adapted to new types of observations. In addition to the data currently used by OI, the first operational implementation of 3D-Var will use scatterometer wind data and TOVS radiance data. The TOVS data are used in OI in the form of retrieved thicknesses and precipitable water content data.

The addition of scatterometer data has improved the analysis of the low-level wind fields over sea - a tropical cyclone case will be presented below. The better use of TOVS radiance data is one likely reason for the good Southern Hemispheric results.

Another advantage of the new scheme is that the level of noise (gravity waves) is controlled within the analysis itself. 3D-Var thereby combines three tasks: retrieval of TOVS data, analysis and initialisation in one step, rather than in three separate job steps. This should lead to a more optimal combination of the information in the different types of data.

3D-Var is a completely new program code for the upper-air analysis (temperature, vorticity divergence and specific humidity on model levels, and surface pressure). The analysis is performed directly in terms of the forecast model's spectral representation, on model levels. Observation processing, quality control, background error computation and surface analysis remain from the old system, but will soon be replaced by new modules.

The specification of background errors and observation errors are somewhat different and more realistic in 3D-Var, compared to the operational scheme. There are therefore visible differences between analysis increments produced by 3D-Var compared to OI. Wind increments are smaller due to tuning of background errors and broader vertical correlations; temperature increments are smaller, particularly near the surface, and they have a larger vertical scale in the free atmosphere. The stratospheric increments are more large scale and much smoother in 3D-Var.

The implementation of 3D-Var is an important step towards 4D-Var, the four-dimensional variational scheme. The fourth dimension in 4D-Var is time, indicating that the development in time of the atmosphere and of forecast errors will be better accounted for by the use of observations distributed in time over several (up to 24) hours, and by the use of the forecast model itself as a constraint. 4D-Var and 3D-Var share the same program codes. The testing of 4D-Var will intensify on the more powerful Fujitsu super computer that will be available during 1996.

Summary of forecast impact

The new analysis scheme has undergone a very comprehensive programme of testing and assessment of its impact, during the last two years, first at T106 resolution and then during the last year at full operational T213 resolution. 3D-Var has been run in parallel with OI for a total of 149 days:

	From	To	No. of days
1	941206	950117	43
2	950405	950421	17
3	950422	950514	23
4	950824	951028	66
Total			**149**

3D-Var has significantly different characteristics to OI. Many important aspects of the analyses will change, when 3D-Var replaces OI. The parallel tests have shown that there can be sizable differences between the two analyses, especially in the more data sparse oceanic areas. Although the differences on occasion have been large, it has often been difficult or impossible to assess which of the two analyses is more correct. There are few independent data to use for verification.

For individual days, there may also be large differences in the forecast performance in forecasts run from 3D-Var and from OI, respectively. Over short periods one system or the other may produce better forecasts. However, on average over larger samples (50 cases or more), the forecast quality of the two schemes comes out to be very similar. It is equal in the European area (Figure 1a) as well as in the Northern Hemisphere as a whole (Figure 1b); whereas 3D-Var performs clearly better than OI in the Southern Hemisphere, as we can see in Figure 1c. The diagrams show an objective measure of forecast accuracy as a function of forecast time (0 to 10 days), averaged over the last T213 parallel run (i.e. 66 cases). 3D-Var is shown as a full line and OI is dashed.

The large variability from day to day comes across well in scatter plots of the type shown in Figure 2. Here, 3D-Var forecast performance (x-axis) is plotted against OI (y-axis), each circle representing one 5-day forecast - so there are 66 circles in all. The cross represents the mean. The points near, or on, the diagonal represent cases with equal 3D-Var and OI forecast quality. Points above the diagonal indicate cases in which 3D-Var has outperformed OI, and vice versa for points below the diagonal. We see that in Europe, the North Atlantic and the Northern Hemisphere as a whole (panels a, b, c), there is a large number of cases in which one scheme is better than the other, with an approximately equal number of cases on both sides of the diagonal. In the Southern Hemisphere (panel d) we see a significant shift in the cloud of points in favour of 3D-Var.

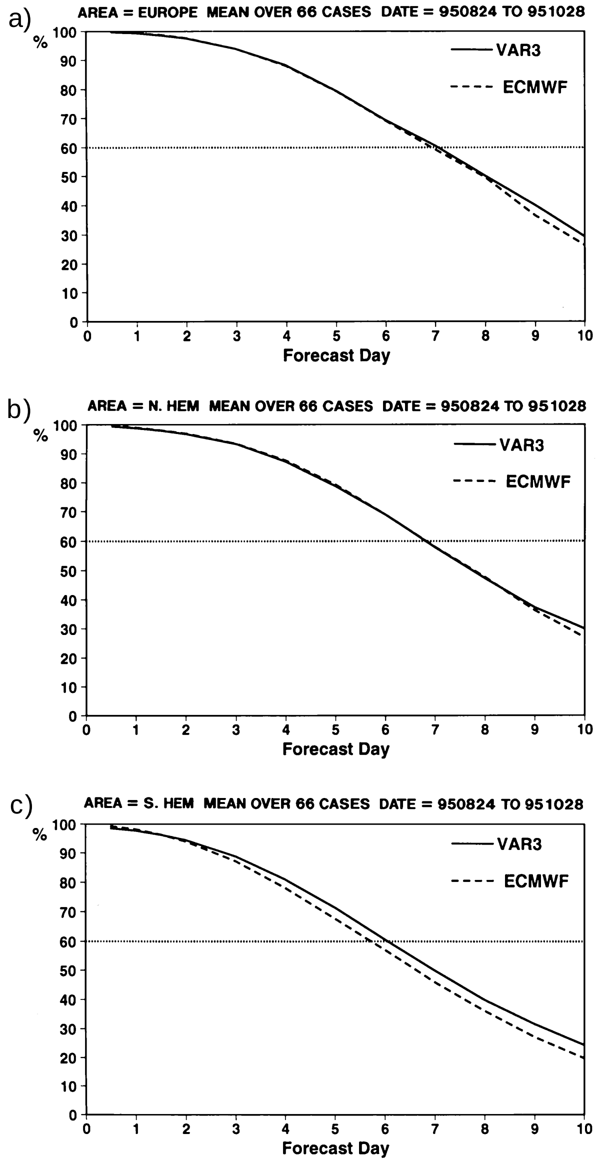

Fig. 1 Average forecast scores (anomaly correlation) for 500 hPa height for the fourth experiment period (950824-951028), 66 cases. The full line, labelled VAR3, shows 3D-Var and the dashed line, labelled ECMWF, shows the forecast performance of the operational (OI) scheme. Panel a) shows Europe, b) shows Northern Hemisphere and Panel c) shows Southern Hemisphere extra tropics.

Although the medium-range Northern Hemisphere forecast scores are neutral compared with operational ones, there is a very small degradation of the short range (1 day) geopotential scores from 3D-Var. This is not noticeable on any charts (1-2 m height) but is systematic in scores. It is likely to be related to different large scale analysis and no initialisation of the full field in 3D-Var. This absence of initialisation is believed to be beneficial for the tropics. The prediction of tropical cyclones is noticeably improved with 3D-Var. Temperature scores are generally better from 3D-Var, especially at low levels and in the stratosphere. Wind scores are better at 200 hPa and above.

Main characteristics of the variational scheme

At ECMWF, global analyses are produced every six hours, using data within the six-hour time window surrounding each analysis time. Approximately 40,000 to 80,000 observations are used in each analysis.

The variational formulation allows the analysis problem to be solved globally but not exactly. An approximate solution is found through a number of iterations of a minimisation algorithm. The scheme minimises a cost-function which measures the global misfit of the current model state to the observations and to the background. The background is a six-hour forecast valid at the analysis time, and based on the previous analysis. For each iteration the minimisation requires the calculation of the cost-function and its gradient with respect to the model variables. Given the value of the cost-function and its gradient, the minimisation algorithm finds an updated model state, with a smaller cost-function. The process is iterated until convergence, or until the predefined maximum number of iterations has been reached. We currently allow a maximum of 70 iterations in 3D-Var. Figure 3 shows an example of the gradual decrease of the cost function during minimisation. The very slow decrease after iteration number 30 indicates that the minimisation could be interrupted earlier with little loss of accuracy.

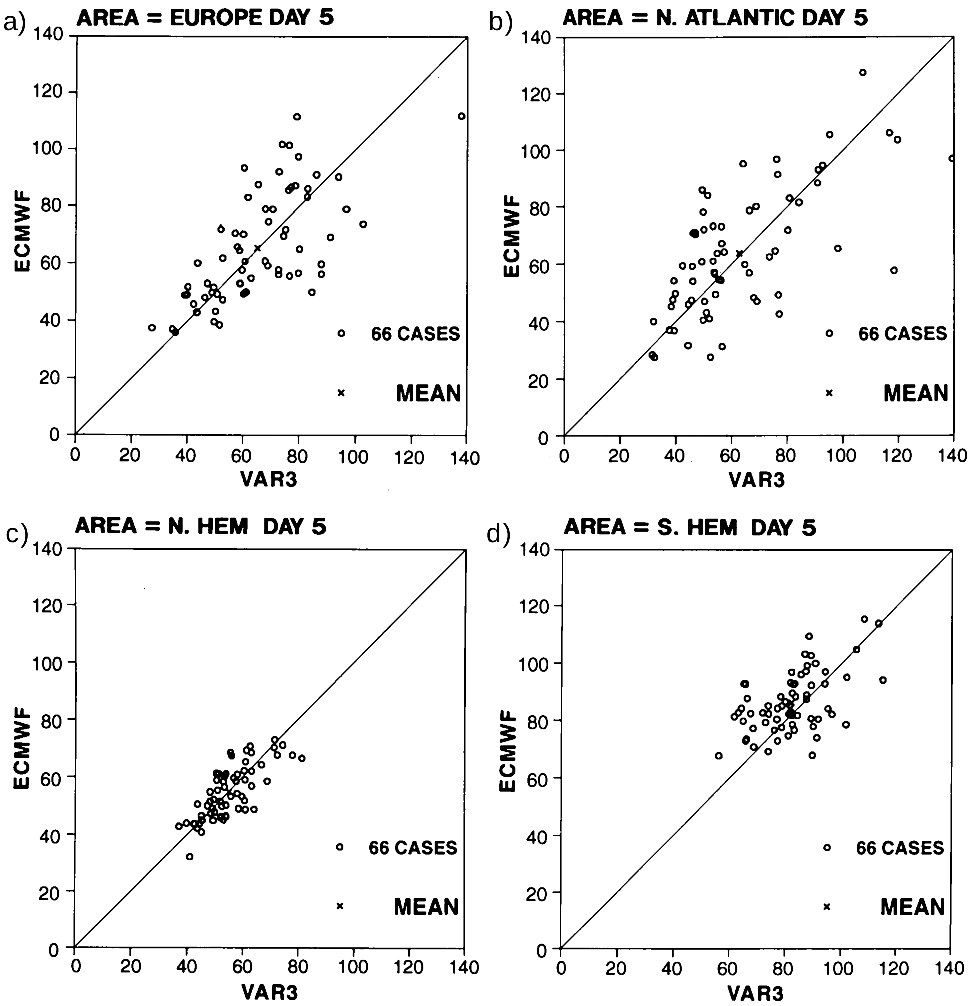

Fig. 2: Scatter-diagram of the 66 individual forecasts in the fourth experiment period (950824-951028). Each case is represented by a circle and the mean is indicated by a cross. VAR3 indicates 3D-Var, and ECMWF indicates the operational (OI) scheme. The measure of forecast skill is here root-mean-square of 500 hPa height forecast error. Panel a) is Europe, b) is North Atlantic, c) is Northern Hemisphere and d) is Southern Hemisphere.

The cost-function consists of just one number (a scalar) but the gradient has the same dimension as the model - in the order of 106. The most efficient way to calculate the gradient is through the application of the adjoint technique. The cost of calculating the gradient with the adjoint technique is typically only two to four times the cost of calculating the cost-function.

The operational OI scheme cannot use all observations in one go. It would be too expensive. Instead the globe is subdivided into so-called “analysis boxes” with dimensions of around 2000 km. OI solves the analysis problem exactly within each box, but since partly different sets of data are used for the analysis of each box some discontinuity between neighbouring boxes is unavoidable. The variational scheme on the other hand is applied globally. All observations are used in one go, so data selection is not an issue. This will produce less small-scale ‘noise’ in the analyses.

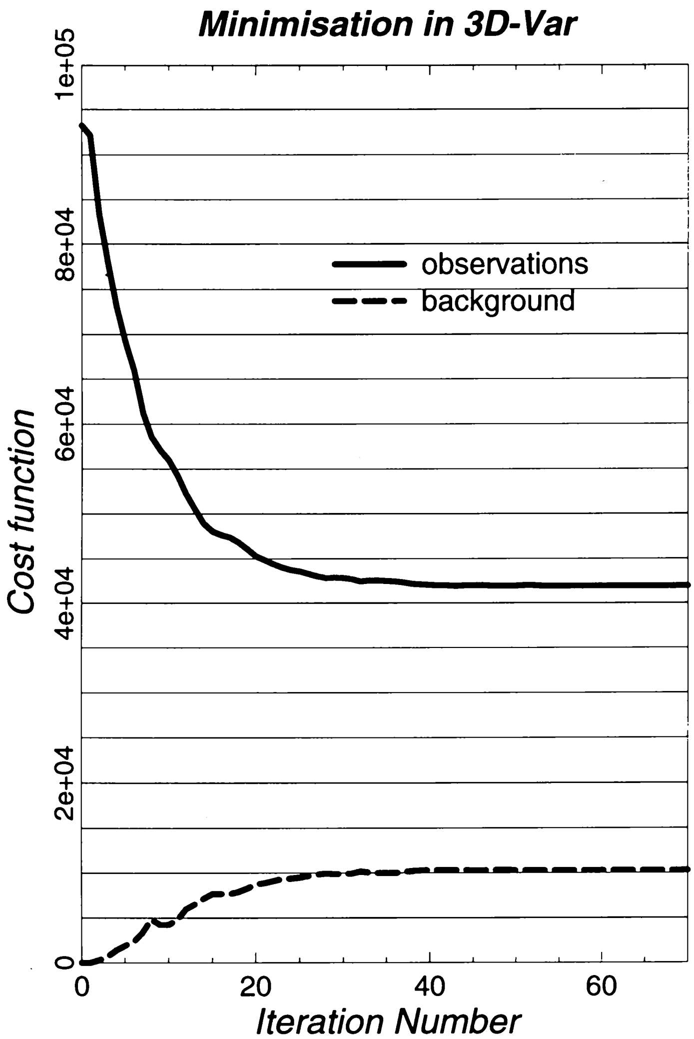

Fig. 3: Evolution of the cost-function during the course of minimisation.

Another important feature of 3D-Var is that observations can be treated multi-variately. By this we mean that one observation can influence the analysis of more than one analysis variable. Radiosonde height observations, for example are a function of both temperature and humidity through the virtual temperature effects of the hydrostatic integration. Thus, in 3D-Var, upper-air height data influence the analysis of both temperature (mainly) and humidity (to a lesser extent). Similar multi-variate effects are more important for the use of some TOVS channels, which depend strongly on both temperature and humidity. Another example is observations of relative humidity.

Examples of analysis impact

Figure 4 shows root-mean-square of the difference between 3D-Var and operational analyses of 500 hPa height, for a 18-day period in October 1995. We see that the two analyses generally are very close over the Northern Hemisphere continents (less than 2.5 m), and that larger differences (up to 7.5 m r.m.s.) occur over the Atlantic and Pacific oceans. The largest

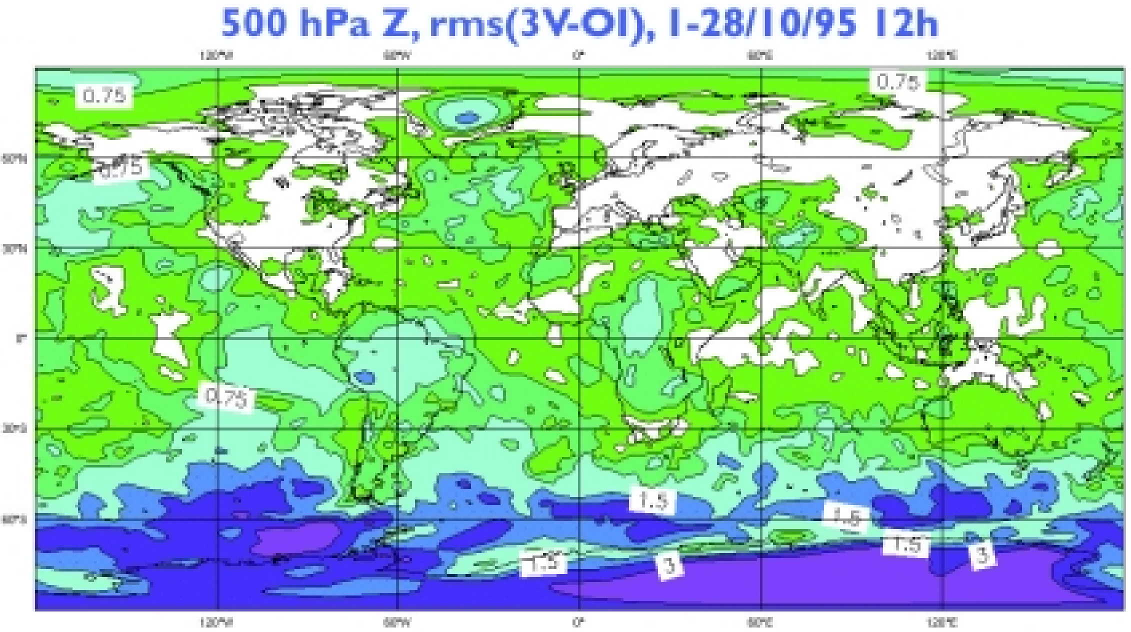

Fig. 4: Root mean square of the 500 hPa height difference between 3D-Var and OI analyses for a 28-day period at the end of the fourth experiment, 951001-951028. The contours are 2.5, 5, 7.5, 10, 15, 20 and 30 metres.

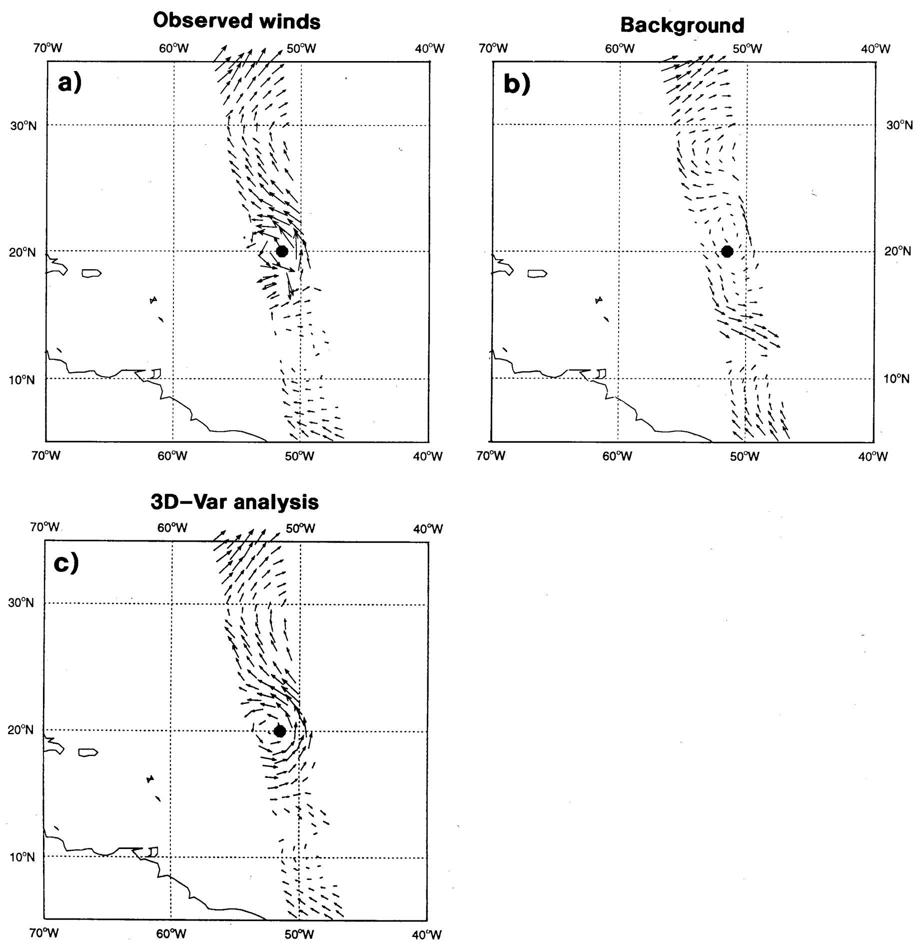

Fig. 5: Tropical cyclone Karen, on the 31 August 1995, 00 UT. Panel a) shows the observed scatterometer winds (ERS-1), b) shows the background winds interpolated to the locations of the scatterometer data and c) shows the 3D-Var analysis, also interpolated to the scatterometer data points.

differences are, as expected, in the Southern Hemisphere mid-latitudes (in excess of 15 m) and over the Antarctic, where the analysis is most uncertain, due to relatively sparse data-coverage.

Figure 5 shows an example of an analysis of a tropical cyclone - in this case tropical cyclone Karen, on the 31 of August, 1995. Panel a) shows the observed scatterometer winds for an orbit which passes directly over the cyclone position (indicated by a large dot, at 20 North, 52 West). Panel b) shows the background (six-hour forecast) valid at the same time, and panel c) shows the 3D-Var analysis. The OI analysis in not shown, but is similar to the background field in this case, because few conventional data exist in this area, and OI does not use scatterometer wind data. This is a very striking example of favourable impact of additional data used in 3D-Var. Statistically, over the whole experiment period we see a significant improvement of the definition of the analysed wind field in and around tropical cyclones, and a small improvement of the day-1 and day-2 forecasts.

Summary

The three-dimensional variational analysis scheme replaced the operational optimum interpolation scheme at the end of January, 1996. This will lead to important changes in many aspects of the ECMWF analyses, but not significantly alter the forecast performance in the Northern Hemisphere mid-latitudes, and Europe in particular. The results from extensive experimentation indicate that we should expect an improvement in the forecast performance in the Southern Hemisphere.

Implementation of 3D-Var provides the basis for future more general improvements in data assimilation and medium range forecasting.

E. Andersson, P. Courtier, C. Gaffard,
J. Haseler, F. Rabier, P. Undén and D. Vasiljevic.