On the impact of frequent data in ECMWF's 4D-Var scheme: Hourly surface pressure data, European profilers and profiling aircraft data.

E. Andersson, C. Cardinali, L. Isaksen and A. Garcia-Mendez

ECMWF, Shinfield Park, Reading, U.K., e.andersson@ecmwf.int

1. Background

We summarise the results from three recent impact studies involving data from frequently reporting data types. Our main interest is to test if a modern data assimilation system (such as 4D-Var) can derive benefits from irregularly distributed a-synoptic data, even in relatively data dense regions. Complete reports describing each of the three studies separately are under preparation, and will appear as ECMWF Technical Memoranda shortly (*Andersson and Garcia-Mendez* 2002; *Cardinali et al* 2002). The three studies are:

(1) The denial of hourly SYNOP, SHIP and DRIBU observations;

(2) The addition of European wind-profiler data;

(3) The denial of profiles from ascending and descending aircraft (AMDAR and ACARS) over America and Europe.

The WMO/CBS Expert Team on data requirements and the re-design of the global observing system in April 2001 suggested seven Observing System Experiments (OSEs) for consideration by Numerical Weather Prediction (NWP) centres. As part of ECMWF's contribution it was proposed to study items 1) and 3) above, with a high-resolution (T511/T159) 12-hour 4D-Var assimilation system.

This short paper provides a brief description of the operational 4D-Var system (Section 2), the definitions of the three experiments and summaries of main results (Section 3), followed by overall conclusions in Section 4.

2. The 4D-Var scheme

The operational 4D-Var scheme (*Rabier et al.* 2000) uses a wide variety of meteorological observations (as outlined in *Courtier et al.* 1998) from both conventional and satellite instruments. Global data within a 12-hour period (the so-called assimilation window running from 03 to 15 UTC and from 15 to 03 UTC the following day) are used simultaneously (*Bouttier* 2001).

2.1 *Incremental formulation*

The 4D-Var estimation problem is solved by minimising iteratively a cost function J with respect to the model state x at the time t_0 at the start of the assimilation window. In the incremental formulation (*Courtier et al.* 1994) the cost function is written in terms of increments δx with respect to the background-state x_b (a short-range forecast), i.e. $\delta x = x - x_b$. The increments are propagated in time using the tangent linear $\mathbf{M} \equiv (\partial M / \partial x)|_{x = x_b}$ of the forecast model M, and compared with the observations by means of the tangent linear $\mathbf{H}$ of the observation operators H:

$$J(\delta x) = \delta x^T B^{-1} \delta x + \sum_{i=1}^{N} (H_i M_i \delta x - d_i)^T R^{-1} (H_i M_i \delta x - d_i)$$

where the summation is over N sub-divisions (or time slots) of the assimilation time window. The length of each time slot was one hour (i.e. $N = 13$) in the operational system before January 2002 when it was halved to 30 minutes ($N = 25$). The vector d_i represents the innovations:

$$d_i = y_i - H_i x_b(t_i) = y_i - H_i M_i x_b(t_0),$$

where y_i represents the observations. Note that the innovations are calculated using the non-linear observation operators, after propagating the model state to the time of the observations using the full non-linear forecast model. This ensures the highest possible accuracy for the calculation of the innovations which are the primary input to the assimilation.

To reduce the computational cost, the increments δx are calculated at a lower resolution than that of the full model. The current forecast model is run at T511 spectral truncation (corresponding to a 40 km resolution, Miller 1999) whereas the analysis increments δx are evaluated at T159 (120 km). The analyses $x_a(t)$ at times t = [0, 6, 12, 18] UTC are formed by adding the increments to the background fields:

$$x_a(t) = x_b(t) + \delta x(t).$$

3. Recent data impact studies at ECMWF

The OSEs were undertaken with the most recent configuration (late 2001) of the operational forecasting system. The data impact was studied through the denial or addition of observations received operationally from the Global Telecommunications System (GTS). Any demonstrated impact will be the incremental effect of the observations denied or added in the experiments. All other data sets used in the operational system entered the data assimilation as usual.

3.1 *Denial of hourly surface data from SYNOP, SHIP and DRIBU*

The impact of hourly surface pressure observations (over land and sea) and also wind observations (over sea) has been evaluated. It is known from earlier experimentation (*Järvinen et al.* 1999) that the temporal information in time-series of surface pressure data can correct the intensification rate of developing cyclones, when used in 4D-Var. This has now been re-investigated as the operational system recently was enhanced with a longer time window (12-hourly cycling, *Bouttier* 2001) and higher analysis resolution: the main forecast model at T511 and the analysis increments at T159 spectral resolution. These enhancements are expected to be particularly beneficial with respect to high density and high frequency data.

The study period was 1-31 May 2001. The globally available observations from the main synoptic hours at 00, 06, 12 and 18 UTC were used in the experiment and the control. Only the data from the intermediate hours were excluded from the experiment. The amount of excluded data is about 22,000 SYNOP/SHIP and 2,300 DRIBU per 12-hour cycle. ECMWF receives such observations mainly from the European-Atlantic region, from moored buoys in the American coastal waters, from Australia and from the drifting buoys in the southern oceans (Figure 1). The impact from the hourly data can be detected in the analysis increment fields and in the 48-hour forecast error of 500 and 1000 hPa geopotential. However, in the medium-range the signal is weak and, considering the relatively short period of the experiment, is largely statistically insignificant. Figure 2 shows scatter plots of the 48-hour forecast error for the North Atlantic and North America. The positive impact of the hourly data, at this forecast range, is statistically significant at the 90% level (t-test).

The results of the study confirm that hourly surface observations have a positive impact when used in the operational 4D-Var assimilation system. The positive impact in the short-range forecast is seen in those areas where such data are available, i.e. the North Atlantic and the southern oceans, and which are otherwise relatively data sparse with respect to in situ surface observations. The global exchange of all hourly surface observations for use in a 4D-Var system appears to be beneficial for NWP, and will be further pursued through the WMO.

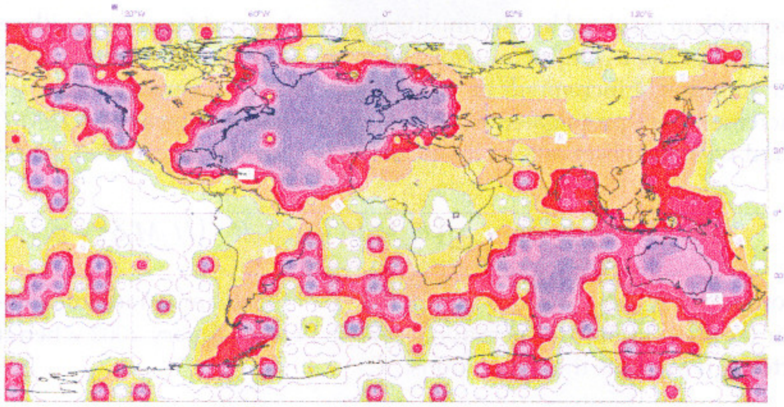

Fig. 1: Frequency of occurrence (per day, per 10°x10° box) of hourly surface pressure data, on average during the month of May 2001. SYNOP, SHIP and DRIBU data have been considered. The contour interval is 2 (per day), with shading starting at 4 per day (I.e. one observation per 6 hours, on average).

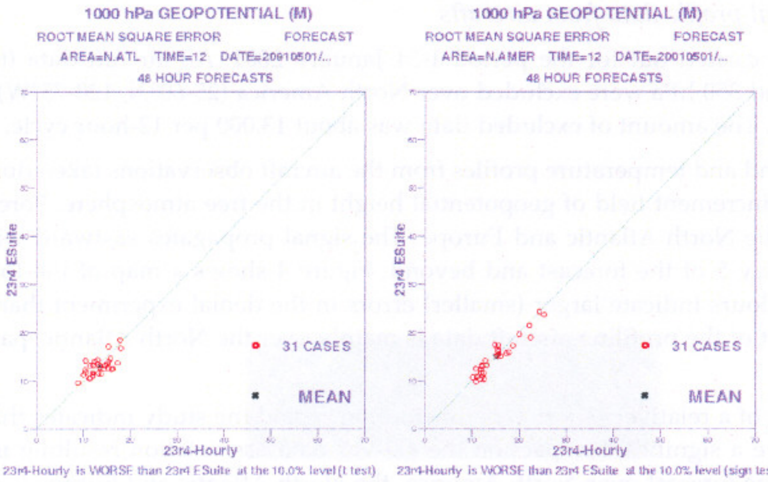

Fig. 2: Scatter plot of the rms of 48-hour forecast error (m) for 1000 hPa geopotential height. Each red circle represents one day in the period 20010501-20010531, 12 UTC, for the North Atlantic (left) and North America (right). The error in the forecasts from the denial experiment is plotted along the x-axis and that of the control along the y-axis. The average forecast error is shown by the x-symbol. Markers plotted below the diagonal indicate larger error in the data denial experiment than in the control. A t-test indicates that the positive result is significant on the 90% level, both panels.

3.2 *Addition of European wind profiler data*

The European research on stratospheric and tropospheric wind radars has in recent years been co-ordinated to form a more coherent operational network. This has been achieved through the COST-76 and CWINDE-99 projects. The current network consists of more than twenty stations in ten countries, with further expansions planned. In the year 2001 ECMWF has been receiving data in real-time from 17 of the stations: in Great Britain (5), Ireland (1), the Netherlands (1), Switzerland (1), France (3), Germany (2), Austria (3) and Italy (1). The real-time data have been monitored and compared against short-range forecasts and analyses, generated by ECMWF's global NWP system. The quality of the observations has been assessed (*Andersson and Garzia-Mendez* 2002).

The data were found to be of generally good quality with some distinct improvements towards the end of 2001. The network is far from homogeneous as instrumentation, quality control procedures and reporting practices differ from station to station. Three of the stations reach the lower stratosphere, and five are exclusively for boundary layer profiling. Most operational stations report every 30 minutes with Aberystwyth (in Wales) reporting even more frequently at ten-minute intervals.

Eight of the network's stations were selected for use in an impact study in May 2001. Approximately 1,800 additional data per 12-hour cycle were used. The analysis and forecast impact was relatively small and confined to parts of Central Europe and the areas around the British Isles (Figure 3). A small but statistically significant positive forecast impact was found at day-3 for the North Atlantic verification area, based on a sample of 31 forecasts in May 2001. The results were considered sufficiently good to go ahead with pre-operational trials. Trials have started from 1 January 2002 in which European profiler time series are now used from twelve stations on a half-hourly basis. Thinning to a minimum vertical separation of 5 hPa (or 5% of pressure above 100 hPa) has been introduced.

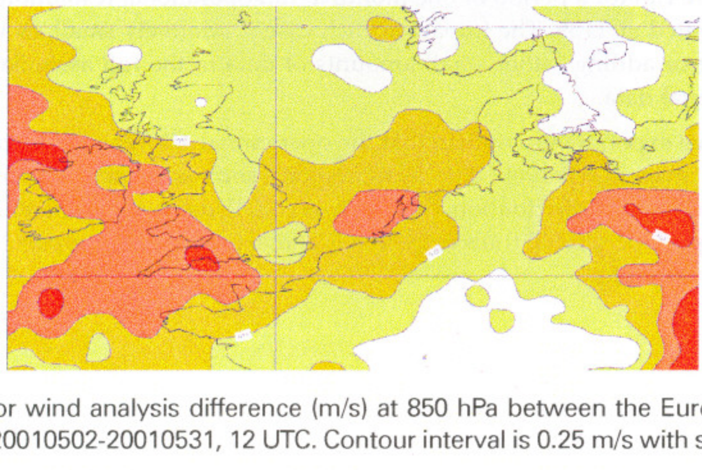

Fig. 3: RMS of the vector wind analysis difference (m/s) at 850 hPa between the European profiler experiment and the control, 20010502-20010531, 12 UTC. Contour interval is 0.25 m/s with shading starting at 0.5 m/s.

3.3 *Denial of vertical profile data from aircrafts*

The experiments were carried out for the period 1-31 January 2001. All aircraft data (temperature and wind) between the ground and 350 hPa were excluded over North America (25-60°N, 120-75°W) and over Europe (35 – 75°N, 12.5°W – 42.5°E). The amount of excluded data was about 13,000 per 12-hour cycle.

The impact of the wind and temperature profiles from the aircraft observations taken during ascent and descent can be detected in the increment field of geopotential height in the free atmosphere. Forecast errors are reduced over North America, the North Atlantic and Europe. The signal propagates eastward with forecast time and is clearly visible out to day 5 of the forecast and beyond. Figure 4 shows a map of the forecast impact at day 5. Red/orange (green) colours indicate larger (smaller) errors in the denial experiment than in the control. We see that the positive impact of the profiling aircraft data is mainly over the North Atlantic, parts of Europe and parts of the Arctic.

Even with the caveat of a relatively short experimentation period the study indicates that the atmospheric profiles from aircrafts have a significant impact on the 4D-Var data assimilation resulting in improvements of the short and medium range forecast over North America, the North Atlantic and Europe (*Cardinali et al.* 2002). The results are statistically significant (t-test) and support the expansion of the coverage of aircraft observations including the observations taken during ascent and descent from other parts of the globe.

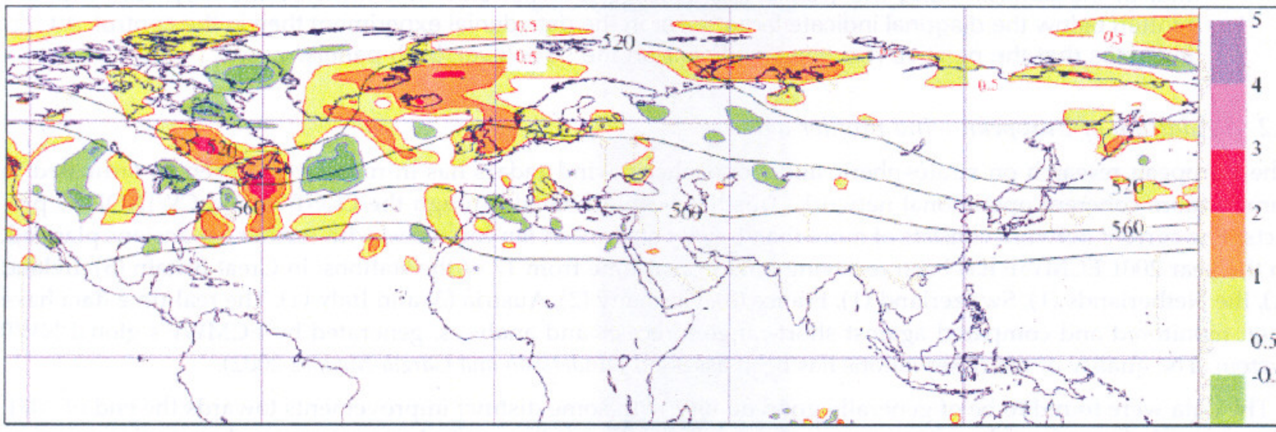

Fig. 4 Difference in rms of 120-hour forecast error, between experiment 3 and the control, for 500 hPa geopotential height, 20010102-20010131, 12 UTC. The shading starts at +- 0.1 decameters, with yellow (positive) shading indicating larger analysis increments in the data denial experiment. Area integrated values are: Europe 3.02 m, North Atlantic 2.90 m, North America 0.31 m and the Northern Hemisphere extra tropics 1.35 m.

4. Conclusions

Three data impact experiments have been conducted, assessing the benefit of frequent observational data in the context of ECMWF's operational 4D-Var system. The three experiments were 1) denial of hourly surface pressure data, 2) addition of European profiler data and 3) denial of profiling aircraft data over America and Europe. In each case experiments were run for a period of one month at full operational resolution: T511 forecast model with analysis increments computed at T159. The impact in each case is relatively small, as we have studied the incremental impact of denying or adding rather small amounts of data to the full assimilation system, which uses a wide variety of observational data.

The results from the three experiments confirm that the operational 4D-Var scheme can obtain benefit from frequent and a-synoptic data, even in relatively data dense regions, such as Europe and North America. Results from experiment 1) support the global distribution of hourly SYNOP data, experiment 2) led to the inclusion of European profiler data in pre-operational trials at ECMWF, and experiment 3) support the continued expansion of the AMDAR and ACARS networks.

References

Andersson, E. and A. Garcia-Mendez, 2002: Assessment of European wind profiler data, in an NWP context. ECMWF Tech. Memo. In press.

Bouttier, F., 2001:The development of 12-hourly 4D-Var. ECMWF Tech. Memo. 348.

Cardinali, C., L. Isaksen and E. Andersson, 2002: Use and impact of automated aircraft data in 4D-Var. ECMWF Tech. Memo. In press.

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

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

Järvinen, H., E. Andersson and F. Bouttier, 1999: Variational assimilation of time sequences of surface observations with serially correlated errors. *Tellus*, **51A**, 469-488.

Klinker, E., F. Rabier, G. Kelly and J-F. Mahfouf, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. III: Experimental results and diagnostics with operational configuration. *Q. J. R. Meteorol. Soc.*, **126**, 1191-1215.

Miller, M., 1999: Resolution studies. ECMWF Tech. Memo. 299.

Rabier, F., H. Järvinen, E. Klinker, J.F. Mahfouf and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: experimental results with simplified physics. *Q. J. R. Meteorol. Soc.* **126**, 1143-1170.