371

Use and impact of automated aircraft data in 4D-Var

Carla Cardinali, Lars Isaksen
and Erik Andersson

Research Department

Submitted for publication in Mon.Wea.Rev

June 2002

For additional copies please contact

The Library
ECMWF
Shinfield Park
Reading, Berks RG2 9AX

library@ecmwf.int

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications.html

Abstract

The use of automated aircraft data (AMDAR and ACARS) has recently been extended in ECMWF's operational 4D-Var data assimilation system. We report on a modified data selection procedure that allows the use of more aircraft profiling data during the aircraft's ascending and descending phase, and more of the most frequent reports at cruise level. We show that the accuracy of analysed jet streams is improved through these changes, as verified against independent (non-real time) aircraft data that had not been used in the experiments. The modifications are shown to have a clear positive impact on the short and medium-range forecast performance. The revised aircraft usage was implemented operationally in January 2002.

The impact in 4D-Var of profiles from American and European automated aircraft in ascending and descending phase has been tested in a data denial impact study, for January and July 2001. This particular impact study was run partly on the request of the WMO/CBS Expert Team on Data Requirements and the Redesign of the Global Observing System. Their interest is in testing whether a modern data assimilation system (such as 4D-Var) obtains substantial benefit from the aircraft profiles, which sample very irregularly in space and time, given that America and Europe are relatively well covered by radiosondes and wind profilers. The results show a substantial positive impact of the profiling aircraft data on analysis and forecast accuracy. The short-range forecast performance is improved over North America, the North Atlantic and Europe. In the medium-range a clear positive impact is found in the North Atlantic, the European and Arctic areas, in the winter period, and beyond day-6 in the summer period. These results are statistically significant, and support the ongoing WMO initiative for further expansion of the AMDAR/ACARS coverage. The results also illustrate the effectiveness of 4D-Var with respect to observations that are irregularly distributed in space and time.

1. Introduction

Over the past few years the number and the coverage of automated aircraft data has increased very significantly - in particular the ACARS (Aircraft Communication Addressing and Reporting System) and AMDAR (Aircraft Meteorological Data Report) systems (WMO 1996). In 2001 the WMO AMDAR programme completed its second year of operations having implemented a data distribution plan for the co-ordinated network of air-routes. The reconfigured data distribution makes the AMDAR programme more cost effective by reducing the number of redundant data and by improving the spatial and temporal coverage. Due to this effort more aircraft observations become available in otherwise data sparse areas.

AMDAR reports are often produced at the specified frequency of one report per 7.5 minute at cruise level and one per minute during ascending and descending phases, thus providing data at altitude roughly every 100 km along the flight path, as well as detailed profiles in the near vicinity of airports. The telecommunications cost to collect the data in real-time can be as low as 1 US cent in some countries, and up to 50 US cents over some oceans, with a global median value of 5 US cents (Jeff Stickland, personal communication).

The ECMWF data assimilation system is a 4D-Var scheme (Rabier *et al.* 2000). One of the strengths of 4D-Var is its ability to assimilate frequent data, as demonstrated by Järvinen *et al*. (1999). All available observations within a 12-hour period are used simultaneously (Bouttier 2001a) in one global estimation problem. The observations are compared with a short-range forecast on an hourly basis (or half-hourly, since January 2002). The differences between observations and a short-range forecast are analysed to obtain a corrected model state (the analysis) which evolves during the 12-hour assimilation period in better agreement with the observations. The short-range forecast and the comparison with observations is carried out at full resolution, currently T511 spectral truncation (40 km), whereas the analysis increments are evaluated at T159 (120 km). The vertical domain extends from the surface to 0.1 hPa and is discretised on 60 model levels.

Given the relatively high spatial resolution of the current 4D-Var and that the distribution in time of the observations is accurately accounted for, it is of particular interest to demonstrate that the system can make effective use of frequent and dense data. In this paper we assess the impact of data from European and American automated aircraft data in ascending and descending phase, and report a significant positive impact on analysis and forecast. Other sources of frequent data have also been subject to similar investigations recently, e.g. water vapour radiances from geostationary satellites (Köpken, 2001), European radar wind profiler (Andersson and Garcia-Mendez, 2002; Bouttier 2001b) and hourly data from surface stations (Järvinen *et al*. 1999).

The AMDAR and ACARS measurements are of higher quality than the traditional AIREP measurements. Wide-bodied aircraft provide automated reports of wind and temperature measurements with an accuracy comparable to that of radiosondes: 1-2 m/s for wind and 0.7-1.2 K for temperature (Benjamin *et al*. 1999; WMO 1996). The assumed observation errors have been reduced in many operational forecast centres (e.g. Dalby and Berney, 1999) to reflect the increased quality of the observations. At ECMWF aircraft observation errors were reduced by about 30% in 1999 as part of a general revision of the assimilation scheme at the time. The assumed errors for AMDAR and ACARS are set 10-15% lower than for AIREPs.

At ECMWF many data types are thinned before use to avoid potential imbalances between data types with very different densities. Moreover, the effects of error correlation for certain data types are reduced by means of data thinning (Bormann *et al.* 2002, Liu *et al.* 2001). For aircraft data the observation error correlation is thought to be very small, enabling the data to be used at a resolution similar to that of the assimilating model. Aircraft data are thinned along-track considering one flight at a time (Järvinen and Undén 1997), i.e. where there are several flights in an area, the data from different flights will be used as mutually independent measurements. Given that the resolution of the 4D-Var assimilation system has increased significantly in recent years (Untch *et al.* 1996; 1998; Teixeira 1999; Miller 1999, Bouttier 2001a), the possibility of using far more of the high-density aircraft data has become a subject for renewed investigation. In this paper we describe revised thinning and redundancy checks that allow about 60% more of the aircraft data to be used at some levels; these changes were implemented in the operational system in January 2002.

The 4D-Var scheme and the methods for quality control are briefly described in Section 2. In Section 3 we describe the recent changes to aircraft data usage in the ECMWF 4D-Var assimilation system. A revised redundancy check is also described. In Section 4, the analysis and forecast impact of the revised data usage is investigated. An impact study (over two separate one-month periods) on the analysis and forecast impact of profiles from European and American aircraft data in ascending and descending phase, is presented in Section 5. The revision of aircraft data usage (Section 4) and the impact study (Section 5) have been carried out in the context of ECMWF's operational 4D-Var data assimilation scheme (Rabier *et al*. 2000). Conclusions are drawn in Section 6.

2. The 4D-Var scheme

The scheme 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 the following day) are used simultaneously (Bouttier 2001a).

2.1 Incremental formulation

The 4D-Var estimation problem is solved by minimising iteratively a cost function J with respect to the model state $\mathbf{x}$ 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 $\delta\mathbf{x}$ with respect to the background-state $\mathbf{x}_b$ (a short-range forecast), i.e. $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$. The increments are propagated in time using a the tangent linear $\mathbf{M} \equiv (\partial M / \partial \mathbf{x})|_{\mathbf{x}=\mathbf{x}_b}$ of the model M, and compared with the observations by means of the tangent linear $\mathbf{H}$ of the observation operators H:

$$J(\delta\mathbf{x}) = \delta\mathbf{x}^T \mathbf{B}^{-1} \delta\mathbf{x} + \sum_{i=1}^{N} (\mathbf{H}_i \mathbf{M}_i \delta\mathbf{x} - \mathbf{d}_i)^T \mathbf{R}^{-1} (\mathbf{H}_i \mathbf{M}_i \delta\mathbf{x} - \mathbf{d}_i)$$

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 $\mathbf{d}_i$ represents the innovations:

$$\mathbf{d}_i = \mathbf{y}_i - H_i \mathbf{x}_b(t_i) = \mathbf{y}_i - H_i M_i \mathbf{x}_b(t_0),$$

where $\mathbf{y}_i$ 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.

For computational cost reasons the increments $\delta\mathbf{x}$ 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) whereas the analysis increments $\delta\mathbf{x}$ are evaluated at T159 (120 km). The analyses $\mathbf{x}_a(t)$ at times $t = [0,6,12,18]$ UTC are formed by adding the increments to the background fields:

$$\mathbf{x}_a(t) = \mathbf{x}_b(t) + \delta\mathbf{x}(t).$$

2.2 Quality control

All data are quality controlled through comparison with the background (BgQC) and through variational quality control (VarQC) as described by Andersson and Järvinen (1999). In BgQC data are rejected if the departure from the background $d = y - Hx_b(t)$ exceeds a multiple α of its expectation, i.e. rejection occurs if $d > \alpha\sqrt{\sigma_b^2 + \sigma_o^2}$. Here σ_b is the background error standard deviation in terms of the observed quantity (Andersson *et al.* 2000) and σ_o is the assumed observation error. The value of α is set to 5 for most data types including aircraft wind and temperature data.

In VarQC a probability of gross error is computed for each observation at each iteration of the variational minimisation, based on Bayesian probability theory. The weight of the observations in the assimilation is smoothly reduced with increased probability of gross error. Effectively VarQC rejects data that cannot be fitted by the analysis, often because of marked inconsistency with surrounding observations. See Andersson and

Järvinen (1999) for further details. Rejection counts for the aircraft data will be shown as a diagnostic, in the following.

3. Use of aircraft data

All aircraft data received via the Global Telecommunications System (GTS) are considered for use by the assimilation after duplicates have been removed. Daily data coverage maps and monthly time series may be inspected through the internet, on http://www.ecmwf.int/. Daily data counts are currently (March 2002) in the order of 130,000 reports per 24 hours: 20,000 AIREP, 30,000 AMDAR and 80,000 ACARS.

3.1 Revised vertical and horizontal thinning

3.1.1 Vertical thinning of aircraft profiling data

We wish to avoid data densities in excess of the resolution of the assimilating model, as correlated representativeness errors (Lorenc and Hammon 1988) would then potentially become an issue. Other sources of profiling data (PILOT, radiosondes and American profilers) are used at all reported levels without thinning, as data selection has been performed by the data producers prior to distribution on the GTS. In the assimilation system each flight is considered an independent source of data. Thinning is thus performed independently for each flight. In the pre-2002 thinning algorithm (Järvinen and Undén, 1997) the standard pressure levels determine the minimum vertical separation between used data from each flight. With the recent increase in vertical resolution of the ECMWF forecast model, from 31 to 60 levels, the thinning on standard pressure levels no longer provide a satisfactory data distribution in the vertical. The 60-level model has about five times more levels in the troposphere than there are standard pressure levels. Consequently there is the suspicion that a good source of information, such as the wind and temperature profiles from ascending and descending aircraft, is not used to full effect. We have devised and studied a less strict thinning algorithm where (within a small horizontal domain) one report per model level may be selected from each flight.

3.1.2 Horizontal thinning at cruise level

Horizontal thinning is applied to the aircraft data at cruise level. The thinning is performed by selecting one report per flight within a certain distance. In ECMWF's pre-2002 operations the minimal separation between selected data is 125 km. The horizontal analysis resolution has recently been increased from T319/T63 to T511/T159. This paper includes experiments on reducing the minimum separation to 60 km. With the 60 km thinning, all data reported with the stipulated AMDAR frequency (one report per 7.5 minute) will be used. The 60-km thinning will thus only affect those flights that report more frequently than required in current WMO guidelines.

3.2 Revised redundancy checks

With the increasing volumes of AMDAR and ACARS observations it happens increasingly frequently that the same aircraft report is received on the GTS both as a traditional AIREP (in character code) and as an AMDAR or ACARS (in BUFR code). In the present data processing system, aircraft measurements are checked for such redundancy. If two reports have the same time, latitude, longitude and flight level and the measured values are identical, the measurements are considered redundant and only one is used. It turns out, however, that the AIREP and AMDAR/ACARS measurements with identical time/space location often have slightly different observed values due to different reporting practice and reporting accuracy. In the revised redundancy procedure used here, such data are considered redundant and only one of the reports will be used, provided the reports agree to within a specified tolerance.

Each datum should be used only once by the assimilation system. This improved redundancy check typically reduces the number of used aircraft data by 10%, mainly over North America. The rejected observations are not wrong, but duplicates. If two identical observations were used mistakenly and treated as independent they would effectively support each other in the VarQC quality control checks and bias the quality control decisions in an undesirable way - in addition to wrongly increasing the weight of the observation in the analysis.

We have also introduced a stricter consistency check on the reported flight track for each aircraft. We check that it is physically possible for the aircraft to have travelled the distance between consecutive reported locations. We assume that a normal aircraft has a flight speed not exceeding 1200 km/hour (and a super-sonic flight not exceeding 2400 km/hour). The new check rejects those locations that imply unrealistic flight speeds. If more than half of the locations are suspect, the whole flight is rejected. The most typical reason for rejection of this type is that an aircraft incorrectly reports the same time during the whole flight.

3.3 Experiment definitions

For the purpose of our experiments, the 2001 operational 4D-Var data assimilation system has thus been modified to produce analyses with the revised aircraft data thinning and revised redundancy checks, as described above. We studied the impact of the revisions by running two assimilation experiments, each in two separate periods. The two experiments are labeled: CA and RA. CA is the control experiment in which all data types have been assimilated as in the pre-2002 operational system at ECMWF, including the aircraft data. In the RA experiment the revised vertical and horizontal aircraft thinning (Section 3.1) has been applied as well as the revised redundancy check (Section 3.2). Two separate periods have been studied: a 31-day summer period (15 June-15 July 2000) and an 18-day winter period (1-18 December 2000).

4. Analysis and forecast impact of the revisions

The analysis impact has been evaluated in terms of differences (departures) between the observations and the background (a short-range forecast) and between the observations and the analysis. All departures are evaluated at the correct time of the observations using the forecast model at full resolution (T511), and the observation operators to compute the observed quantities from the model fields. Statistics of analysis differences are also shown.

4.1 Departure statistics

Figure 1 shows the root mean square (rms, left) and bias (right panels) of the background (solid) and analysis departures (dashed lines) for the used aircraft wind observations (u-component) accumulated over the whole 49-days of experimentation, summer and winter. Statistics for CA (the control, red lines) and RA (revised thinning and redundancy check, black lines) are shown for the Northern Hemisphere extra-tropics (top panels) and the Tropics (lower panels). The number of observations used in the RA-experiment are printed in black whilst the number differences between RA and CA are printed in red.

The positive sign means that more observations were used in RA, as expected. Throughout most of the troposphere the RA-experiment has slightly smaller departures, i.e. an improved fit to both the background and the analysis. This positive impact is particularly clear at lower levels in the Tropics. There appears to be no systematic change in biases. The temperature statistics (Figure 2) show similar results. In the mid-troposphere, the modified vertical thinning increases the number of used observations by up to 60%, and in the upper troposphere by about 15%.

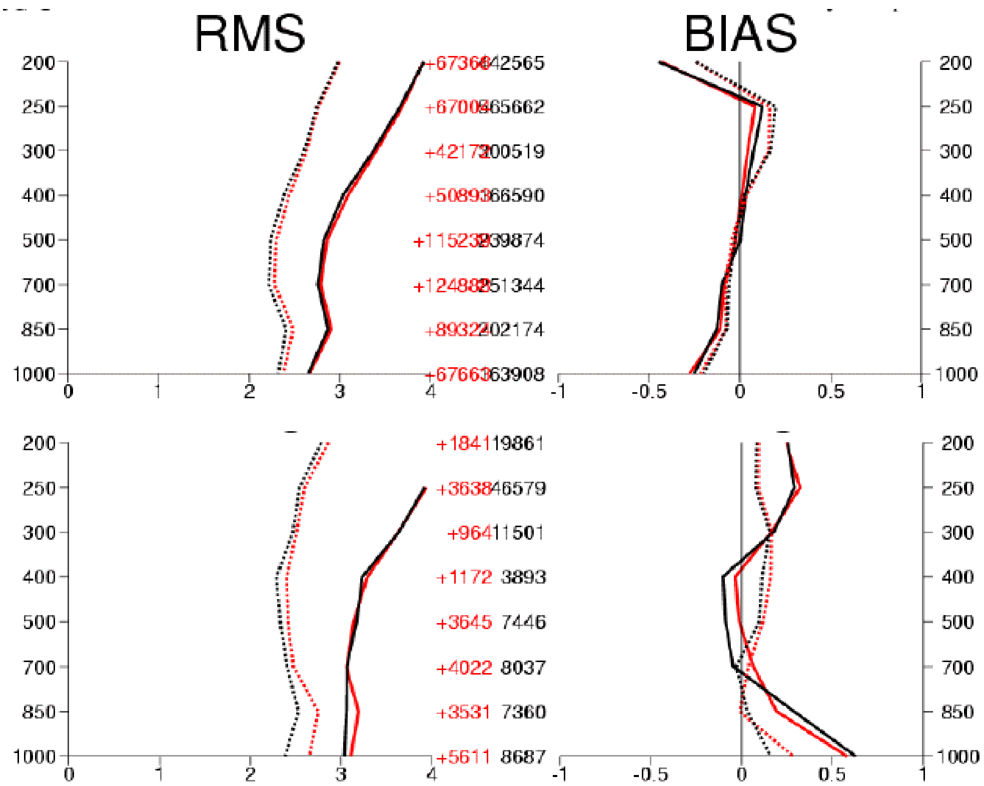

Figure 1 Background (solid) and analysis (dashed) departure statistics (r.m.s. to the left and bias to the right) for the u-wind component of all used aircraft observations, in the Northern Hemisphere extra tropics (top) and the Tropics (lower panels). The RA experiment (revised thinning and redundancy checks) is shown in black and CA (the control) in red. The black numbers indicate data counts in the RA experiment, the red numbers indicate their difference with respect to CA.

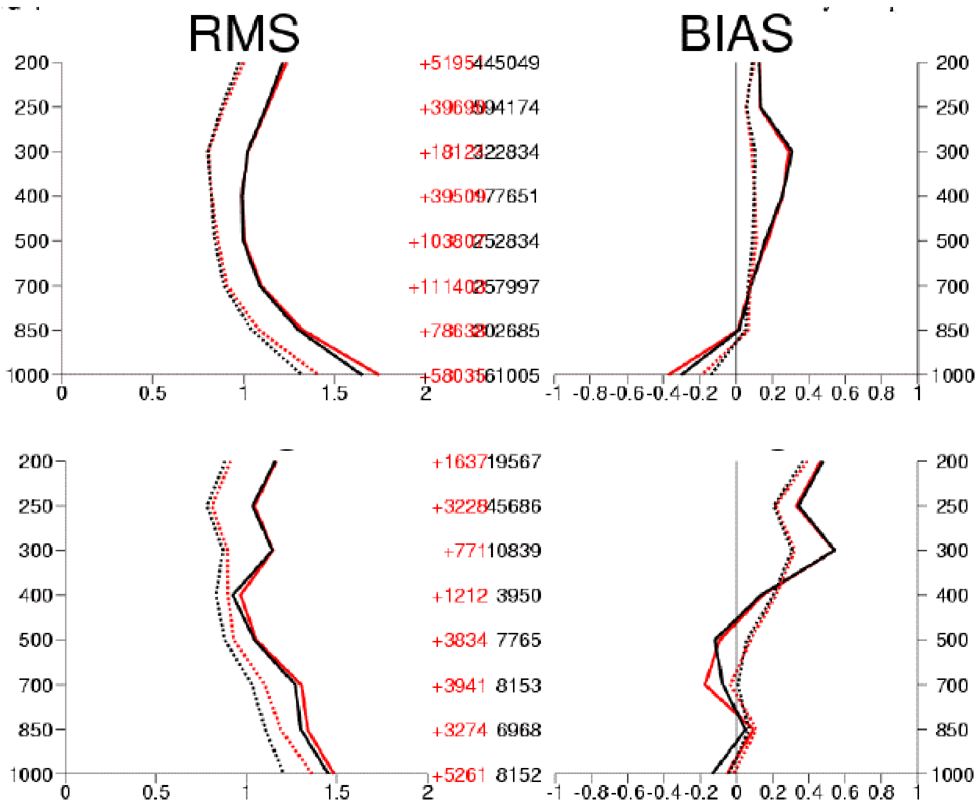

Figure 2 As Figure 1, but for temperature.

4.2 Analysis differences

In Figure 3 we show a latitude-pressure cross-section of the root mean square (rms) of the wind speed analysis differences between the RA-experiment and CA. The North American area has been averaged in the east-west direction from 125° to 70° West. The figure shows that the largest analysis impact is over the central part of North America from 45° to 55° North, in the upper troposphere, with a maximum rms difference of about 1.2 m/s at 300 hPa. A similar effect is found also for temperature (not shown). Analysis differences of order 1 m/s rms are also found in the eastern part of the North Atlantic, around 500 hPa (not shown), which likely is due to increased usage of aircraft data in the ascending and descending phase e.g. in the U.K. and Ireland. Only relatively small analysis differences are found over continental Europe (not shown).

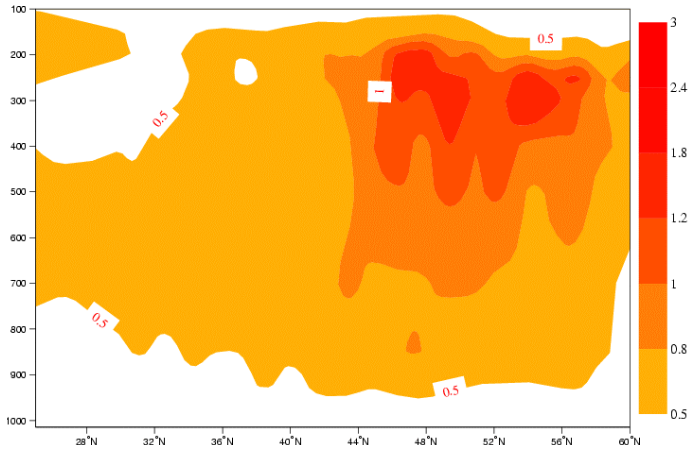

Figure 3 Cross section of the RMS of wind speed analysis difference between the RA-experiment (revised redundancy and thinning) and CA (the control), averaged in the longitudinal direction from 125° to 70° West, in the period from 20000615 to 20000711.

4.3 Validation of analyses against GADS data

To make a more detailed validation, the analyses and forecasts were compared with an independent data set, the Global Aircraft Data Set (GADS, Rukhovet *et al.* 1998). The comparisons were done in practice by including the GADS data in the experiments without being used, i.e. in a passive mode. This enables the calculation of background and analysis departures, which can then be compared between the experiment and the control. The GADS data set consists of wind and temperature measurements collected from the British Airways fleet of 747-400 aircraft since 1996. The GADS sampling is at 128 s intervals and the absolute instrument accuracy for wind is in the order of 1 m/s (Shapiro and Kennedy 1975). The data are not distributed via GTS but recorded on tapes.

On the 24th of June 2000 a long-haul westbound aircraft with a GADS recorder was reporting frequently as it crossed the jet stream between 00:10 and 00:40 UTC in the area between 49°-53° North and 99°-95° West. Figure 4 (left panel) shows the analysed wind speed interpolated to the aircraft locations at 238 hPa for the RA-experiment (green squares) and CA (blue triangles), which can be compared with the GADS wind speed observations (red circles). The GADS observations show several separate peaks in the jet streams that are not reproduced by either of the two analyses. These marked oscillations in wind speed may be the signatures of gravity waves, which are detectable in very high-frequency aircraft data (Tenenbaum 1996). The GADS wind speeds are significantly stronger than both analyses in this case, by up to a -22% discrepancy for the highest wind speeds in the jet core. The RA-experiment is in better agreement with the GADS observations than the control, as its analysed wind speed is higher by 1-2 m/s, in this case.

A winter case for the 18th of December 2000 at 00 UTC is shown similarly in the right-hand panel (Figure 4b). A GADS aircraft was flying westward at 228 hPa over the North Atlantic between 00:00 and 00:50 UTC as it crossed the jet stream. The figure shows the analysed wind speed interpolated to the aircraft position as before. The GADS data show two local peaks in the jet. These peaks are reproduced slightly better in the RA-assimilation than in CA, the control. We have seen in both cases that the analyses show a much smoother variation of wind speed over short distances than the GADS measurements do. This may be because the data from any one particular flight are likely to be affected by transient small scale features (such as gravity waves) or random observation errors, whereas the analyses provide a consensus of the information from several flights in the vicinity, as well as from the background. The analyses are therefore smoother than individual data. Also, the effective resolution of the analysis may be insufficient to represent the jet streaks accurately. The degree of smoothing in the analysis is primarily determined by the specified background error covariance matrix. Altogether, 20 cases were examined similarly to those shown in Figure 4. The RA analysed wind speed was in all cases in better or equal agreement with the GADS observations than CA. However, an underestimation of the jet maxima was often seen in both experiments (as in Figure 4b). Tenenbaum (1996) and Cardinali *et al.* (2002) have identified the underestimation of analysed jet streaks as a general problem, which is common to several NWP centres.

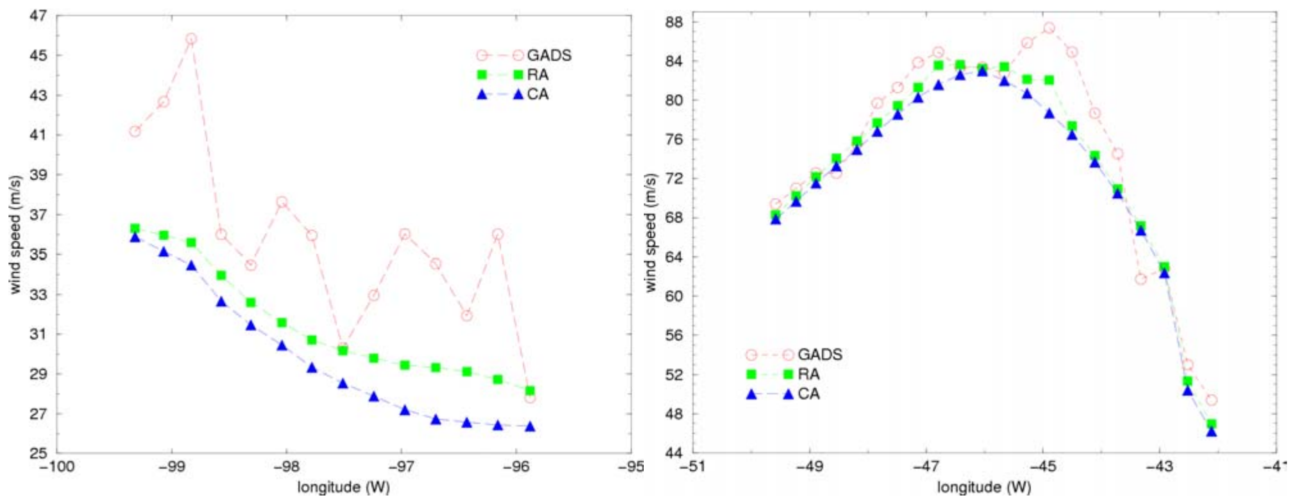

Figure 4 Wind speed versus longitude for two GADS-flights crossing a jet-core region over North America (left) between 00:10 and 00:40 UTC 24th of June 2002, and the North Atlantic (right) between 00:00 to 00:50 UTC, 18th of December 2000. The GADS measurements, which were not used by either assimilation, are shown in red (circles), the RA analysis in green (squares) and the CA analysis in blue (triangles).

4.4 Forecast impact of the revisions

Figure 5 shows a latitude-pressure cross section (averaged in the east-west direction) of the 24-hour geopotential height difference of rms forecast error between the RA and CA experiments over North America (left) and the North Atlantic (right). The figure shows that the experiment has smaller errors than the control (blue shading) in the upper troposphere, by up to 12 m. The area with larger errors (yellow shading) is over central Canada (48° – 58° N, 97° - 103° W) which is less well covered by aircraft routes. For the North Atlantic, the RA experiment shows an error reduction (by up to 10 m) in most parts. Over Europe, however, the 24-hour forecast impact of the revision is relatively small (not shown).

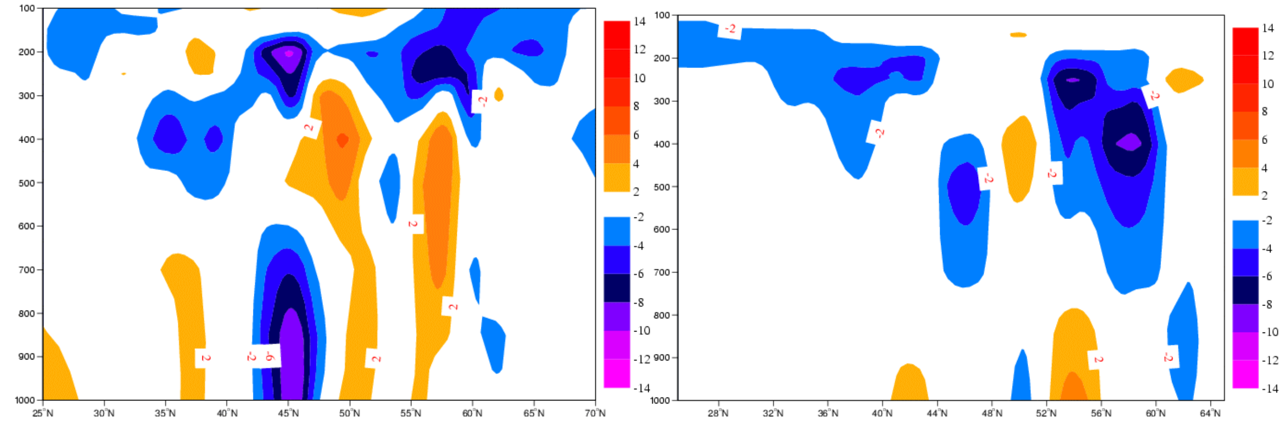

Figure 5 Latitude-pressure cross sections of the difference in rms of the 24-hour geopotential height (m) forecast error, between the RA-experiment (revised redundancy and thinning) and CA (the control). The cross sections are averaged in the longitudinal direction from 125° to 70° West over North America (left) and from 70° to 10° West over the North Atlantic (right hand panel). The contour interval is 2 m with shading starting at +- 2 m.

For the medium range forecast, Figure 6 shows the 200 hPa geopotential height anomaly correlation score (%) verified against operational analyses for the Northern Hemisphere (left) and the North Atlantic (right). The RA-experiment is shown in red and the control (CA) in blue. The verification is averaged over the summer and winter periods (31+18 cases). The positive result for the Northern Hemisphere at Day 5 is 98% significant. The positive result for the North Atlantic is 95% significant from Day 1 to Day 5. For North America 95% significant (but small) positive impact was found for Day 1, only. For Europe, the Tropics and the Southern Hemisphere (not shown) the forecast impact of the revised aircraft usage was found to be essentially neutral. Figure 7 shows a map of the difference in rms of Day-1 forecast error, between the RA and CA experiments, and Figure 8 shows Day-3. These maps indicate that the reduction in short-range forecast errors (green shading) over North America propagated into the western North Atlantic at Day-1, and further downstream into parts of Europe and the Arctic at Day-3. Each of the three main components of the revision (the horizontal thinning, the vertical thinning, and the redundancy checks) have been tested separately over shorter periods (not shown). Based on these results the revised aircraft usage was implemented in ECMWF's operational system in January 2002.

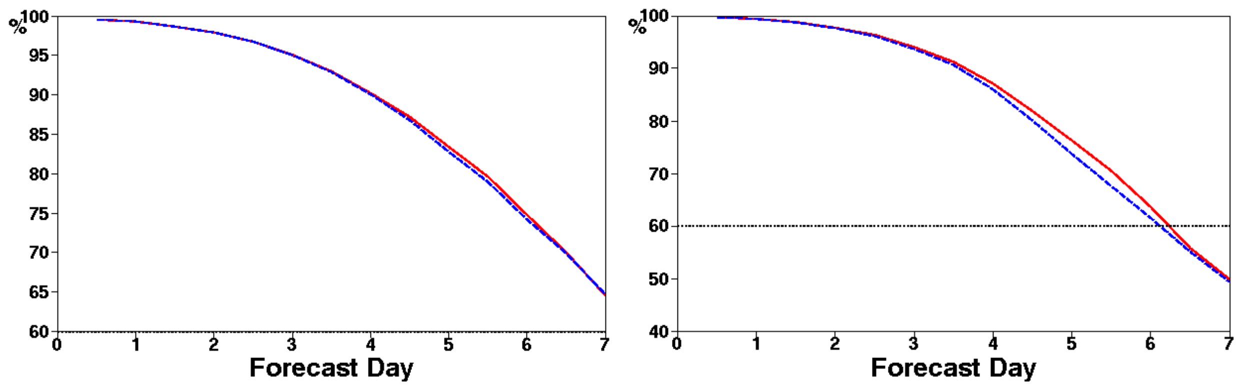

Figure 6 Anomaly correlation (%) for the 200 hPa geopotential height forecast error averaged over the two test periods (31 days in summer and 18 days winter). The verification is against operational analyses, for the Northern Hemisphere extra tropics (left) and the North Atlantic (right). The experiment with revised AIREP usage (RA) is shown in red, and the control (CA) in blue.

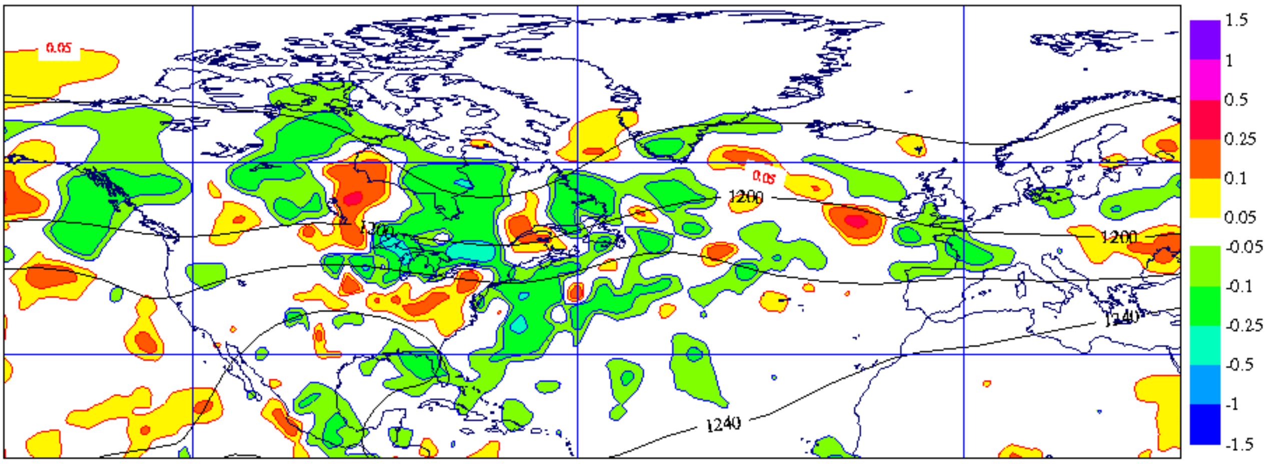

Figure 7 Difference in rms of 24-hour forecast error, between the RA- and CA-experiment, for 200 hPa geopotential height, 20000616-20000715, 12 UTC. Shading starts at +- 0.05 deca meters, with yellow (positive) shading indicating larger errors in RA than in CA. Area integrated values are: Europe –0.02 m, North Atlantic –0.23 m and North America –0.25 m.

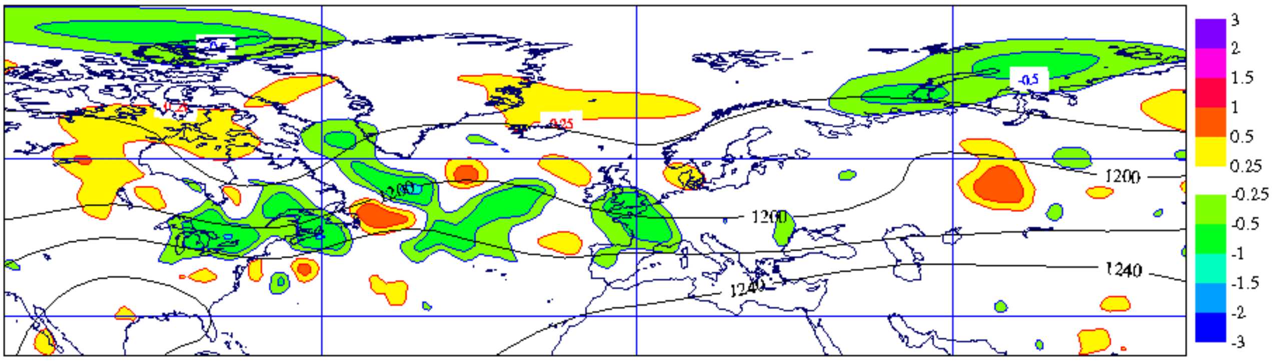

Figure 8 Like Figure 7, but for 72-hour forecasts. Area integrated values are: Europe –0.27 m, North Atlantic –0.76 m and North America 0.29 m.

The North Atlantic forecast impact of the revision has furthermore been verified against observations. The 3-day forecasts from the CA and RA experiments were verified against observations in the region (40° - 65° N, 70°-10° W). Data for a two-week period in June 2000 were used. The most significant results, as determined by a P-squared test on standard deviations, are presented in Table 1. The results show that RA experiment performed better than the control. In general the improvements are largest for the wind field with a near-neutral impact for the temperature field. The most significant results are seen for the validation against aircraft wind data at cruise level, but improvements are also seen in verification against other observing systems. The wind biases are also improved, but the temperature biases are neutral or larger (not shown). This could be an indication that the aircraft wind observations are more valuable than the temperature data, perhaps because the ATOVS radiance data control the temperature analysis.

Obs type	Variable	Pressure hPa	CA st.dev	RA st.dev	RA Count	Significance (%)
Aircraft	Wind	200	8.16 m/s	7.83 m/s	23494	99.95
Aircraft	Wind	300	8.37 m/s	8.18 m/s	28880	99.95
Radiosonde	Temperature	400	1.59 K	1.53 K	1670	88
Radiosonde	Wind	600	4.88 m/s	4.72 m/s	2084	85
SATOB	Win	200	4.71 m/s	4.56 m/s	2364	87
SATOB	Wind	300	4.16 m/s	4.11 m/s	14452	80
SATOB	Wind	500	2.95 m/s	2.82 m/s	932	84
SATOB	Wind	900	2.30 m/s	2.24 m/s	5094	94
DRIBU	MSLP	-	2.79 hPa	2.74 hPa	7618	83

Table 1 Statistics for CA and RA experiments of Day-3 forecasts verified against observations. Data from the North Atlantic region (40N-65N 70W-10W) in a two-week period, have been used. The statistical significance as given by a P-squared test on the standard deviations, is also shown.

5. Impact of profiling data from ascending/descending aircraft

An impact study in the form of a data denial experiment has been performed in order to test the analysis and forecast impact of profiling data from aircraft in America and Europe. It is of interest to test if 4D-Var extracts significant information from the aircraft data, which are irregularly distributed in space and time, given that in these areas there is good coverage of PILOTs, radiosondes and profilers. This study is one of several recommended by the WMO/CBS Expert Team on Data Requirements and the Redesign of the Global Observing System.

5.1 Definition of the impact study experiment

The data denial experiment was run for two one-month periods: January and July 2001. All aircraft data below 350 hPa were removed over North America (25° – 60° N, 120° – 75° W) and Europe (35° – 75° N, 12.5° W to 42.5° E). This resulted in approximately 13,000 less data (temperature, u and v wind components) being used in the experiment, per 12-hour data assimilation cycle. The experiment was compared with a control assimilation, which used aircraft data according to the revisions outlined above, i.e. like the earlier RA-experiment.

5.2 Analysis and forecast impact of the denial of profiling aircraft data

One aspect of the analysis impact is illustrated in Figure 9. It shows the difference in rms of analysis increments for 300 hPa geopotential height, for the winter period. Yellow (green) shading indicates larger (smaller) analysis increments in the experiment without profiling aircraft data. We show 300 hPa here because at levels above 350 hPa similar numbers of data have been used in both experiment and control. It may at first seem counter-intuitive that the increments are larger in the assimilation that uses less data. The explanation is that the exclusion of accurate and useful data can make the errors in the short-range forecasts larger. When a less accurate 12-hour forecast is used as background in the next assimilation cycle, observation minus background departures are larger, resulting in larger analysis increments. A similar behaviour with respect to the non-use of hourly surface data was reported by Järvinen *et al*. 1999. With this interpretation in mind we can conclude from Figure 9 that the denial of profiling aircraft data has had a detrimental effect on the assimilation over the Eastern United States extending into the Western parts of the North Atlantic storm track region.

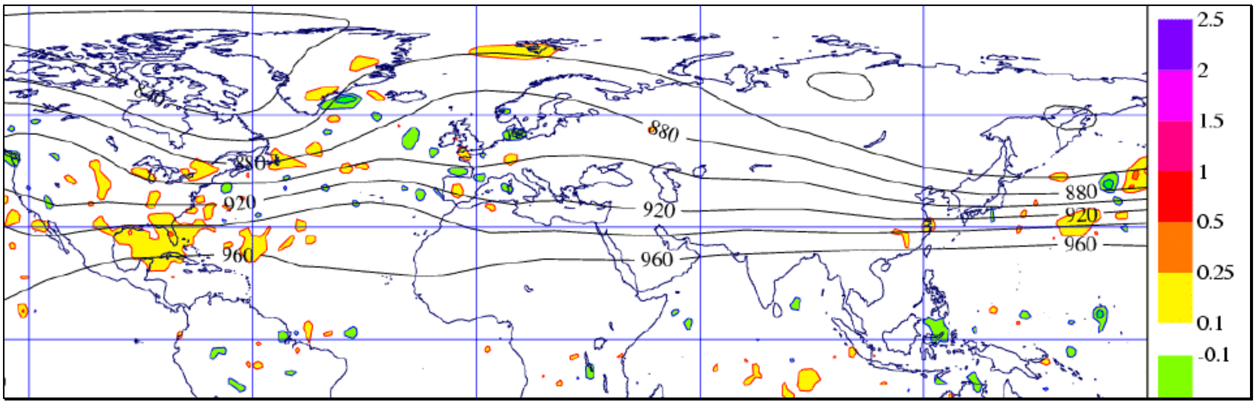

Figure 9 Difference in rms of analysis increments, between the data denial experiment and the control, for 300 hPa geopotential height, 20010102-20010131, 12 UTC. The shading starts at +- 0.1 deca meters, with yellow (positive) shading indicating larger analysis increments in the data denial experiment. Area integrated values are: Europe 0.0 m, North Atlantic 0.11 m, North America 0.33 m and the Northern Hemisphere extra tropics 0.15 m. The mean 300 hPa geopotential height analysis is contoured with an interval of 20 deca meters.

The deterioration at analysis time amplifies rapidly during the early stages of the forecasts in this experiment. Figure 10 shows the difference in rms of 48-hour forecast error, at 500 hPa. We can see that the deterioration due to the denial of the profiling aircraft data has amplified during the first two days of the forecasts, and spread to Northern Canada and parts of the Arctic. The impact over Europe and its surroundings is also negative on average, with large variations.

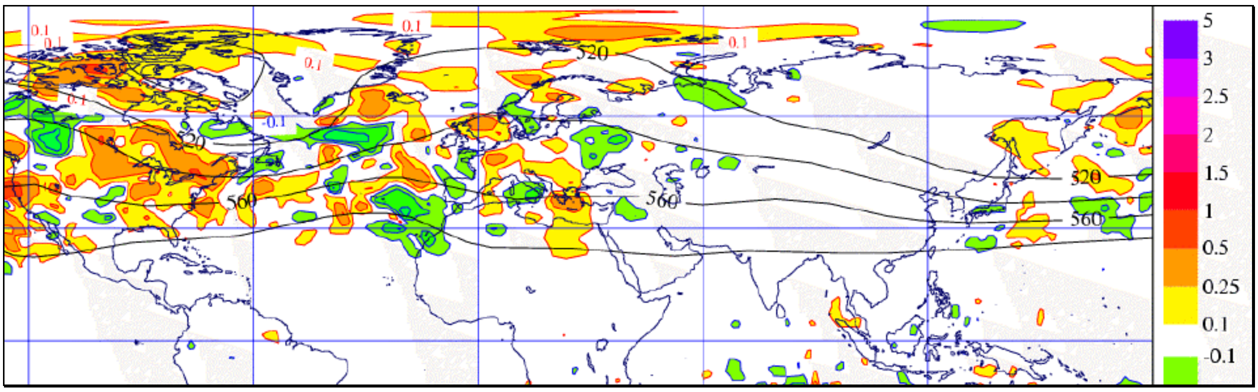

Figure 10 Difference in rms of 48-hour forecast error, between the data denial experiment and the control, for 500 hPa geopotential height, 20010102-20010131, 12 UTC. Otherwise like figure 6. Area integrated values are: Europe 0.20 m, North Atlantic -0.02 m, North America 0.89 m and the Northern Hemisphere extra tropics 0.30 m.

The impact remains significant also in the medium range, as seen in Figure 11. The impact has shifted predominantly down stream from North America, and to higher latitudes. There is a clear deterioration in the denial experiment, for the forecast performance in the northern parts of the North Atlantic, the British Isles, parts of Scandinavia and the Arctic. The precise location of the areas of forecast impact is likely to depend strongly on the synoptic situation during the test period.

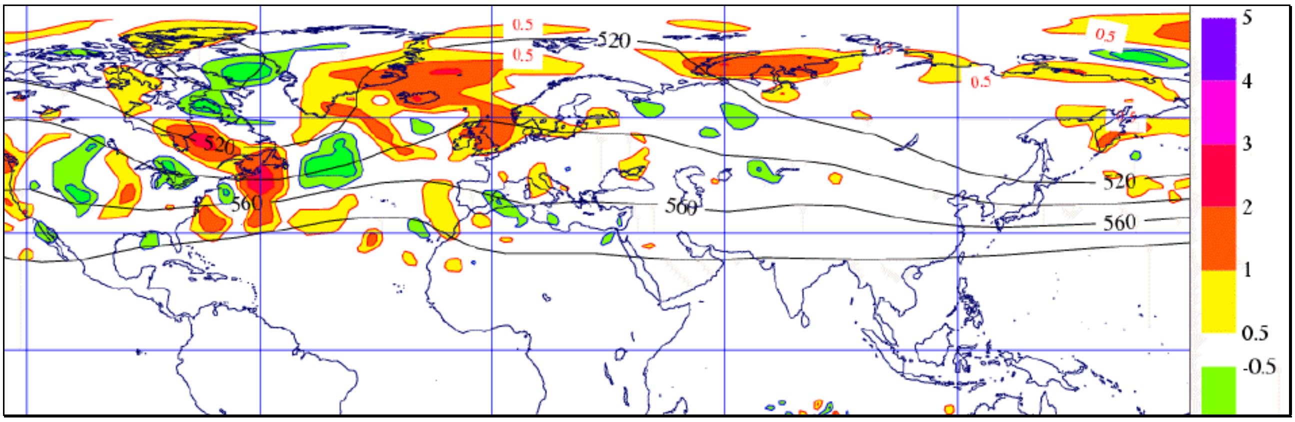

Figure 11 Difference in rms of 120-hour forecast error, between the data denial experiment and the control, for 500 hPa geopotential height, 20010102-20010131, 12 UTC. Otherwise like figure 6. Area integrated values are: Europe 3.02 m, North Atlantic 2.90 m, North America 0.31 m and the Northern Hemisphere extra tropics 1.35 m.

The forecast verification anomaly correlation scores for the Northern Hemisphere (Figure 12) shows a clear deterioration from day-4 onwards in winter, and from day-6 onwards in summer, due to the removal of data from ascending and descending aircraft. In winter, this is a larger impact than that due to the revisions tested in Section 4. In the revision-experiment we added data to what was already a fairly good coverage of aircraft data, whereas in this experiment we left two important regions totally without aircraft data below 350 hPa. The results are consistent in that most of the positive impact is seen in the eastern parts of North America in the short range, propagating eastwards and northwards into the North Atlantic, Europe and the Arctic in the medium range.

The significance of the forecast impact is shown more clearly in scatter diagrams, such as those shown in Figure 13, for the Northern Hemisphere (left) and Europe (right), for the two experiment periods. In the diagram for the Northern Hemisphere a majority of cases fall below the 45-degree line, indicating consistent positive forecast impact. A t-test on the statistical significance of the results gives 98 % for the Northern Hemisphere. For Europe the impact is essentially neutral over all, as most of the day-5 forecast impact appeared over the North Atlantic, the North Pacific and the Arctic.

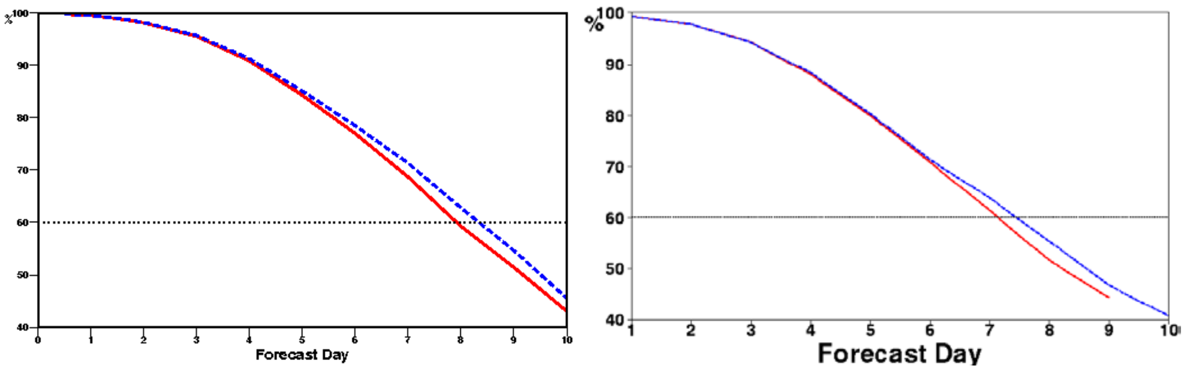

Figure 12 Northern Hemisphere anomaly correlation (%) of 500 hPa forecast error verified against operational analyses, each averaged over 31 cases, January 2001 (left) and July 2001 (right), 12 UTC. The data denial experiment is shown in red and the control is the blue dashed line.

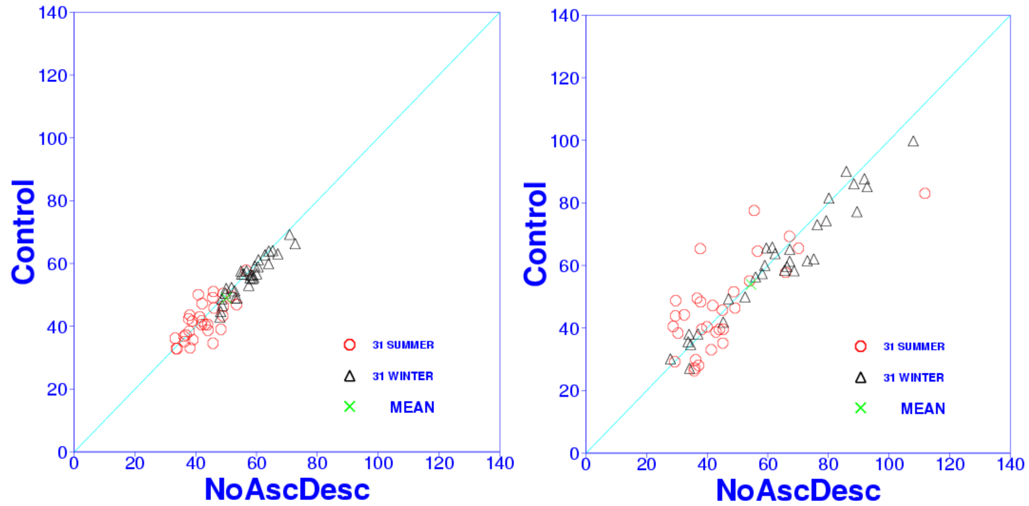

Figure 13 Scatter plot of the rms of 120-hour forecast error (m) for 500 hPa geopotential height. Each marker represents one day in the January period 2001 (black triangles) and July 2001 (red circles), for the Northern Hemisphere extra tropics (left) and Europe (right). The error in the forecast from 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 green x-marker. Markers plotted below the diagonal indicate larger error in the data denial experiment than in the control.

5.3 Consistency with American profilers and radiosondes.

The results so far indicate that the denial of profiling data from aircraft had a clear negative impact on data assimilation and forecast performance. Our experiment also provides statistics with respect to all other data types that were used in both experiments, and these can be compared. The departure statistics with respect to the American wind profilers (left) and the bias with respect to temperature profiles from the American radiosondes (right panel) are shown in Figure 14. The plots show the denial experiment in blue and the control in black. The panel on the left shows that the background is slightly worse in the denial experiment than in the control, between 300 and 400 hPa, at the wind profiler locations. The aircraft data have thus in the control assimilation made a positive contribution to the accuracy of the background (which is consistent with the results presented earlier in this section). The panel on the right, however, indicates that the temperature bias in the background is slightly

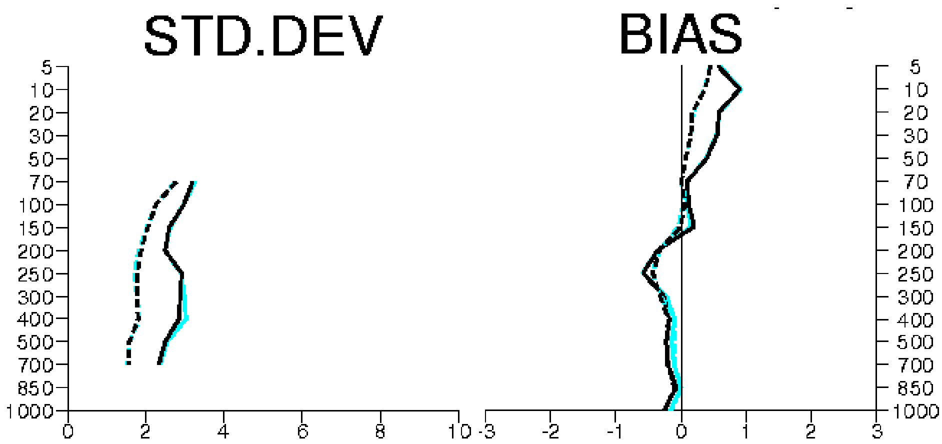

Figure 14 Background (solid) and analysis (dashed lines) departure statistics. Left: standard deviation for

wind speed (m/s) from American profilers. Right: bias for temperature data (K) from American radiosondes. The denial experiment is shown in blue and the control in black.

smaller in the denial experiment than in the control, below 300 hPa, at radiosonde locations. This result implies that there are some biases between temperature measurements from radiosondes and from ascending/descending aircraft. This is an issue that will require further detailed study in order to identify the source of the bias.

5.4 Occurrence of gross errors

The occurrence of gross errors as detected by BgQC and VarQC (described in Section 2.2) is shown in Table 2. This is for the data whose impact was tested in the data denial experiment, i.e. AIREP, ACARS and AMDAR wind and temperature data below 350 hPa, over the European and American areas (as defined above). We can see that gross errors are extremely rare for these data types – between 0.1 % and 0.7 % in total. It is curious that AMDAR have a ten times higher frequency of BgQC rejections for temperature (0.57 %) than for wind (0.05), whereas ACARS show nearly equal rejection rates for both: 0.33 for temperature and 0.30 for wind. The automated aircraft data (ACARS and AMDAR) have smaller random errors than the AIREPs, as mentioned earlier in the Introduction. The occurrence of gross errors (as detected by BgQC and VarQC) is nevertheless very low for AIREPs. It is as low or even lower than for the automated data. Gross error rejection rates for aircraft data at cruise level (not shown) are generally similar to those in Table 2.

	AIREP	ACARS	AMDAR
Data Count-T	20680	92942	24035
BgQC rejections-T	0.16 %	0.33 %	0.57 %
VarQC rejections-T	0.01 %	0.13 %	0.13 %
Data Count-wind	20600	91623	23873
BgQC rejections-wind	0.06 %	0.30 %	0.05 %
VarQC rejections-wind	0.05 %	0.12 %	0.05 %

Table 2 Data counts and occurrence of gross errors (%) as detected by BgQC and VarQC, respectively, for data below 350 hPa over the European and American areas (as defined for the data denial experiment), for the period 20010110-00 to 20010119-18 UTC. Vector wind and temperature are shown separately.

6. Conclusion

The usage of aircraft data in ECMWF's operational 4D-Var data assimilation system has been revised and expanded, through modification of the data selection procedures (horizontal and vertical thinning) and redundancy checks. The revised thinning method admits more of the profiling aircraft observations in ascending and descending phase by selecting up to one report per model level, per flight. The use of high frequency data at cruise level has also been enhanced by allowing data up to a minimum along-flight separation of 60 km. The increased usage of these (mostly automated) aircraft data was shown to result in clear positive impacts on analyses, short-range and medium-range forecasts. In terms of wind speed, the analysed jet core is better depicted as shown in comparisons with high quality independent observations from the non-real time Global Aircraft Data Set (containing data from the British Airways 747-400 fleet). These revisions to the aircraft data usage were implemented operationally at ECMWF in January 2002.

Following these revisions a data denial impact study was performed testing the impact in 4D-Var of profiles from American and European AMDARs and ACARS in ascending and descending phase (below 350 hPa) over two separate months: January and July 2001. This particular impact study was run partly on the request of the

WMO/CBS Expert Team on Data Requirements and the Redesign of the Global Observing System. The study addresses the question whether a modern data assimilation system (such as 4D-Var) obtains substantial benefit from the aircraft profiles, which sample very irregularly in space and time, given that America and Europe are relatively well covered by radiosondes and wind profilers. The study does not provide a comparison of the impact of profiling aircraft data with that of other profiling data, and we cannot comment on their relative importance or cost effectiveness based on the results presented here.

The results show a substantial positive impact of the profiling aircraft data on the analysis and forecast accuracy. As always in this type of study, the exact locations and magnitudes of the impact partly depend on the synoptic situation in the test period. The short-range forecast performance is improved over North America, the North Atlantic and Europe. In the medium-range a clear positive impact is found in the North Atlantic, the European (winter only) and Arctic areas. These results are statistically significant, and support the ongoing WMO initiative for further expansion of the AMDAR/ACARS coverage. The results also illustrate the effectiveness of 4D-Var with respect to observations that are irregularly distributed in space and time.

Acknowledgements

We thank Bruce Truscott and Jeff Stickland for informative discussions on the development of the AMDAR network, and Joel Tenenbaum for providing the GADS data. We are grateful to Anthony Hollingsworth, Adrian Simmons and Horst Böttger for comments on the manuscript.

References

Andersson, E. and H. Järvinen, 1999: Variational quality control. *Q. J. R. Meteorol. Soc*. **125,** 697-722.

Andersson, E., M. Fisher, R. Munro and A. McNally, 2000: Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Q. J. R. Meteorol. Soc*. **126,** 1455-1472.

Andersson, E. and A. Garcia-Mendez, 2002: Assessment of European wind profiler data, in an NWP context. ECMWF Tech. Memo. 372.

Benjamin, S.G., B.E. Schwartz and R.E. Cole, 1999: Accuracy of ACARS wind and temperature observations determined by collocation. *Weather and Forecasting*, **14,** 1032-1038.

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

Bouttier, F., 2001a: The use of profiler data at ECMWF. *Meteorologische Zeitschrift*, **10,** 497-510.

Cardinali, C., L. Rukhovets and J. Tenenbaum, 2002: Jet stream analysis and forecast errors using GADS observations in the DAO, ECMWF and NCEP models. 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.

Dalby, T.D., Berney A., 1999: Modification to aircraft thinning and observation errors. NWP Forecasting Research, UKMO Tech. Memo. 276.

Fisher, M. and E. Andersson, 2001: Developments in 4D-Var and Kalman Filtering. ECMWF Tech Memo 347.

Järvinen, H. and P. Undén, 1997: Observation screening and first guess quality control in the ECMWF 3D-Var data assimilation system. ECMWF Tech. Memo. 236.

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.

Köpken, C., 2001: Monitoring of METEOSAT WV radiances and solar stray light effects. EUMETSAT/ECMWF Fellowship programme, Research Report No 10.

Lorenc, A.C. and Hammon, P., 1988: Objective quality control of observations using Bayesian methods: theory and practical implementation. *Q. J. R. Meteorol. Soc.,* **114,** 515-543.

Mahfouf, J-F., 1999: Influence of physical processes on the tangent linear approximation. *Tellus*, **51A,** 147-166.

Mahfouf, J-F. and F. Rabier, 2000: The ECMWF operational implementation of four dimensional variational assimilation. Part II: Experimental results with improved physics. *Q. J. R. Meteorol. Soc.*. **126,** 1171-1190.

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.

Rukhovet, L., J. Tenenbaum and M. Geller, 1998: The impact of additional Aircraft data on the Goddard Earth Observing System Analyses. *Mon. Wea. Rev.,* **126,** 2927-2941.

Shapiro, M. A. and P. J. Kennedy, 1975: Aircraft measurements of wave motion within frontal zone system. *Mon. Wea. Rev.,* **103,** 1050-1054.

Teixeira, J., 1999: The impact of increased boundary layer vertical resolution on the ECMWF forecast system. ECMWF Tech. Memo. 268.

Tenenbaum, J, 1996: Jet stream winds: Comparisons of aircraft observations with analyses. *Weather and Forecasting*, **11,** 188-197.

Untch, A., J.-J. Morcrette and M. Hortal, 1996: The ECMWF model with increased resolution in the stratosphere. *Research activities in atmosphere and ocean modelling.* Ed. A. Staniforth.

Untch, A., A. Simmons, M. Hortal, C. Jakob and colleagues, 1998: Increased stratospheric resolution in the ECMWF forecasting system. Proc. workshop on chemical data assimilation. KNMI.

WMO, 1996: Guide to meteorological instruments and methods of observation. Sixth Edition. WMO-No.8. Geneva, Switzerland.