TECHNICAL MEMORANDUM

244

Observing system experiments with the 3D Var assimilation system

P. Undén, G. Kelly, D. Le Meur
and L. Isaksen

Research Department

December 1997

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

Observing System Experiments with the 3D-Var Assimilation System

Per Undén, Graeme Kelly, Didier Le Meur and Lars Isaksen

1. Introduction

The cost of making observations is by far the largest single cost in the whole process of making a forecast. The cost–efficiency of the composite observing system is uppermost in the minds of those concerned with funding and operating the World Weather Watch. The requirements of a composite observing system vary with the forecast range, forecast area, and forecast weather element. The requirements also vary with the required accuracy for each forecast element. For example, precipitation forecasts are less accurate than pressure forecasts, because of modelling problems and a lack of data on moisture and clouds. Equally, requirements for wind forecast accuracy are more stringent for ocean wave– forecasting than for general marine forecasting, because a wave model is very sensitive to wind forcing. Here we assess through observing system experiments (OSEs) the contribution made by the main ground–based and satellite based operational systems to medium range forecasting. OSEs are data assimilation experiments where one can assess the impact of an operational observing system by deleting its observations from the operational network and then running extended data assimilation and regular forecasts with the reduced system to assess the contribution of the deleted system to the total operational system. Equally OSEs can be used to assess the value of a new or experimental observing system by running extended data assimilation and regular forecasts with and without the new system, to assess its overall impact on forecast and analysis skill.

Given the vast cost of observations it behoves NWP centres to make effective use of observations. Kelly et al. (1993) performed a series of observing system experiments in 1991 with the then-operational ECMWF system based on optimal interpolation (OI, Lorenc, 1981). They used a baseline observational system consisting of in–situ data only and comprising radiosondes, aireps, synops, ships and buoys. They used OSEs to assess the impact of adding SATEMs only, SATOBs only, and SATEMs plus SATOBs to the baseline system. The satellite observing systems were not additive in terms of their forecast impact. Northern Hemisphere forecast scores were best with SATOBs without using SATEMs. Also SATEMs without using SATOBs was slightly better than using both SATEMs and SATOBs in the assimilation. These unsatisfactory results triggered a critical appraisal of the use of data in the OI system. The experiments revealed problems with the use of NESDIS TOVS satellite thicknesses in the Northern Hemisphere (Andersson et al. 1991, Kelly et al. 1991) and as a consequence TOVS data were removed from the Northern Hemisphere and tropical troposphere until the introduction of 1D–VAR in 1992 (Eyre et al.,1993).

Furthermore, it helped to motivate the development a 3D-Var system which could make a direct analysis of TOVS sounding radiances together with all other data. The purpose of the present paper is to repeat and extend the earlier experiments, but this time using the 3D–Var assimilation system, which became operational in early 1996. The present results are much more satisfactory than the earlier results, and demonstrate that the main observing systems are contributing in important ways to medium range forecasts in the Northern Hemisphere, in the tropics, and in the Southern Hemisphere. Equally the results are a validation of the performance of the 3D-Var system, and show that in broad terms it is working as intended in all three areas.

OSEs are normally performed globally at ECMWF in order to study global impacts of observing systems as well as of the data assimilation system itself. With a global medium range forecasting system it becomes natural to study the effects on the global scales since impact spreads over large parts of the hemispheres during medium range forecasts. Such observing system experiments have been carried out recently for the satellite observing systems and for some of the conventional systems. Impact of scatterometer data has been studied separately in both 3D- and 4D-Var.

Another set of experiments has been performed to study impacts of the extra FASTEX data. The enhancement of the North Atlantic radiosonde network during the period was considerable and probably to the highest level of data coverage ever for that area. In addition the impact of the dropsondes has been studied separately from that of the land- and ship-based radiosondes. First, a long period was assimilated with the dropsondes and the impact gauged. Then, five particularly interesting cases were analysed with targeted dropsondes assimilated over only one or two analysis cycles. This was in order to enable the local direct impact of targeted dropsondes to be determined.

The interpretation of OSEs is not always straight–forward. The value of an observing system is most easily shown when energetic events occur during the test period, and when only one observing system sees the event (Uppala et al. 1985). Thus an observing system's value is easier to demonstrate by adding it to a minimal base–line system, than by deleting it from a maximal base–line system. The stand–alone value of new system may be masked if there are overlaps in the observations. This can happen directly if two systems observe the same variables in the same area. It can also happen indirectly if a multi–variate data assimilation system uses multi–variate relations from observations of other variables in the same area, or if a 4D–Var data assimilation uses the time– history of other observations in the same area or upstream of the area of interest. Equally it is possible that observations critical for one purpose or area may be redundant for another purpose or area. For these reasons we have assessed synoptic aspects of the results as well as assessing forecast scores.

2. Global observing system experiments

Two series of Observing System Experiments (OSEs) were run for periods in December 1996 and February 1997. OSEs are designed to study the forecast impact of different observation types and detect any problems with their combination. The first set used the ECMWF data assimilation and forecasting system with the operational version as of December 1996 (Andersson et al., 1994, 1996). The period chosen was 19961205 00 UTC until 19961219 12UTC and ten-day forecasts were run from 12 UTC every day. Thus there were 15 forecasts in this set. Soon after this time a number of revisions of the 3D-Var analysis were made. A completely new background constraint (J_b) formulation was introduced operationally on 15 May 1997 (Bouttier et al.,1997). A second set of experiments, with the same configurations of observing systems, was run using the data assimilation system as it became operational in May 1997. This second period was from 19970201 00 UTC until 19970214 12 UTC with ten-day forecasts from 12 UTC every day. Combining these 14 new forecasts with the first set of 15 showed very similar average forecast impact of the different observing compared with the first period on its own. Therefore both periods have been combined into one 29-case sample of ten-day forecasts in order to enhance the significance of score averages and distributions.

2.1 Satellite based systems

The first configurations of the OSEs to be described here involve the satellite based systems.

2.1.1 TOVS and SATOB data

The control assimilation used both SATEM retrieved thicknesses in the stratosphere, TOVS radiances in the troposphere and SATOB cloud and water vapour winds in addition to conventional in-situ observations and scatterometer data from ERS-2. Three experiments were then run where a) both TOVS and SATOBs were removed b) TOVS removed and c) SATOBs removed. This allows us to see the impact of TOVS on their own, TOVS without SATOBs, SATOBs on their own, SATOBs without TOVS and TOVS and SATOBs together. Fig. 1 shows geopotential anomaly correlations at 500 hPa for the two hemispheres in the extratropics. The scores are averaged over all the 29 cases. In the Northern

Hemisphere there is a clear (but small) degradation without any of the TOVS or SATOBs data. This is an important finding, as there is often the suggestion that it is not possible to show the impact of the satellite data in the Northern Hemisphere. The SATOBs have a positive impact in the Northern Hemisphere. This impact must come mostly from the tropics, where there are plenty of SATOBs in addition to conventional data and scatterometer data from ERS-2. Withdrawing the TOVS additionally gives a slight further deterioration. Just running without the TOVS shows almost neutral impact in the Northern Hemisphere. The fact that withdrawing TOVS when SATOBs are not used shows a degradation indicates that TOVS only has a positive impact in the Northern Hemisphere in the absence of SATOBs.

The impact of the TOVS is striking in the Southern Hemisphere; its main impact being up to 1 1/2 days in the medium range. With TOVS used there is little impact of the SATOBs and without them the SATOBs do have a small but noticeable effect in terms of forecast score improvement.

The effect on the scores of withholding both satellite systems is significant in both hemispheres, as shown in Fig. 2 for 500 hPa heights. In the Northern Hemisphere the differences in forecast scores are small, but the distribution is well clustered on the lower score side for the NOSAT configuration. Two forecasts are better from the NOSAT assimilation; all the rest are either better or very close to neutral. In the Southern Hemisphere there are no better forecasts without satellites.

The tropics have been verified in terms of RMS errors of wind vectors at 850 and 200 hPa. The satellite systems show a very large impact on the tropical scores (Fig. 3). (In the beginning of the forecasts the discrepancies are however exaggerated by the fact that the operational analysis, which used all the data, was used for verification). Nevertheless, in the medium range, withholding of TOVS shows a considerable degradation compared with the control. Additionally removing the SATOBs shows an ever larger degradation. When TOVS are used the withdrawal of SATOBs has a small negative impact in the medium range. In the shorter range SATOBs have a similar impact to TOVS. Without SATOBs, the removal of TOVS shows a large degradation, especially at 200 hPa.

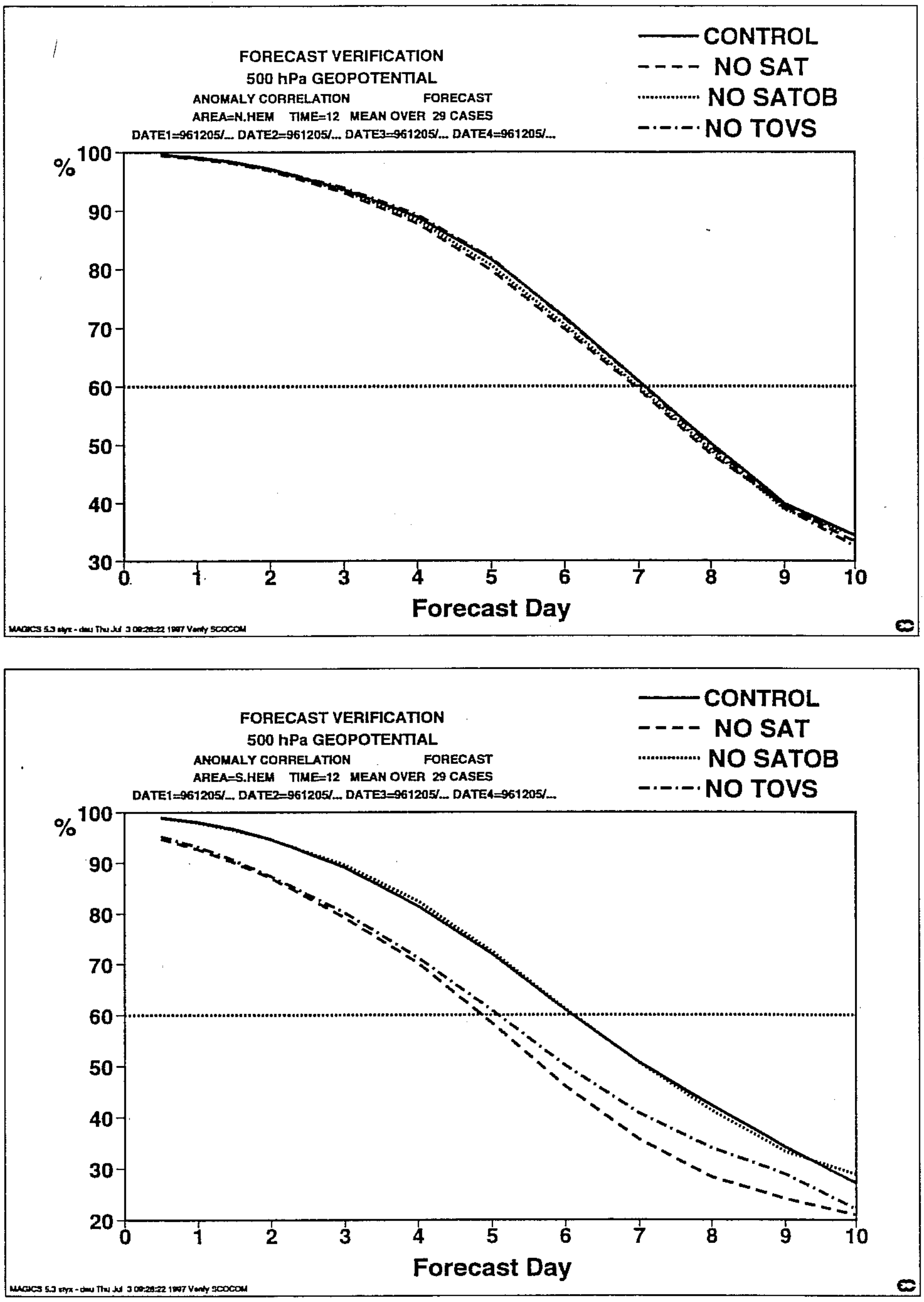


Figure 1. Mean 500 hPa geopotential forecast anomaly correlations in the Northern and Southern Hemisphere extratropics

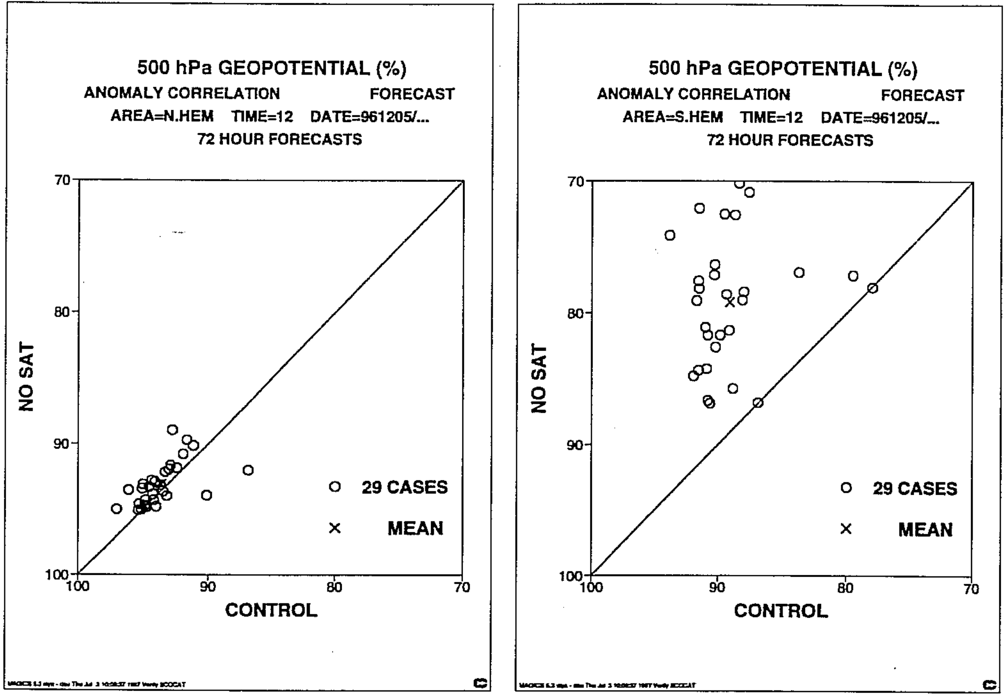
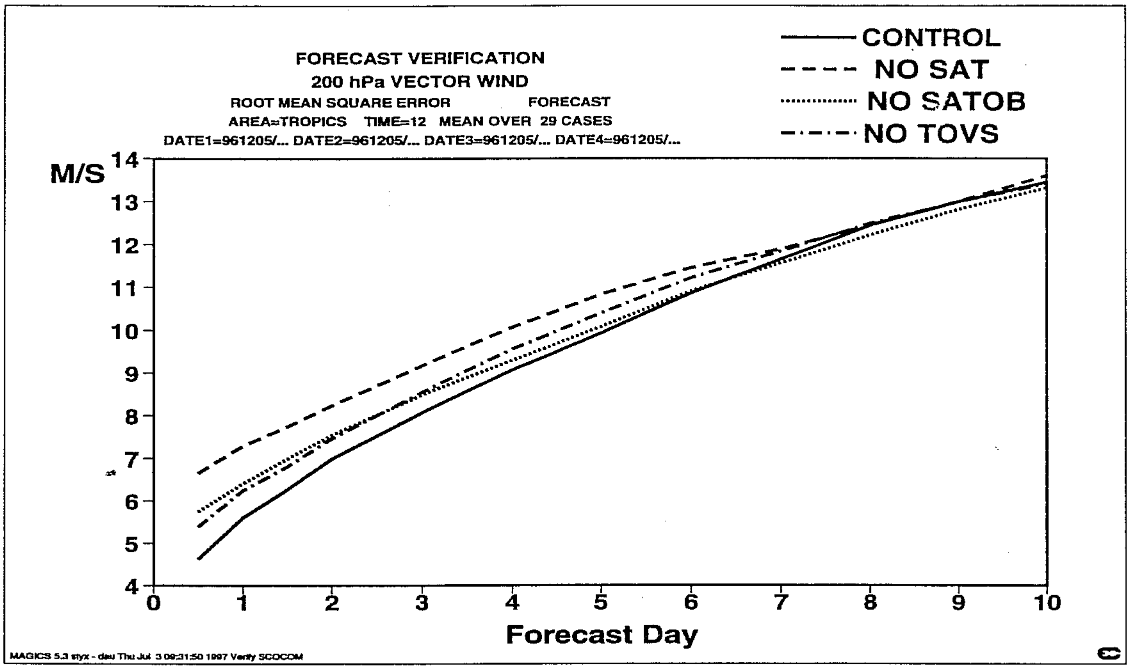


Figure 2. Scatter diagrams of anomaly correlations between forecasts without satellite data (TOVS and SATOBs) and control at 500 hPa and 72 hours forecast range for the Northern and Southern Hemispheres.



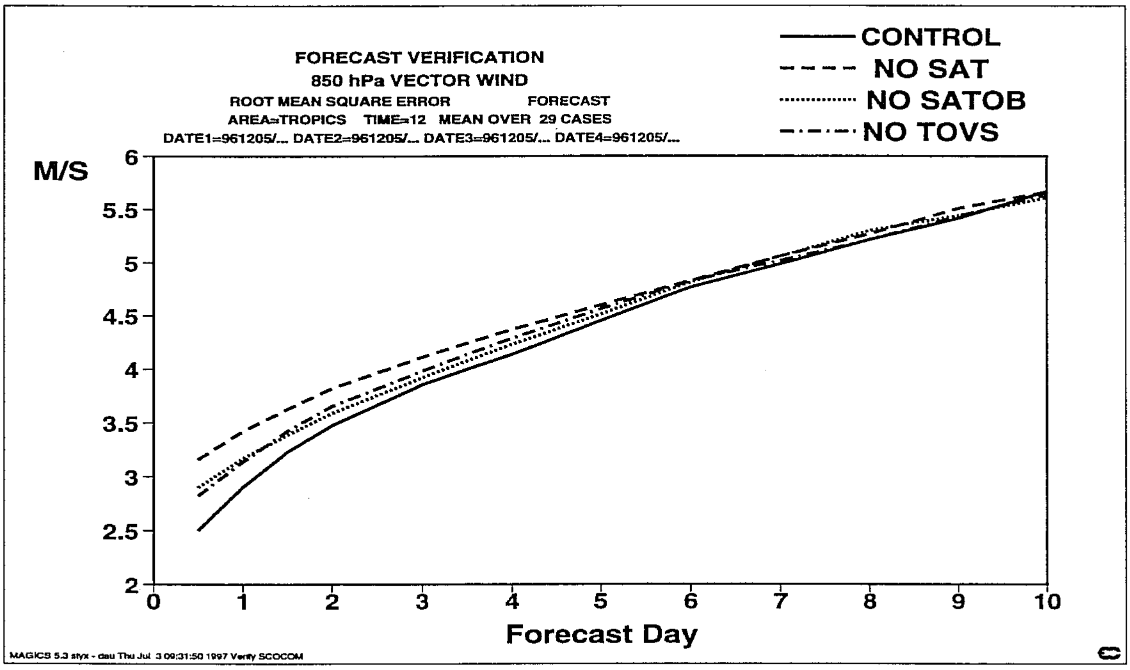


Figure 3. Mean 200 and 850 hPa wind forecast RMS errors in the tropical belt 20° N to 20° S.

2.1.2 Scatterometer data

In order to study the impact of ERS-1 scatterometer data in 3D-Var and 4D-Var a set of assimilations was performed for the last two weeks of January 1996. Reference runs were done with IFS CY15R7 (“old J_b”) using four updates for the 4D-Var assimilation. The comparison assimilations were performed using similar configurations except for using scatterometer data.

The results shows that scatterometer data improve the scores for the Northern Hemisphere in both 3D-Var and 4D-Var but that it is most pronounced in 4D-Var. The improvements are largest near the surface but extend throughout the troposphere in 4D-Var. The positive impact in 4D-Var comes from better assimilation of some low pressure systems in the North Pacific. The analysis improvements are maintained by the forecast model leading to better medium-range forecasts for the test period investigated.

For the Southern Hemisphere the use of scatterometer data has a neutral impact on the scores both in 3D-Var and 4D-Var. We believe this is due to less intense weather systems during the Southern Hemisphere summer period.

In addition to the previous studies, the tandem operations of the ERS-1 and ERS-2 scatterometers in April-May 1996 were taken as an opportunity to investigate the impact of the increased data coverage provided by using them together. For that, four parallel assimilation experiments were performed in 3D-Var over the first week of April, using either no scatterometer data (NOSCAT), ERS-1 or ERS-2 data only (ERS1, ERS2), and both ERS-1 and ERS-2 data (ERS1+2). The same version of the ECMWF model as above was used in all cases (CY15R7).

The study was first focused on the impact on the surface wind analysis, comparing the departures between first guess and observations for each experiment. Table 1 shows the average results obtained taking the observations from both scatterometers as a common reference. The vector RMS difference between first guess and observations is reduced respectively by 10 cm/s and 8 cm/s using the ERS-1 and ERS-2 instruments separately, and by 15 cm/s using them together.

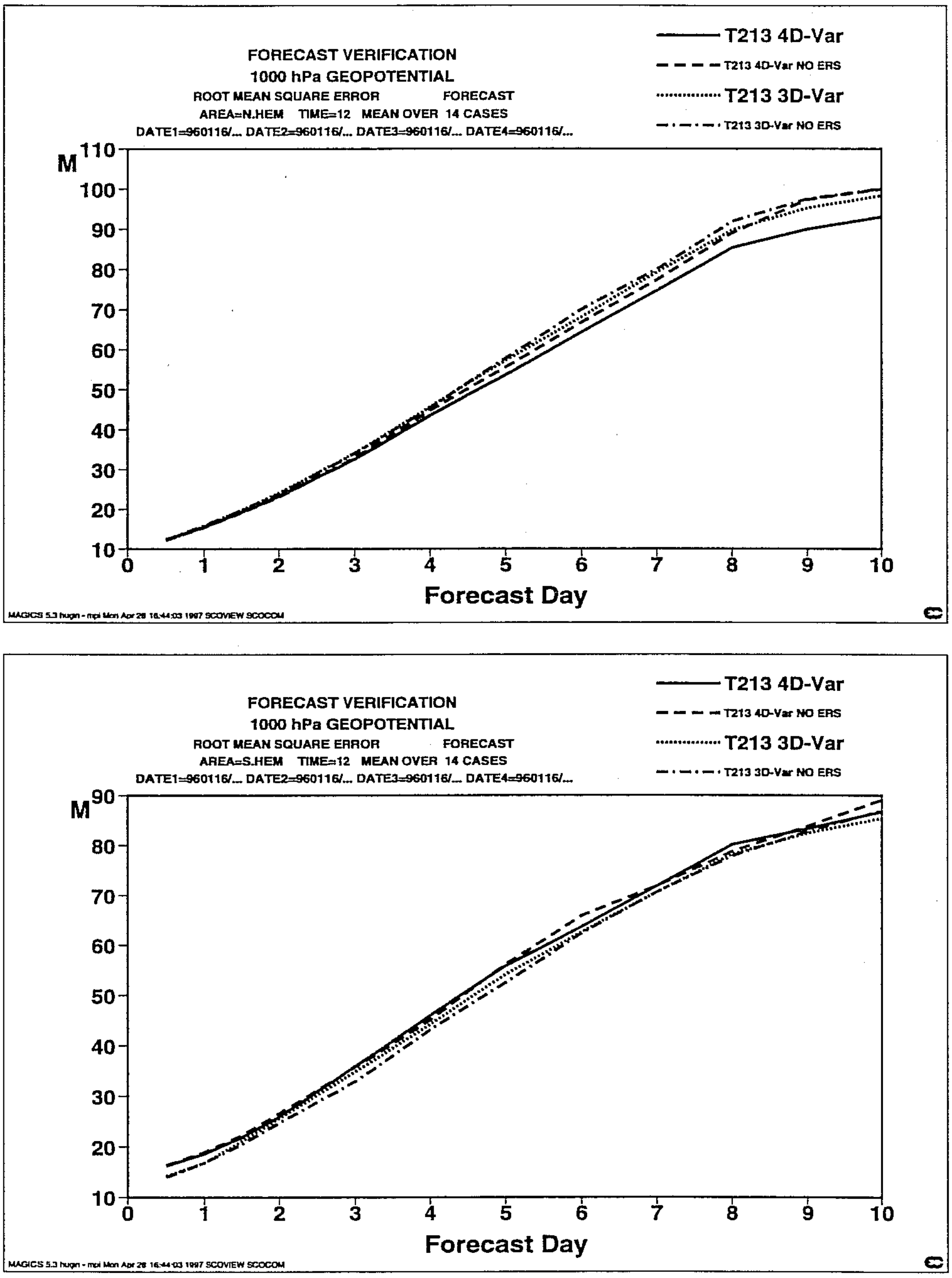


Figure 4. 1000 hPa height forecast RMS errors for Northern and Southern from 4D-Var and 3D-Var assimilations with and without scatterometer data.

Experiment	vector RMS (m/s)	difference wrt NOSCAT
NOSCAT	3.40	/
ERS1	3.30	-0.10
ERS2	3.32	-0.08
ERS1+2	3.25	-0.15

Table 1. RMS departures between first guess and scatterometer observations obtained using no scatterometer data (NOSCAT), ERS-1 or ERS-2 data only (ERS1, ERS2) and both ERS-1 and ERS-2 data.

These average results were then studied further by limiting the comparisons with the NOSCAT experiment to the cases with most significant impact, through the application of a minimum "impact threshold" on the vector difference between the ERS1, ERS2 or ERS1+2 first guess and the NOSCAT first guess at each scatterometer observation. Figure 5 shows the vector RMS first guess minus observation differences thus obtained for the ERS1 and NOSCAT experiments as a function of the "impact thereshold" in the assimilation using ERS-1 data. Very similar RMS reductions were observed for comparable impact thresholds when evaluating the ERS2 and ERS1+2 cases in the same way. The only significant difference was in the number of data exceeding a given impact threshold, which in the tandem assimilation case was roughly the sum of its values in both single assimilation cases. A good complementarity was thus demonstrated between both scatterometers, their separate benefits being juxtaposed without particular overlap when using them together. This was not obvious, since a redundancy could on the contrary be expected due to the fact that the ground tracks of ERS-1 and ERS-2 were following each other with a 24-h delay during their tandem operations.

A similar complementarity was found again when extending the study to the ten-day forecasts. Here the separate improvements obtained in each single assimilation case tended to be systematically added in terms of anomaly correlation scores in the tandem configuration in most of the verification areas. The results were however more clearly positive in the Northern Hemisphere, where a kind of synergy could even be noticed on average, the ERS1+2 experiment exhibiting a gain of 6 hours in the reliability of the forecast around day 7, whereas both ERS1 and ERS2 on their own have a nearly neutral overall impact (Fig. 6).

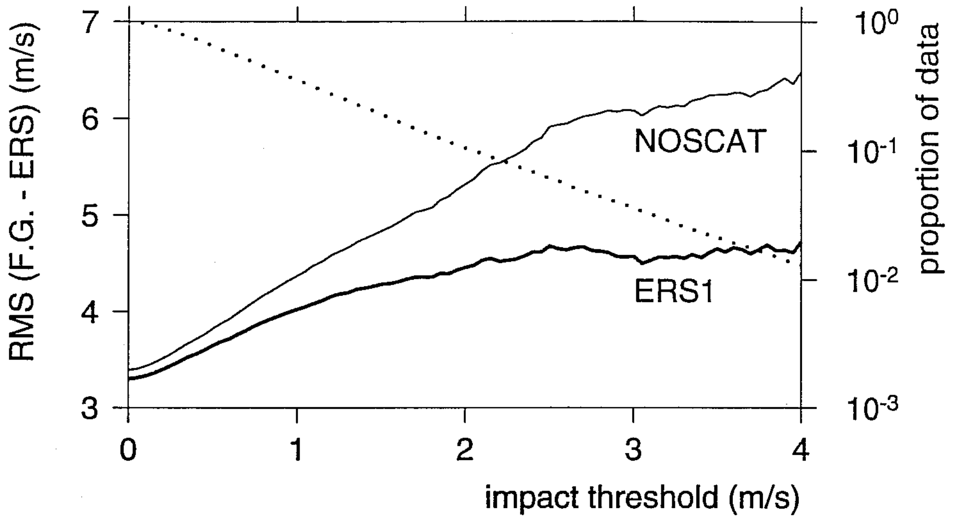


Figure 5. RMS differences between first guess and scatterometer observations using no scatterometer data (NOSCAT) and ERS-1 data only (ERS1) as a function of the "impact threshold" in the assimilation using ERS-1 data. The dotted line indicates the proportion of data exceeding a given threshold in logarithmic scale.

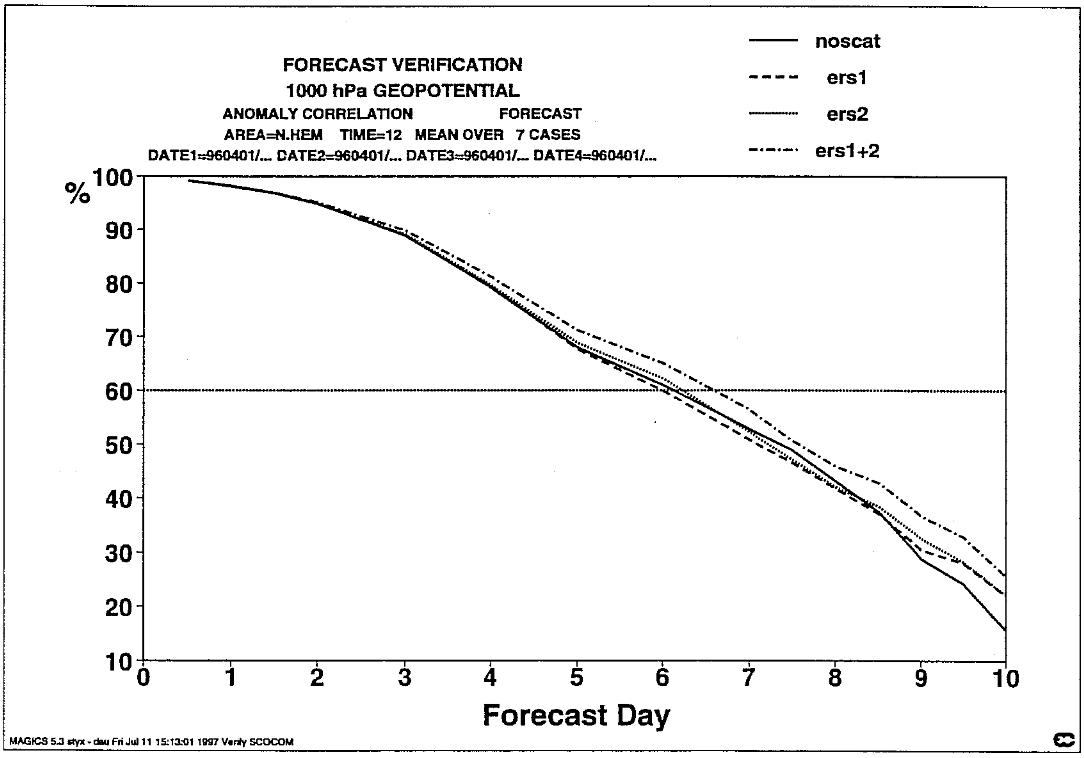


Figure 6. Mean 1000 hPa geopotential forecast anomaly correlations obtained over the Northern Hemisphere using no scatterometer data (NOSCAT), ERS-1 or ERS-2 data only (ERS1, ERS2) and both ERS-1 and ERS-2 data (ERS1+2).

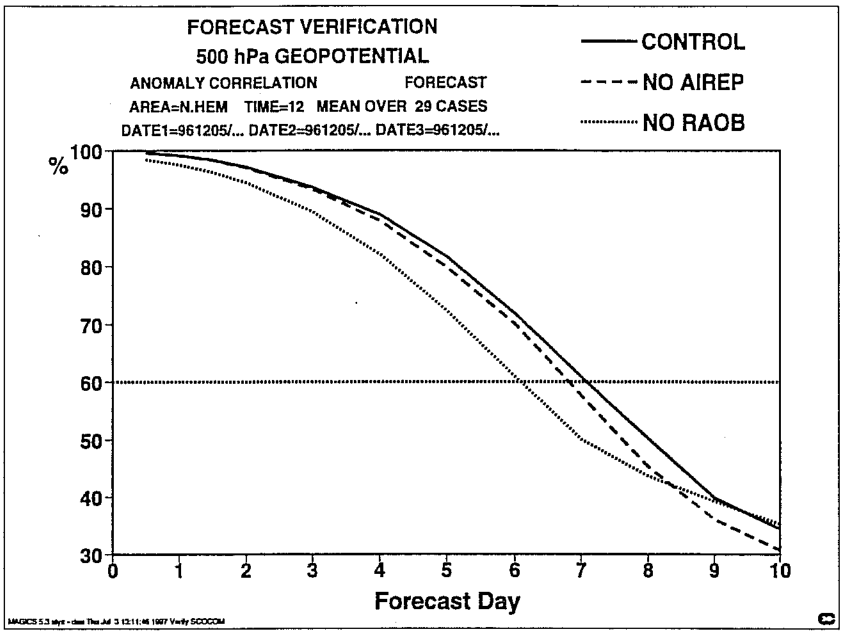
2.2 In-situ observing systems

The second group of experiments involve the impacts of some of the in-situ observing systems. One pair of assimilations was done without TEMP/PILOTs to show the impact of the upper air network globally. The other assimilations were done without AIREP data (all kinds of aircraft observations) and can be compared with the same control as above, which used all the observations.

The impact of the radiosondes and PILOTs is very large in the Northern Hemisphere (Fig. 7a). In the Southern Hemisphere the scarcity of these data compared with the satellite data shows very small impact of the radiosondes in the medium range (Fig. 7b). The radiosondes do however show a slight positive impact in the day 1 -3 range of the Southern Hemisphere forecasts. After day 6 there is even some degradation when radiosondes have been used, but it is at a range when the Southern Hemisphere forecasts are of low quality anyway and when a large sample of forecasts may be needed for reliable results. It may however indicate that some large scale information is not used correctly from the sparse network of radiosondes in the Southern Hemisphere and there may be a problem in the data assimilation system itself as well as with the quality of the observations.

Aircraft observations show a clear positive impact on the Northern Hemisphere forecasts. Their positive impact is somewhat larger at 200 hPa, but still much less then the radiosonde impact. For the different areas, the North Pacific shows the largest sensitivity to radiosondes (Fig. 8a). All the other areas in the Northern Hemisphere also show a significant and large positive impact of radiosondes in the medium range. The aircraft data show particularly large positive impact over North America (Fig. 8b), but still less than the radiosondes do. Over the North Atlantic and Europe (Fig. 9a and b) their impact is also positive but is, especially over Europe, small compared with the radiosonde impact.

Tropical wind scores are mainly affected by the radiosondes, where they have a large positive impact. Aircraft data show a much smaller positive impact in this area (Fig. 10), probably due to a the relatively low number of data in the tropics.



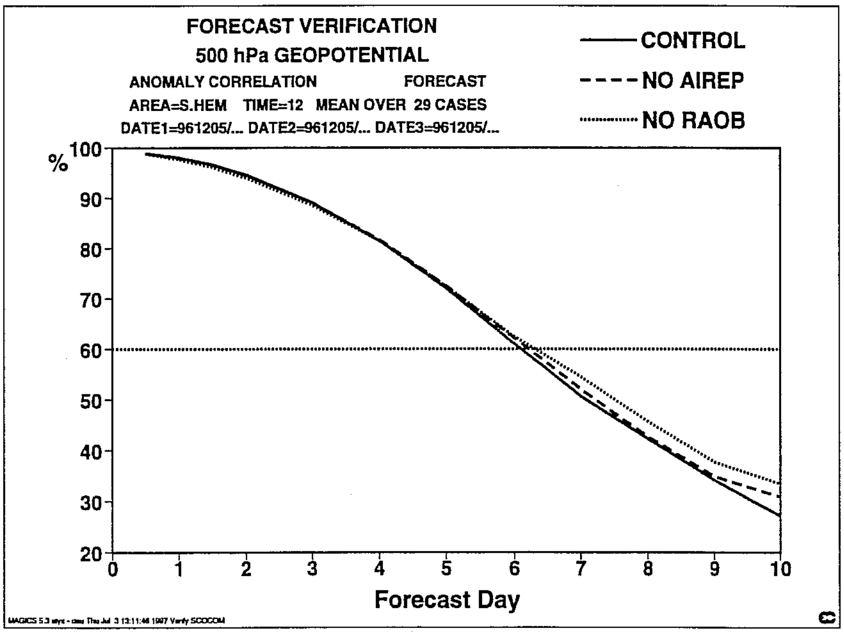
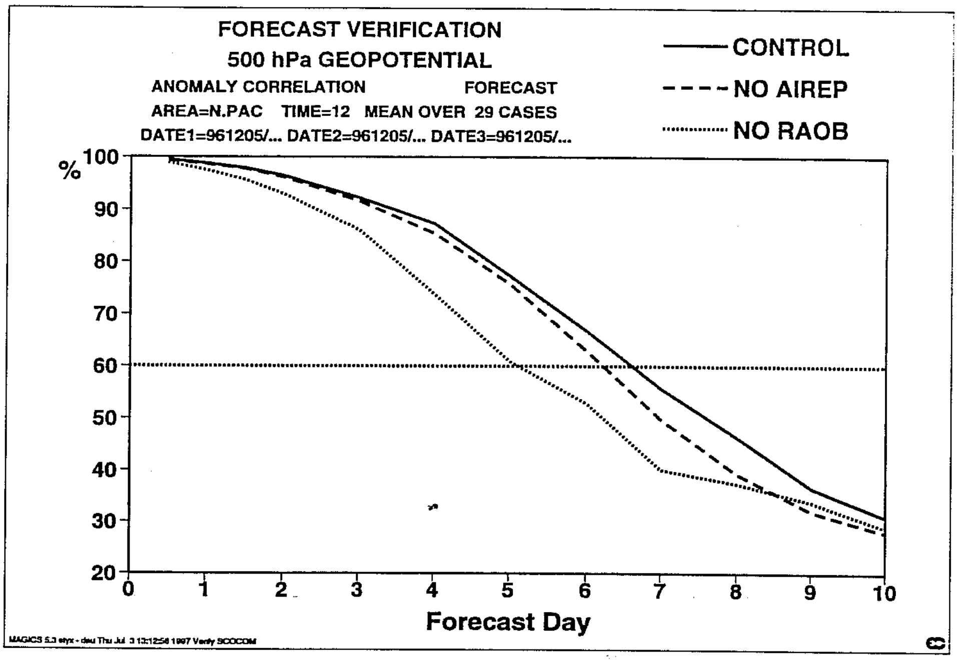


Figure 7. Mean 500 hPa geopotential forecast anomaly correlations for the Northern Hemisphere (top) and Southern Hemisphere (bottom).



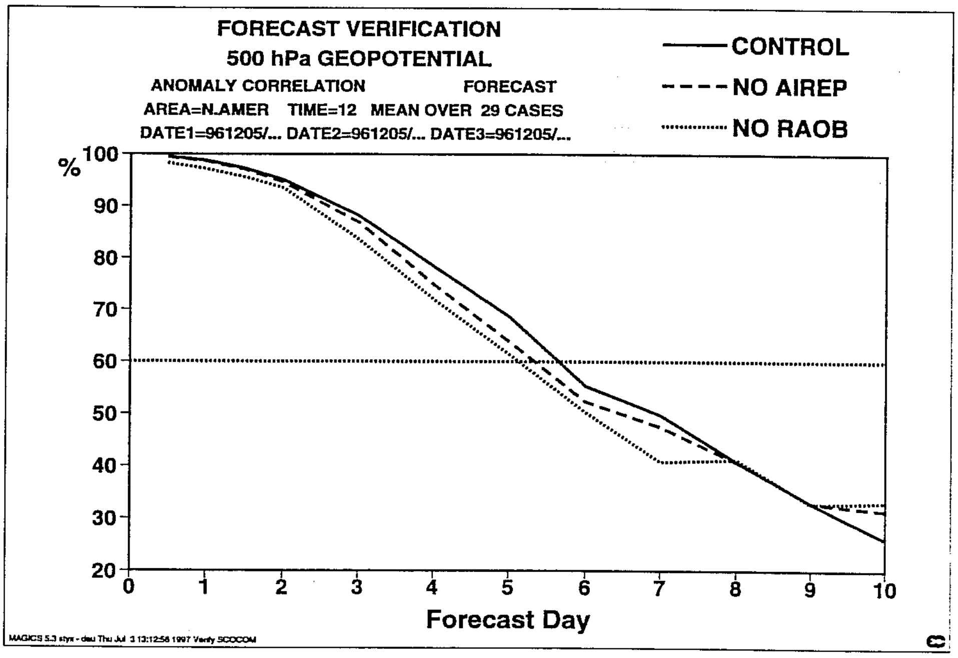


Figure 8. Mean 500 hPa geopotential forecast anomaly correlations for North Pacific (top) and North America (bottom).

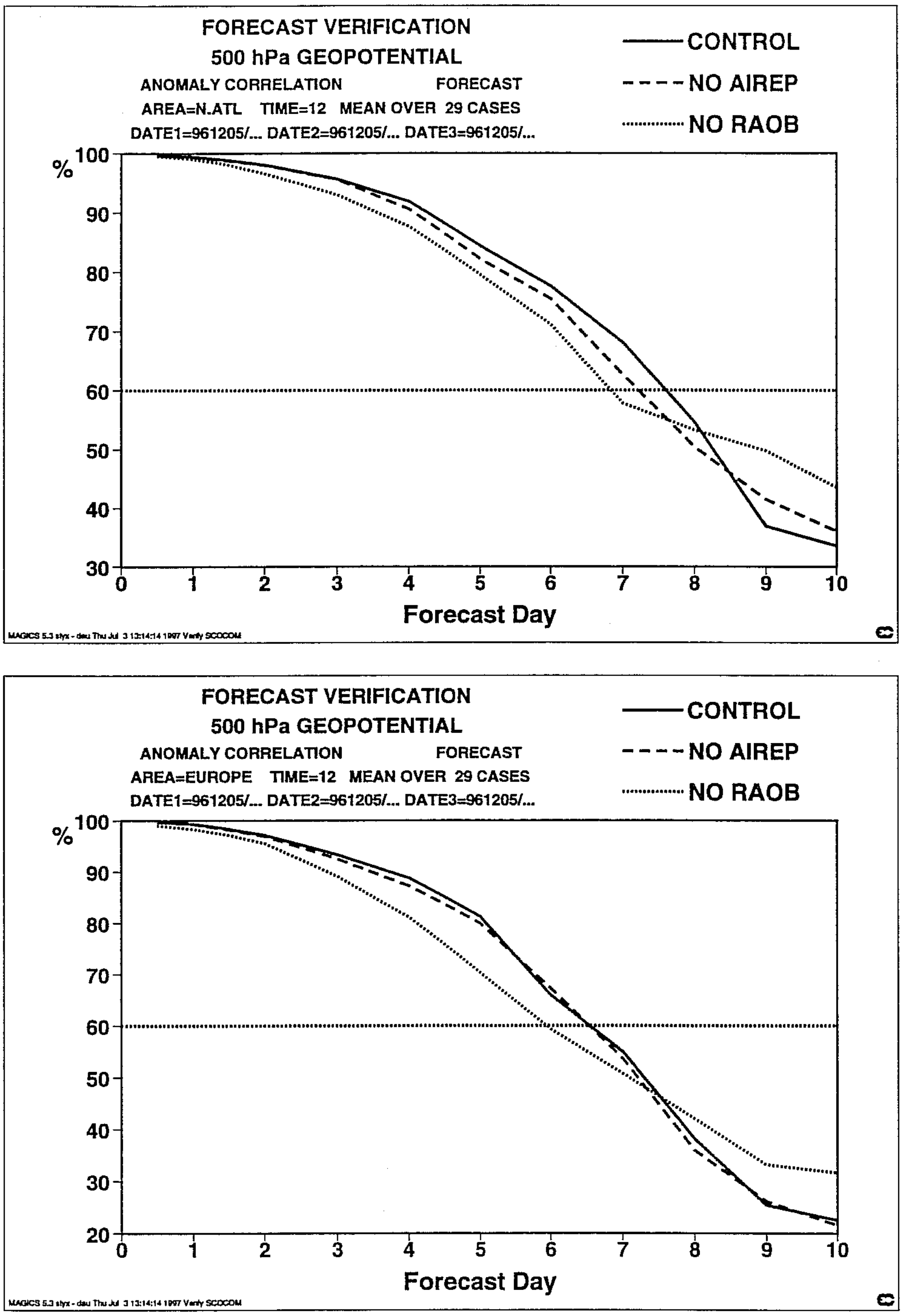
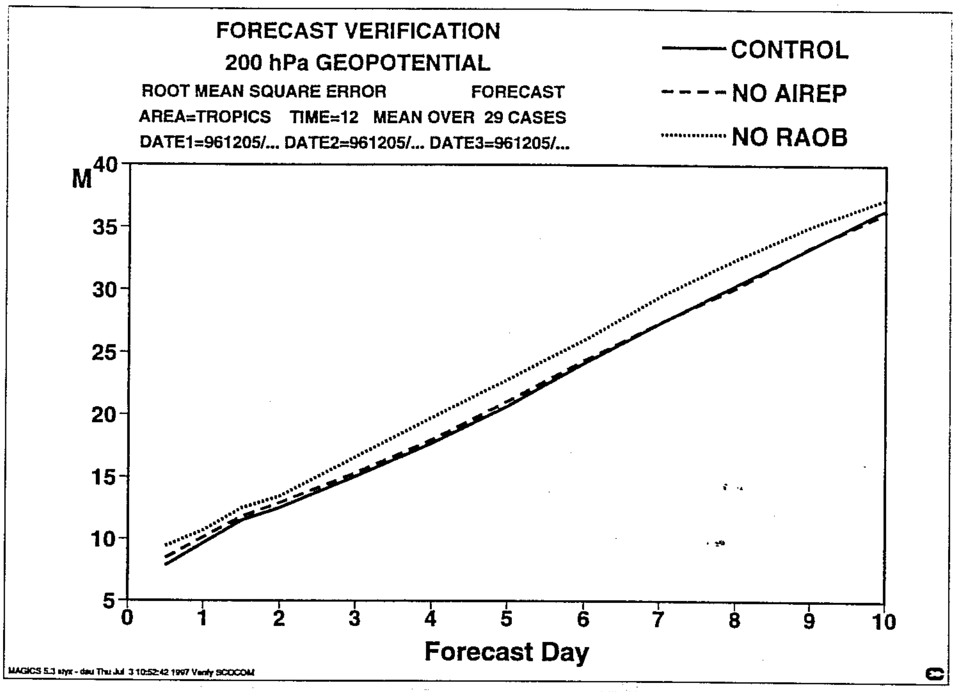


Figure 9. Mean 500 hPa geopotential forecast anomaly correlations for North Atlantic (top) and Europe (bottom).



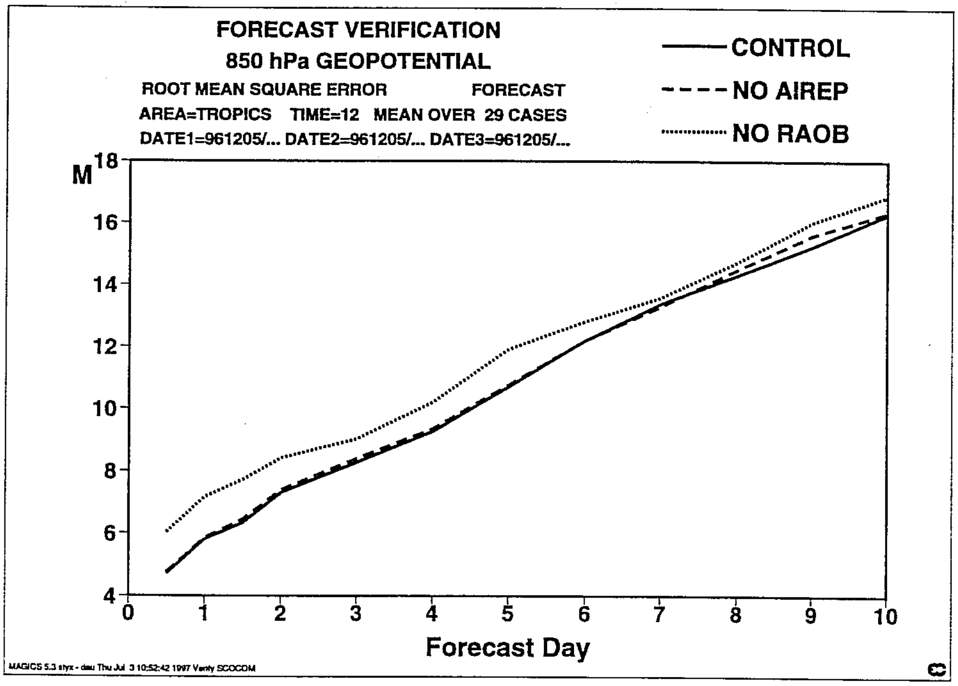


Figure 10. Mean 200 and 850 hPa wind RMS forecast errors in the tropics.

2.3 Synoptic impacts

Impacts from not using radiosondes are very large in the Northern Hemisphere, as seen from the forecast scores. The largest impacts are over the Pacific and large degradations from not using the radiosondes are very evident in almost all synoptic cases. There are very large position errors of major cyclones, troughs and ridges with large phase errors (sometimes completely out of phase) and minor lows which are completely missing or spurious.

Over the North Atlantic and Europe the results of withholding radiosondes are less dramatic but still clearly visible in most synoptic cases. It is more a question of different strengths of the cyclones and errors in their structure and phase errors in troughs and ridges rather than complete misrepresentations. Impacts of not using satellite data are smaller but still important. Such a case is shown in Fig. 11 for four-day forecasts from 19970210 12 UTC. The verifying mean sea level pressure analysis can be compared with the forecast not using radiosondes in the assimilation (NORAOB, b), the control forecast from assimilation with all data (c) and the forecast from the no satellite assimilation (NOSAT, d). Large errors in the position and depth of the cyclone east of Newfoundland can be seen in both the NORAOB and NOSAT forecasts. The NOSAT forecast develops the cyclone much too weakly and has a large south-westerly position error. Also the NORAOB forecast has a westerly position error whereas the control managed to capture both the depth and position of this system better.

Another region in which the NORAOB experiment is further degraded compared with the others is in the complex low pressure system extending from the mid-Atlantic through Scotland to east of the Baltic. The structure of this system is very different from the control or NOSAT experiments. It has a much deeper central part over southern Norway with a strong pressure gradient east of Iceland. The Icelandic part of the low is almost lost. The flow over Scandinavia is very much in error.

The low south of Ireland has only a small position error in the control. The NORAOB forecast has a large south-westerly position error whereas the NOSAT experiment has failed to develop the system, just showing a trough in the area.

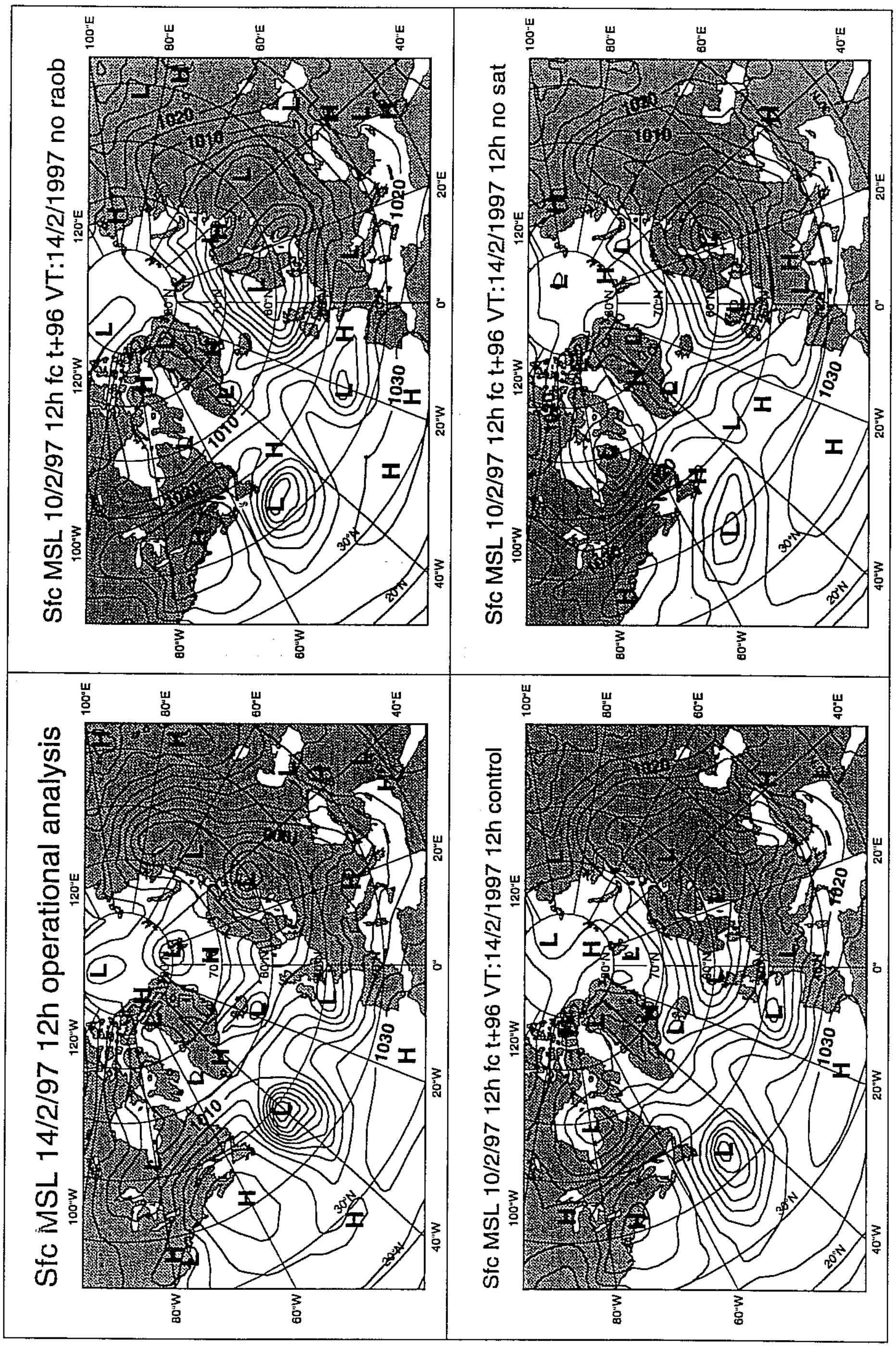


Figure 11. Verifying operational analysis 19970214 12 UTC (a) and 4 day forecast from 19970210 12 UTC from assimilations without radiosondes (b), control (c) and without satellite data (TOVS and SATOBs) (d). Shows mean sea level pressure with isolines for every 5 hPa.

3. Fastex experiments

During the FASTEX experiment the upper air network was considerably enhanced around and over the North Atlantic. 29 normal land stations plus 3 ASAPs in the area increased their launch frequency from twice a day to four times per day. Additionally there were 4 extra FASTEX ASAPs launched four times per day. Thus there was a very good coverage at 06 and 18 UTC during FASTEX in the North Atlantic area. In addition, on a large number of occasions dropsondes were launched from special flights mainly in targeted areas. Hence the combined observation coverage in the area was greatly enriched.

This period gives us a unique opportunity to test the impact of such an enhanced observing system. In fact two experiments have been done. The first measured the impact of the enhanced radiosonde network (but without dropsondes) compared with a reduced, normal network. The second experiment tested the impact of the dropsondes by using them in addition to the enhanced radiosonde network. Furthermore, five individual cases have also been reanalysed using targeted dropsondes for only one or two analysis cycles each.

3.1 Enhanced radiosonde network

The radiosonde network enhancement experiment was run from 1997-01-27 00 UTC until 1997-02-22 12 UTC. In practice the **operational** data assimilation used the enhanced network (without dropsondes) and the experimental assimilation blacklisted (withheld) all the extra radiosonde stations and ASAP ships at the relevant times of the day, when they were supplemental to the normal network. A complete list of the stations that were withheld can be found in Appendix 1. For stations that reported more frequently than every 6 hours, only the closest one to the analysis time was used.

The assimilations have been monitored in terms of observation minus background departures, numbers of used observations, analysis differences and forecast differences. The operational assimilation used more TEMP data than the reduced (= normal) network assimilation did. The observation fits were however almost identical in the two assimilations. Analysis differences between the two assimilations were of course introduced at 06 and 18 UTC near the positions of the extra observations. These differences were then propagated in the short range forecasts, but 6 hours later, at 00 or 12 UTC, when the two assimilations used the same network except for the extra 4 ASAPs, the analysis differences were normally small due to the similarity of the networks. There were however a few occasions when analysis differences could propagate beyond this point and could be traced a bit further in time.

Ten-day forecasts were run from 12 UTC for each day of the experiment. Forecast verification scores were computed and compared with the ECMWF operational ones. Fig. 12a shows that the impact on the Northern Hemisphere average scores is very slight but there is a small improvement over Europe in the scores until forecast day 5 (Fig. 12b). Figure 13 contains scatter plots of the individual scores and shows that at the five-day range there is small positive impact of the enhanced network. There is a fair deal of scatter, so the significance of the improvent is probably small. Similar average positive improvements in forecast scores are however seen for all verification regions in the Northern Hemisphere.

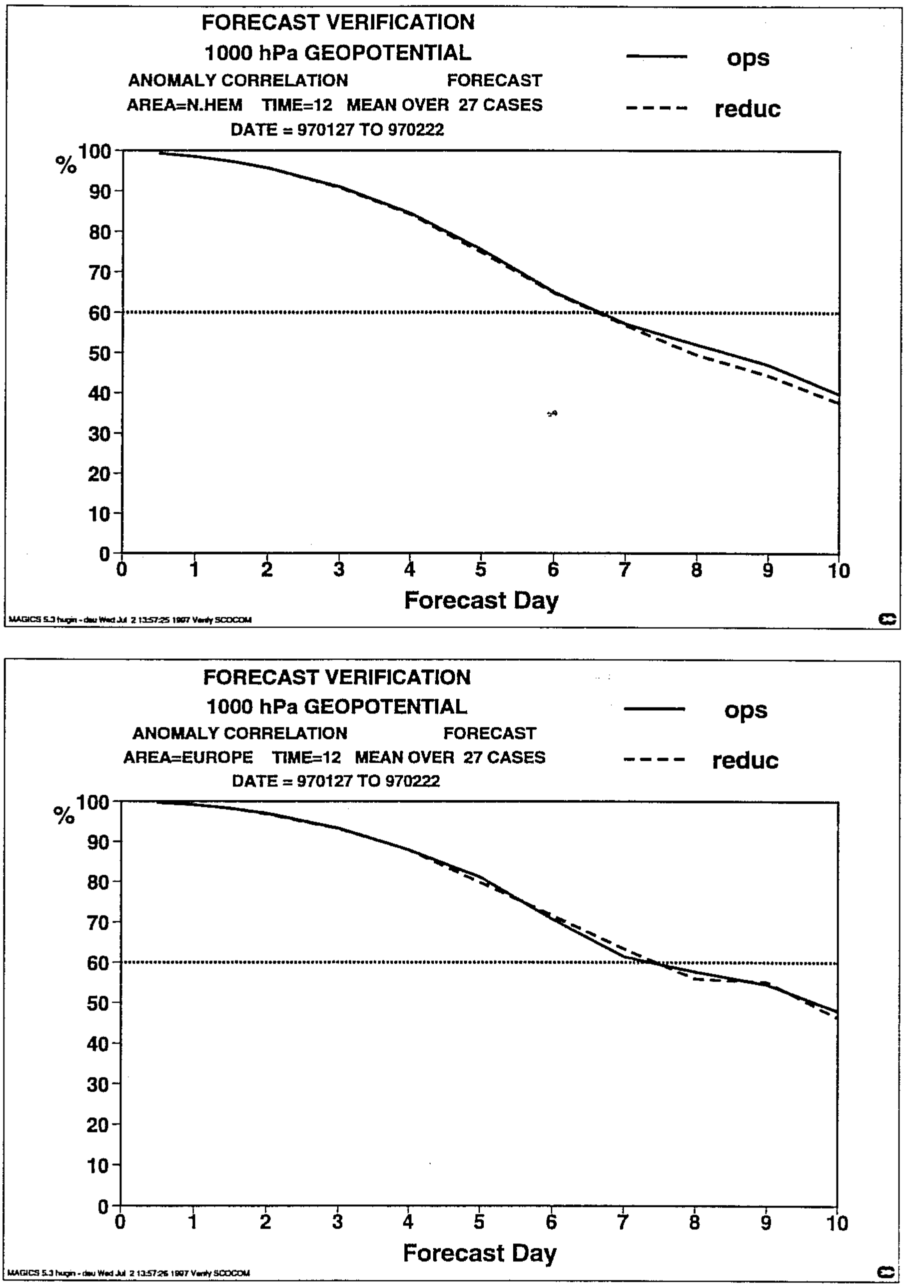


Figure 12. Mean of 1000 hPa geopotential forecast anomaly correlations in the Northern Hemisphere and Europe.

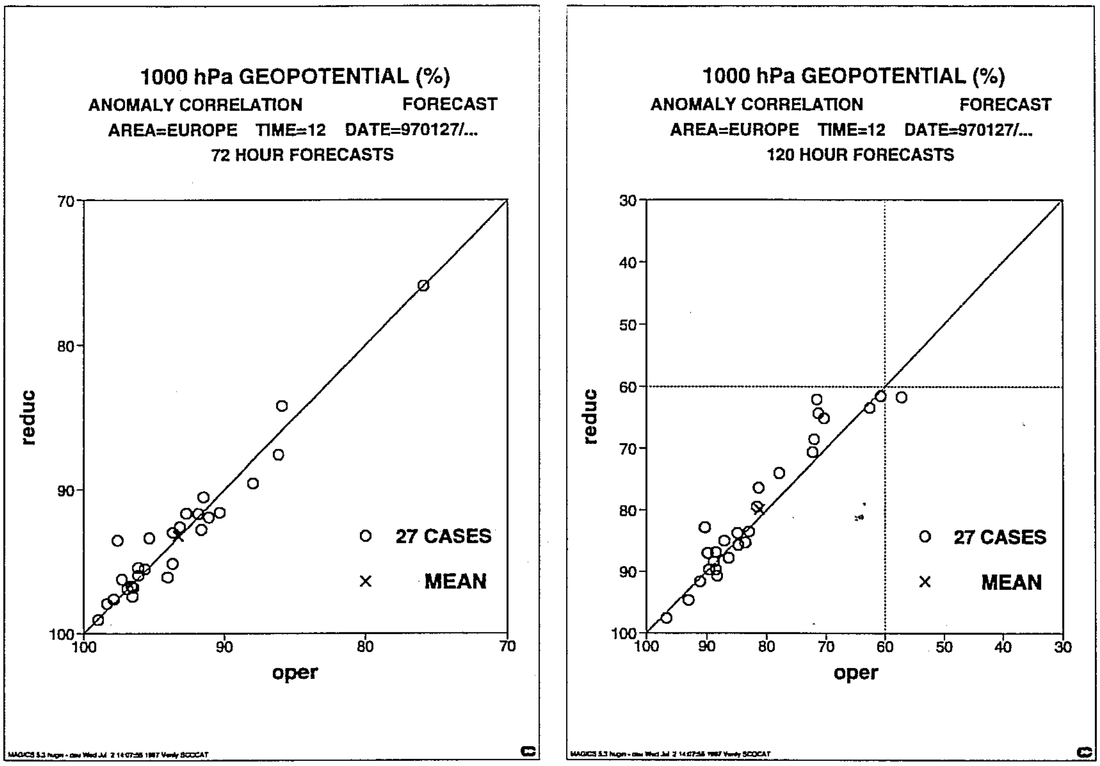


Figure 13. Scatter plots of day 3 and 5 1000 hPa forecast anomaly correlations over Europe.

The differences in forecast scores are fairly small and this is at levels where both forecast sets are very good. Individual forecasts have been scanned for interesting synoptic impacts. Most of the time the synoptic improvements are small in terms of the flow pattern, but a few of the cases show significant improvements. One particular case can be seen in Fig. 14.

The day 4 forecast using the enhanced network from 19970204 (upper left) shows a much better developed 500 hPa low east of Iceland compared with the reduced network (upper right) and this verifies rather well (lower right). The difference of absolute error between the enhanced and reduced network forecasts is shown in the bottom left figure. These forecast error differences can be traced back towards west Greenland in the 24 hour forecast from 1997-02-03 and in the analysis somewhere over southern Greenland on 1997-02-03 06 UTC.

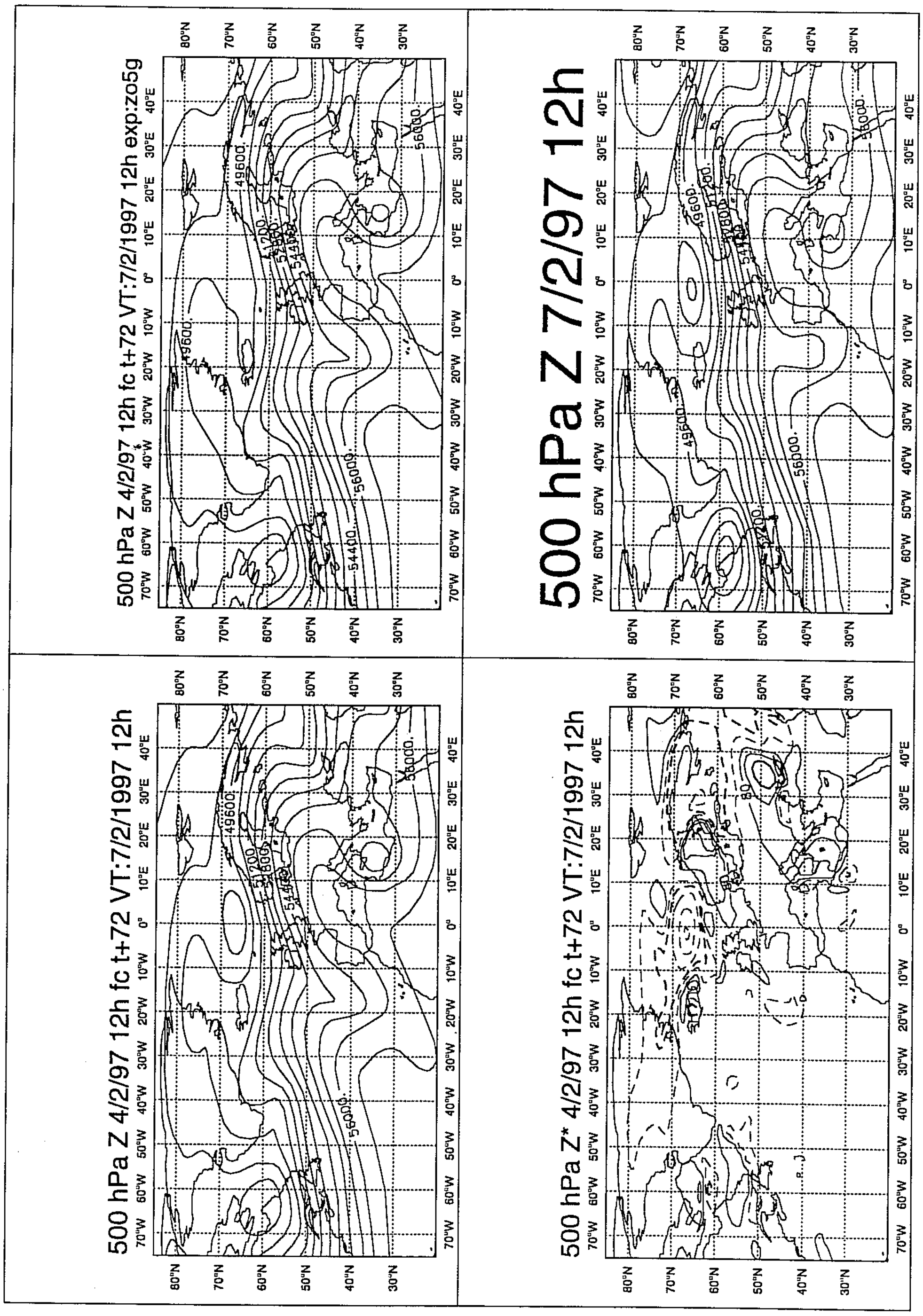


Figure 14. Day 3 500 hPa geopotential forecast with enhanced network from 1997-02-04 12 UTC (lower left), reduced network (upper left), verifying analysis (upper right) and difference of absolute forecast errors (enhanced - reduced, lower right). In geopotential units (J/kg).

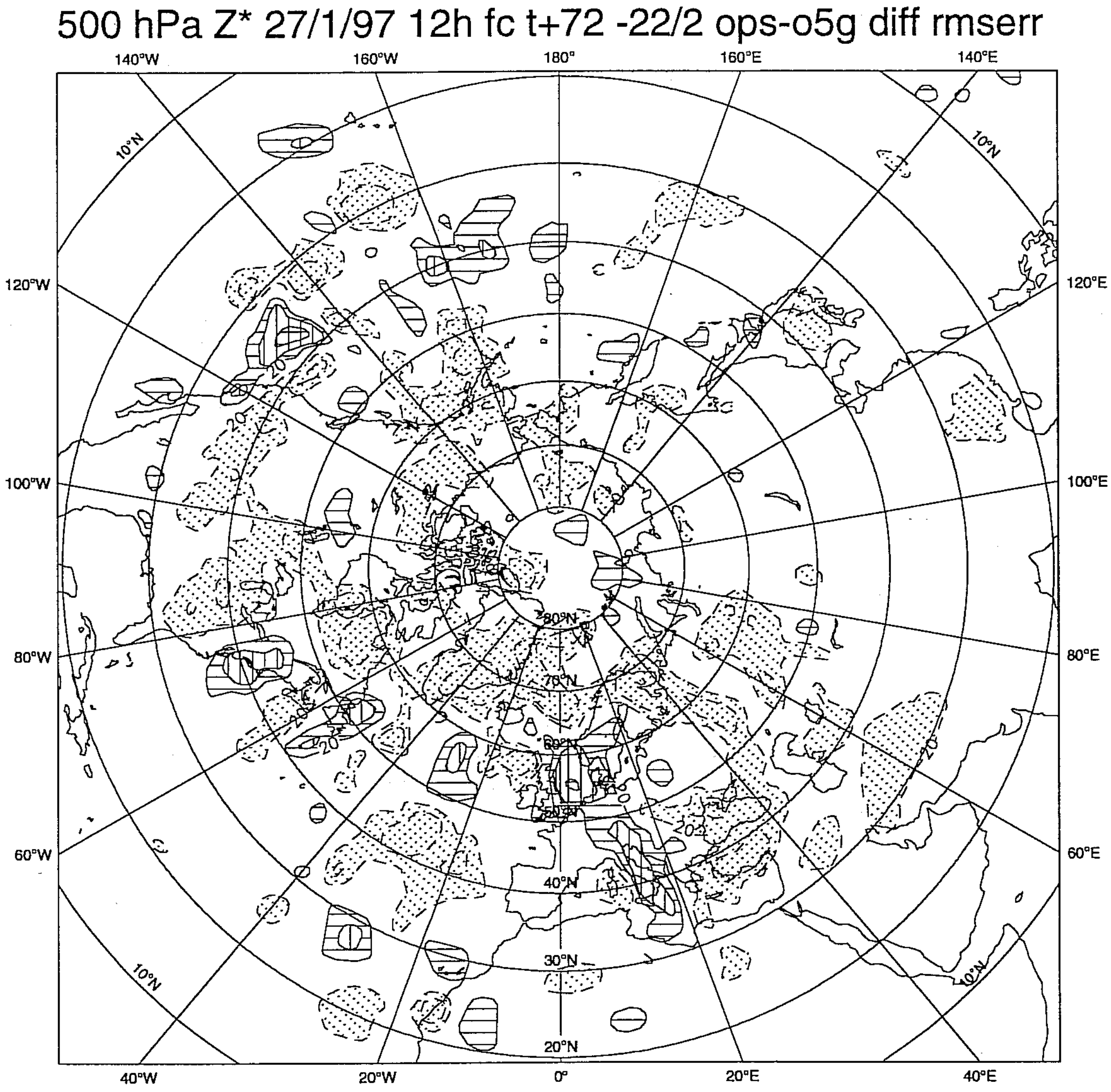


Figure 15. Difference of RMS errors in forecasts with enhanced network compared with reduced. Negative areas dotted with dashed isolines +- 20, 40, 100, 200, 400, 800, 1200, 1600, 2000, 2400, 3200 and 4800 J/kg (geopotential units).

The differences in RMS errors of the two forecast sets has been computed for 500 hPa geopotential and are shown in Fig. 15 for day 3 forecasts. It shows that the reduced errors with the enhanced network are mainly over the northern part of the North Atlantic, Greenland, northern Scandinavia and Russia. Presumably the fact that the flow pattern and dominating activity was at these higher latitudes in the beginning of the period had an influence on these results. Further south in the Atlantic and over central Pacific there are also areas with worse scores. This must be due to random sampling effects in the system.

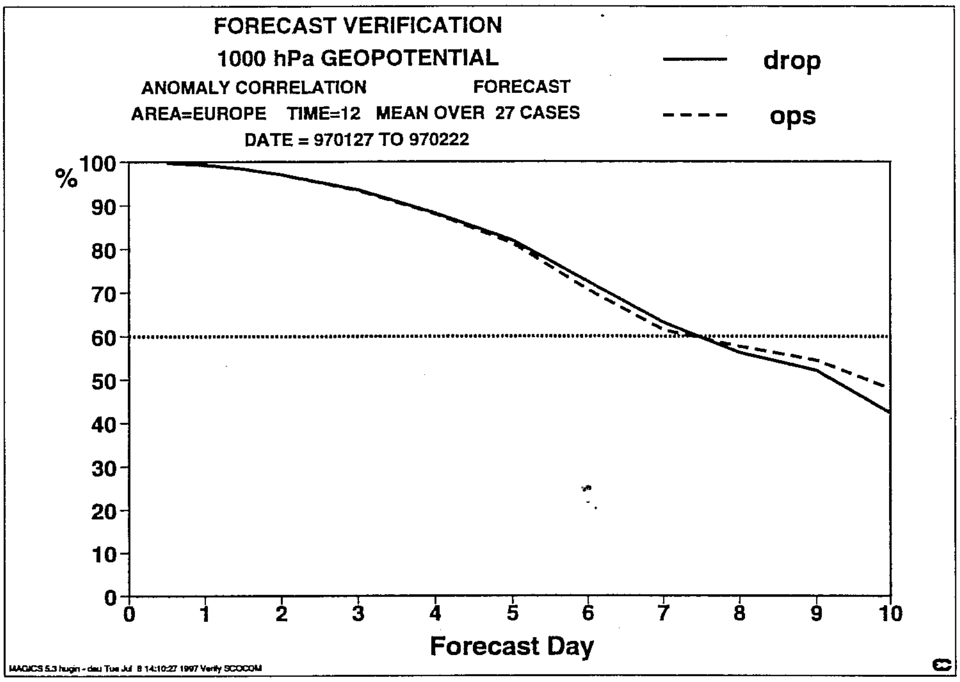
3.2 Dropsonde impact

A second experiment was then run in order to test the impact of the dropsondes in addition to the already enhanced FASTEX network. This was done by running an assimilation with dropsondes activated and comparing it with ECMWF operations. The dropsondes which were used in these experiments were the ones received in real time at ECMWF from the GTS. We are aware of some more observations that were taken but not inserted onto the GTS and consequently not used in this study.

The dropsondes monitored at ECMWF seemed to have unreliable geopotential data. Most of them displayed large biases against ECMWF background fields. Presumably the reference/sea level pressure is not well known in the case of dropsondes. Because of the unsatisfacory results of the monitoring they were not used operationally at ECMWF. So far only the geopotential has been used operationally at ECMWF as the mass information from radiosondes. In order to extract useful mass information from the dropsondes, in this experiment the observed **temperature** data instead of geopotential data were used from the dropsondes only (geopotentials were still used from all other radiosondes). The use of the temperatures works well in the analysis in terms of drawing to the data and producing mass and wind increments around the dropsondes. This configuration has however not been tested before, so the tuning of observation errors, quality control and horizontal thinning of observations may not be optimal. Another feature of the ECMWF 3D-Var analysis system is that the background departures for observations not on the exact analysis time are still computed using the background forecast valid at the analysis time.

The assimilation of these data, in addition to the other FASTEX data used in operations, was run for the same period as the reduced network experiment discussed earlier (19970127 until 19970222). Analysis impacts of these data are in general not as large as in the previous network experiment since the number of times with extra data is lower. The observations were taken at what are supposedly more optimal positions and might have larger impacts on the ensuing forecasts.

The forecast scores from the dropsonde experiment show very little average impact over the Northern Hemisphere as a whole. Over Europe and North Atlantic there is a slight positive impact (Figs. 16 and 17). The scatter plots of the scores (Fig. 17) show that almost all the three-day forecasts were neutral or slightly improved over Europe and the North Atlantic. Over other areas the scores show less systematic impact and average out to be neutral.



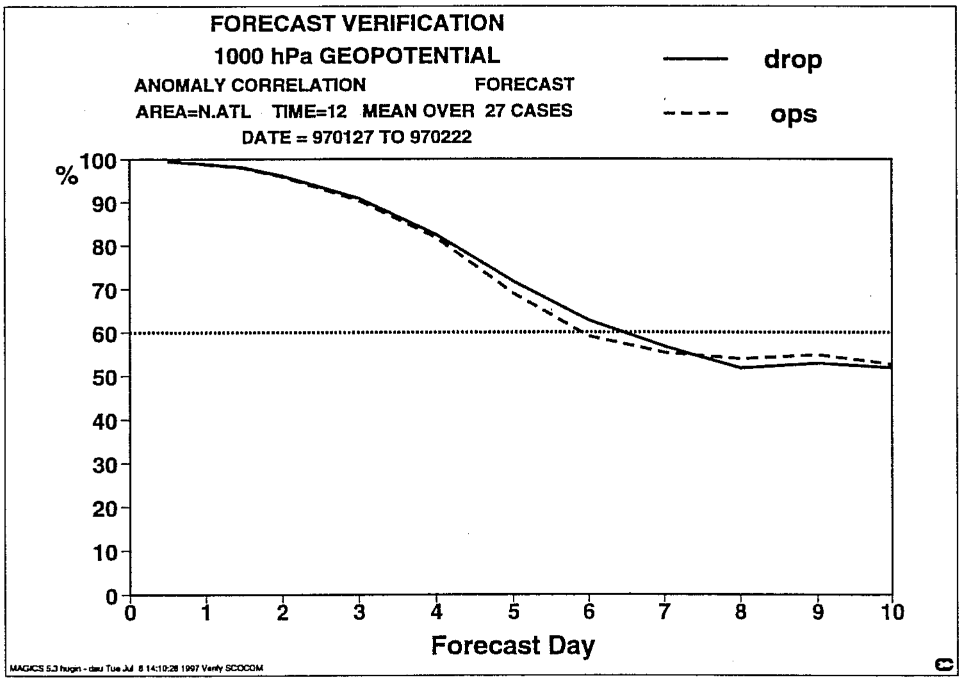


Figure 16. Mean of 1000 hPa geopotential forecast anomaly correlations for Europe and North Atlantic.

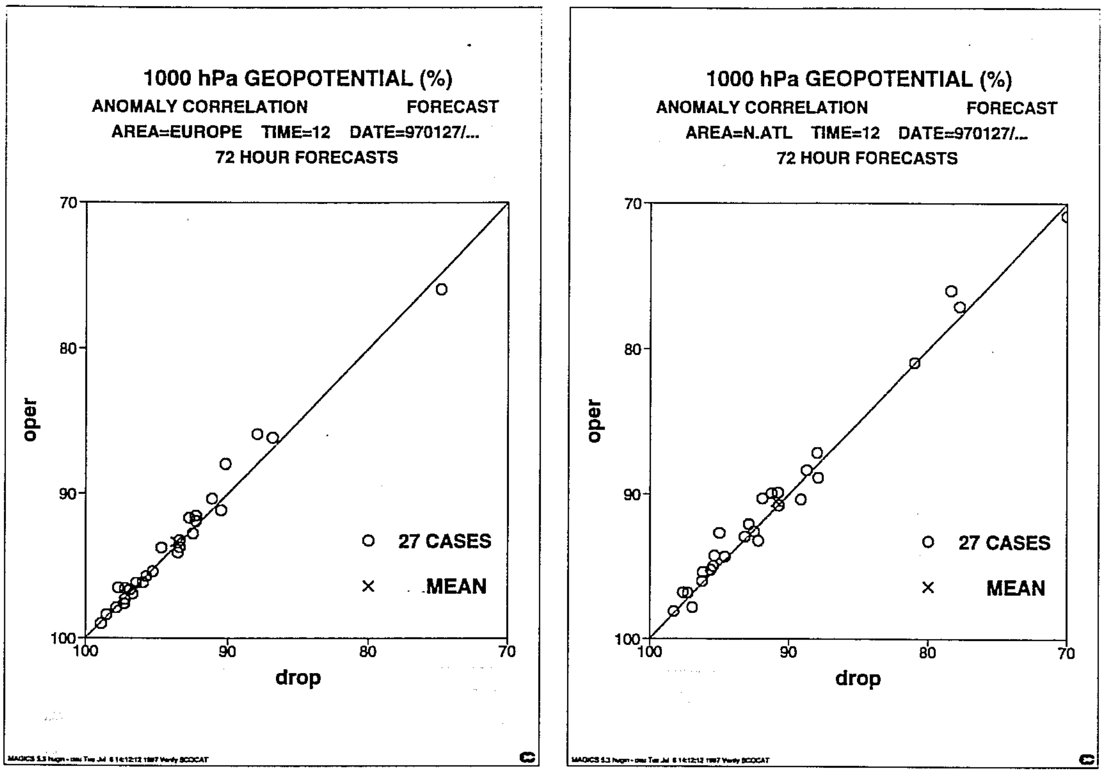


Figure 17. Scatter plots of day 3 1000 hPa geopotential forecast anomaly correlations over Europe and North Atlantic.

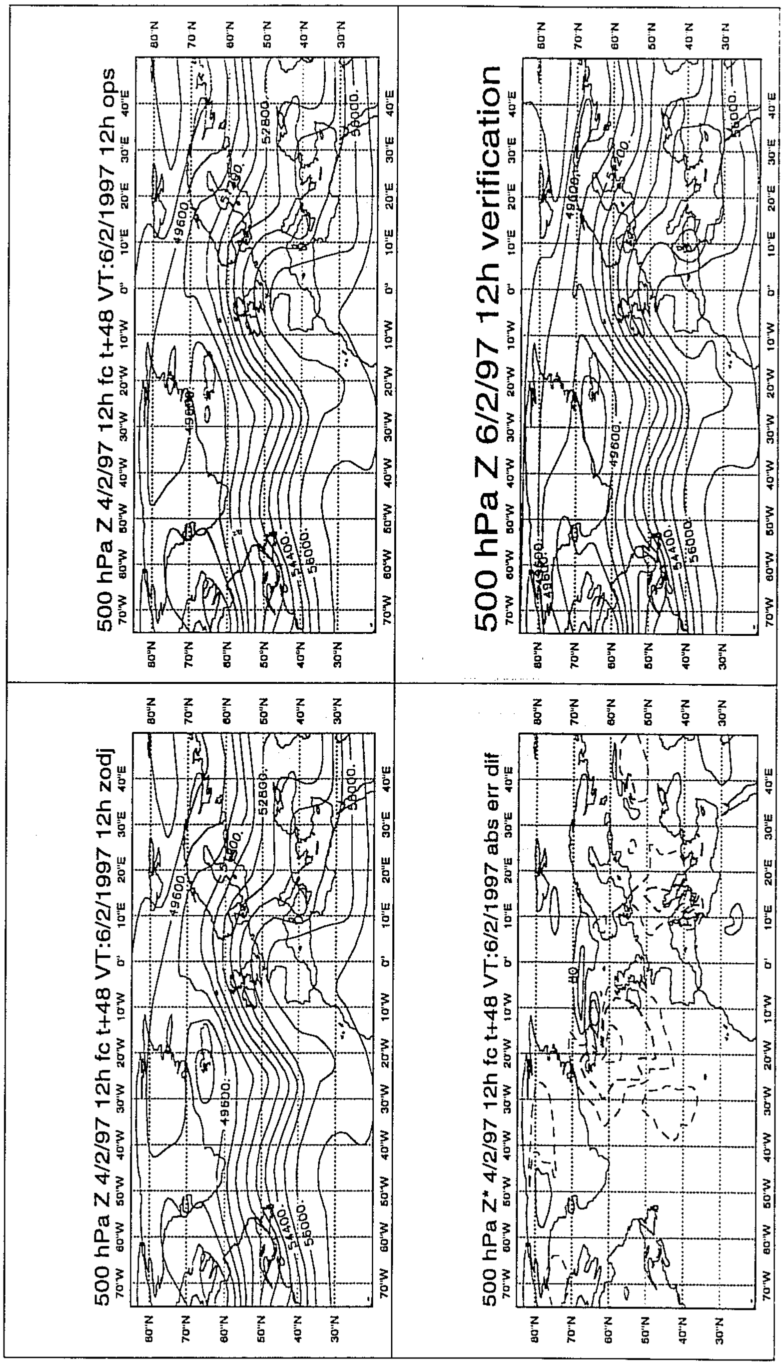
Most forecasts showed only small synoptic differences. One case for which there was a noticeable synoptic improvement is shown in Fig. 18. The dropsonde two-day 500 hPa forecast from 19970204 12 UTC shows an improved definition of the low southwest of Iceland (upper left, when you have turned the page) compared with operations (upper right). The verifying analysis is at he bottom right. Absolute forecast error differences at the bottom left. Negative isolines show where the dropsonde experiment has lower forecast errors. Also the through at 25° W is in a more advanced (easterly) position, in better agreement with the verification. Another interesting aspect is the cut off, or Genoa cyclone, north of Sicily, which is better defined in the experiment.

Figures 19a and b show difference of RMS 500 hPa two-day forecast errors for the first 13 days and all the 27 days, respectively. The experiment has lower RMS errors over most of the North Atlantic and Europe with particular improvements in two stripes over the North Atlantic and over central Europe. Intermingled areas of larger errors are also apparent, but at lower magnitude. Figure 19a shows a reduction of up to 7 m in RMS height error over the Atlantic west of the British Isles. For the whole period the mean improvement is reduced to areas of about 3 m. Later on in the forecast range the signals become more mixed and difficult to interpret.

A problem with this type of impact studies with localised observing systems is that the effects of the extra data are mixed with many very non-linear effects in the data assimilation system, which introduce quite large random noise in the forecast results. A more direct way of measuring the impact of the dropsondes is to perform single analyses and have an almost exclusively local analysis impact of the data. Another reason for doing single analyses was to include only dropsondes which were in the area of sizeable amplitude of the ECMWF singular vectors. These observations could then be regarded as the targeted ones based on the singular vectors. (For a discussion of this, see Palmer et al., 1997.) This approach emerged as a result of discussions with A. Thorpe and A. Montani of Reading University, who have been following this experimentation. Five cases of such targeted dropsondes (one or sometimes two 6 hour periods analysed) were selected by A. Montani (pers. communication), see Table 2. They were then analysed or assimilated without the enhanced radiosonde network. Forecasts were run out to 3 days and compared with forecasts from the assimilation without the enhanced network and without dropsondes. Most of the cases had improved forecast scores over the North Atlantic and Europe at day 2 and 3 (see Fig. 20). Synoptically the improvements were mainly in terms of slightly better positioning of systems but in areas where the errors were already large. The most striking improvement is the forecast from 19970217 18 UTC for 19970219 12 UTC. With the dropsondes the deep low north-west of Scotland (Fig. 21) has been deepened to 961 hPa compared with 968 without dropsondes. The position is also closer to the analysis and the ECMWF operationally analysed depth is 961 hPa; the before mentioned dropsonde assimilation (aodj) 957 hPa and manual Deutscher Wetterdienst (DWD) analysis about 954 hPa (Europäisher Wetterbericht).

Experiment	Date and time
app3/zpp3	19970218 00 UTC
appp/zppp	19970217 18 UTC
apq3/zpq3	19970208 12-18 UTC
apqe/zpqe	19970204 12-18 UTC
apqz/zpqz	19970201 12-18 UTC

Table 2. Experiment names and analysis times/periods for the 5 targeted dropsonde analyses.

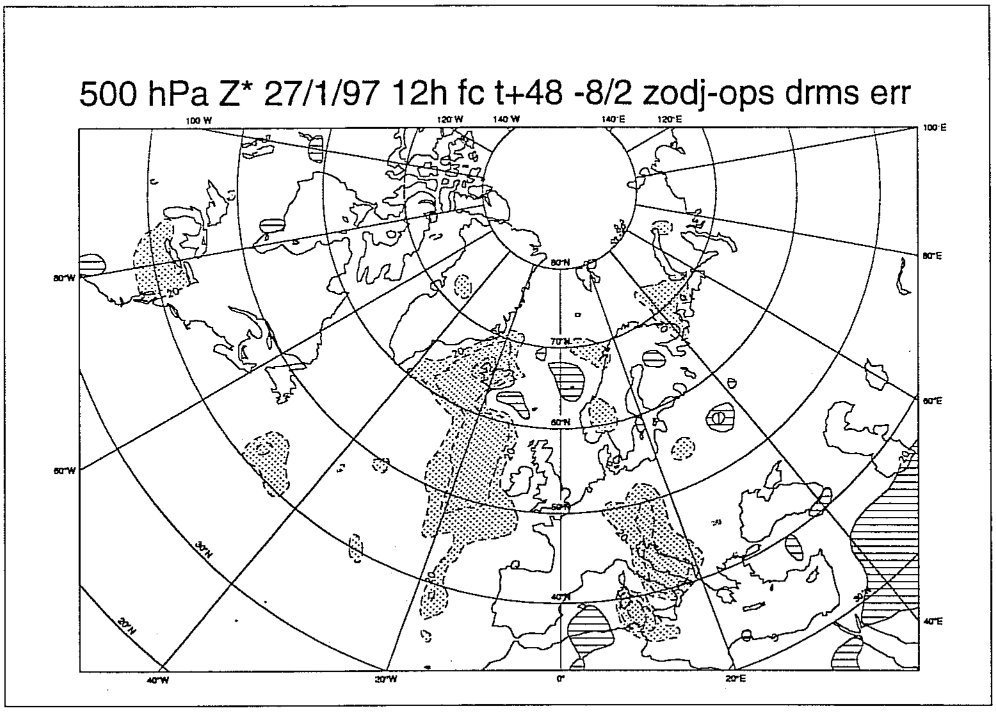

500 hPa Z 4/2/97 12h fc t+48 VT:6/2/1997 12h ops

500 hPa Z 6/2/97 12h verification

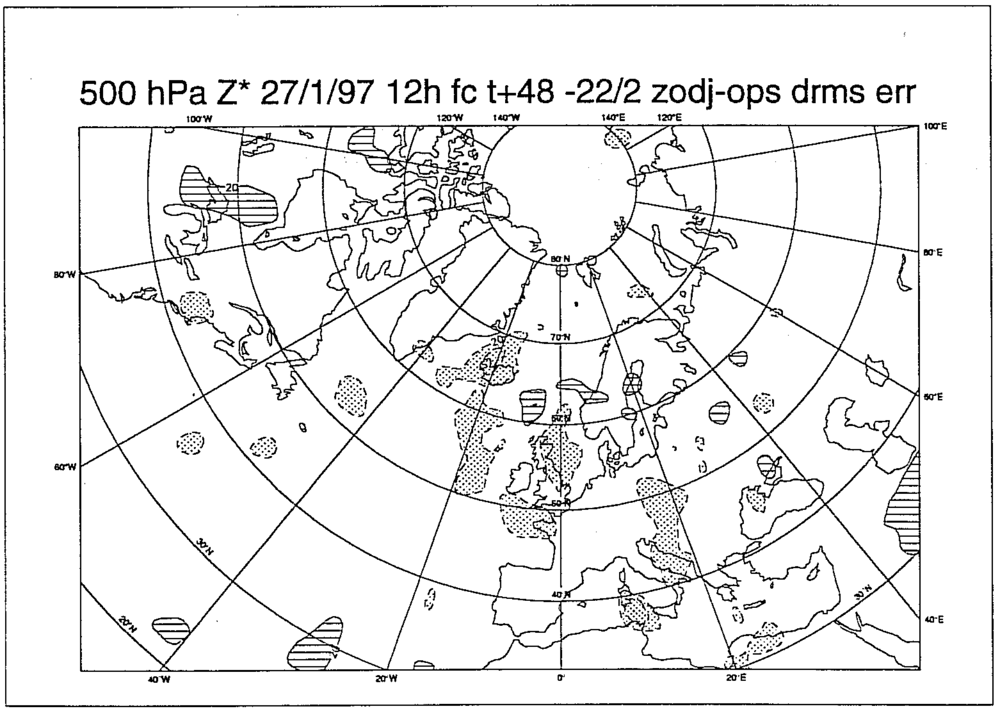
500 hPa Z 4/2/97 12h fc t+48 VT:6/2/1997 12h zodj

500 hPa Z* 4/2/97 12h fc t+48 VT:6/2/1997 abs err dif

Figure 18. 500 hPa day geopotential forecast from dropsonde experiment (lower left), operations (upper left), verifying analysis and differences of absolute error between dropsonde forecasts and operations. In geopotential units (J/kg).



500 hPa Z* 27/1/97 12h fc t+48 -8/2 zodj-ops drms err



500 hPa Z* 27/1/97 12h fc t+48 -22/2 zodj-ops drms err

Figure 19. Difference between 500 2 day height forecast RMS errors with dropsondes and without. Isolines as in Fig. 15.

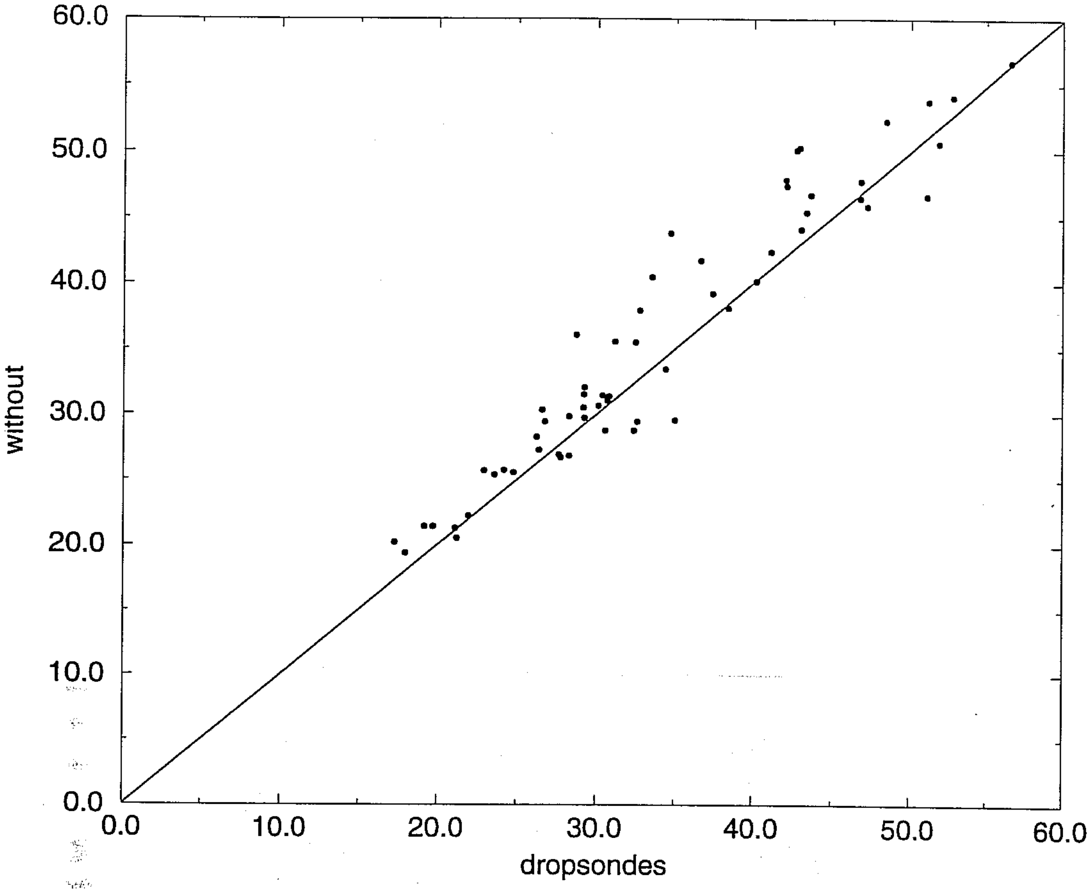


Figure 20. Scatter plot of 1000 and 500 hPa height RMS forecast errors at 48, 60 and 72 hours forecast range from targeted dropsonde analyses and control without dropsondes.

4. Conclusions

In general the results of the observing system experiments are quite encouraging since they show that the current operational ECMWF data assimilation system is able to gain benefits from both conventional and satellite based systems. Global observing system experiments show a very large impact from radiosondes and PILOTs in the Northern Hemisphere and the tropics. The aircraft data also have quite a large positive impact, particularly over the North Pacific and North America. The TOVS data have very large impact over the Southern Hemisphere and in the tropics. This observing system seems to have at least as large impact as SATOBs in the tropics. The SATOBs show also a small but significant positive impact in both hemispheres. In the Northern Hemisphere the major impact from the satellite based systems comes from the SATOBs.

The forecast impact of the enhanced FASTEX radiosonde network is at most marginally positive. Over Europe and the North Atlantic there is a slightly clearer positive impact from using dropsondes in addition to the FASTEX radiosondes. It is easier to demonstrate positive impact of the dropsondes in single (or two consecuetive) analyses for interesting cases when dropsondes are available in targeted areas. These analyses showed a significant improvement of the two- and three-day forecasts for the North Atlantic and Europe.

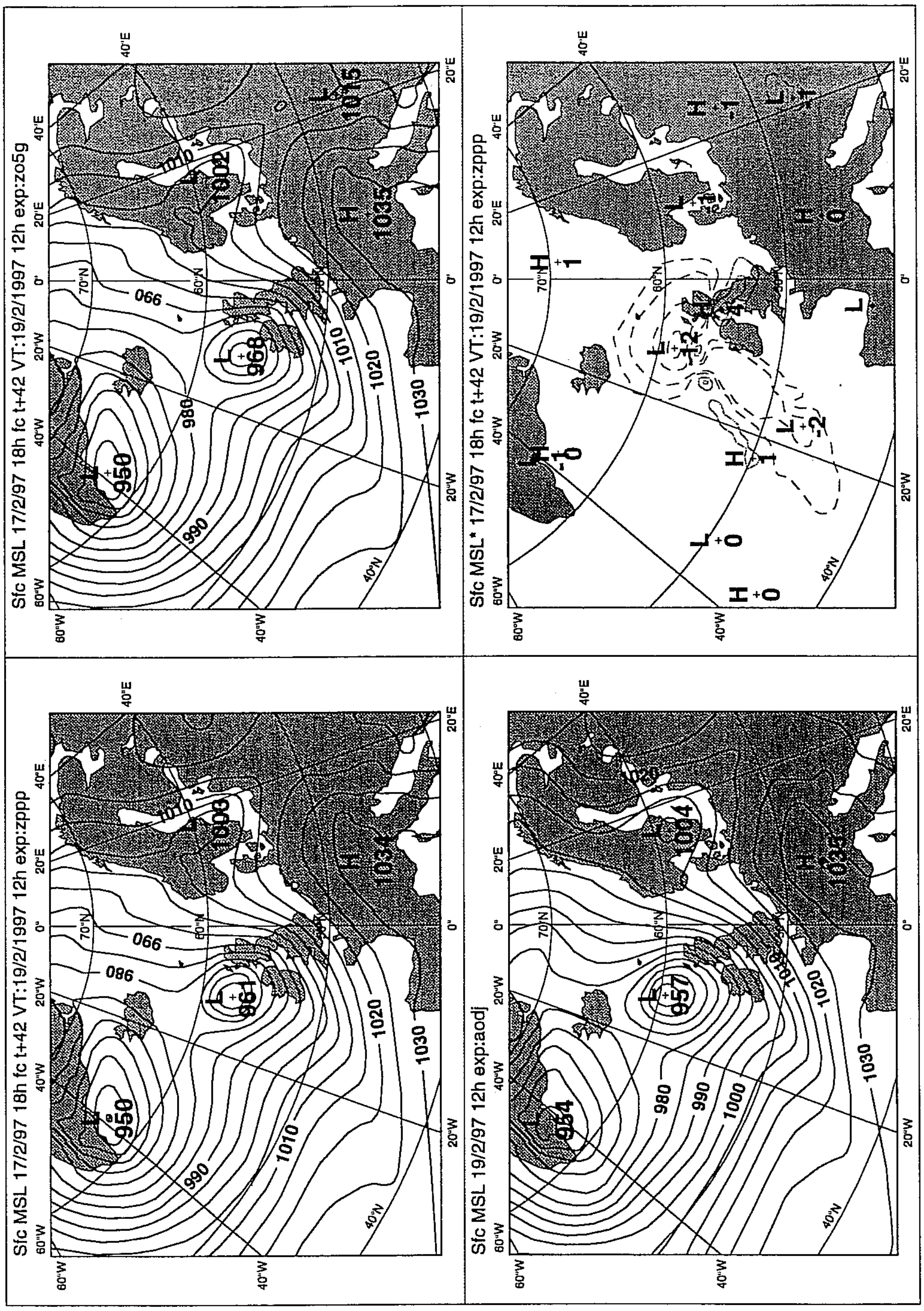


Figure 21. 42hour mean sea level pressure forecasts from 19970217 18 UTC from analyses using dropsondes (a, upper left), without dropsondes (b, right) and verifying analysis (c, lower left). Lower right panel (d) shows differences of absolute errors with isolines at +- 1,2,5,10, 15 and 20 hPa.

There is a small but measurable average positive impact in the early medium range forecasts from the FASTEX observing system enhancements. Only occasionally is it possible to find significant synoptic improvements. It should be pointed out that the enhancements of the observing systems were an area which is the most well observed of any of the ocean areas in the world. It is probably more cost effective to improve the observing network in other less well observed areas of the globe in order to get larger medium range forecast impacts.

A significant reduction of the radiosonde network in the Northern Hemisphere or the tropics would have a profound impact on forecast quality. At least with today's data assimilation systems the loss of radiosondes cannot be compensated by aircraft data (with current coverage and instruments) although they would become increasingly important. TOVS data still seem to show very small average benefits for the Northern Hemisphere forecasts, but again they might play a bigger role in a reduced radiosonde configuration.

References

Andersson, E., A. Hollingsworth, G. Kelly, P. Lönnberg, J. Pailleux and Z. Zang 1991: Global observing system experiments on operational statistical retrievals of satellite sounding data. Mon. Wea. Rev., Vol 119, 1851-1864.

Andersson, E., J. Haseler, P. Undén, P. Courtier, G. Kelly, D. Vasiljevic, C. Brankovic, C. Cardinali, C. Gaffard, A. Hollingsworth., C. Jakob, P. Janssen, E. Klinker, A. Lanzinger , M. Miller, F. Rabier, A. Simmons, B. Strauss, J-N. Thépaut and P. Viterbo 1996: The ECMWF implementation of three dimensional variational assimilation (3D–Var). Part III: Experimental results. Accepted to appear in Quart. J. Roy. Met. Soc.

Andersson E., J. Pailleux, J–N. Thépaut, J.R. Eyre, A.P. McNally, G. Kelly and P. Courtier 1994: Use of cloud–cleared radiances in three/four–dimensional variational data assimilation. Q.J. R. Meteorol. Soc., 120, 627–653.

Bouttier F., J. Derber and M. Fisher 1997: The 1997 revision of the J_b term in 3D/4D–Var. ECMWF Tech. Memo. No. 238.

Eyre, J.R., G. Kelly, A.P. McNally, E. Andersson and A. Persson, 1993: Assimilation of TOVS radiance information through one–dimensional variational analysis. Quart. J. Roy. Met. Soc., 119, 1427–1463.

Kelly, G. and J. Pailleux, 1988: Use of satellite vertical sounder data in the ECMWF analysis system, ECMWF Tech. Memo. No. 143.

Kelly, G., E. Andersson, A. Hollingsworth, P. Lönnberg, J. Pailleux and Z. Zang 1991: Quality control of operational physical retrievals of satellite sounding data. Mon. Wea. Rev., Vol. 119, 1866-1880.

Kelly, G., J. Pailleux, F. Rabier and J-N. Thepaut 1993: Observing System Experiments made with the ECMWF System. World Weather Watch Tech. Report.16. WMO/TD No. 594.

Lorenc, A.C. 1981: A global three–dimensional multivariate statistical interpolation scheme. Mon. Wea. Rev., 109, 701–721.

Palmer, T. N., R. Gelaro, J. Barkmeijer and R. Buizza 1997: Singular vectors, metrics and adaptive observations. To appear in J. Atmos. Sci.

Uppala, S., A. Hollingsworth, S. Tibaldi and P. Kållberg 1985: Results from two recent observing system experiments. ECMWF Seminar/Workshop on "Data assimilation systems and observing system experiments with particular emphasis on FGGE", Reading, 165-202.

Appendix 1

```
if (OBSTYP = temp) then

     if   (( 030100 <= TIME <= 090000 )

     or ( 150100 <= TIME <= 210000 )   )

and  STATID in ("03354", "04220", "04270", "04339", "04360",

          "06011", "07110", "07145", "07510", "04018",

          "03953", "08508", "08522", "08001", "03005",

          "03026", "03496", "03502", "03808", "03240",

          "03920", "71801", "71816", "71906", "71600",

          "78016", "72402", "72208", "74494", "FNOR",

          "FNOU", "FNPH", "FNRS", "OXVH2", "OXYH2",

          "V2EZ", "KCEJ", "FZVN", "EOGW", "TFTA",

          "V2LV", "V2LX", "DBBH", "V2GH", "EHOA")

then fail(CONSTANT); endif;

     if   (( 090100 <= TIME <= 150000 )

     or ( TIME >= 210100 or TIME <= 030000 ) )

and  STATID in ( "KCEJ", "FZVN", "EOGW", "TFTA")

     then fail(CONSTANT); endif;
```