TECHNICAL REPORT No. 47

IMPACT OF AIRCRAFT WIND DATA ON ECMWF ANALYSES AND FORECASTS DURING THE FGGE PERIOD, 8-19 NOVEMBER 1979

by

A.P.M. Baede[1], P. Kållberg[2] and S. Uppala

March 1985

[1] K.N.M.I.
De Bilt, The Netherlands

[2] Visiting Scientist from
S.M.H.I.
Norrkoping, Sweden

Abstract

This report describes the results of an observing system experiment, set up to assess the impact of aircraft wind data on ECMWF analyses and forecasts. Two data assimilations were run for an 11 day period on FGGE II-b data, with and without aircraft data, followed by two similar assimilation runs from which SATOBs and SATEMs were excluded.

From an analysis of these experiments it is concluded that the quality of the automatically transmitted aircraft data (ASDAR, AIDS) is high. Moreover the ECMWF assimilation system is able to extract valuable information from these data by directing this information to the external and graver internal Rossby modes of the model. During the period studied in this report, the impact of aircraft data was significant in the tropics but outside the tropics a considerable redundancy appeared to exist between these data and satellite data. It is shown however that this may be a consequence of the particular synoptic situation during that period.

It is concluded from this assessment that a future system of aircraft data, based on ASDAR, will be a valuable contribution to the global observing system.

For the aviation community it is interesting to note that a positive impact of aircraft wind data was found on both wind analyses and 6 hour forecasts along flight tracks. On average an improvement of about 1 m/s is found for both wind components but locally and in individual cases this may be considerably larger.

CONTENTS

1. INTRODUCTION

In recent years new meteorological observing systems have been developed which in the future could become potentially valuable components of an improved Global Observing System. When a sufficient number of observations with such a new system, preferably made under operational conditions, is available, the impact may be assessed by carefully comparing a series of numerical analyses and forecasts based on data sets with and without these observations. Such an experiment is called an Observing System Experiment (OSE). The FGGE level II-b data set provides an excellent opportunity for conducting such OSE's because during FGGE several newly developed observing systems were operational.

In this report we present the results of an OSE on the impact of aircraft wind data on numerical analyses and forecasts. During FGGE two new automatic aircraft meteorological data systems were operational: ASDAR and AIDS. Because the number of these data during FGGE was still small compared to the number of conventional AIREPs it was decided to assess the impact of all aircraft wind data. Earlier work by Nitta et al. (1979) suggested that this impact may be small but positive.

The OSE consisted of two parallel data assimilations with the ECMWF system for the period from November 8, 1979, 00GMT, up to and including November 18, 1979, 18GMT. In Section 2 we describe the experiment and the data in detail.

The two sets of analyses were compared extensively both from a synoptic point of view and objectively. This is described in detail in Section 3. When it was found that the impact was rather small it was suspected that the abundance of satellite data might have obscured the impact. Therefore two new assimilation runs were made without satellite winds and temperatures. The results of these runs will be published elsewhere, but occasionally we shall make use of them in this report.

One way to assess the impact of data on analyses is to compare the quality of forecasts based on these analyses. Consequently we ran four 10-day forecasts on selected cases and evaluated them by comparisons with the FGGE III-b analyses of ECMWF and with observations. In Section 4 we report not only on the impact on medium range forecasts but also on very short range forecasts, in which the aviation community will be interested.

Results of an OSE are likely to depend on the assimilation system used. It was therefore very fortunate that a similar study was carried out simultaneously by Barwell and Lorenc (1985) at the U.K. Meteorological Office. In Sections 3 and 4 the result of their study and ours are compared. This leads to valuable conclusions about the relative merits of different assimilation systems.

We realize that the outcome of OSE's may depend not only on the system used but also on the synoptic period studied. A thorough and conclusive assessment of the impact would require more runs than are practicable. This means that conclusions should be handled with care. Further studies are in progress at ECMWF and will be published in due course.
Summaries of the results of this study have been published earlier (Gilchrist, 1982 and Baede et al., 1983).

2. THE EXPERIMENT

2.1 Introduction

In order to evaluate the impact of aircraft wind data on numerical analyses and forecasts it was decided to use the FGGE level II-b data set and to perform an OSE consisting of two parallel assimilation cycles, one with and one without aircraft wind data. In Section 2.2 we first discuss in detail both the experiment and the data used. Some specific aspects of the numerical systems are then described in Section 2.3.

2.2 Description of Experiment and Data

Aircraft wind data in the FGGE level II-b data base consist of both conventional AIREP data, i.e. wind data transmitted orally by the aircrew, and AIDS/ASDAR data, produced automatically by the aircraft inertial navigation system and transmitted in real time via satellite (ASDAR) or (during FGGE) in delayed mode via tape cassettes (AIDS). For a description and evaluation of these automated aircraft observation techniques we refer to De Jong (1981, 1982). At ECMWF in 1981 (unpublished) a minor pilot study was performed to test the impact of AIDS and ASDAR. It consisted of 3 days of parallel data assimilation and one pair of forecasts up to 6 days. Both the analyses and forecasts were very similar. It was felt that one of the reasons for this small impact may have been the abundance of conventional AIREP's in the same areas as the AIDS/ASDAR flights, making the latter contain redundant information. It was therefore decided to aim at a study of the overall impact of aircraft wind data on global analyses and forecasts.

The period selected is November 8-19, 1979. This period was chosen for several reasons. In the first place the level II-b data coverage for that period is very good; two polar orbiting satellites were available. The excellent data coverage makes this period a popular one for other OSE's, thus allowing intercomparison of experiments. Synoptically the period is characterized by an active winter circulation over the northern hemisphere, with marked activity over the Pacific.

The AIDS and ASDAR data present in the level II-b data set for this period have been subjected to a careful manual quality control at KNMI to eliminate possible erroneous data. It is known that the II-b data base contains some AIDS with erroneous data. By carefully comparing wind data from all AIDS and ASDAR flights with existing ECMWF level II-b analyses, suspect data were identified and excluded. In total about 3% of all AIDS data points were excluded in this way. The AIREPs used were those available from the II-b data base, although many apparent errors were identified during the evaluation of the experiment.

Fig. 1 shows a typical example of the data coverage for a particular analysis. Clearly aircraft data are limited to certain specific areas, mainly over the sea, in the northern hemisphere; southern hemisphere data are scarce. Because the present ECMWF assimilation system has a 6 hour cycle, data within plus or minus 3 hours of the analysis time are accepted. This leads to, typically, around 850 AIREPs, 450 AIDS and 150 ASDAR reports for every analysis.

Vertically the aircraft wind data are strongly peaked around 250 mb. A small percentage of the observations is taken during climb and descent. The data are typically of a "single level" type. For further details see De Jong (1981).

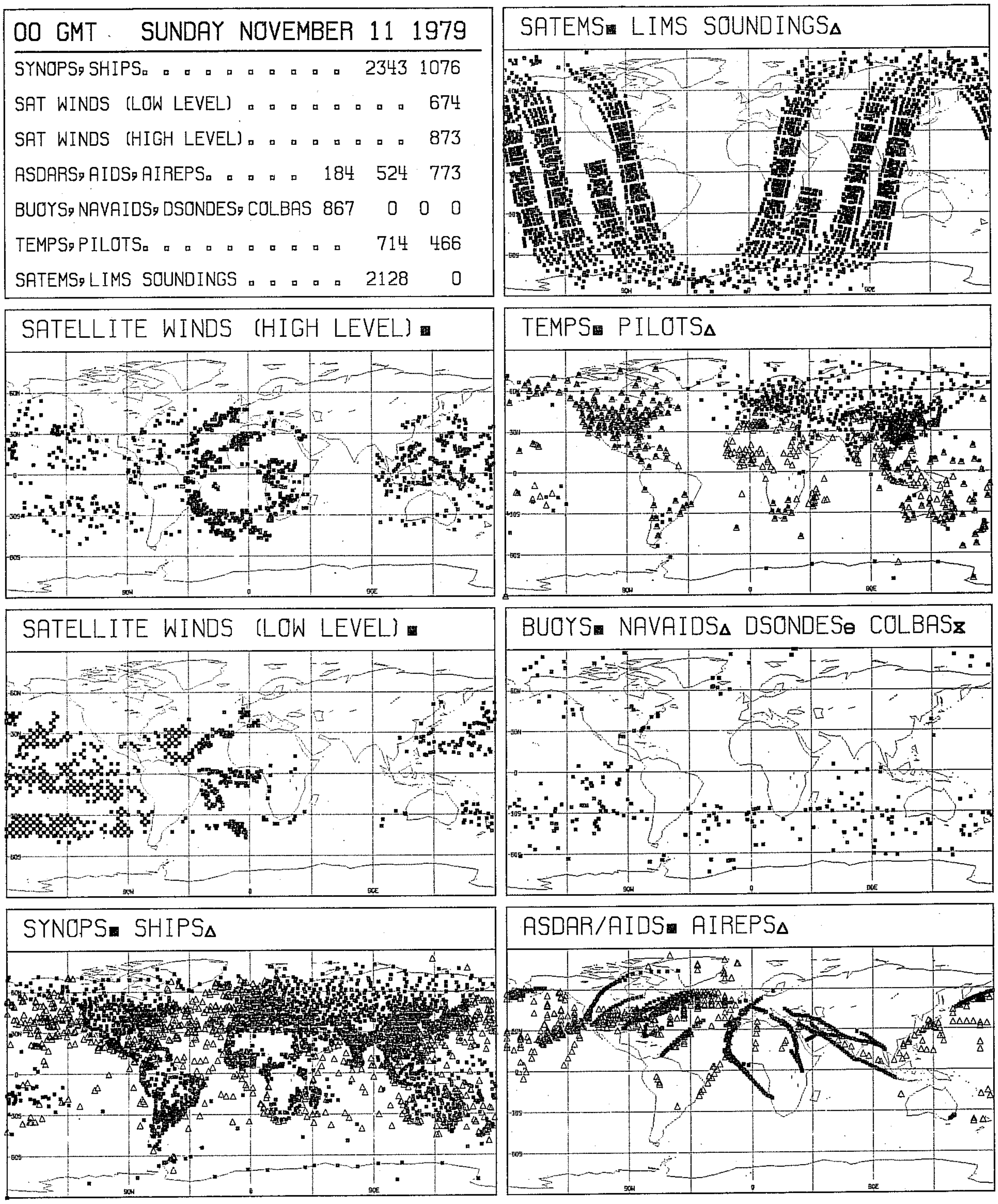

Fig. 1 Data coverage for November 11 1979 00GMT ± 3 hrs.

Apart from the aircraft wind data all FGGE level II-b data were used in both assimilation cycles except the satellite temperatures (SATEMs) below 100 mb over land and all satellite winds (SATOBs) over land. Experience gained at ECMWF and from daily operational forecasting at NMC has indicated that these data are liable to contain large errors. Aircraft temperature data are not used in this experiment either.

Starting from a climatological guess field (a so called "cold start") the ECMWF data assimilation system, with some minor changes described in the next paragraph, was run twice for a period of 11 days from 8 till 19 November 1979. In the first assimilation run (AI) AIREP and AIDS wind data were used, whereas these were excluded from the second run (AO). The rather small number of ASDAR reports was excluded from both runs to serve as an independent reliable data set for measuring the quality of the analyses. During the assimilation procedure observations may be rejected because they do not match the other data or the first guess field sufficiently well within certain preset limits. This led to rejection of about 2% of the AIREPs and 0.5% of the AIDS reports, reflecting the higher reliability of the automatic transmission. This does not mean that all accepted AIREPs are reliable. An example of the contrary will be presented later.

From the two series of analyses produced from the parallel assimilation cycles, four sets of parallel forecasts were run out to ten days using the ECMWF operational grid point model. One case (11 November, 00GMT) was selected because the two analyses showed a considerable difference over the Pacific. The other three cases (10, 13 and 16 November, 00GMT) were selected in a rather arbitrary fashion.

During and after the performance of the experiment two errors were detected but not corrected because they were thought to have no impact on the outcome of the experiment. Due to a coding error in one of the parallel assimilation runs SATEMs over land were used only below 100 mb, instead of above during the first three cycles. This caused a difference between both sets of analyses over Antarctica through the whole depth of the atmosphere during the initial period; however the difference had disappeared after three days. Because the affected area lies well outside the area influenced by aircraft data no measures were taken to correct the error. A second error was made by accidentally assimilating ASDAR reports in all the runs with aircraft data during the eight assimilation cycles of 10 and 11 November. This error cannot possibly have had any impact on the outcome, but the ASDAR data of these two days cannot be considered independent for verification purposes.

When during the evaluation of the results of these experiments it was found that the impact of aircraft data on analyses and forecasts was rather small, it was suspected that the abundance of satellite data might have obscured the impact. Therefore two new assimilation runs and corresponding forecasts were made with (SX) and without (SO) aircraft data, both runs excluding satellite winds (SATOB) and temperatures (SATEM). Occasionally we shall refer to these runs, but a full account of them will be given by Källberg et al. (to be published). In Table 1 the characteristics of all four runs and their acronyms are given.

Table 1 Data used in the four assimilation runs

(+ data used, - data not used)

	AI	AO	SX	SO
satellite data	+	+	-	-
aircraft winds	+	-	+	-
remaining FGGE II-b data	+	+	+	+

2.3 The Numerical System

The numerical system used in this experiment is the same as that used for the production of the FGGE level III-b analyses at ECMWF with some modifications. The system will not be described in detail but we concentrate on the modifications and some details relevant to this experiment. For a detailed description we refer to Bengtsson et al. (1982).

The most important difference between the system used in this experiment and the one used for the level III-b production is that only the pressure level analysis increments to the first guess are interpolated to σ-levels. This method guarantees that the first-guess forecast is carried unaltered through the analysis in data void areas. Furthermore, since the vertical resolution in the boundary layer is much higher in the σ-level system than in the pressure level system, interpolation of the analysis increments results in less destruction of the boundary layer, compared to the full field interpolation. A second difference is that a more detailed orography was used in the present system.

The optimum interpolation analysis scheme requires a prescription of observational errors of the different data systems. Again the same values were used as in the level III-b production system except for the wind errors assigned to TEMP/PILOT, ASDAR/AIDS and AIREP. Table 2 shows the modified values for these systems for several pressure levels. The difference between the error assigned to AIREPs and that assigned to ASDAR and AIDS reflects the larger reliability of the latter.

Table 2 Observational wind component errors (m/s) used in this study. For other data systems see Bengtsson et al. (1982).

Pressure level (mb)	1000	850	700	500	400	300	200
TEMP/PILOT	1.8	1.8	2.5	3.0	3.5	4.0	4.0
ASDAR/AIDS	1.8	1.8	2.5	3.0	3.5	4.0	4.0
AIREP	5.2	5.2	5.2	5.5	5.5	6.0	6.0

A question pertinent to this study is how the assimilation scheme extracts information from the high-density and essentially one dimensional single level aircraft data. Particularly relevant in this respect are the horizontal and vertical forecast error correlation functions which are modelled in the assimilation scheme. No attempt has been made to optimize this aspect of the assimilation scheme or tune it to the special characteristics of aircraft data; this would deserve a special study. The horizontal and vertical forecast error correlation functions of the wind components were the same as in the ECMWF operational system.

The extent to which single level information is distributed vertically depends to a large extent on the vertical prediction error correlation functions. These functions have been tabulated in a N x N symetric matrix, N being the number of levels (N = 15 in our case). Fig. 2 shows the vertical auto-correlation between the forecast errors of the wind components at 250 mb and those at other levels. It indicates that single level 250 mb wind information will be felt throughout most of the troposphere.

3. IMPACT OF AIRCRAFT DATA ON ANALYSES

3.1 Synoptic situation during the period 8-19 November 1979

Rather than presenting a detailed description of the synoptic development during this period we restrict ourselves to the general characteristics of the circulation in those areas where aircraft data are likely to have an impact, i.e. the northern hemisphere Pacific and Atlantic areas around 250 mb. For a description of the circulation during this period the reader is referred to the 300 mb geopotential height maps in Bjørheim et al. (1982).

Figs. 3 and 4 show the mean vector wind at 250 mb and its variability for the AI analyses during the period 10-19 November 1979. These quantities are defined as follows:

$$\text{Variability} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\vec{v}_i - \overline{\vec{v}})^2}$$

where $\overline{\vec{v}} = \frac{1}{N} \sum_{i=1}^{N} \vec{v}_i$ is the mean vector wind and N is the number of analyses. Both figures indicate intense activity over the Pacific connected with the movement of a trough-ridge system towards the North American continent. It is this system that features prominently over the Pacific during the assimilation period. Over the Atlantic the circulation is relatively blocked with considerable activity at high latitudes.

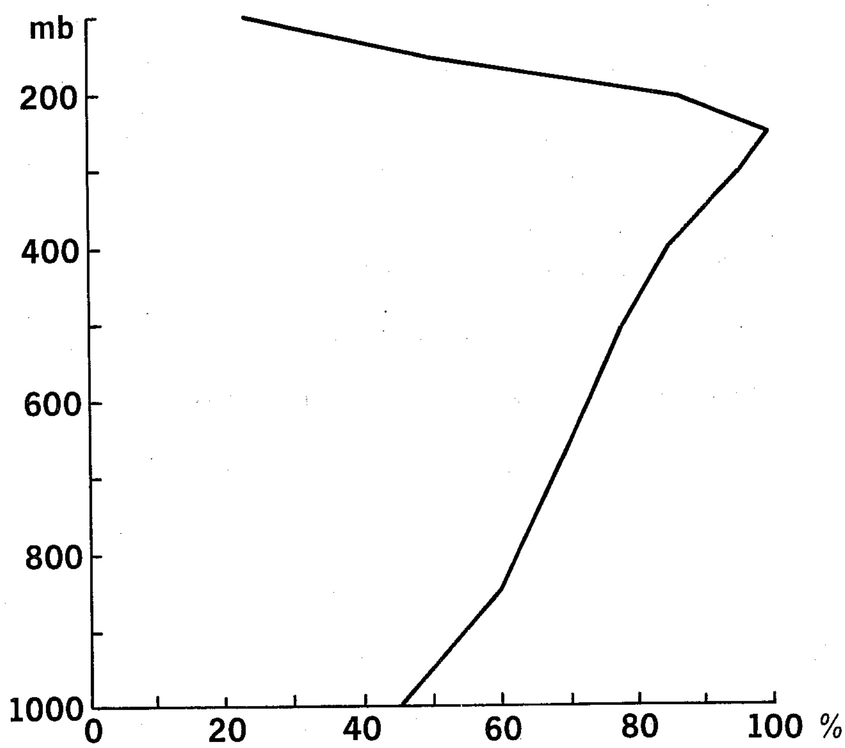

Fig. 2 Assumed vertical correlation between wind first guess errors at 250mb and those at other levels.

Summarizing, the period is characterized by considerable activity over both oceans. This is important because it should be realized that the conclusions drawn from OSEs may depend sensitively on the synoptic situation. The impact of an observational system may be larger when the main activity occurs in those areas where this system is the main source of information. This should be borne in mind when drawing conclusions from this and similar studies.

3.2 Synoptic impact of aircraft data on analyses

In this section the impact of aircraft wind data on analyses is examined from a synoptic point of view. A superficial inspection of the difference maps revealed that in the lower troposphere the differences between both sets (AI and AO) of analyses were very small. At 500 mb local differences of 2 dam were found but over most of the globe differences were much smaller than 1 dam. The largest differences were found at flight level and therefore we will only consider the impact on the 250 mb analyses. We do this in two ways. First we look at the impact on the mean circulation during the period. This reveals the geographical distribution of the impact. Next we study the impact on a particular analysis: the case of November 11, 00GMT. We will also investigate the apparent partial redundancy between aircraft and satellite data.

During the first few days a considerable difference between both sets of analyses was found over Antarctica through the whole depth of the atmosphere caused by an unfortunate coding error, already referred to in Section 2.2. This difference had disappeared after about three days. Because the impact of aircraft data is negligible in that area, this error had no detrimental effect on our conclusions.

(a) Geographical distribution of aircraft impact

The geographical distribution of aircraft data during FGGE is very uneven, with good coverage over the North Atlantic and to some extent the North Pacific. The rest is concentrated in narrow bands following the most frequent air routes. Only a minor fraction of the data comes from the southern hemisphere (Fig. 1). The geographical distribution of the aircraft data impact may be expected to reflect this uneven distribution. This is confirmed by Fig. 5 which shows the RMS difference between the vector wind field, analyzed with (AI) and without (AO) aircraft data, averaged over 40 analyses from November 9, 00GMT till November 18, 18GMT. The largest differences of up to 10 m/s are found over the North Pacific and North Atlantic, the Arabian Sea, the Gulf of Guinea and the equatorial Pacific. In the southern hemisphere the impact is very small almost everywhere and limited to the aircraft lanes. In order to interpret the root mean square wind differences further, they must be related to the mean wind field and its variance over the period. Fig. 3 and 4 show the mean vector wind at 250 mb and its variability (as defined in the previous section) for the 40 AI analyses.

The large RMS differences over the Sahara and Arabian Sea are clearly connected with the subtropical jet which is intense and shows great variability over these regions.

By the very nature of the problem of analyzing wind fields from aircraft data in otherwise completely data void areas (SATEMS were not used over land in this experiment), it is impossible to determine unambiguously whether AI or AO is the better analysis of the two. The very good agreement between aircraft data and collocated rawinsonde data (De Jong, 1982) ensures that the marked impact of aircraft data on the subtropical jet analysis in assimilation AI is beneficial.

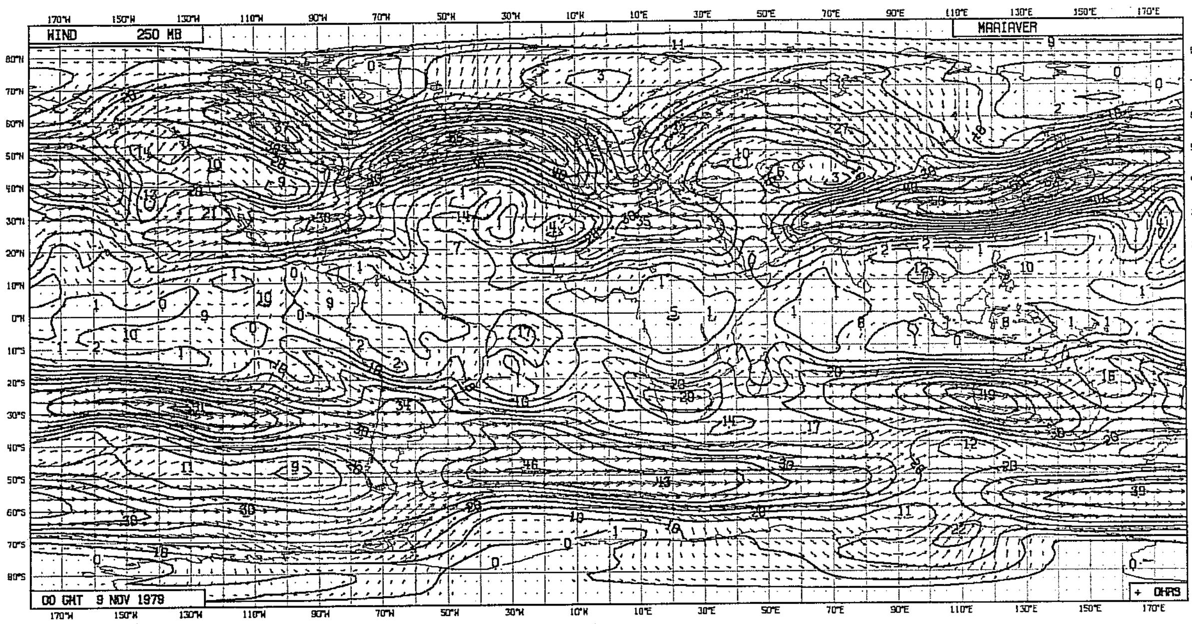

Fig. 3 Time averaged vector wind at 250 mb in assimilation AI for the period 10-19 November 1979.

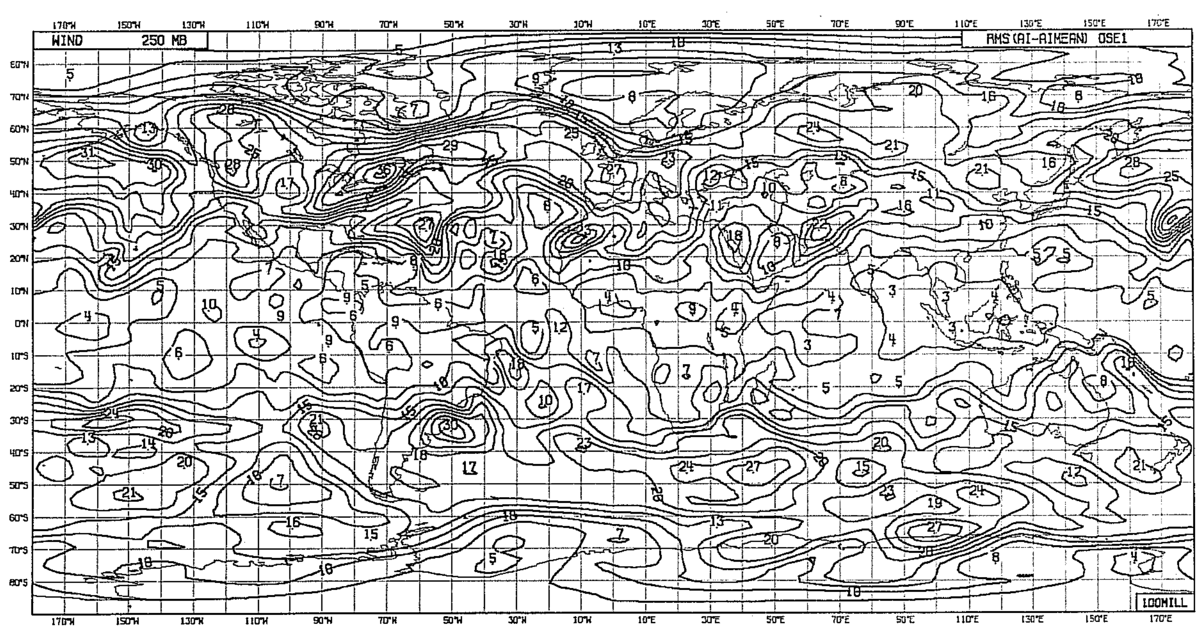

Fig. 4 Time variability of 250 mb vector wind in assimilation AI for the period 10-19 November 1979.

AI-AO

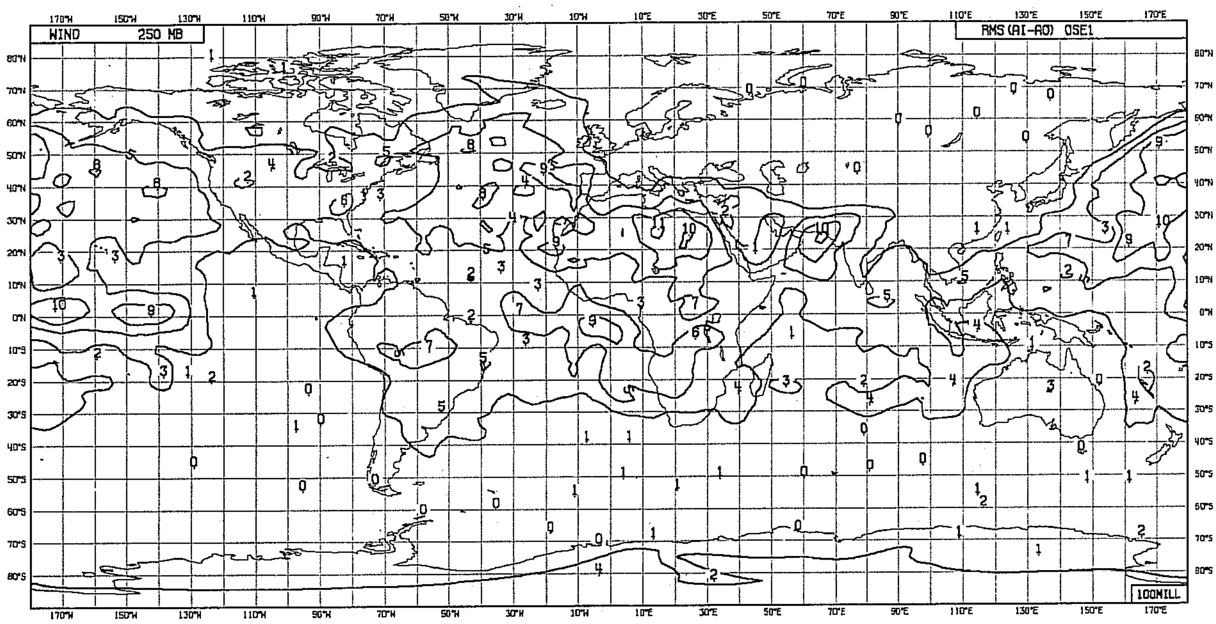

Fig. 5 Time averaged RMS difference of vector wind fields AI-AO.

SX-SO

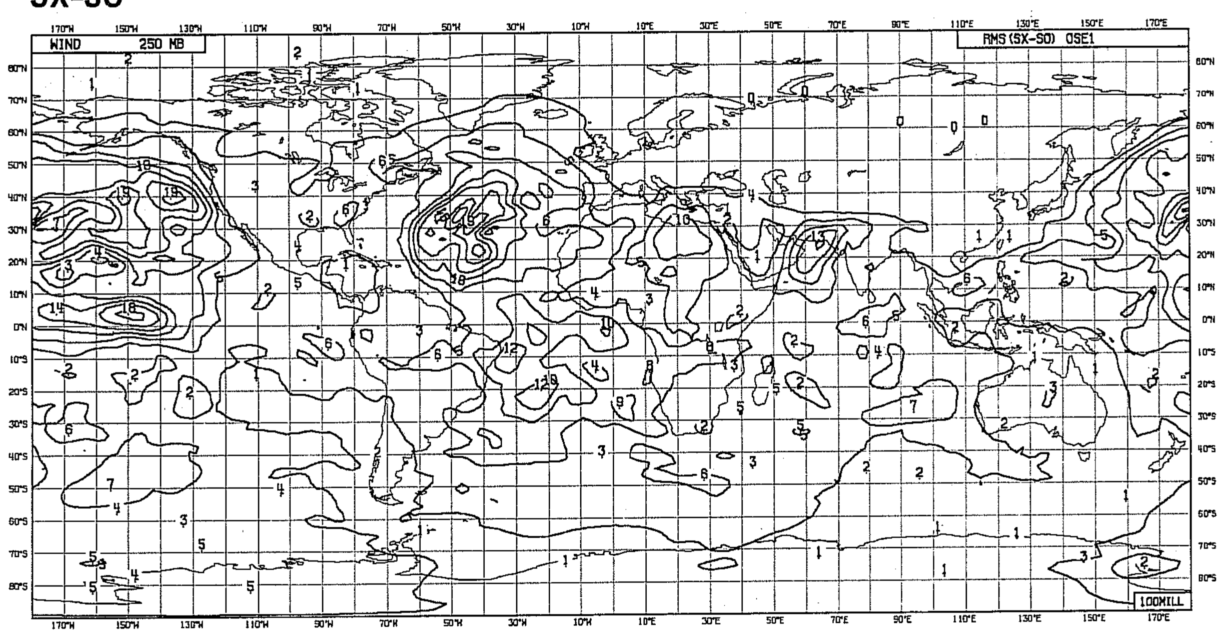

Fig. 6 Time averaged RMS difference of vector wind fields SX-SO

At higher latitudes, i.e. the North Atlantic and North Pacific regions, the RMS differences are smaller in spite of strong and highly variable jets. Over major parts of these regions the RMS differences are of the order of 8 m/s or less, while the variability often exceeds 25 m/s. Thus the aircraft impact at higher latitudes is less significant than in the subtropics.

This conclusion is further emphasized by the very small aircraft wind impact to be seen over the Arctic area. Although the average windfield was weak during the period, there was a considerable variability, in excess of 8 m/s over large portions. The RMS differences between the AI and the AO runs on the other hand are negligible in spite of a relatively good data coverage.

In equatorial regions, finally, the aircraft impact is very marked in areas with data. In particular in the equatorial Pacific Ocean between 160°E and 130°W the AI analysis are greatly modified by the relatively frequent AIREP observations there. The area is characterized by highly persistent winds, mainly from an easterly direction. It is clear that in the AO run, the first guess forecast, together with a very small number of rawinsonde data and some cloudwind data, produces a very different analysis from that obtained from the aircraft winds.

From those findings it seems clear that the aircraft wind data have a larger impact at low than at high latitudes. The polar jets, at least their synoptic scale features, are quite well analyzed over the northern oceans when the aircraft data are excluded.

In order to investigate whether this could be attributed to data redundancy between aircraft and satellite data, we undertook a similar analysis of the two assimilation runs SX and SO from which the satellite data were omitted.

In Fig. 6 the RMS vector difference field between runs SX and SO is shown for the 250 mb wind field. Now the differences are considerably larger over the northern oceans, compare for instance the central Atlantic around 35°N, 45°W or the eastern Pacific around 40°N. The differences are comparable to the total time variability in many areas. In the tropics, the removal of SATOB data has increased the aircraft impact over the central Pacific even further.

It is clear from the second experiment that the somewhat limited impact of aircraft data on mid-latitude analyses is due to the presence of satellite data, in particular high resolution SATEM profiles. There is to some extent a redundancy in the data over the northern oceans such that only limited additional synoptic information is provided by the aircraft winds when SATEM data are used. This additional information may still be very valuable as will be shown below and in Section 4 where we discuss the impact on forecasts. Moreover it should be remembered that the satellite data density during this FGGE period was exceptional. Under normal non-FGGE operational conditions less satellite data are available and consequently a larger beneficial impact of aircraft data may be expected. Finally one must realize that data redundancy is necessary to a certain extent. Without it data checking procedures and quality control would be impossible.

(b) The case of November 11, 00GMT

A systematic day-by-day comparison of the height and wind analyses at 250 mb with and without aircraft data reveals no consistent pattern. Quite often a jet is stronger in AI but also the opposite happens, depending on the available aircraft data. Rather than discussing all analysis differences in detail we concentrate on one particular case, November 11 00GMT. This case was chosen because it exhibited a large impact over the Pacific and was studied also by Barwell and Lorenc (1985).

Fig. 7 shows the vector wind differences between AI and AO for this case. Locally differences of 10 to 20 m/s may be found and in one area even more than 25 m/s. A comparison with Fig. 1, which shows the aircraft data distribution for this case, reveals that the larger differences are found over aircraft lanes. The Asian continent and most of the southern hemisphere are void of aircraft data and show correspondingly minor wind differences, most of the difference being restricted to northern hemisphere Atlantic and Pacific regions. A closer examination of the difference fields reveals that quite often the differences show vortex-pair patterns. This is due to the assumed wind-wind correlation in the modelling of the prediction errors in ECMWF's analysis scheme.

The largest impact is found in the Pacific around 35°N, 175°W. This difference is associated with a trough in this area which in the subsequent days will dominate the flow over the Pacific. In Section 4 we shall show that the analysis difference has a very beneficial impact on the forecast of the development of this system.

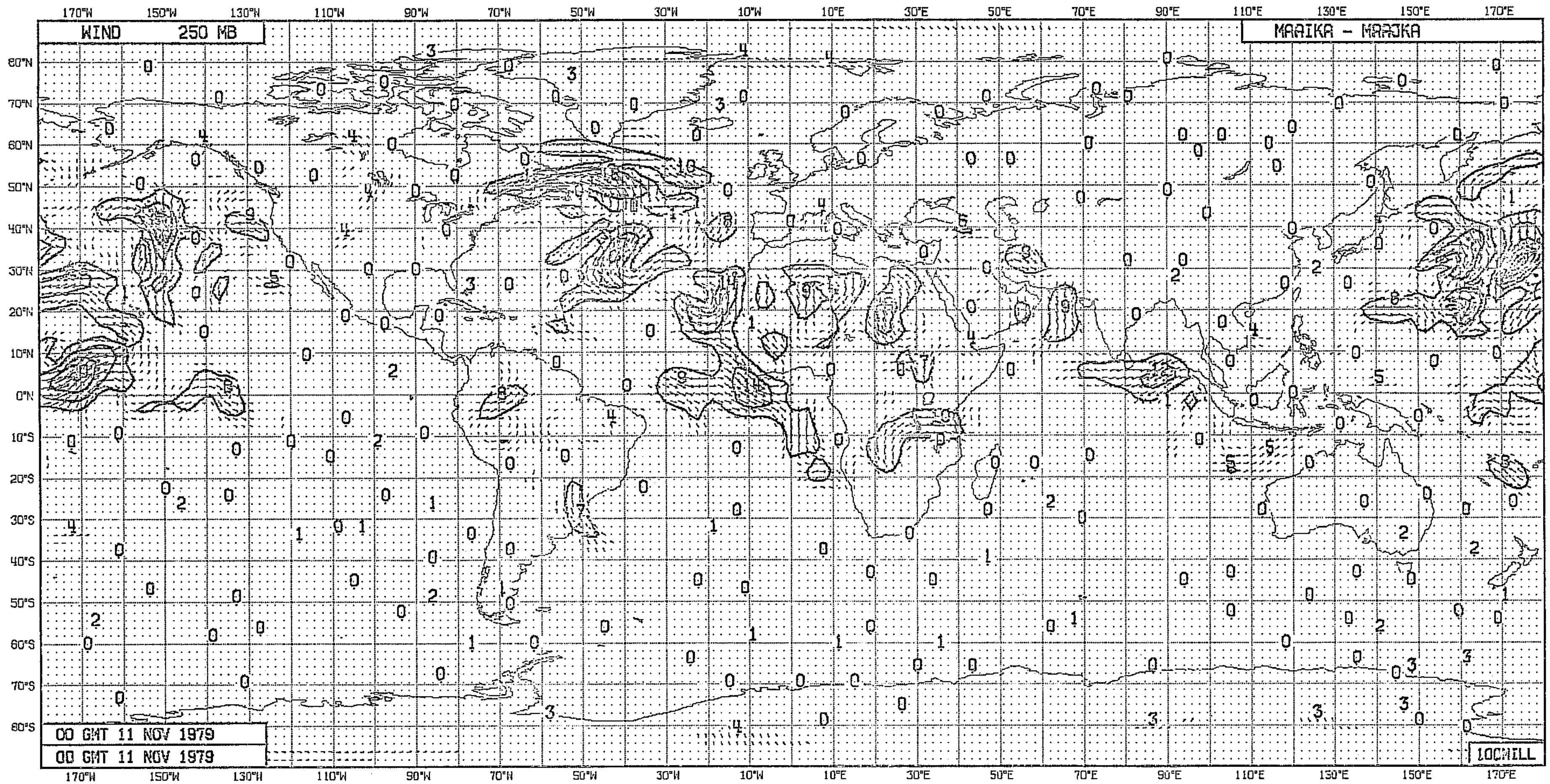

Fig. 7 Vector wind difference between AI and AO for November 11,00 GMT.

Figs. 8a and 8b show the AI and AO 250 mb height analyses of this trough. The AI analysis shows a much sharper trough than AO with much stronger jets along both sides. The western flank shows a maximum windspeed of 59 m/s in AI at 35°N, 161°E against 47 m/s in AO. On the south and east flank the difference is even more marked with a maximum of 58 m/s in AI at 32°N, 176°E against only 32 m/s at the same place in AO. Fig. 9 shows the same trough along with the observations used in each analysis. The differences are apparently caused by AIREPs from a flight along roughly 35°N between 150°E and 180°E and also some AIREPs along 170°E and 180°E. Examination of the individual wind differences between the AIREPs and the analyzed winds shows that on average the AIREP winds agree much better with the AI than with the AO analysis, thus proving that indeed information has been derived from them. It is noteworthy that satellite winds at 33°N, 177°E and 36°N, 159°E also fit the AI analysis better than the AO analysis. Apparently the AO analysis makes only limited use of these satellite winds, due to the rather large wind error (13 m/s) and therefore the small weight assigned to them. In order to rule out the possibility that the analysis difference in this data sparse area might have been caused by differences in the guess field, we repeated the AO-analysis on the basis of the AI guess field. No essential difference between this and the original AO-analysis was found, indicating that possible guess field differences played no role.

As already mentioned in Section 2.2, in the course of this study it was found desirable to perform two assimilation runs without satellite data: SX (with aircraft data) and SO (without aircraft data). Figs. 8c and 8d show the two resulting analyses for this case, together with the corresponding observations.

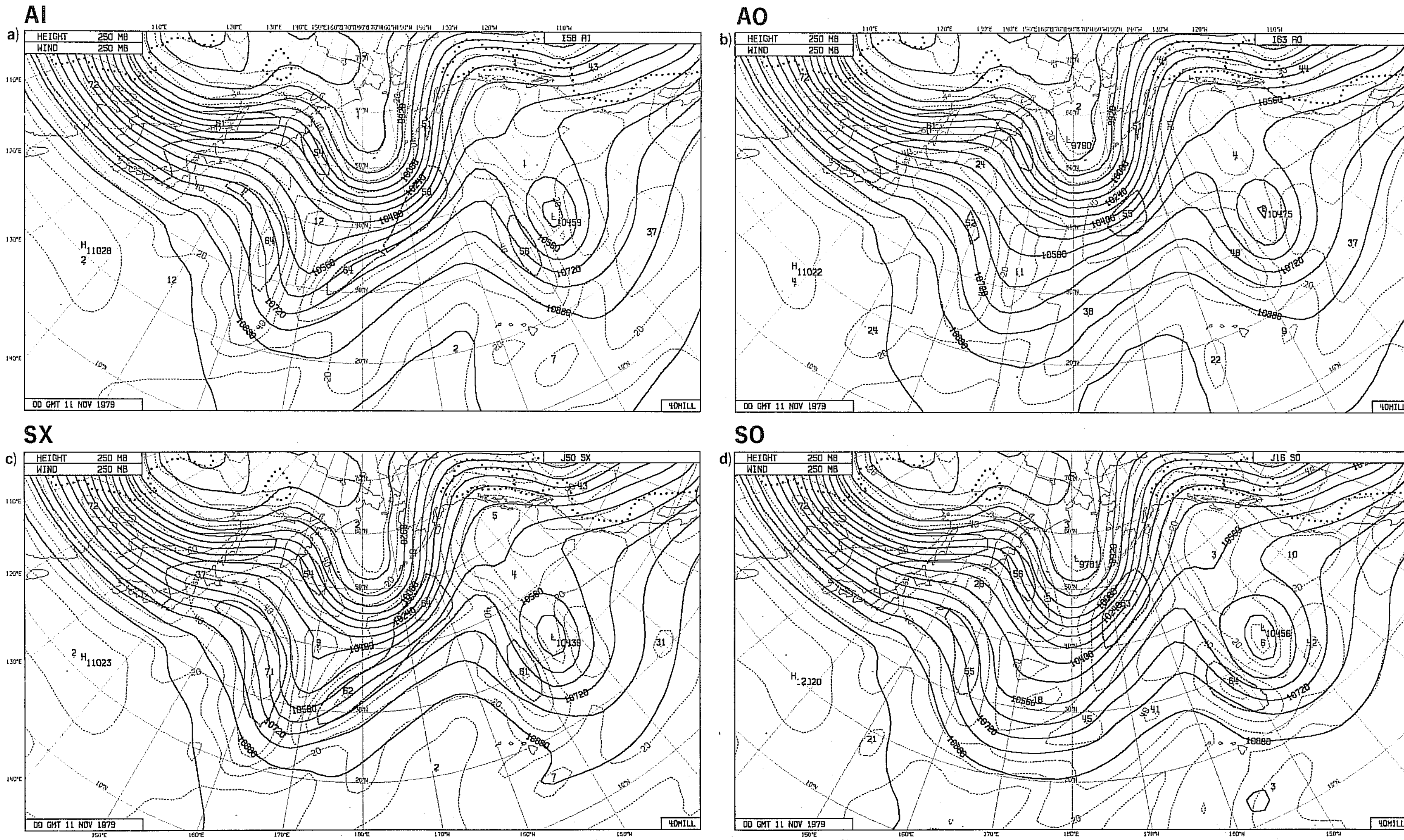

Fig. 8 Analyses of 250 mb height (full lines) and wind speed (dashed lines) for November 11, 00 GMT. Units are in meters and meters per second respectively. a)assimilation AI, b) AO, c)SX, d)SO.

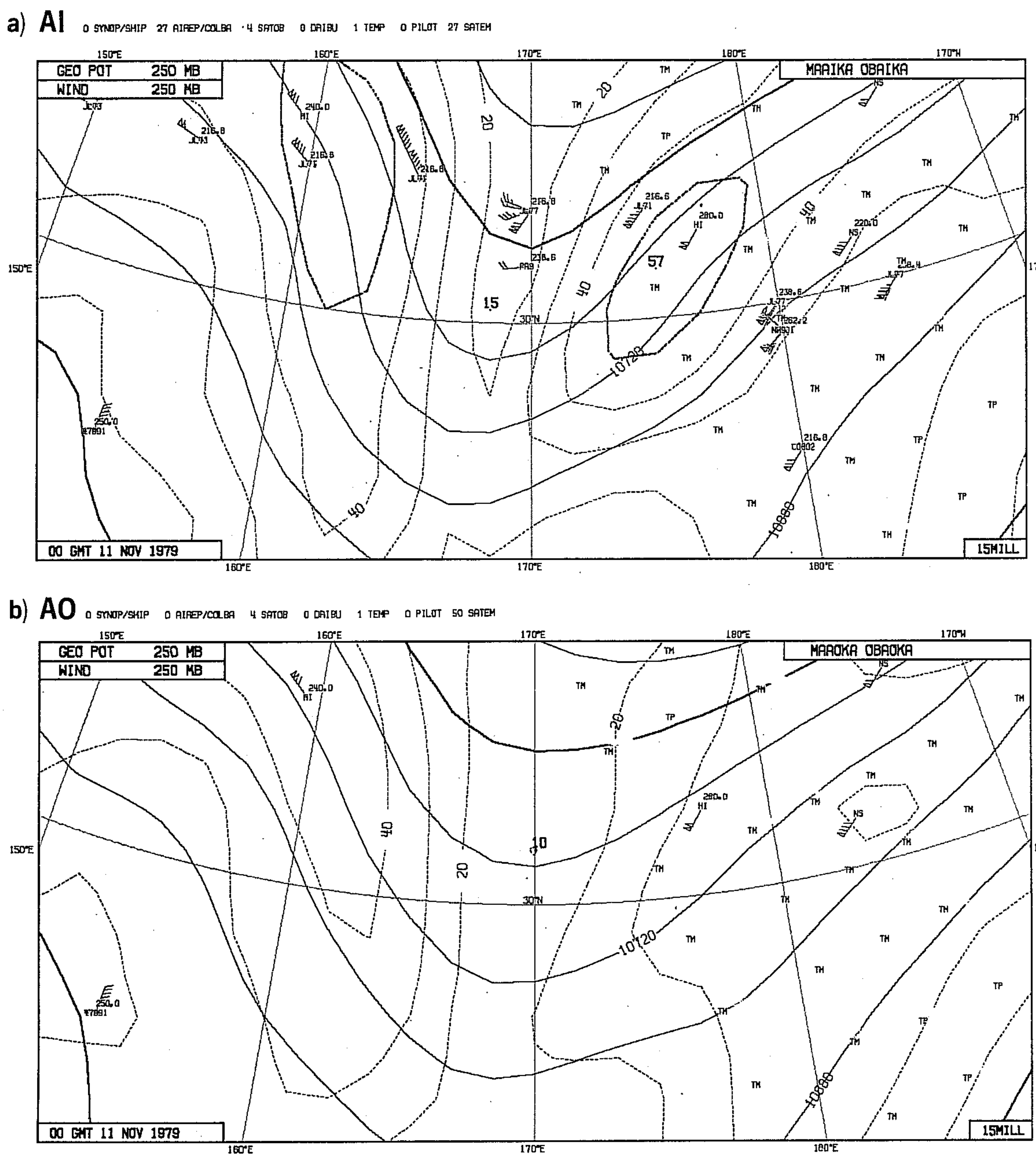

Fig. 9 Analyses AI (top) and AO (bottom) of 250 mb height (full lines) and wind speed (dashed lines) of Nov. 11, together with the wind observations and SATEM positions.

The SX analysis exhibits even stronger jet maxima than AI, indicating that the satellite data in the area, and in particular the SATEM's have a moderating influence on the analysis. It is of course difficult to tell which of the two, AI or SX, is the better. In Section 4 we shall see that, based on the quality of the forecast, there is not much to choose between the two.

The SO analysis is clearly a rather poor one, based as it is on just one TEMP observations outside the trough area. We shall see later that this analysis produces a very poor forecast indeed.

Summarizing we may conclude that in this particular case aircraft data played an essential role in the analysis of an important trough in the Pacific. The only other data sources in the area, satellite winds and temperatures, were unable to properly define the strong jets associated with the trough. Forecasts, based on these analyses and described in Section 4 confirm this.

(c) Analysis differences over North Africa

A very persistent feature in almost all 250 mb wind difference maps is a maximum wind difference over North-East Africa, between 15-30°E and 15-30°N, apparently associated with aircraft data in the area. The maxima are in the range of 10 to 30 m/s. We decided to try and trace the origin of this feature for the two strongest cases: 00GMT 13 and 18 November. Surprisingly the origin was quite different in both cases, although both were related to aircraft data.

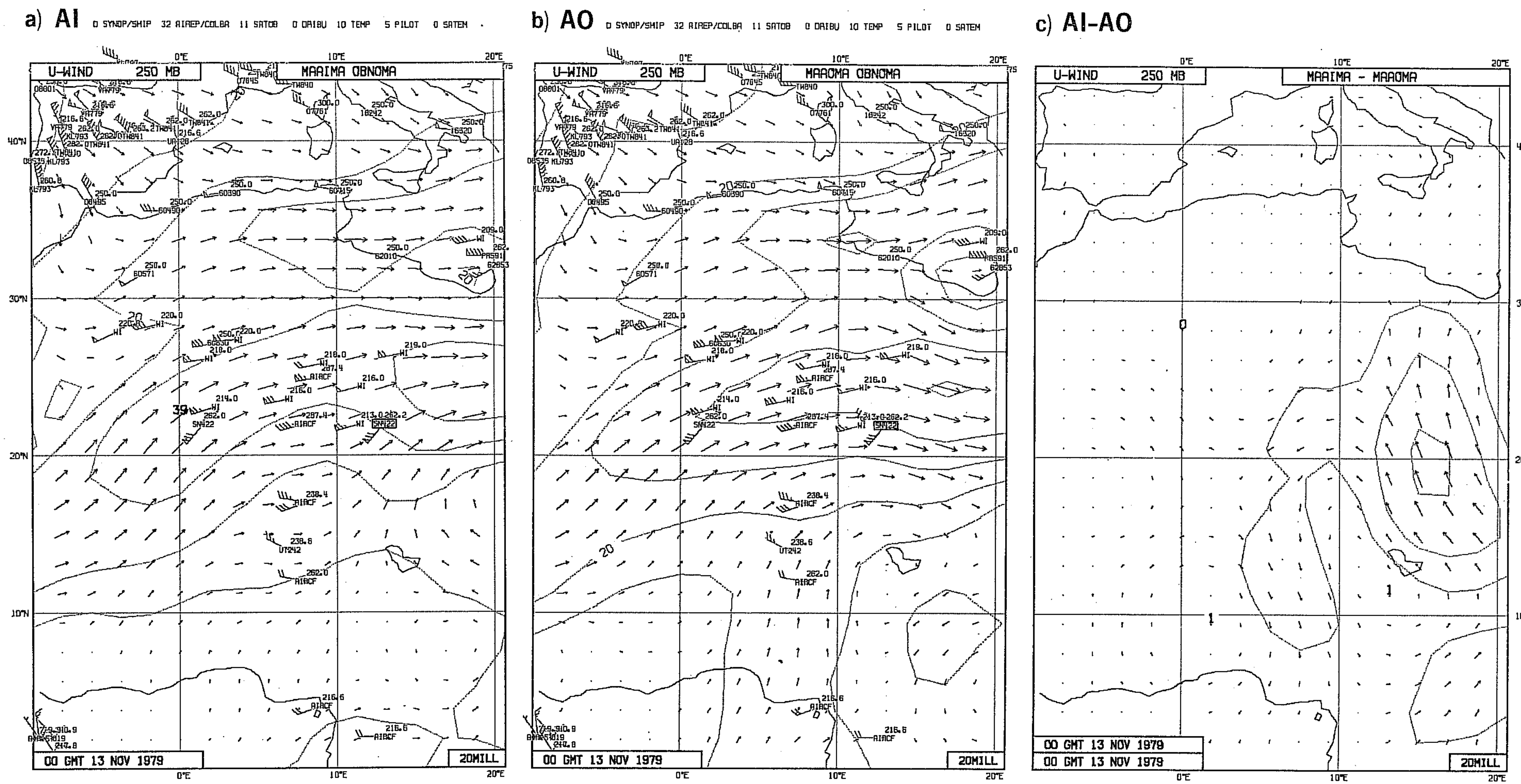

Fig. 10 Wind analyses AI (a) and AO (b) and their difference (c) at 250 mb over Africa for 13 November 00Z. The erroneous AIREP responsible for the analysis difference, SN 422 at 12°E, 22°N, is boxed.

On 13 November 00GMT (Fig. 10) the AI and AO 250 mb wind analyses differed by as much as 32 m/s around 20°N, 15°E. The westerly flow over North Africa is fairly zonal with shallow ridges and troughs. The maximum windspeed in this area is in the order of 40 m/s. The large difference in this case is apparently caused by one single AIREP at 13°E, 22°N from flight SN422, reporting a 70 kts south westerly wind, causing a backing of the wind by as much as 45° compared to the AO analysis over a large area. There are however, doubts about this AIREP. It was reported by an aircraft flying from Abidjan to Brussels. The position of this AIREP is therefore off the expected route. Moreover, there is an AIREP from the same flight about a thousand kilometers to the west in much better agreement with the expected flight path. Therefore the correctness of the first AIREP is suspicious, suggesting that the AO analysis may be better than the AI analysis.

On 18 November 00GMT (Fig. 11) both analyses show a maximum vector difference of about 30 m/s at 27°N, 15°E and at 26°N, 22°E. In AO the flow in that area is characterized by a sharp trough around 18°E. In the AI analysis, however, the trough is much broader and several degrees to the west around 15°E. The information in this case is clearly derived from AIDS data from flight SK963, extending from 50°N, 10°E to 17°N, 28°E. The AI analysis is further supported by some other aircraft data in the area. In this case the AI analysis seems better than the AO analysis.

3.3 The fit of observations to the analyses

We have seen that aircraft data have a considerable impact on the 250 mb wind analyses in the tropics. The good agreement between these data and collocated rawinsonde data, reported by De Jong (1982), suggests that

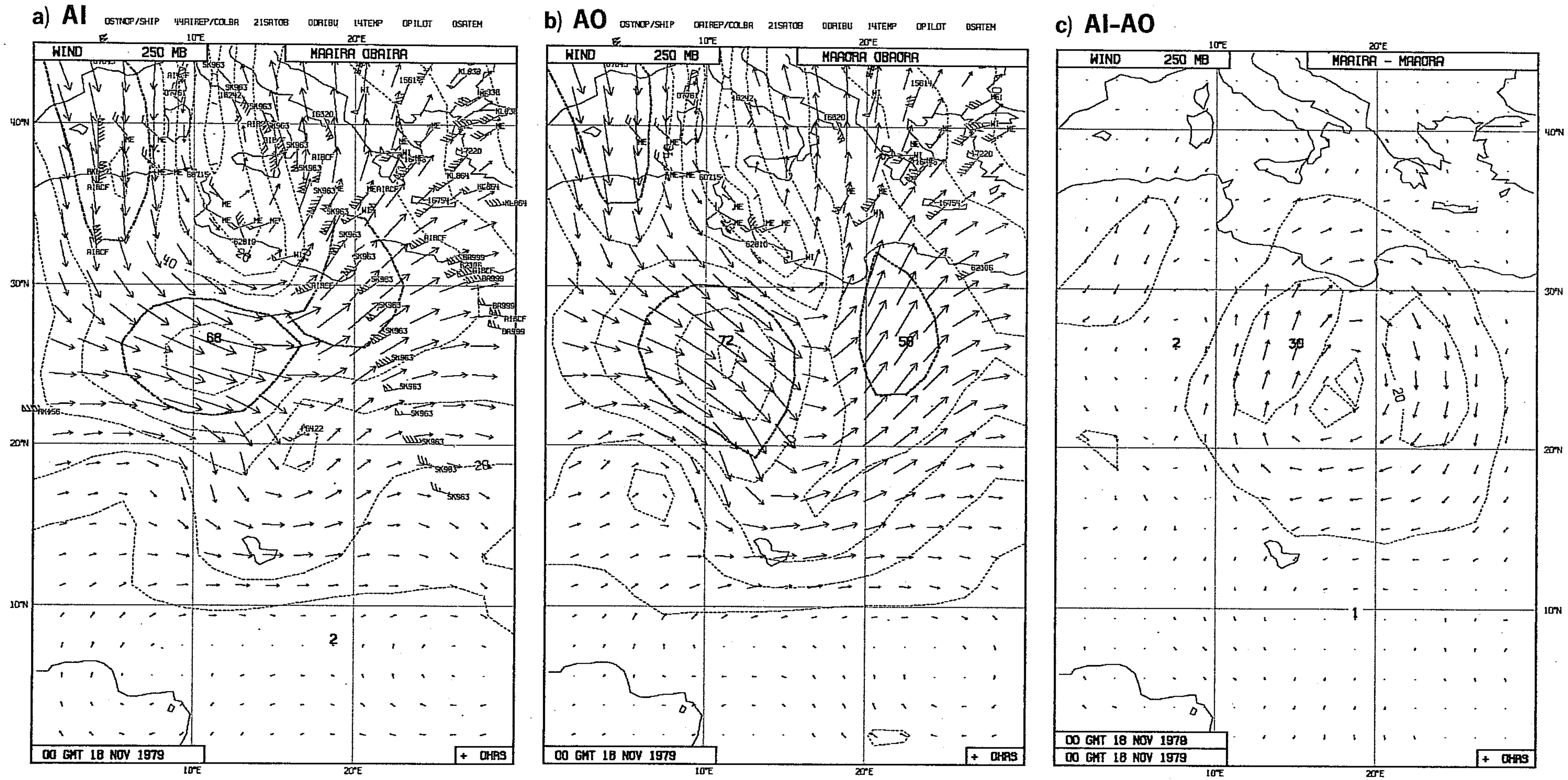

Fig. 11 Wind analyses AI (a) and AO (b) and their difference (c) at 250 mb over Africa for 18 November 00Z.

this impact is indeed beneficial. There are in principle two ways to prove this: the fit of observations to the analyses and the quality of forecasts based on these analyses. In Section 4 we shall investigate the forecast quality. In this section we look at the extent to which the analyses fit the observations.

A multi-variate analysis scheme gives an analysis value which over many realisations minimizes the error of the analysis in a least-squares-sense if the assumed horizontal and vertical correlations are accurate. How closely each observing system fits to the analysis is partly determined by the observational errors assigned to them, and partly by the first guess error characteristics. In the following the fit of observations to the analyses is studied by means of root mean square (RMS) statistics. At each analysis time we computed the RMS difference between analyses and a selected set of observations. From these individual RMS-values, weighted according to the number of observations involved, the mean over the assimilation period is computed. It is this figure that we call "fit". Fig. 12 shows the fit as a function of pressure of three different observing systems to both the AI and AO analyses. Clearly the difference between the fits to AI and AO is negligible. This implies that aircraft data have no detrimental effect on the assimilation of temperature data from polar orbiting satellites and of height and wind data from TEMPs and PILOTs.

On the other hand, as may be expected, AIREP and ASDAR/AIDS winds do fit the AI-analyses better than the AO. This is shown in Fig. 13, again as a function of pressure. First it is clear that ASDAR/AIDS winds have a closer fit than AIREP winds due to the smaller assigned observational

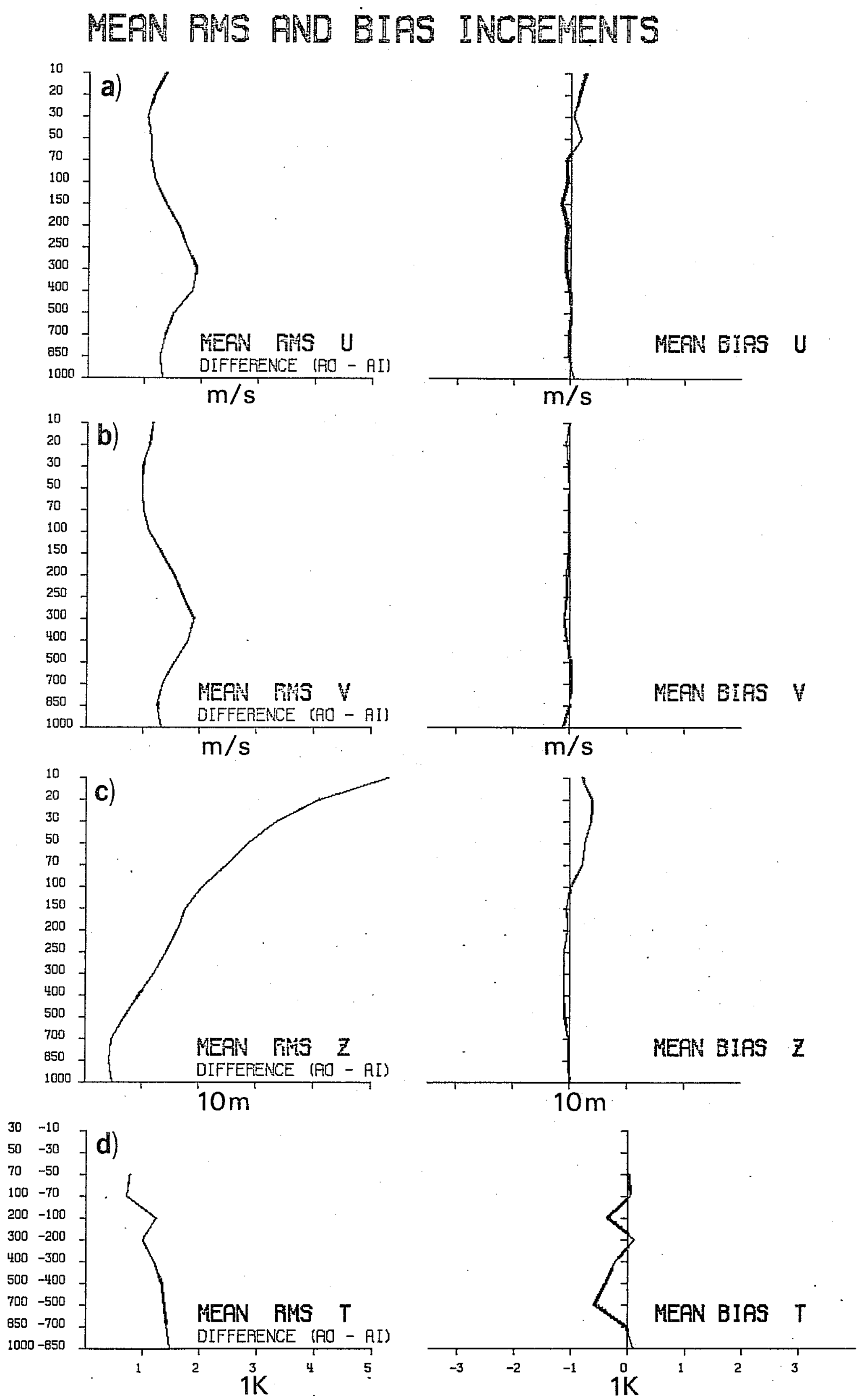

Fig. 12 The fit of AI and AO as a function of pressure against
a) and b) TEMP/PILOT winds (u and v component, Northern Hemisphere only).
c) TEMP heights (global)
d) SATEM layer mean temperatures (global)
Note that the fits of AI and AO can hardly be distinguished.

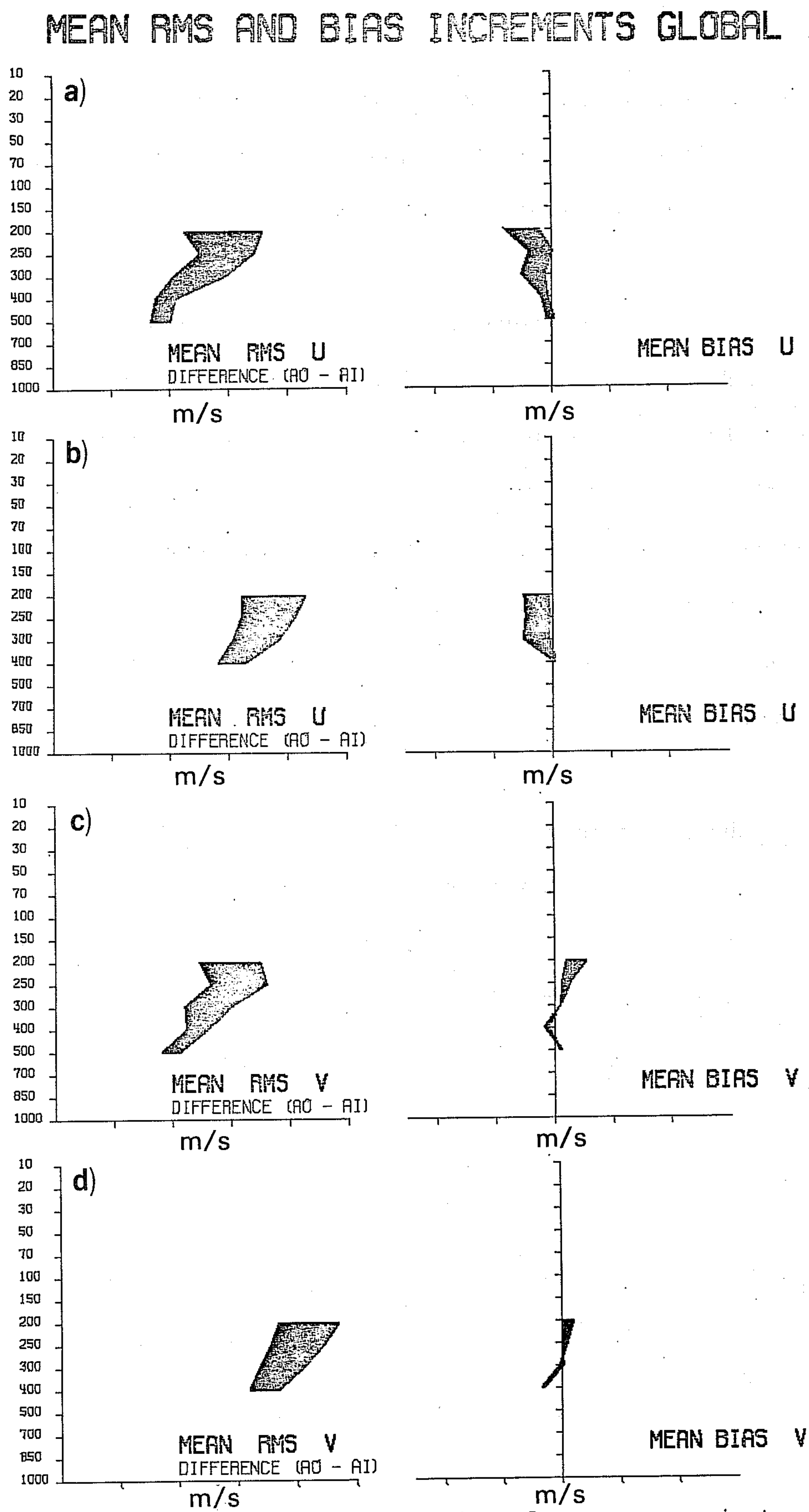

Fig. 13 The fit of AI and AO as a function of pressure against
a) ASDAR/AIDS u-component
b) AIREP u-component
c) ASDAR/AIDS v-component
d) AIREP v-component
The right hand side boundary of the black area is the AO-fit,
The left hand side boundary is the AI-fit. The black area indicates the difference.

error. From the mean bias it may be noticed that if aircraft data are not used, the analysed u-component is underestimated along both AIREP and ASDAR/AIDS tracks at 250 mb by about 1 m/s.

This in itself does not prove that the impact is beneficial. As already noticed in Section 2 we have set aside the relatively small number of ASDAR data for verification purposes. Fig. 14 shows the RMS differences between the analysed and ASDAR u- and v-components for the last 7 days of both series of assimilation. This figure shows clearly a consistent beneficial impact of aircraft data for this period on the analyses along the flight paths. On average the improvement is about 0.8 m/s for both wind components.

3.4 Impact on the Rossby modes

A data system can only improve the specification of the initial state of the atmosphere for numerical weather prediction if it has an impact on the Rossby modes of the model. An impact on the high-frequency inertial gravity waves has no meteorological significance and will either be damped in the course of the subsequent integration or removed in advance by a normal mode initialization.

Daley (1980) concluded from a study on the optimal specification of the initial state for deterministic forecasting, that in general the best variable to measure is the rotational wind component. For horizontal scales representative of synoptic systems this was true in particular for the external and lower internal modes of the model.

Barwell and Lorenc (1985) come to the same conclusion on the basis of arguments derived from simple linear geostrophic adjustment theory. This

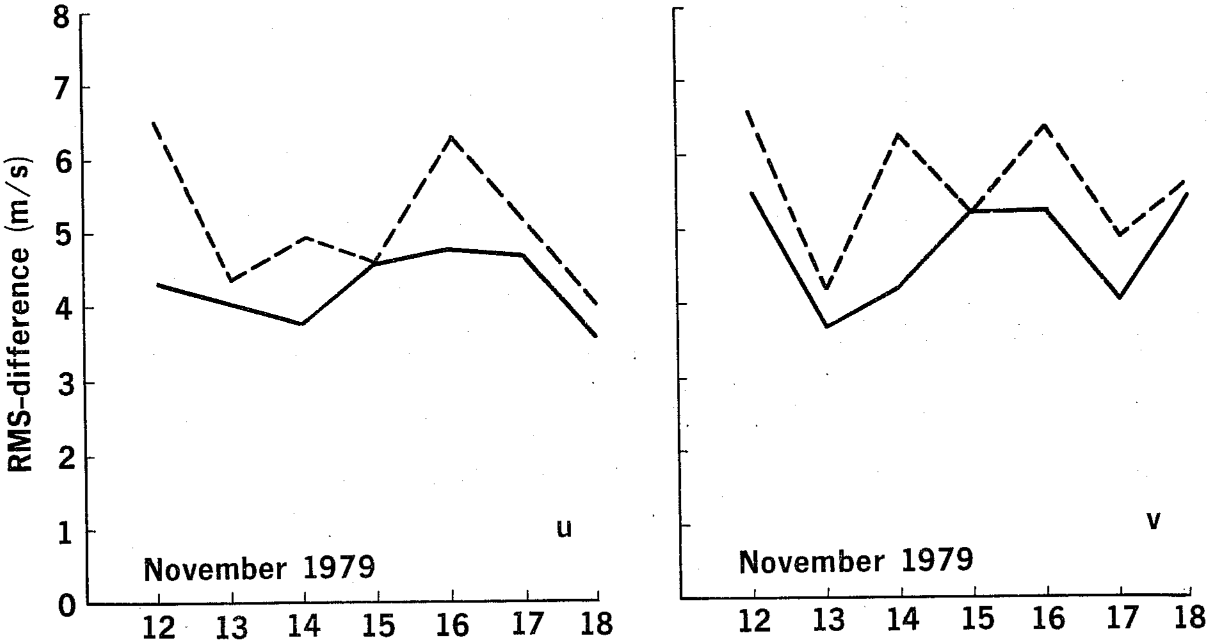

Fig. 14 RMS difference between analysed and observed ASDAR wind components for the period 12-18 Nov. 1979. Full line with aircraft data; dashed- without aircraft data.

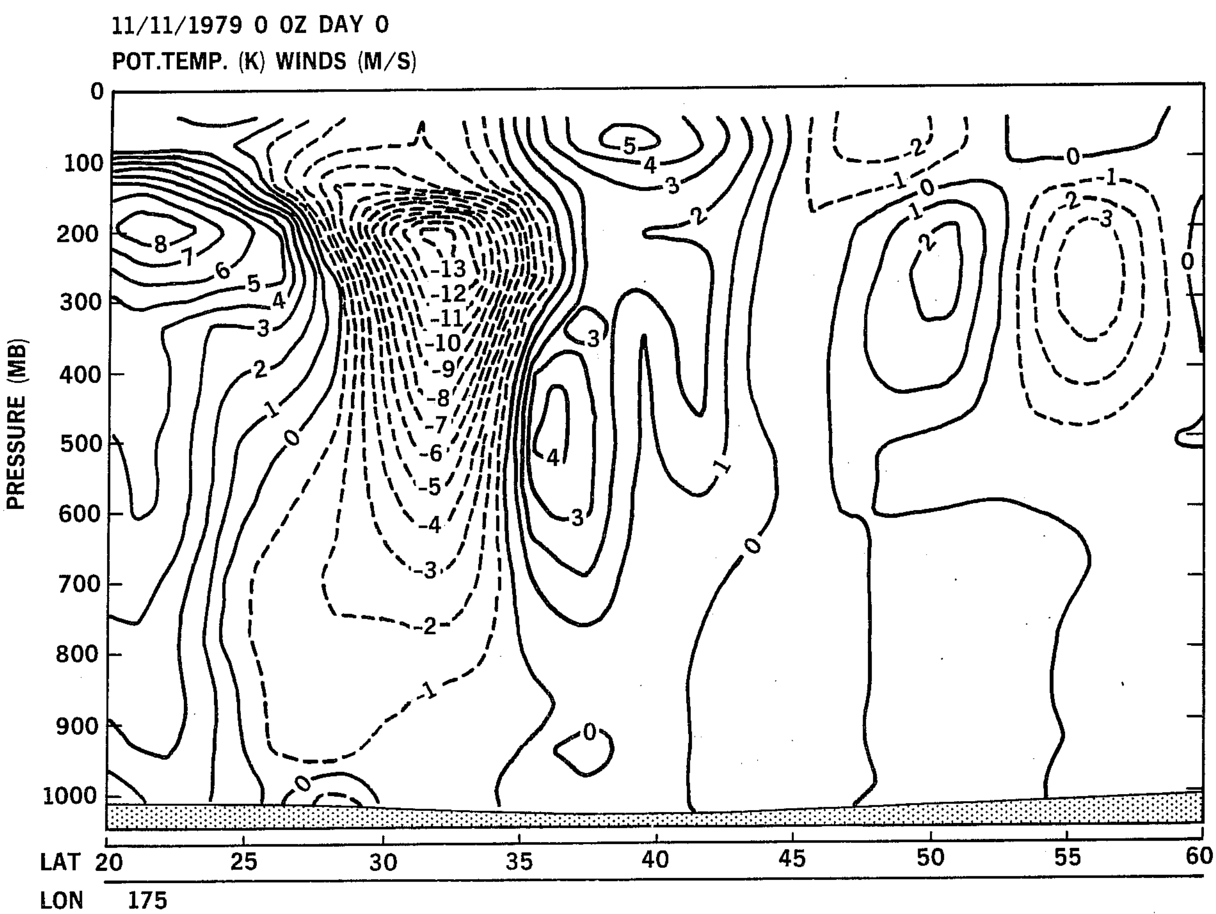

Fig. 15 Latitude-pressure errors section along 175^{o}E through the analysis difference AI-AO of 11 Nov. 1979,00Z. Full lines are the isolines of the zonal wind component.

implies that, in order to benefit optimally from aircraft or other single level wind data, the data assimilation system should be designed so that the information contained in these wind data is directed towards the external or lower internal Rossby modes. This is particularly true in those areas where single level wind data are the only data source available. Sumi (1982), in a recent study on the impact of satellite wind data, found that, in order to achieve this, a proper estimation of the vertical prediction error correlation is essential. In our study the system was not especially redesigned for single level data-assimilation. In particular no special effort was put into the proper specification of the vertical prediction error correlation. We simply took the correlation used in the operational ECMWF system (Fig. 2).

In order to study the impact of the aircraft winds on the different modes of the model, a normal mode analysis of the impact would have been desirable, but for practical reasons we followed a more pragmatic approach. Initially we looked at the vertical structure of the impact on the analysis and then we examined the part of the impact which is left after the normal mode initialization has been carried out.

Fig. 15 shows a vertical cross-section along 175°E through the analysis difference in the Pacific in the case of 11 November, described in Section 3.2. Clearly the difference is essentially barotropic, the vertical distribution rather accurately reflecting the distribution of the correlations used in the analysis scheme. This suggests that, due to the broad structure of the assumed vertical prediction error correlation, the impact is mainly on the external Rossby mode; the more so because it was found that the height and wind differences are in approximate geostrophic balance.

In order to study to what extent the impact survives the subsequent initialization, we computed the RMS differences between the ASDAR wind components and the 00GMT guess fields, and both the uninitialized and the initialized analyses; the results were then averaged over the last 7 days of both assimilation runs. These RMS differences are shown in Table 3, from which it becomes clear that the impact is also present to a large extent in the initialized analyses.

Table 3 Time averaged RMS difference (m/s) between observed ASDAR wind components and the guess field, and both the uninitialized and initialized analysis.

	guessfield		unitialized analyses		initialized analysis	
	u	v	u	v	u	v
AI	5.20	5.91	4.22	4.75	4.38	4.82
AO	5.56	6.29	5.09	5.55	5.15	5.63

These results suggest that the ECMWF assimilation scheme is able to direct information from the aircraft wind data to the proper modes of the model, although perhaps not in the best way possible. Further work on the optimal design of assimilation schemes for this purpose is desirable. However conclusions concerning the importance of certain details of the assimilation scheme may be drawn already from an intercomparison between our results and those from a similar study, conducted by Barwell and Lorenc (1985) at the U.K. Met. Office. Such an intercomparison was carried out by them and is summarized in the next section.

3.5 Intercomparison with the U.K. Met. Office experiment

In this section we present a summary of an intercomparison by Barwell and Lorenc (1985) between their results and ours; details are given in their paper and earlier reports (Barwell, 1982 and Lorenc, 1982). Barwell and Lorenc used an experimental and non-operational assimilation scheme developed at the U.K. Met. Office (UKMO). Data are assimilated by repeated interpolation and insertion into the model at each timestep for a 6-hour period. The data are interpolated univariately and horizontally to the model's grid using a one timestep forecast from the previous estimate as first-guess giving steadily increasing weight to the observations. Unlike the ECMWF system, there is no vertical coupling in the UKMO system, the data being inserted only at the nearest level. The analysis is not followed by an initialization. The model used is an 11-layer General Circulation Model described by Saker (1975).

Parallel assimilations with and without aircraft data were run for 5 days. They were started from the ECMWF level III-b analysis for 00GMT 9 November 1979 with the first observational data being assimilated at 06GMT. Subsequently 4-day forecasts were run from 11 November 00GMT. A comparison of the UKMO and ECMWF analyses showed that the impact of aircraft data in both experiments on the 250 mb analyses was substantial, in particular on the analysis of the trough near 170°E. The effect on the UKMO analyses was somewhat larger than the ECMWF analyses. As noted by Barwell and Lorenc, however, the impact of aircraft data on the 48h forecast was larger and more beneficial in the ECMWF experiment. A forecast with the UKMO model from the ECMWF AI analysis showed more impact of aircraft data than one from their analysis, indicating that the loss of information was due to differences between the analysis techniques. These findings led Barwell and Lorenc to conclude that forecasts from their analyses do not retain all the initial information derived from the aircraft wind observations, substantial part of it is lost during the initial stages of the forecasts. A further analysis of this phenomenon demonstrated that the difference between their wind analyses with and without aircraft is almost entirely ageostrophic, to a substantial extent divergent, and mainly affects the higher vertical modes. By contrast, the ECMWF wind analysis differences are in geostrophic balance and in lower order vertical modes. Based on arguments derived from linear geostrophic adjustment theory, Barwell and Lorenc concluded that these differences explained the fact that in our experiment the analysis improvements were retained in the subsequent forecast, whereas up to half of the improvements was lost in theirs. An analysis of the forecast differences over Europe corroborated this conclusion.

These results are consistent with our findings concerning the effect of the initialization on the impact, described in the previous section, and confirms Sumi's (1982) conclusion that a data assimilation scheme should be designed so as to direct the information to the graver Rossby modes.

We may therefore tentatively conclude that a multi-variate three-dimensional analysis scheme with a properly adjusted vertical structure function is advantageous for a successful assimilation of single level wind data.

4. IMPACT OF AIRCRAFT DATA ON FORECASTS

4.1 Introduction

From the two parallel sets of analyses four pairs of forecasts were run out to ten days. As already explained in the previous section, later on two new parallel sets of analyses were produced, this time without satellite data. Again forecasts were run on the same four pairs of analyses.

The four selected pairs of analyses from which forecasts were run are: 10, 11, 13 and 16 November 1979, 00GMT. The 11 November 00GMT was chosen deliberately because the two analyses for that time showed a considerable difference around a trough over the Pacific as already described in Section 3. The other three dates were selected rather arbitrarily.

In sections 4.2 and 4.3 we shall describe the impact from a synoptic point of view thereby paying special attention to the case of 11 November. Furthermore, in section 4.4, we shall pay attention to the peculiar way the initial analysis differences propagate through the forecasts. In section 4.5 we compare the forecasts in an objective way by examining the RMS differences between them. The next two sections 4.6 and 4.7 will be devoted to objective verification of the forecasts, both against analyses and against observations. Finally in section 4.8 we shall summarize the discussion and draw some general conclusions.

4.2 Synoptic Description of the Impact

In this Section we evaluate the synoptic impact of the aircraft wind data on the four 10 day forecasts. We shall do this by evaluating the differences between both initial analyses and by comparing the 2 and 5 day forecasts with the verifying analyses. We restrict ourselves to the northern hemisphere analyses and forecasts of the 250 mb height and the sea level pressure. For the cases of 10, 13 and 16 November this will be done briefly, but we shall finish with a comprehensive intercomparison for the case of 11 November. Where necessary we have also used information from the runs without satellite data.

(a) 10 November 1979

The two sets of height analyses, both near the surface and at 250 mb (Fig. 16), are very similar and this is typical of most of the cases we studied. A closer examination of the differences reveals that the sea level pressure differences are limited to ± 1 mb, whereas at 250 mb the largest differences of more than 40 m are found over the west Pacific and North-east Africa. In the course of the forecast the initial difference over Africa disappears but those over the Pacific develop into substantial forecasting differences which affect Europe after five days.

An intercomparison of the 2-day sea level pressure forecast (Fig. 17) shows that the only noticeable differences are associated with a low near 170°E, 25°N; but both forecast fail to predict this system properly. At 300 mb the largest 2-day forecast differences are associated with the trough at 170°W. In AI the trough is predicted slightly further to the east. It is not easy to establish which of the two is the better. Of the two forecasts without satellite data, SX (with aircraft data) is again slightly retarded compared to AO, but in SI (without aircraft data) the phase of the system is clearly wrong.

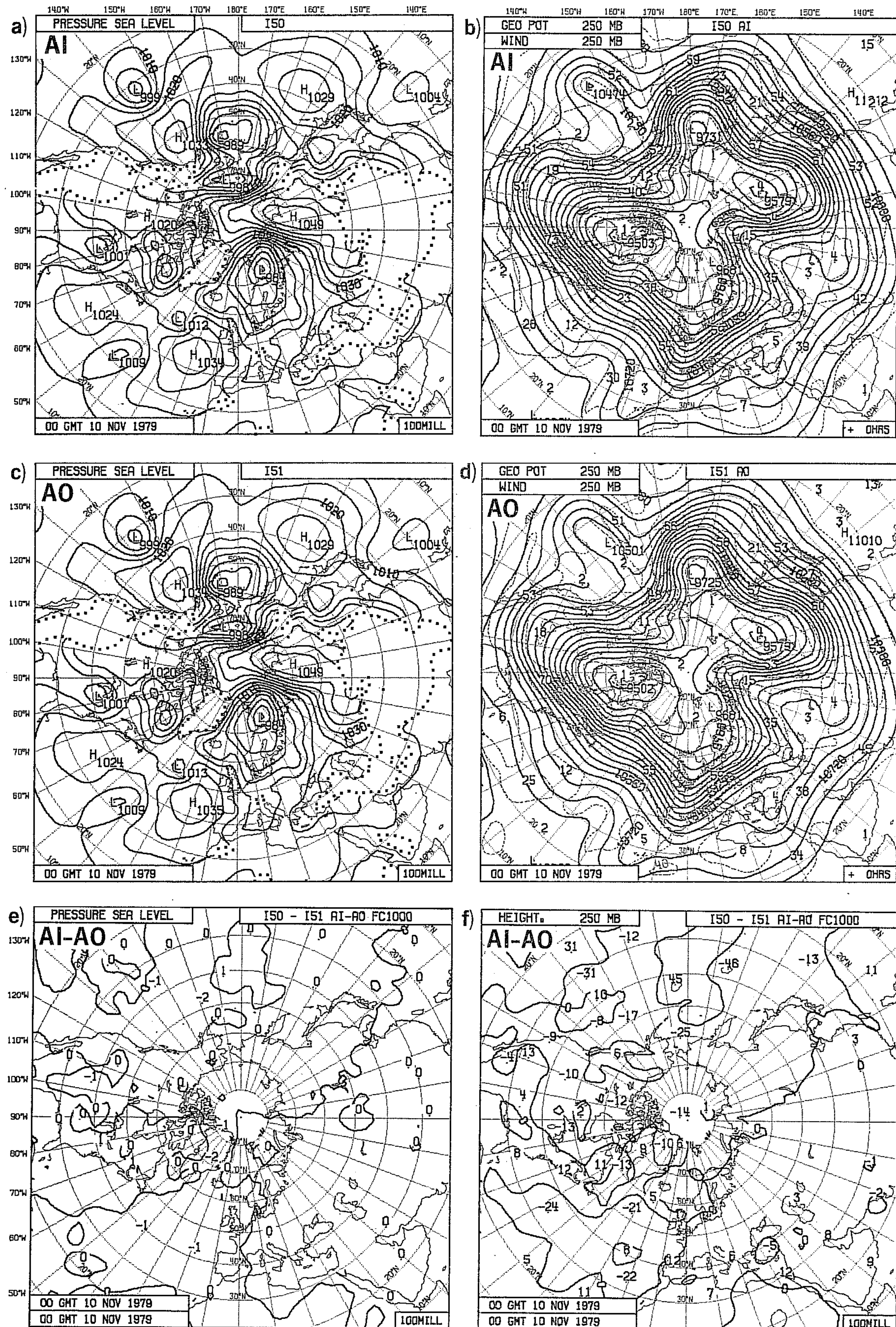

Fig. 16 Mean sea level pressure (left) and 250 mb height (right) analyses for 10 Nov. '79,00Z. Top: AI analyses; middle: AO analyses; bottom: analysis differences.

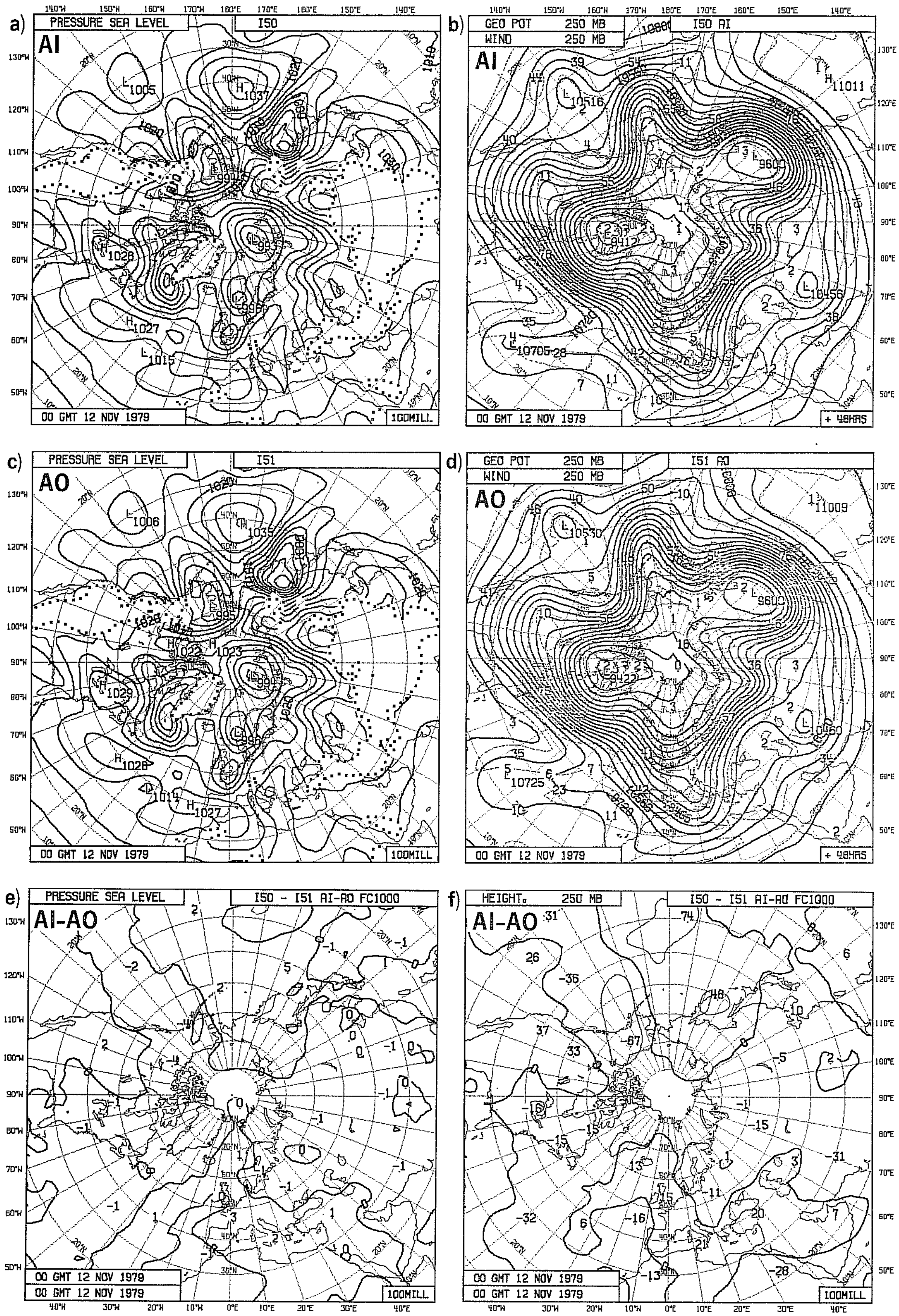

Fig. 17 Idem for day 2 forecast, valid for 12 Nov.,00Z.

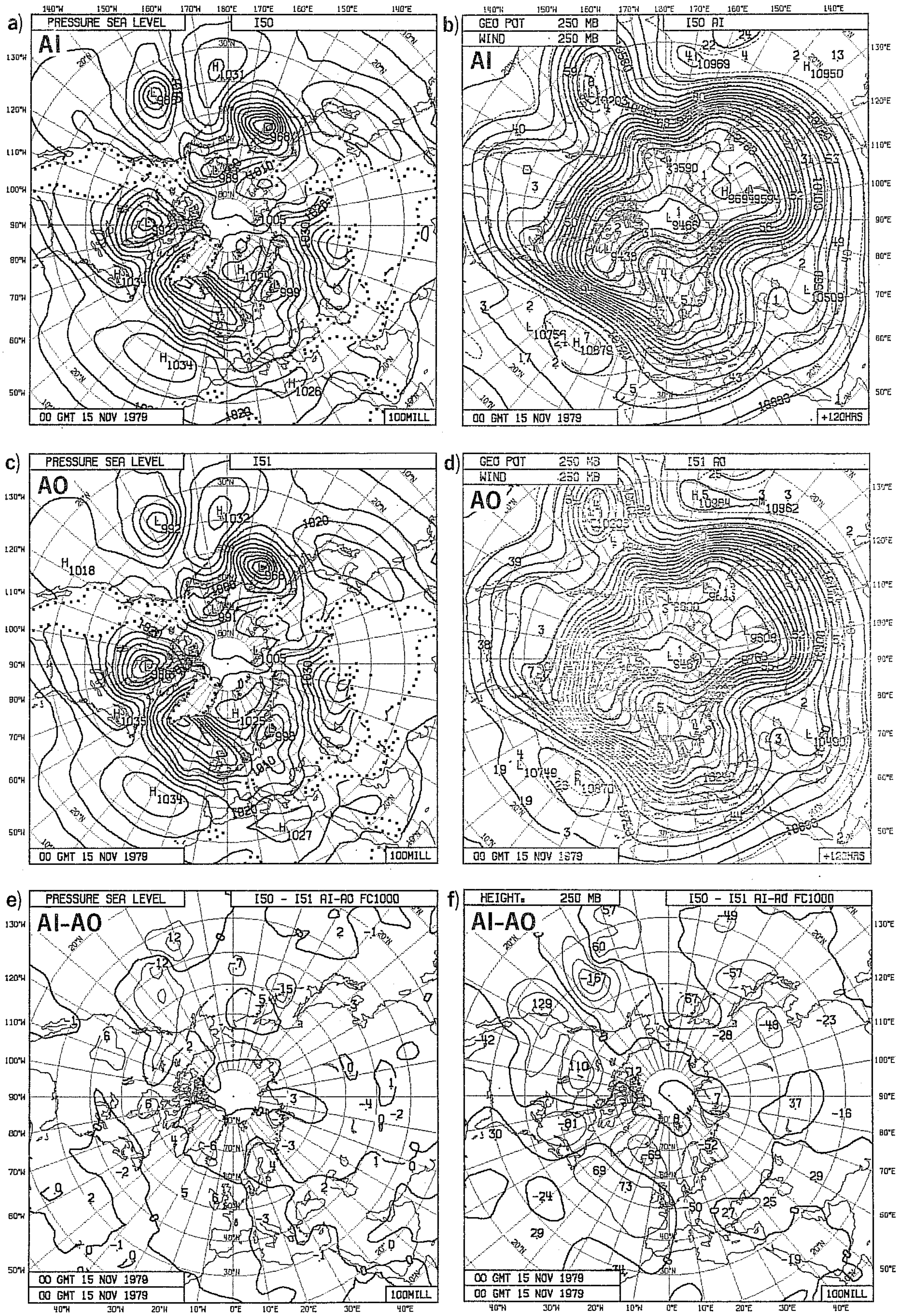

Fig. 18 Idem for day 5 forecast, valid for 15 Nov.,00Z.

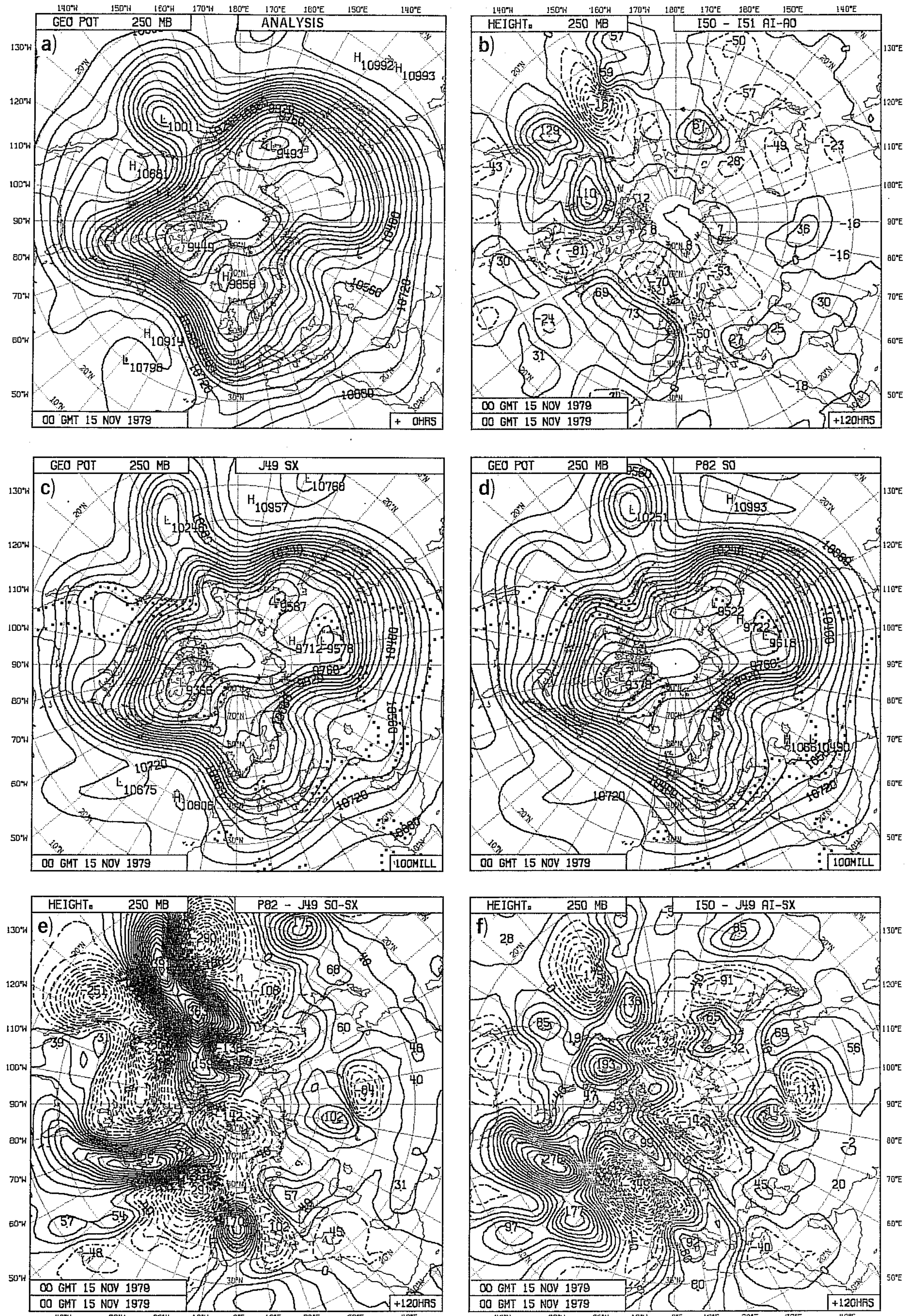

Fig. 18a 250 mb height forecast for day 5, valid for 15 Nov., 00Z.
a) verifying analysis, b) AI-AO, c) SX, d) SO, e) SO-SX, f) AI-SX.

After five days (Fig. 18) the sea level pressure forecast differences are still mainly limited to the low near 140°W. Neither of the two runs predicts its position very well but AI is better. At 250 mb we find the same behaviour: the phase of the cut-off low near 150°W is certainly best predicted by AI. At this level the forecast differences now extend to the European area. In particular the mid-Atlantic ridge is slightly better predicted by AI although it must be admitted that SX is even better. In all cases SO is by far the worst forecast.

This case study shows that for the forecasts starting from 10 November the position of some of the systems is predicted slightly better with the aircraft data than without, although the differences between both runs are smaller than the difference between either of the two forecasts and the verifying analyses. When the satellite data are left out the impact of the aircraft data is much larger, SX doing almost as well as AI but SO producing very poor forecasts. This is shown convincingly in Fig. 18a for the day 5 forecasts.

(b) 13 November 1979

As in the previous case the two surface analyses are virtually the same (Fig. 19), the differences being limited locally to 1 or 2 mb. Also in this case the most important differences at 250 mb are associated with the two pacific cut-off lows. From this area a wave train of height differences develops in the course of the forecast which affects the European area after about five days.

The 250 mb two day forecasts (Fig. 20) show much the same picture as in the previous case. Again a slight phase difference is noticeable of the trough near 145°W with a marginal advantage for AI. The forecast with aircraft data also builds up a slightly stronger ridge behind this trough, again in closer agreement with the verifying analysis. Apart from these minor advantages of AI, the differences are very small, in any case much smaller than the differences with the verifying analysis. Also near the surface there is not much to choose between both forecasts except for a slight negative impact near 180°E, 55°N, where the low is better forecast by AO. Even SO appears slightly better here.

By day 5 there are major forecast differences over the Labrador/Greenland area (Fig. 21). Neither of the two forecasts is very good but perhaps AO is slightly better. AI fails to develop a second generation low at the surface south of Greenland contrary to AO. The surface low north of Scotland is forecast poorly by both runs. The phase, structure and intensity of the 250 mb trough over Western Europe is, however, better in the AI run.

When satellites were excluded (SX and SO; Fig. 21a) phase errors were much larger; this is particularly apparent in the SO run which has gone badly wrong over eastern Canada as well as over western Europe.

Finally we found that the AI analysis error, which was caused by a faulty AIREP over Africa (see Section 3), had no significant impact on the forecasts. The analysis difference induced a low-amplitude wave train of forecast differences over eastern Europe into Asia but its effect on the forecasts was hardly noticeable.

The results from this case are not very conclusive- some features, such as the phase of the western European trough being better predicted with aircraft data, other features, in particular a second generation low over

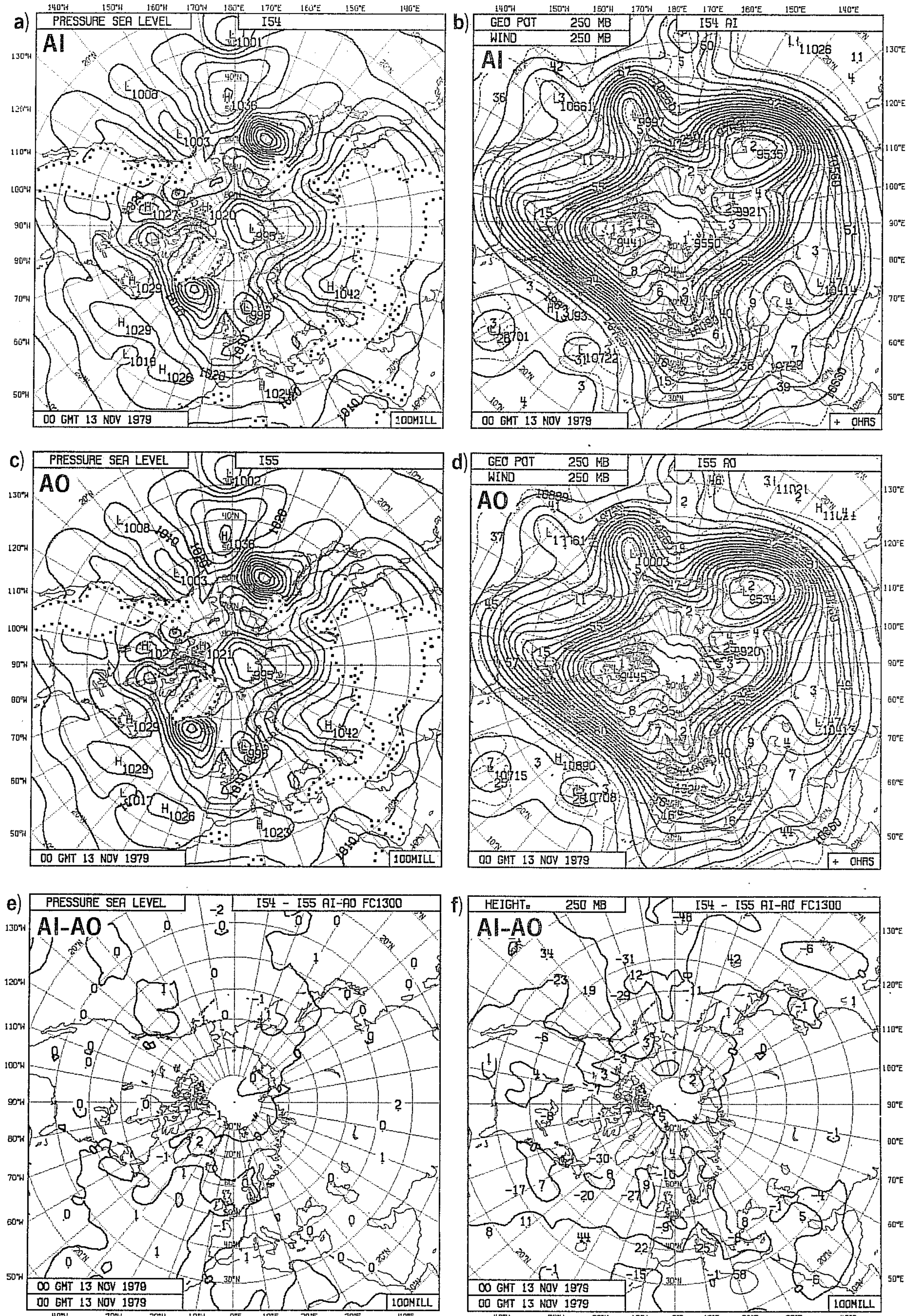

Fig. 19 Mean sea level pressure (left) and 250 mb height (right) analyses for 13 Nov.'79,00Z. Top: AI analyses; middle: AO analyses; bottom: analysis differences.

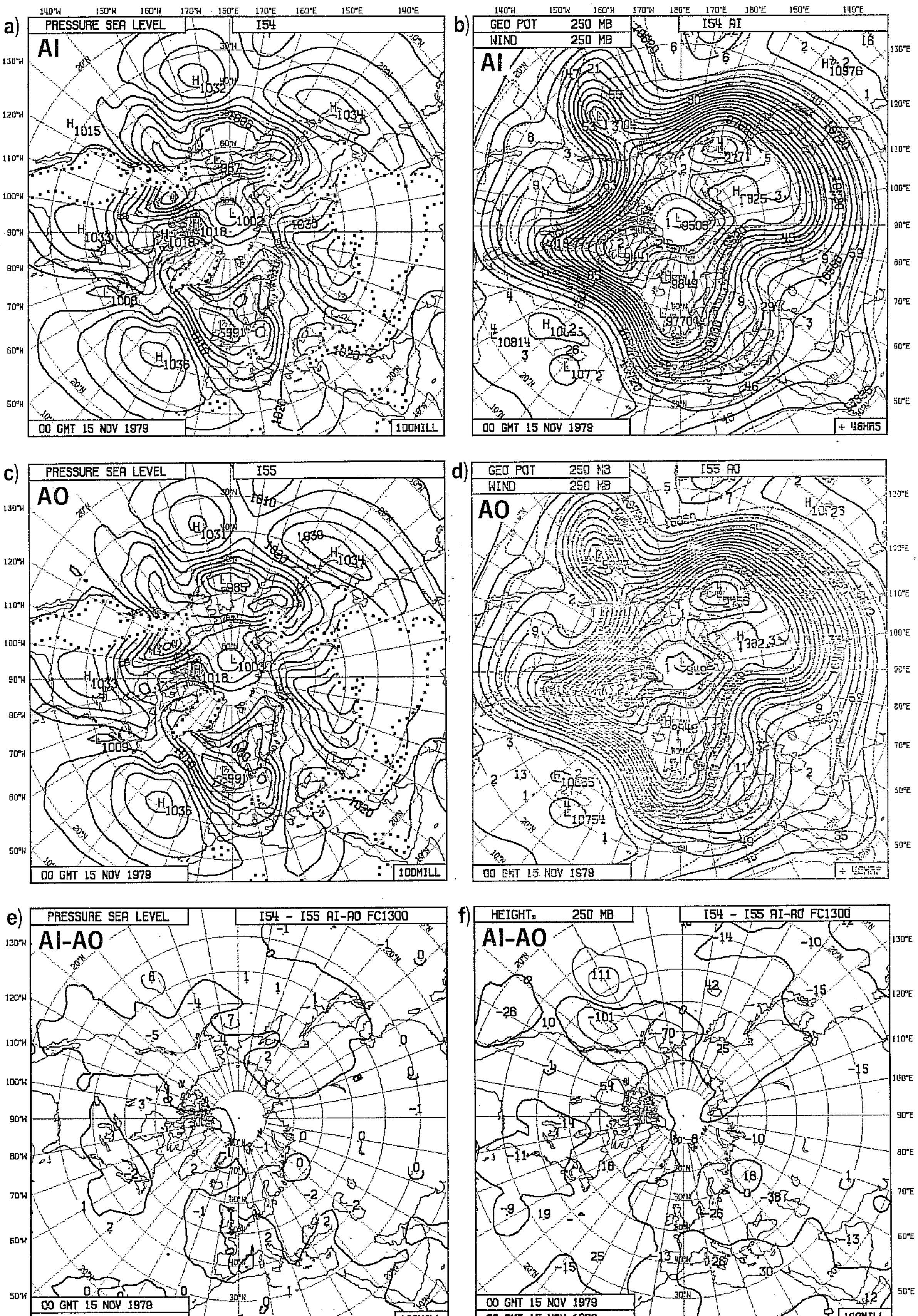

Fig. 20 Idem for day 2 forecast, valid for 15 nov.,00Z.

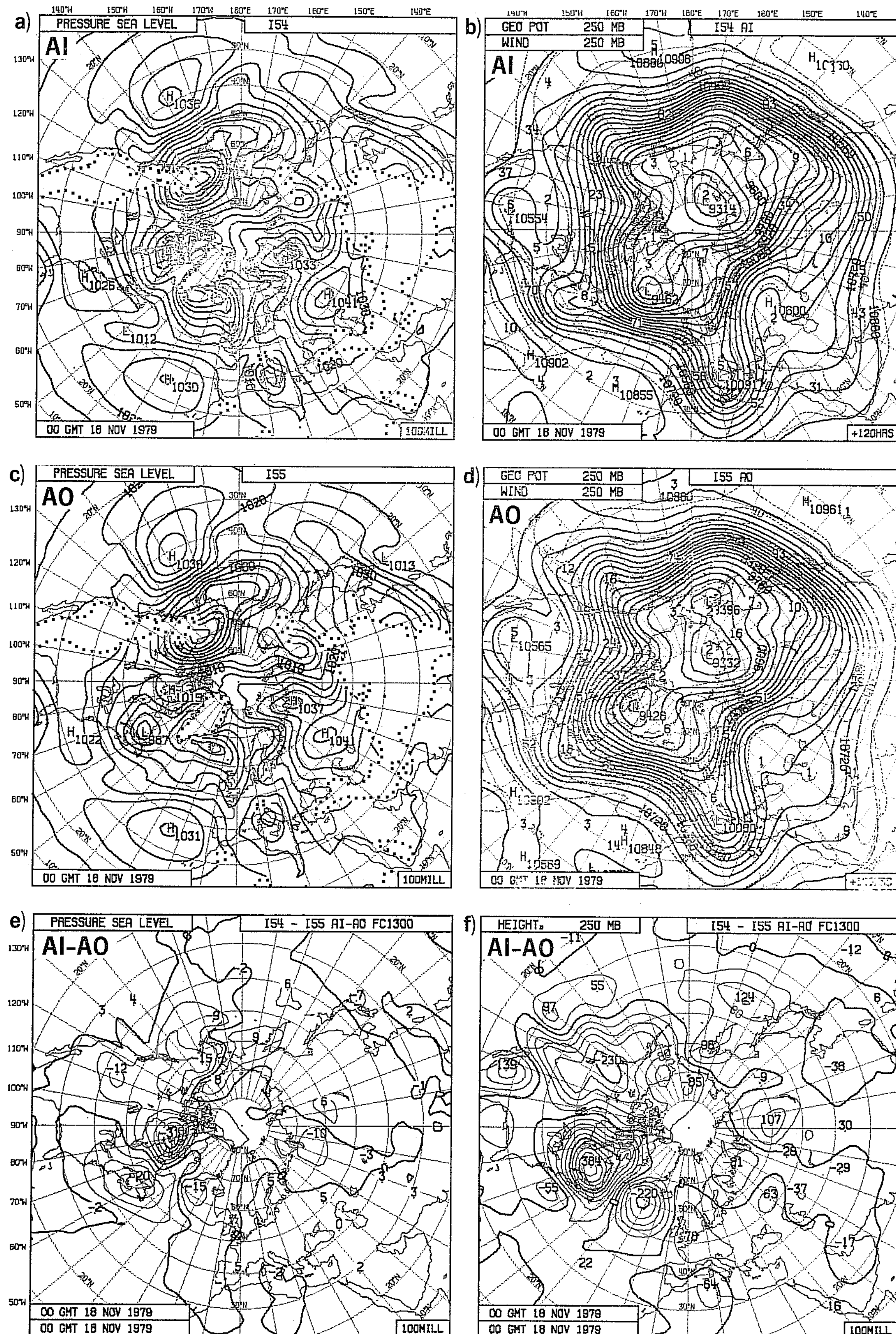

Fig. 21 Idem for day 5 forecast, valid for 18 Nov.,00Z

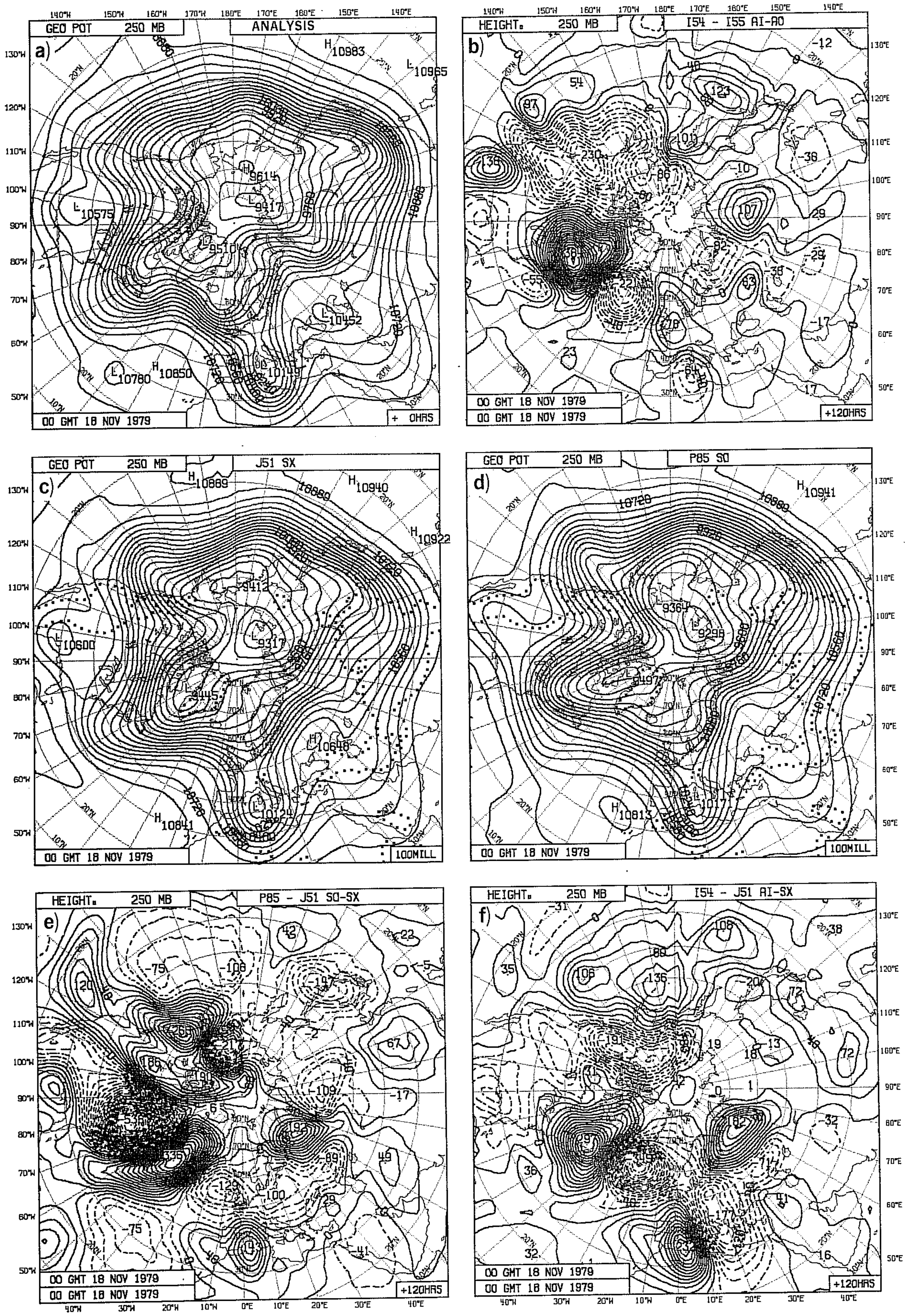

Fig. 21a 250 mb height forecast for day 5, valid 18 Nov., 00Z.
a) verifying analysis, b) AI-AO, c) SX, d) SO, e) SO-SX, f) AI-SX.

the east Atlantic, show a negative impact. However without satellite data the aircraft data impact is very clear and positive.

(c) 16 November 1979

In this case the two forecasts AI and AO are remarkably similar. Even at day 5 the largest differences are less than 200 m at 250 mb and less than 16 mb at the surface. But by then neither of the two forecasts is particularly good. The differences between both are insignificant compared to the forecast errors. The only clearly discernible difference, visible as a phase difference over Karelia on day 4, can be traced to the combined effect of an analysis difference in the Labrador Sea, an area well covered by aircraft data, and a wave train of differences initiated by an analysis difference over Africa.

The same forecast difference is found, with increased amplitude, when satellite data were excluded from the assimilation. In that case, however, there are other large analysis/forecast differences, particularly one over the Central Pacific Ocean. This latter difference propagates downstream causing a large phase error off Newfoundland by day 4, and a marked forecast difference over the British Isles.

The conclusion drawn in this case is that the impact of the aircraft data is hidden by the availability of satellite data over the Pacific, while the Labrador Sea and African analysis differences are caused by aircraft data and these features had not been detected by the satellite data alone.

4.3 The case of 11 November 1979

(a) Description of the impact

In Section 3.2 we described in detail the analysis difference for this case which occurred in the Central Pacific and its relation with aircraft observations in the area. In particular it was found that the slope of a trough near 170° E was considerably different in the two analyses. This was the reason for studying the impact of this analysis difference on the forecasts in some more detail. Moreover this case was also studied by Barwell and Lorenc (1985) who carried out an intercomparison between their results and ours.

The 48 h forecasts of the 250 mb height are shown in Fig. 22 along with verifying analysis. Although both forecasts fail to produce a closed circulation in the trough at 160°W, the shape and position of the trough has definitely improved in AI. This is also true for the shape of the low near 180° E, 25°N. However an even more convincing improvement is evident from the sea level pressure forecasts in Fig. 23. The structure of the high and the depth of the developing low south of it is clearly much better in AI.

Another example of a positive impact of aircraft data in these forecasts can be found over Europe (Fig. 24). Initial analysis differences over the North Atlantic and off the African coast rearrange themselves and develop in such a way as to cause a negative difference between the AI and AO 48 h forecasts near the Straits of Gibraltar. The AI forecast correctly predicts a trough over Spain and north Africa between 0 and 15°W.
We have not attempted to evaluate the impact on the day 5 forecasts because by then both forecasts were already very poor.

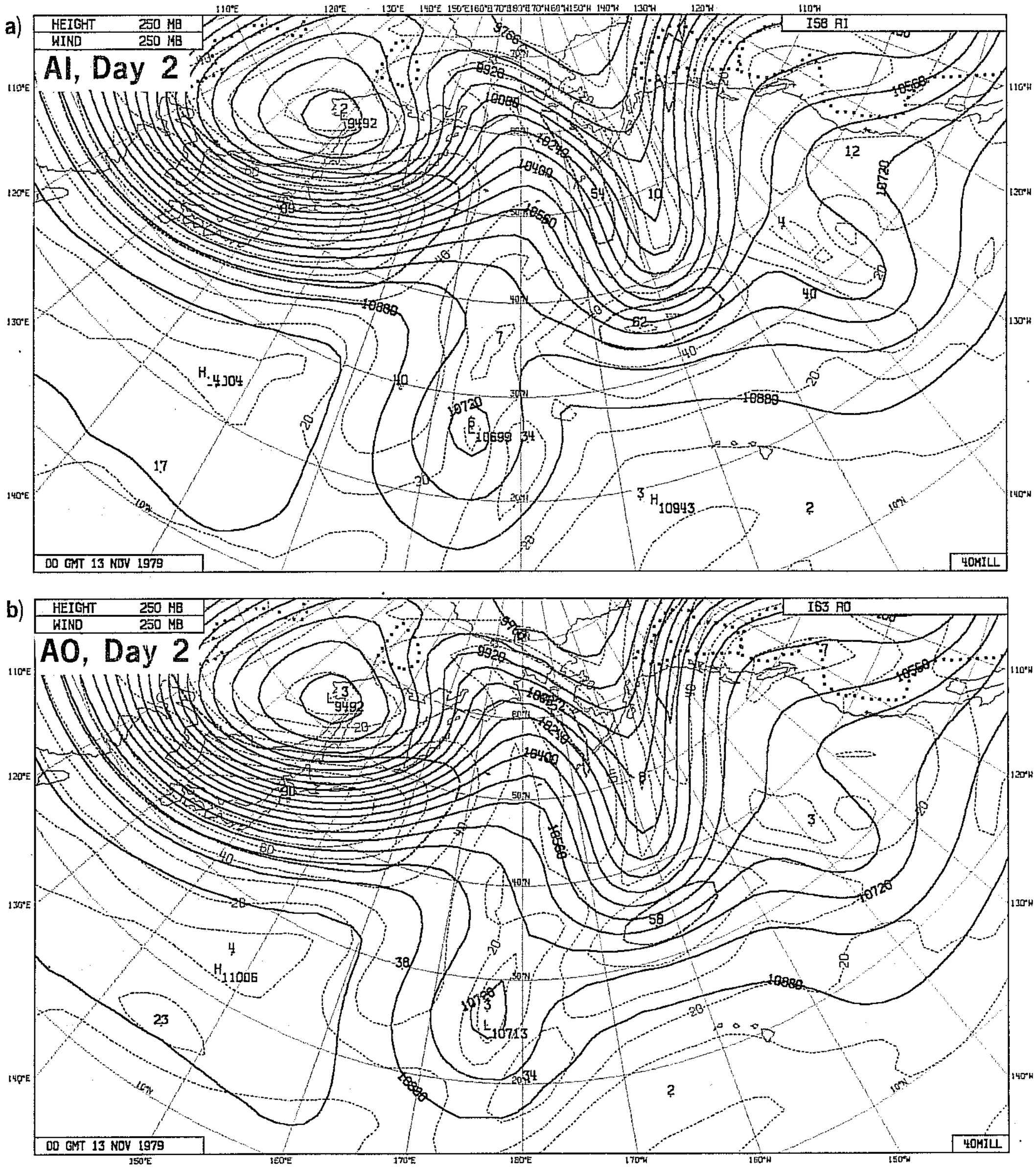

Fig. 22 Day 2 250 mb height forecasts, valid for 13 Nov.'79,00Z, over the Pacific Ocean. Full lines: height (meters); dashed lines: isotachs (meters per second).
a)AI, b)AO, c)SX, d)SO, d) verifying analysis.

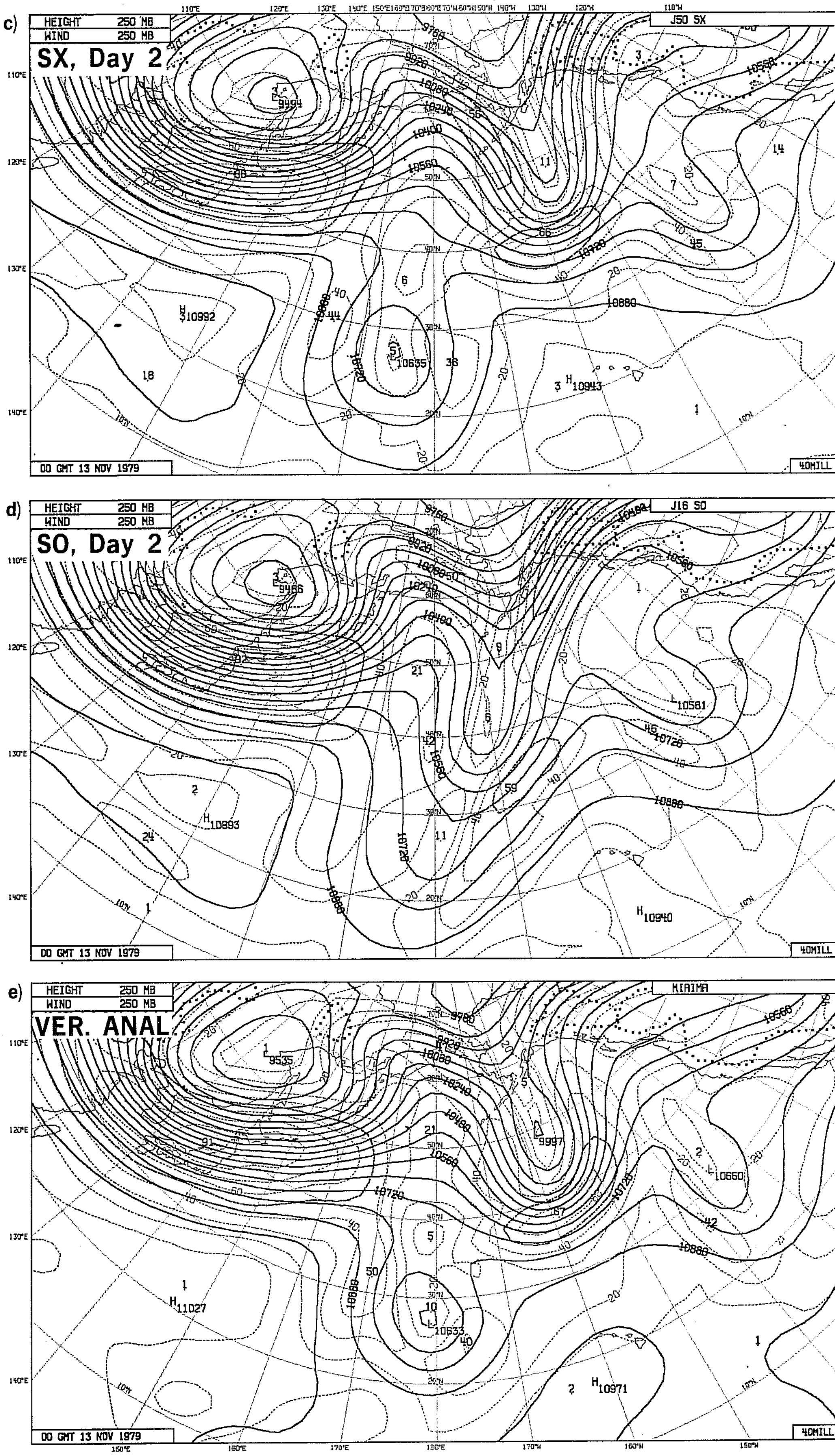

Fig. 22 (continued)

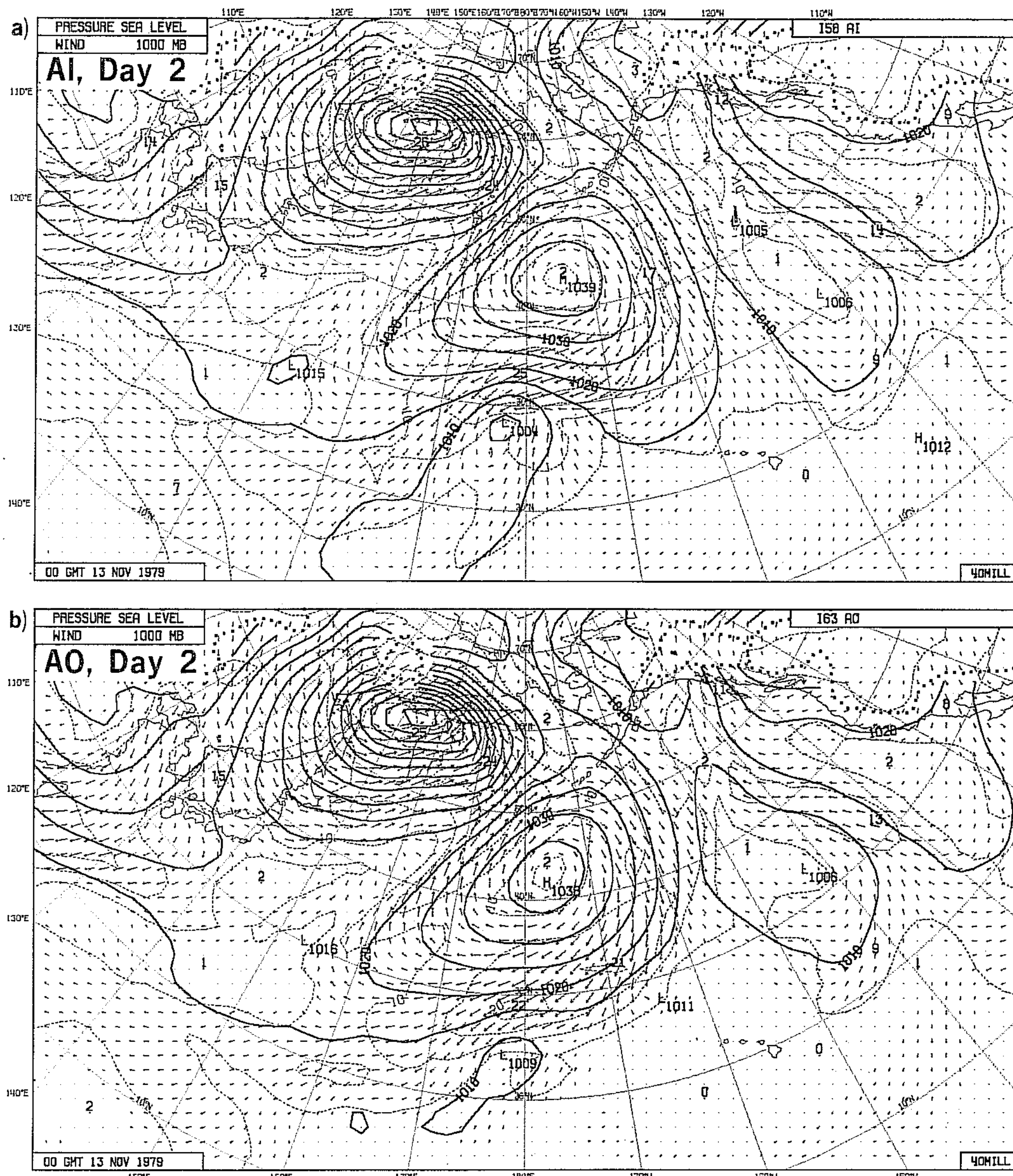

Fig. 23 Day 2 mean sea level pressure and 1000 mb wind forecasts, valid for 13 Nov.'79,00z, over the Pacific Ocean. a)AI, b)AO, c)SX, d)SO, e) verifying analysis.

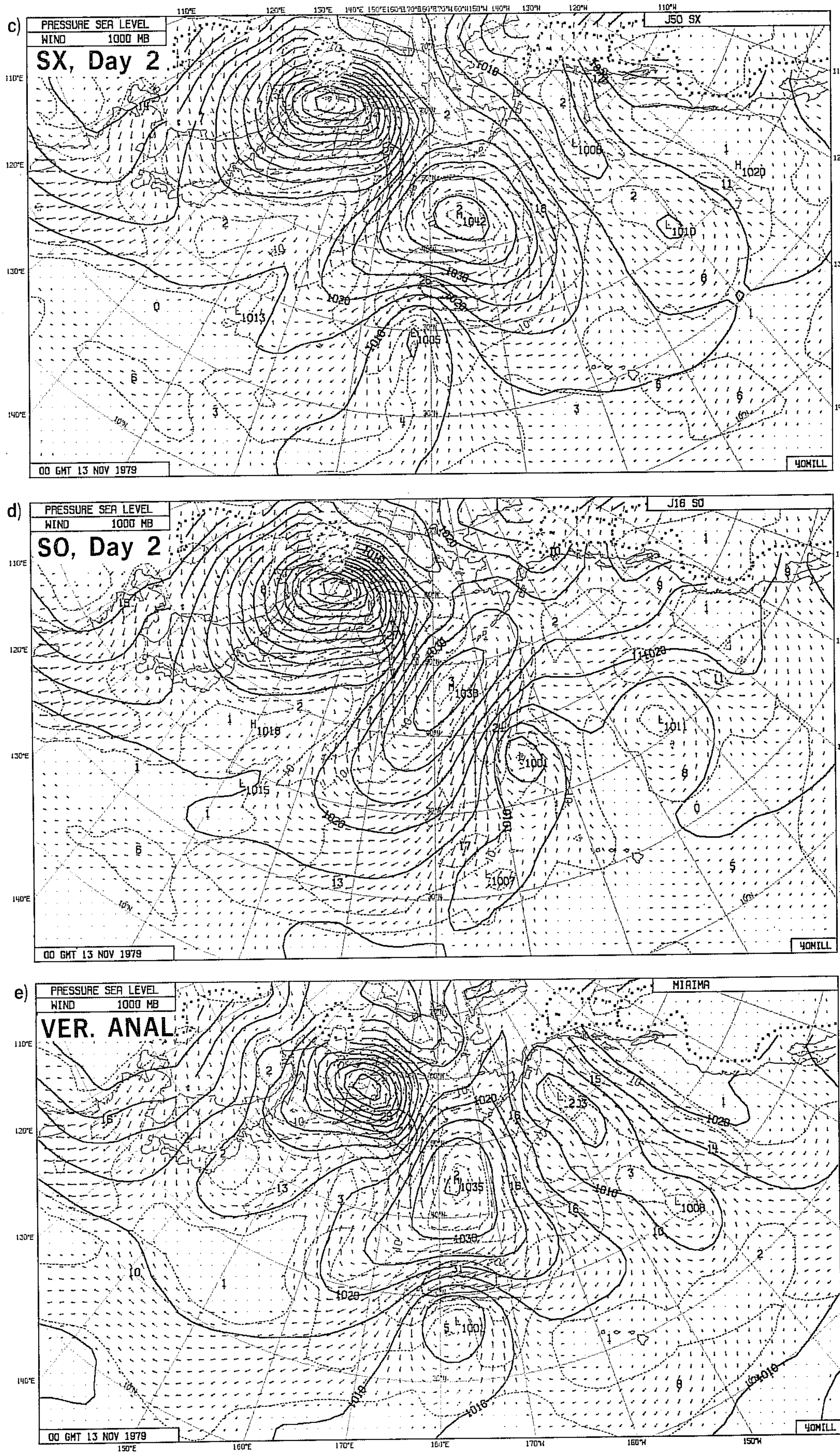

Fig. 23 (continued)

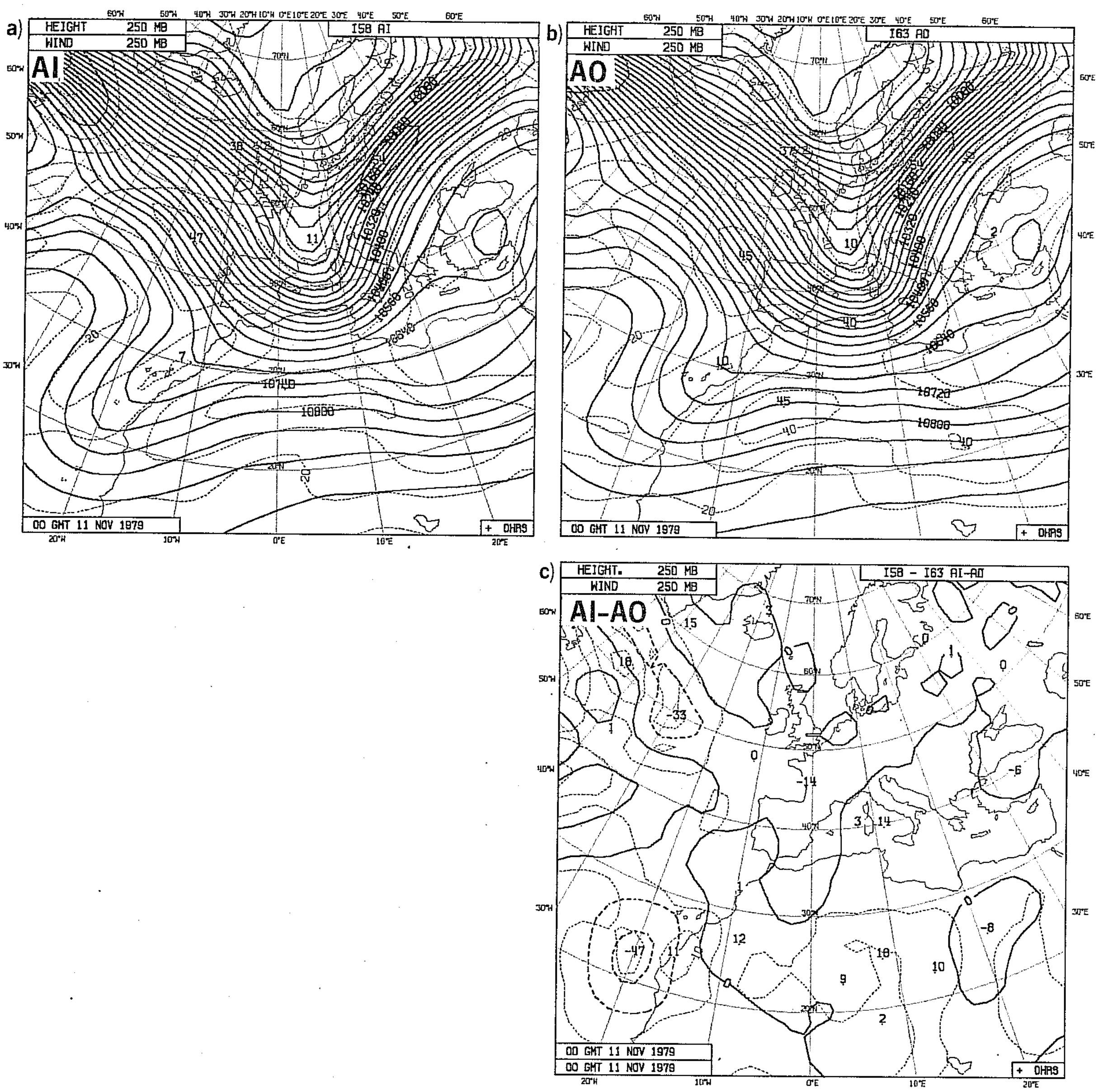

Fig. 24 Analyses of 11 Nov.,00Z (left page) and day 2 forecasts, valid for 13 Nov.,00Z (right page) over Europe and N. Africa. Full lines: height (meters), dashed lines: isotachs (meters per second) a) AI, b) AO, c) difference AI-AO, d)AI, e)AO, f)difference AI-AO, g) verifying analysis.

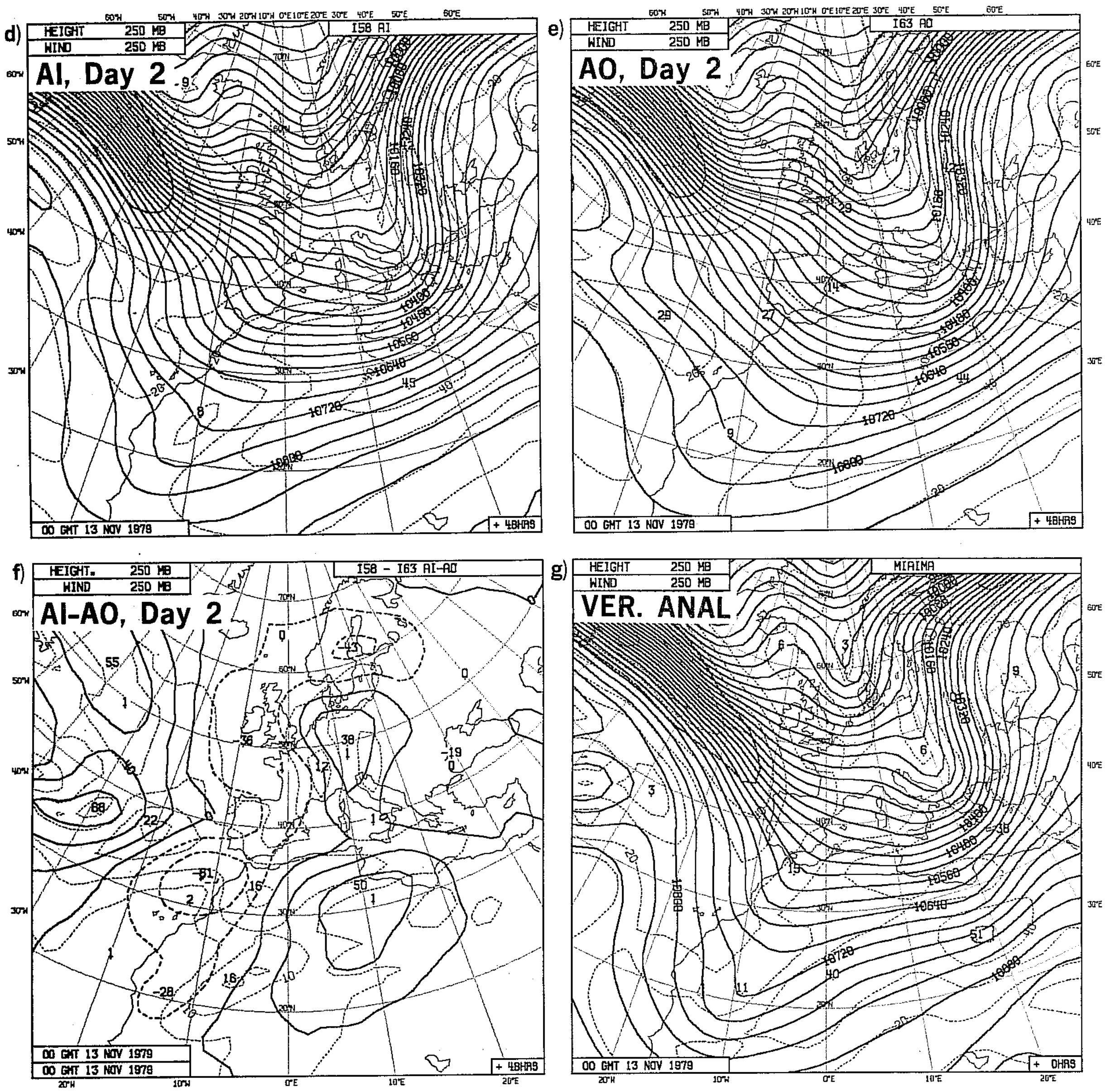

Fig. 24 (continued)

(b) Impact of satellite data

In Section 3.2 we saw that the 11 November 00GMT analysis differences were even larger when satellite data were removed from the data set. In particular we found even sharper jets associated with the Pacific trough in the SX analysis (with aircraft data; without satellite data) than in the AI analysis indicating that the information concerning the structure of the trough was to a large extent due to aircraft data.

Fig. 22 shows all four 48 h forecasts and the verifying analyses. Clearly the SO forecast is by far the worst and is useless. The SX and AI forecasts differ in some minor details but there is not much to choose between them both at 250 mb and near the ground SX and AI are the superior forecasts. A comparison of SX and SO shows very clearly and convincingly the large impact of aircraft data in this particular case, an impact that is only partly masked by the presence of satellite data.

This result underlines a remark made by Bengtsson (1982) that, in assessing the impact of observing systems, one should select cases carefully and purposely. When selecting cases at random it may be difficult to find a significant impact of aircraft data, but this particular case, in which there is a rapidly developing trough in an otherwise data void area, shows the importance of aircraft data very clearly.

4.4 Propagation of initial analysis differences

A closer look at the way initial 250 mb analysis differences propagate through the forecasts reveals a consistent and well organized pattern. After a period of up to 48 hours, during which the initial small scale differences in certain source areas organize themselves in well defined large scale patterns, wave trains are initiated which propagate downstream along the jet axis with the group velocity. The phase velocity relative to the underlying earth is very small if not zero. A similar behaviour was reported by Cats (1982), Hollingsworth et al., (1982), Cats and Åkesson (1983) and by Barwell and Lorenc (1985). These authors noticed the similarity between their results and those of Simmons and Hoskins (1979) who studied the development into a wave train of an isolated perturbation in a baroclinically unstable zonal flow.

In the four cases we studied, three source areas could be identified: the eastern half of the Pacific, the Newfoundland area in the Atlantic and North Africa. The first two of these areas were shown by Bengtsson (1982) to experience the largest reduction of the analysis error from an improvement of the data distribution. This fact, which is due to a combination of an unsatisfactory data coverage under operational non-FGGE conditions and high variance, is consistent with our findings.

As an example we show the 250 mb difference maps out to five days for the 11 November case (Fig. 25). Two wave train are initiated, one from the Central Pacific and one from the Newfoundland area, and both have organized themselves within 24 hours. A third one may be initiated by an analysis difference off the African coast (20°W, 20°N) but it merges rapidly with the Newfoundland wave train. Because the source of the latter is so much closer, Europe is already affected after 24 hours. With some imagination two branches can be distinguished in both wave trains, a feature that was found also in the other cases. It is most clearly observed on the map of November 13, 00GMT.

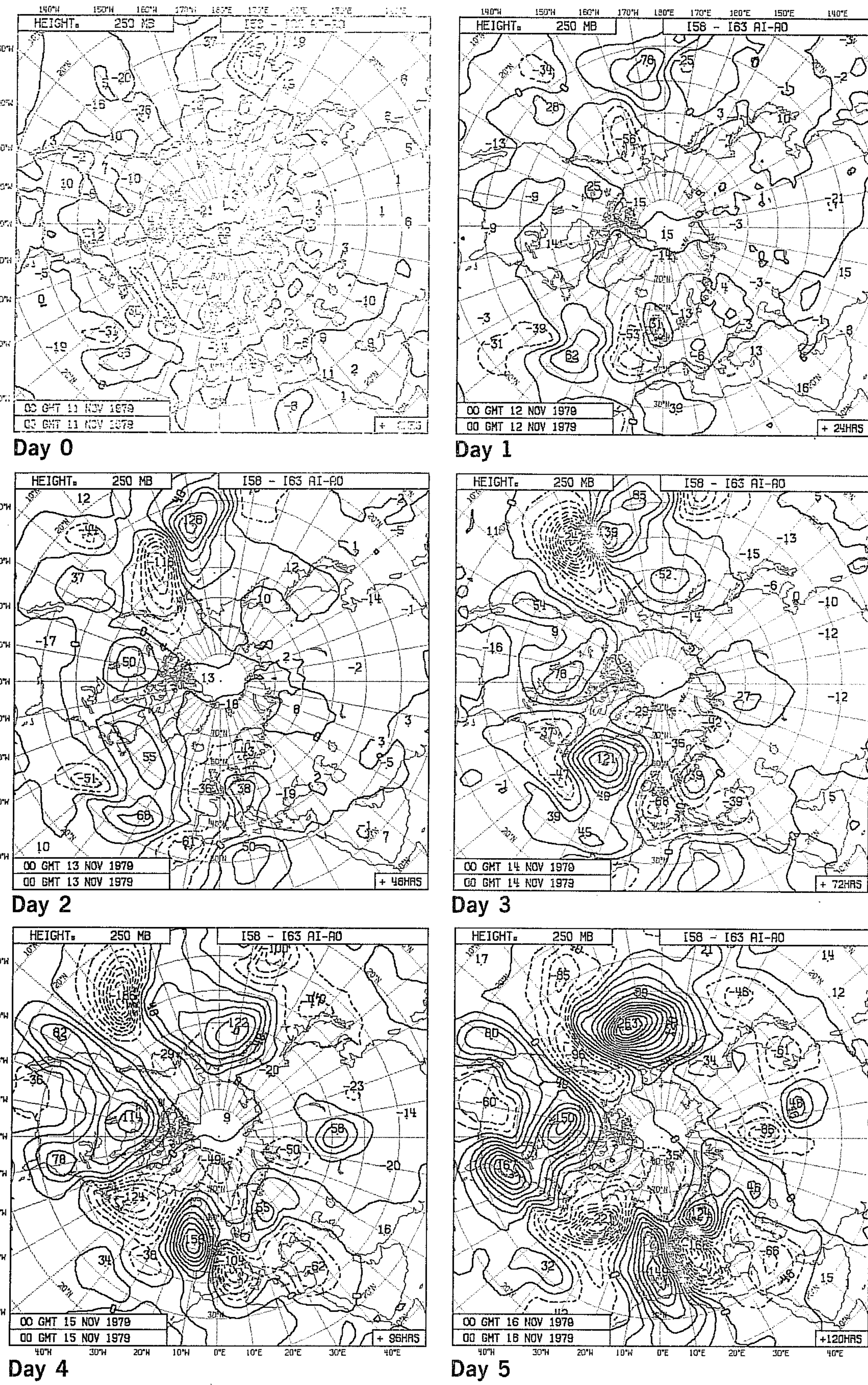

Fig. 25 Differences between 250mb height AI and AO analyses and day 1 till day 5 AI and AO forecasts. The initial analysis is 11 Nov.,00Z. The area is part of the Northern Hemisphere.

A comparison with the other cases shows that the time it takes for the forecast differences to affect the European area depends on the source area. If the differences originate only from the Pacific it takes about four or five days. If however the Atlantic or North African sources are active it takes a much shorter time. An analysis could perhaps reveal under what conditions a source area is likely to be active. This may then shed light on the stability of the atmosphere for analysis errors and thus on the predictability of the atmosphere, particularly in the European area.

4.5 Objective intercomparison of the forecasts

In the previous section, the downstream forecast propagation of initial analysis differences between two assimilations was discussed. Carrying this approach a bit further, the relative forecast differences were evaluated as anomaly correlations and root mean square fits between the three experiments AO, SX and SO on one hand, and the maximum system AI on the other. The procedure is thus similar to that used for verifications against analyses (Section 4.6) but with the verifying analysis replaced by a "control" forecast. In this way any model errors i.e. errors due to the forecast model not being able to predict the verifying analyses perfectly, are eliminated, and all relative differences between the forecasts are due only to differences in the initial states. Since these differences increase with forecast time, we find it convenient to use the expression "forecast divergence", which should not be confused with the divergence of the wind field. The forecast divergences were calculated and averaged for seven forecasts, including the four cases discussed earlier in this report, and three additional cases from the same period that were run at a later time and are not discussed elsewhere in this report.

In the southern hemisphere (Fig. 26, left), which was almost void of aircraft data, the AI and AO forecasts are, in these measures, almost identical up to day 10. In this period the average RMS differences between the AI and AO forecasts increase from a negligible 4 gpm at 250 mb to about 55 gpm at day 10. With the satellites excluded even the initial analyses SX and SO differ more than this amount from the AI analysis, and this difference increases to over 150 gpm at day 10.

In the northern hemisphere the situation is very different (Fig. 26). Now the difference between AO and AI increases from 9 gpm to 100 gpm, indicating a clear impact of the removal of aircraft data from the complete FGGE data set. It is interesting to note that the SX to AI forecast divergence is only slightly larger (increasing from 11 gpm to 115 gpm). This means that removing all satellite data from the FGGE data set only gives a very minor deterioration in this measure, as compared to what happens when only aircraft data are removed. Removing both aircraft and satellites gives a forecast divergence that increases from 25 gpm at 250 mb to more than 140 gpm, i.e. SO is much more different from AI than either AO or SX. These results indicate an apparent redundancy between aircraft data and satellite (primarily SATEM) data. The aircraft data alone seems to be able to define the structure of the large scale flow to a high degree in otherwise data void areas, as does the SATEM data, as was also evident from the case study over the Pacific Ocean discussed earlier. This impact of aircraft data on the large scale flow is, however, not really surprising since the wide structure functions for first guess error correlations (Fig. 2) has the effect of spreading single level data throughout a large part of the tropospheric analysis (Fig. 15). It should thus be stressed that the apparent redundancy may at least to some extent be due to the particular data assimilation system

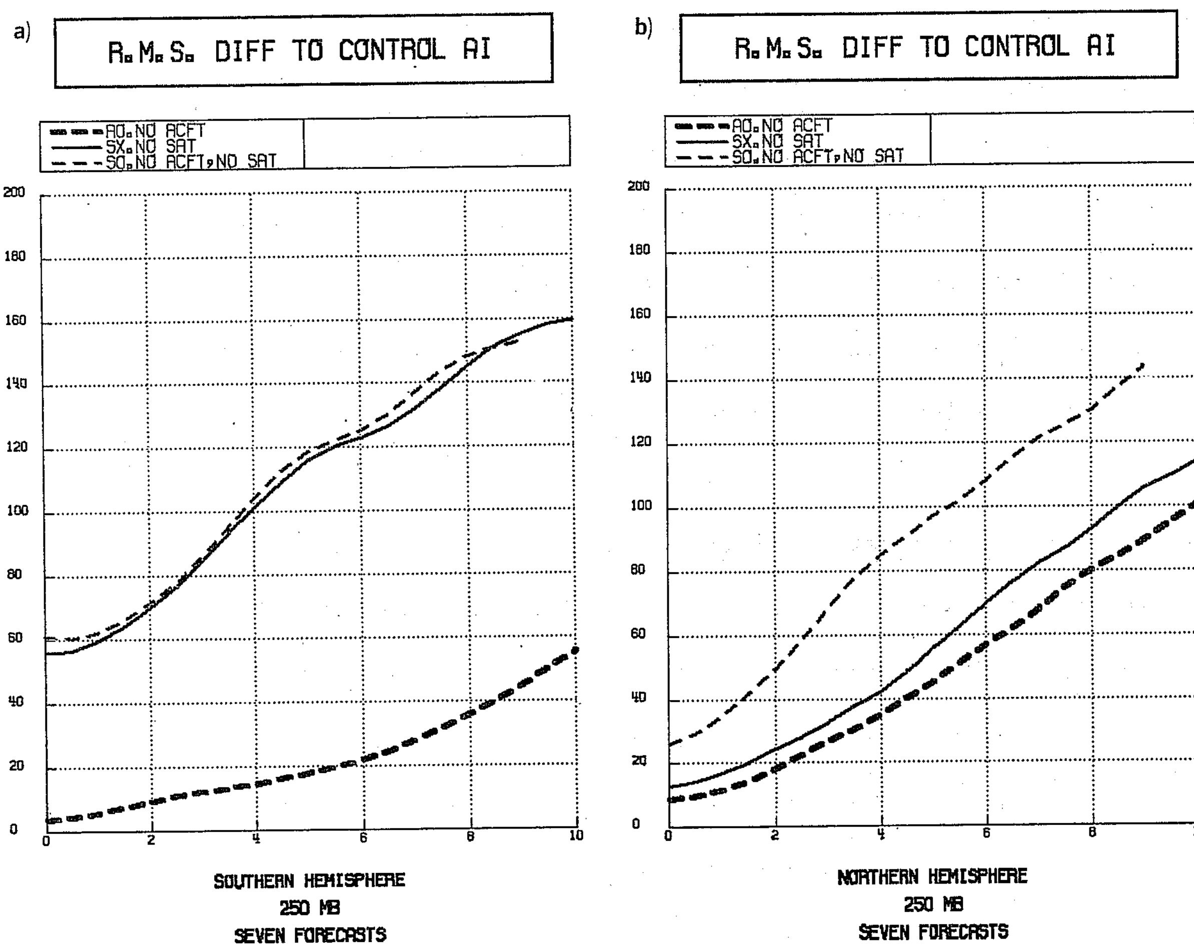

Fig. 26 250 mb RMS height differences (gpm) between control forecast AI and forecasts AO, SX and SO, as a function of forecast period (days). The graphs are averages over seven forecasts. Left: Southern Hemisphere, right: Northern Hemisphere.

used for the experiment. Nevertheless it is encouraging to note that the aircraft wind data do have a high information content and are very useful for numerical weather forecasts in areas with poor satellite coverage (cloudy areas in mid-latitudes, clear areas in equatorial regions).

The vertical distribution of the forecast divergence due to aircraft data is shown in Fig. 27. Between days 2 and 6 a small, but unquestionable difference is seen between 250, 500 and 850 mb. At the top (aircraft) level AO and SX develop very similarly when compared with AI, while at 500 mb and, particularly, at 850 mb the forecast divergence SX to AI larger than that of AO to AI. Thus the aircraft data impact is largest at, or close to, the level of the data, even if their impact is also discernible lower down in the troposphere.

4.6 Objective verification of the forecasts against analyses

In the previous sections we noticed that improvements of analyses and forecasts from aircraft data are locally demonstrable but generally small after two days. The problem is that model error seems to be the dominant error source in the medium range between 2 and 5 days (Arpe et al., 1985). Accordingly it appeared difficult to corroborate the improvements by objectively verifying the forecasts against analyses.

We verified all forecasts run from the four assimilations AI, AO, SX and SO against ECMWF FGGE level III-b analyses, using the standard ECMWF verification package which computes RMS differences and anomaly correlation for different wavelength intervals and for certain predefined geographical areas. A comparison of the scores of individual forecasts turned out to be rather unhelpful. Score differences between corresponding forecasts are generally rather small and sometimes contrary

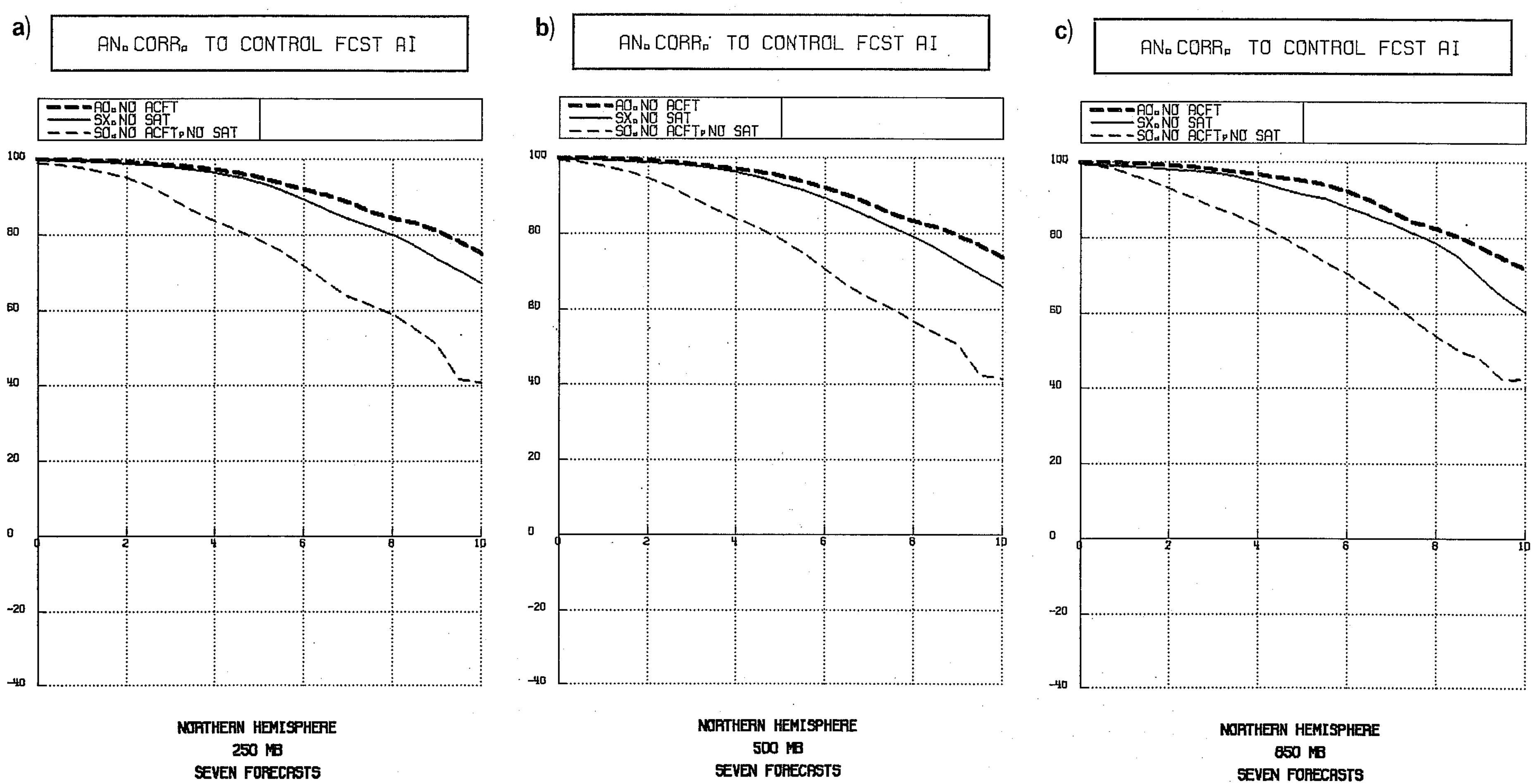

Fig. 27 RMS height differences (gpm) at these different pressure levels between control forecast AI and forecasts AO, SX and SO, as a function of forecast period (days). The graphs are averages over seven forecasts. Left: 250 mb; middle: 500mb; right: 850 mb.

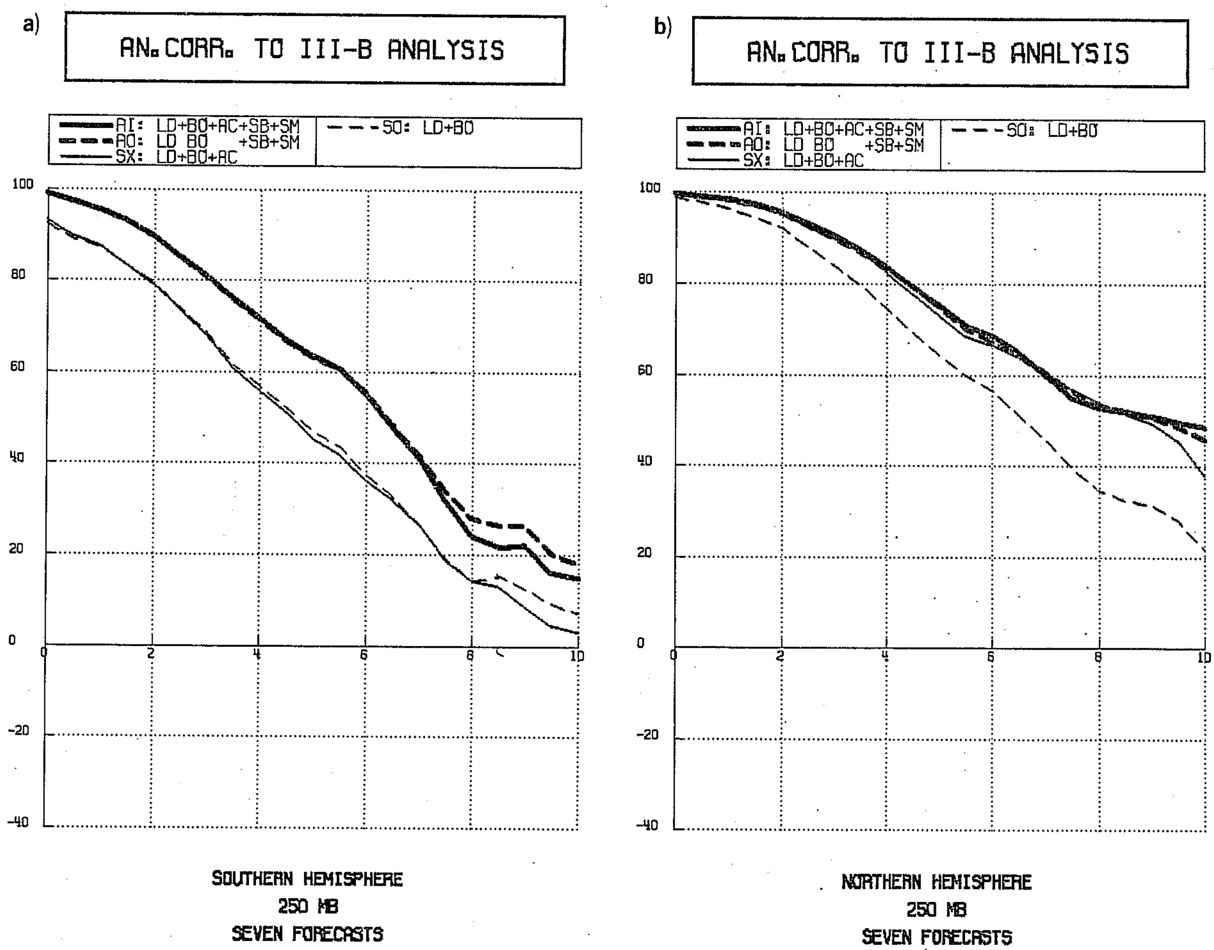

Fig. 28 Anomaly correlations (per cent) between AI, AO,SX and SO forecasts and verifying ECMWF FGGE III-b analyses of 250 mb height, as a function of forecast period (days). The graphs are averages over seven forecasts. Left: Southern Hemisphere; right: Northern Hemisphere.

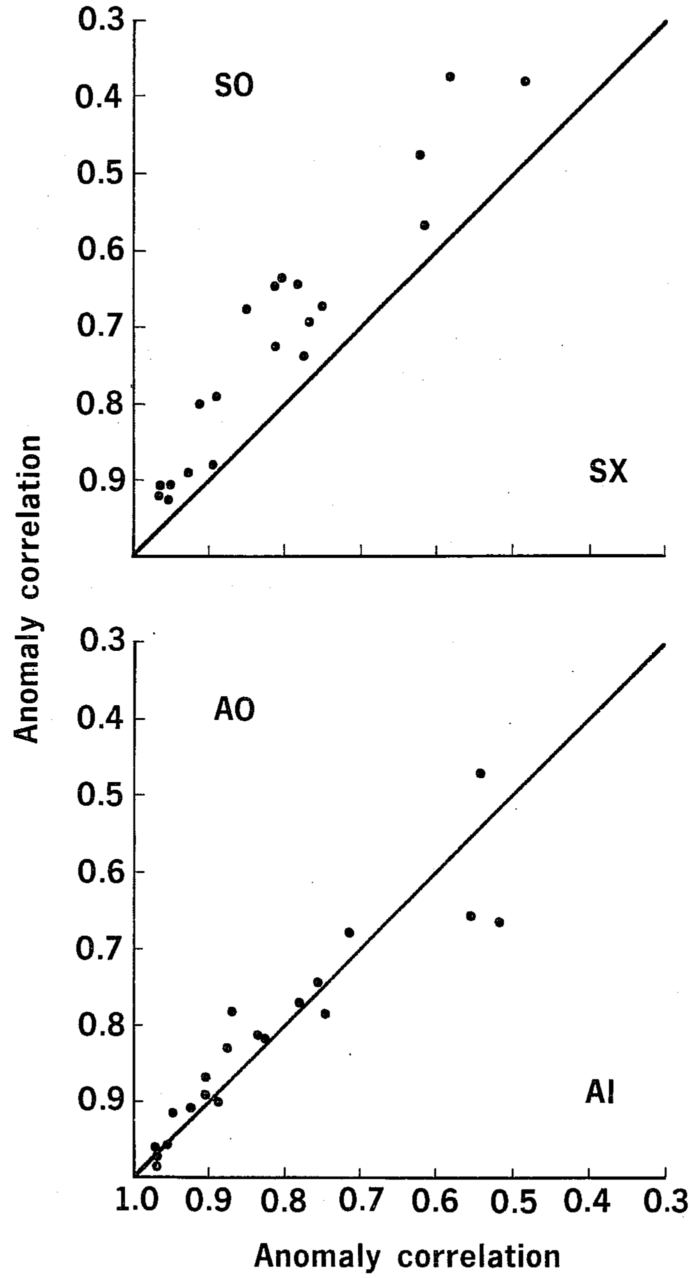

Fig. 29 Comparison of 250mb height anomaly correlations of day 2-6 forecasts with (AI and SX) and without (AO and SO) aircraft data. Lower panel: with satellite data; upper panel: without satellite data.

to what we found synoptically, whereas the quality of the forecasts varies considerably from case to case. We therefore concentrated on mean scores over several forecasts.

Fig. 28 shows the anomaly correlations of the 250 mb height forecast from AI, AO, SX and SO against verifying III-b analyses for the northern and southern hemisphere, averaged over the same seven cases as discussed in the previous section. The difference between both hemispheres is most striking. The northern hemisphere forecasts from all analyses, except SO, verify almost equally well showing that either aircraft data or satellite data are essential but to a large extent mutually redundant. However in the southern hemisphere the satellite data are crucial. Without them even the analyses are quite different. Aircraft data have essentially no impact there.

An easy and compact way to compare bulk scores of a series of forecasts with and without a certain data system is to use a scatter diagram. Fig. 29, lower panel, shows a scatter diagram of the anomaly correlation over the northern hemisphere of the 250 mb height of day 2, 3, 4, 5 and 6 forecasts from the four 00GMT AI and AO analyses. Most points are above the diagonal indicating a slight improvement of the AI forecast over AO. A similar plot, averaged over the troposphere from 1000 to 200 mb, did not show such systematic improvement.

Fig. 29, upper panel, summarizes the results of the SX and SO forecasts. Now a clear and convincing impact of aircraft data is apparent in agreement with the synoptic evaluation.

In summary we may conclude that we have been able to corroborate the positive impact of aircraft data on the northern hemisphere by bulk scores. This statement is true in general but not necessarily in individual cases. A careful synoptic evaluation is certainly more important and valuable than an objective verification.

4.7 Objective verification of the forecasts against observations

Objective verification against analyses, as presented in the previous section, is in a sense an indirect way of verifying forecasts because the analyses themselves are a product of a complicated numerical process in which not only the observations but also assumptions about their error structure and a short range forecast (first guess field) play a role. A more direct way of verifying forecasts, admittedly with its own drawbacks, is by comparing them with observations. In this section we provide some verification statistics of the first guess fields (+ 6 h forecasts) and of the model forecasts out to 5 days, both against observations.

(a) Verification of the first guess

An evaluation of the impact of a data system on 6 hour forecasts is important for two very different reasons. The first is that 6 hour forecasts are used as a first guess in the analysis scheme. The acceptance of observations and therefore the quality of the analysis may crucially depend on the quality of the first guess. In the second place, such an evaluation is of direct importance for the users of short range numerical forecasts. For example the same aviation authorities who provide the meteorological community with aircraft data, are very interested in short range forecasts for their flight planning and thus may wish to know the impact of their data on such forecasts.

In the following the fit of observations to the forecasts is computed in the same way as was done in Section 3.3. We now consider all those 6 hour forecasts which were produced in the data assimilation runs, altogether 33 cases. When comparing only weighted means over all cases, it is important to realise that the impact may be significantly larger in individual cases or sometimes there may be no impact at all, depending on the distribution of observations and on the synoptic situation.

The mean fit of all rawinsonde winds (u-component) to the AI and AO 6 hour forecasts is shown in Fig. 30. Of course these quantities are heavily biased to the fit over continental areas of the northern hemisphere. A small but systematic improvement from aircraft data may be seen at all levels below 150 mb, in particular around the flight level 250 mb, whereas a slightly negative impact may be observed in the stratosphere where the 6 hour forecasts are biased anyway. It may be that the vertical error correlation used in the assimilation scheme does not correctly spread the flight level information into the stratosphere. A similar impact pattern be it about twice as large, is found when satellite data are removed (SO versus SX).

Of greater importance for aviation is the impact of aircraft data on 6 hour forecasts along flight tracks. This impact is measured by computing the fit of aircraft data to the forecasts. This is shown in Fig. 31 and Fig. 32 for ASDAR/AIDS and AIREP data separately, and for both wind components. The maximum of the mean improvement varies between 0.5 and 0.9 m/s per component. When satellite data are removed, the impact is a factor of two larger. These results show the systematic nature of the impact of aircraft data. Again we stress that in individual cases the impact may be considerably larger.

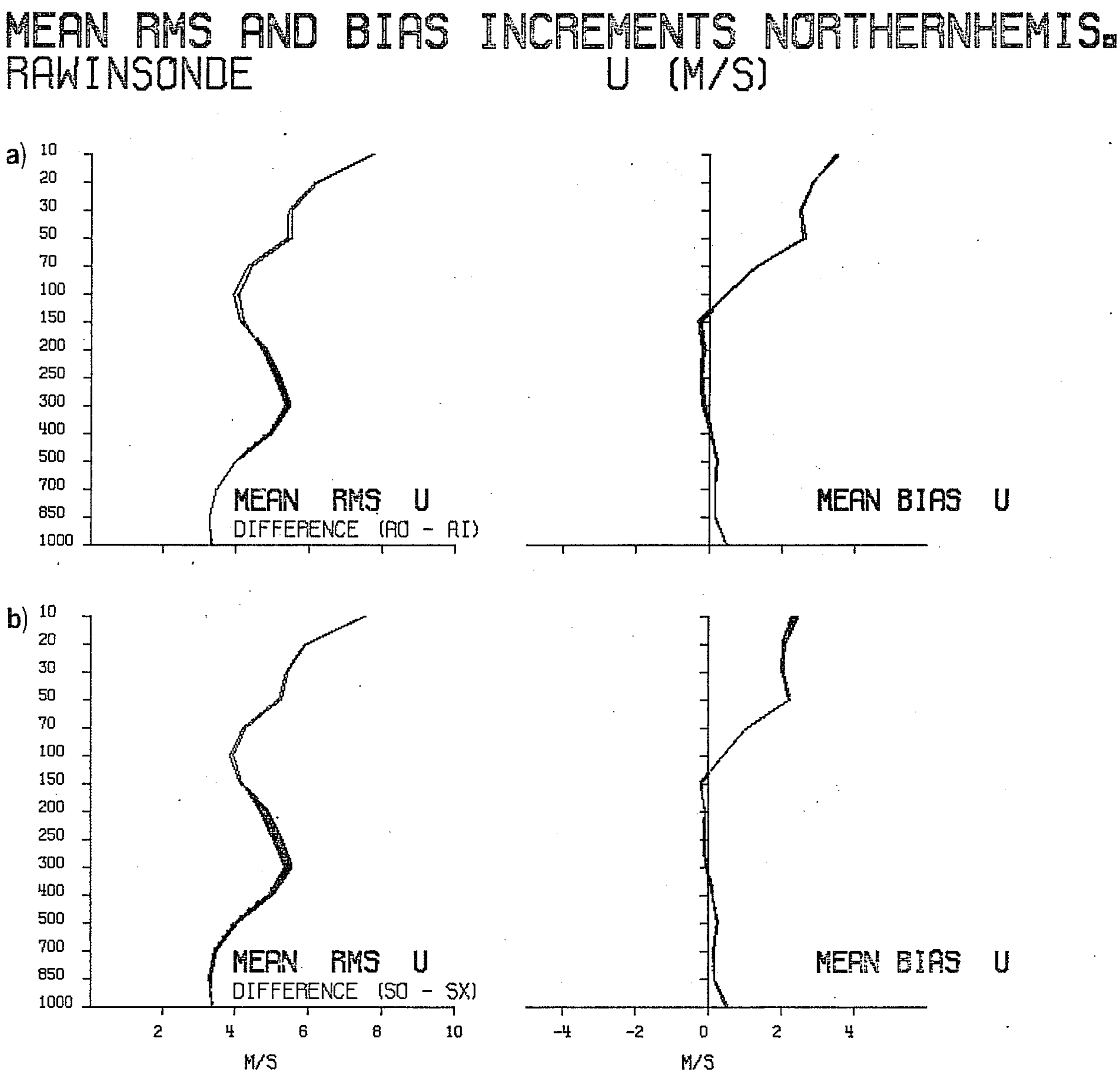

Fig. 30 The fit of rawinsonde wind data (u-component) against the 6-hour forecasts in the Northern Hemisphere. Top: AI and AO; Bottom: SX and SO. Black shading indicates that AI or SX is closer to the origin (improvements by the aircraft data), whereas lighter shading indicates the opposite.

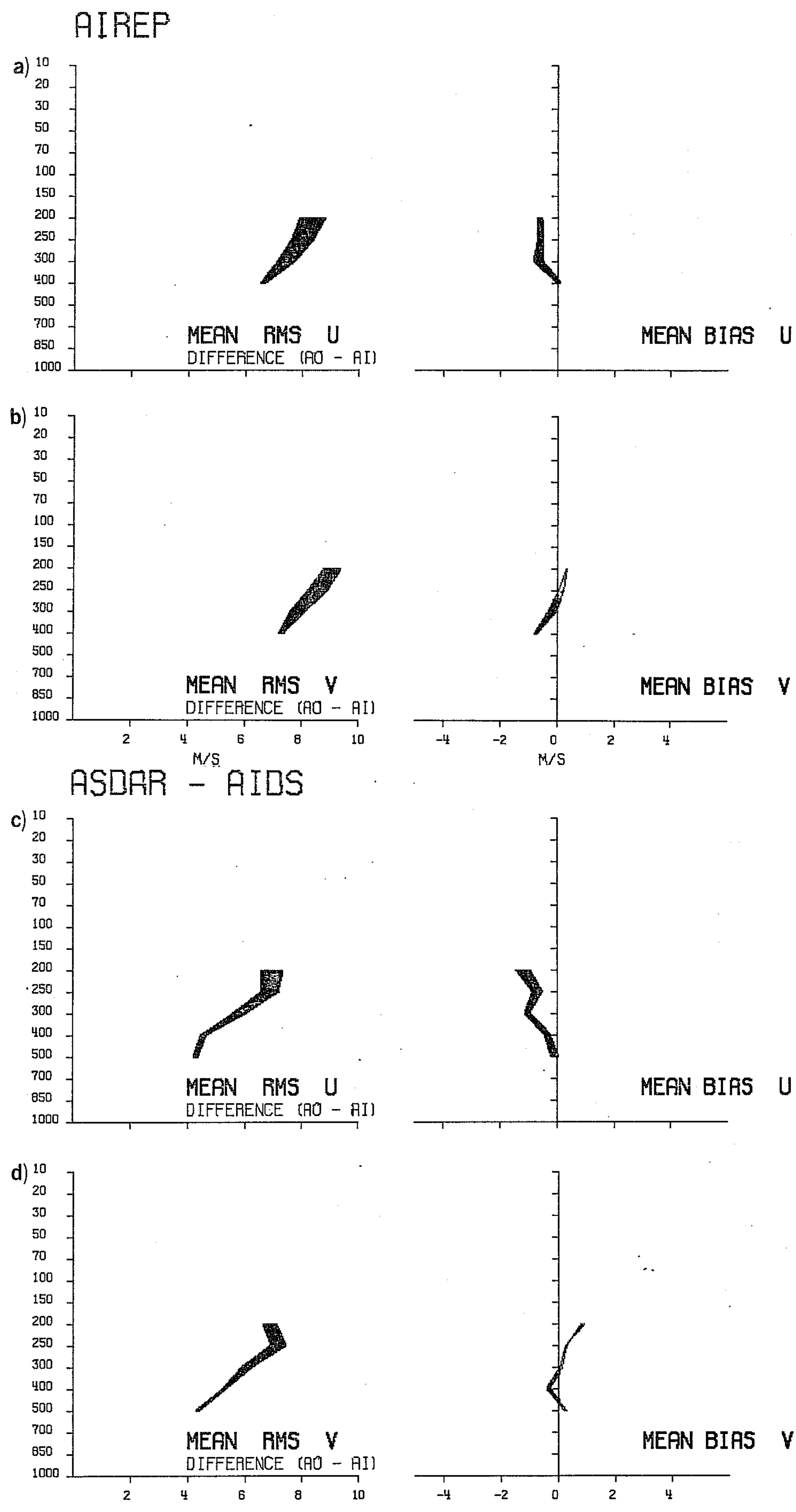

Fig. 31 The 6-hour forecast difference along flight tracks by using aircraft data in the analysis; experiment AO versus AO. The shading convention is the same as in Fig. 30. Top: AIREP; Bottom: ASDAR/AIDS.

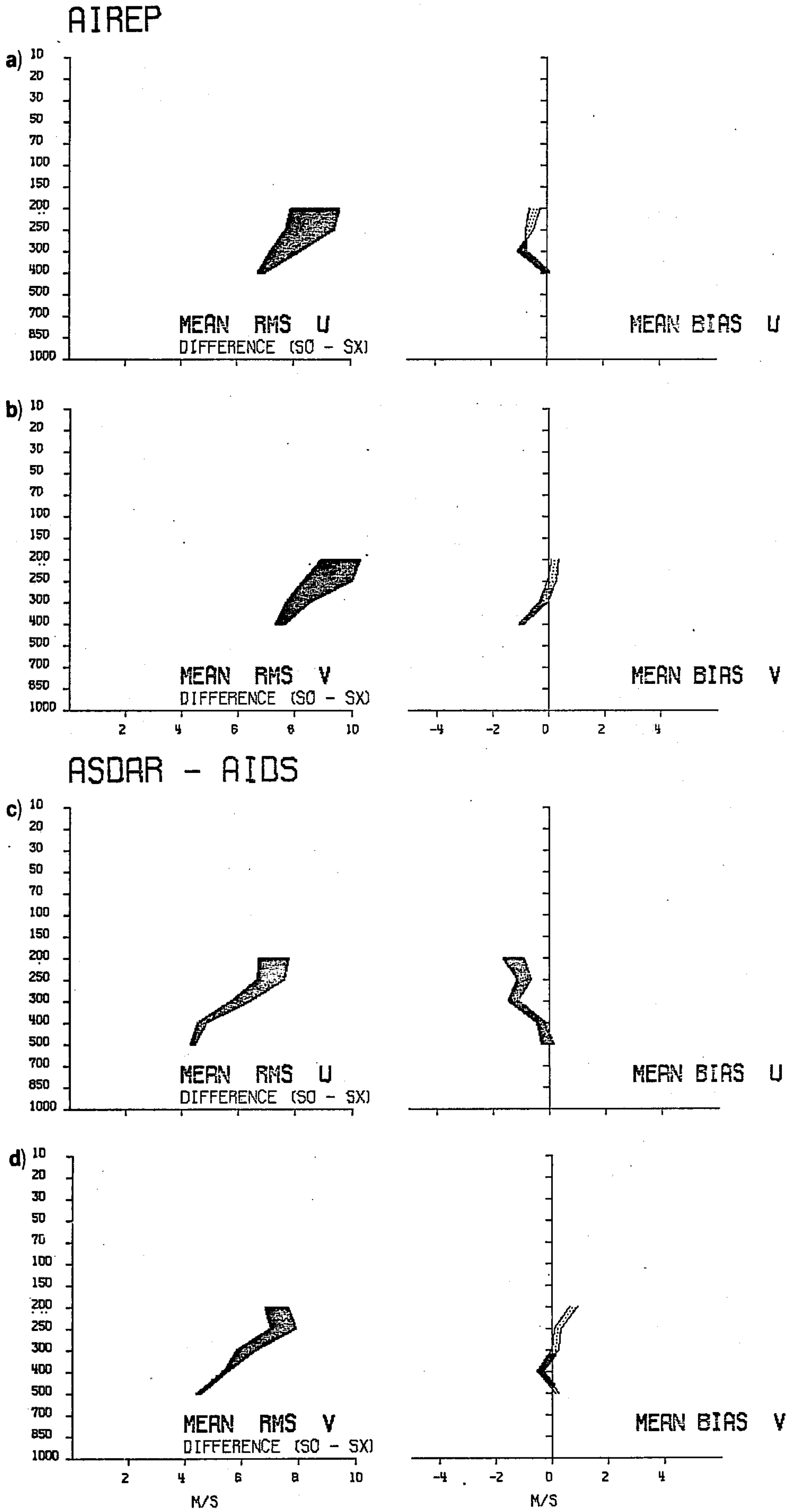

Fig. 32 Idem but for experiment SX vs. SO.

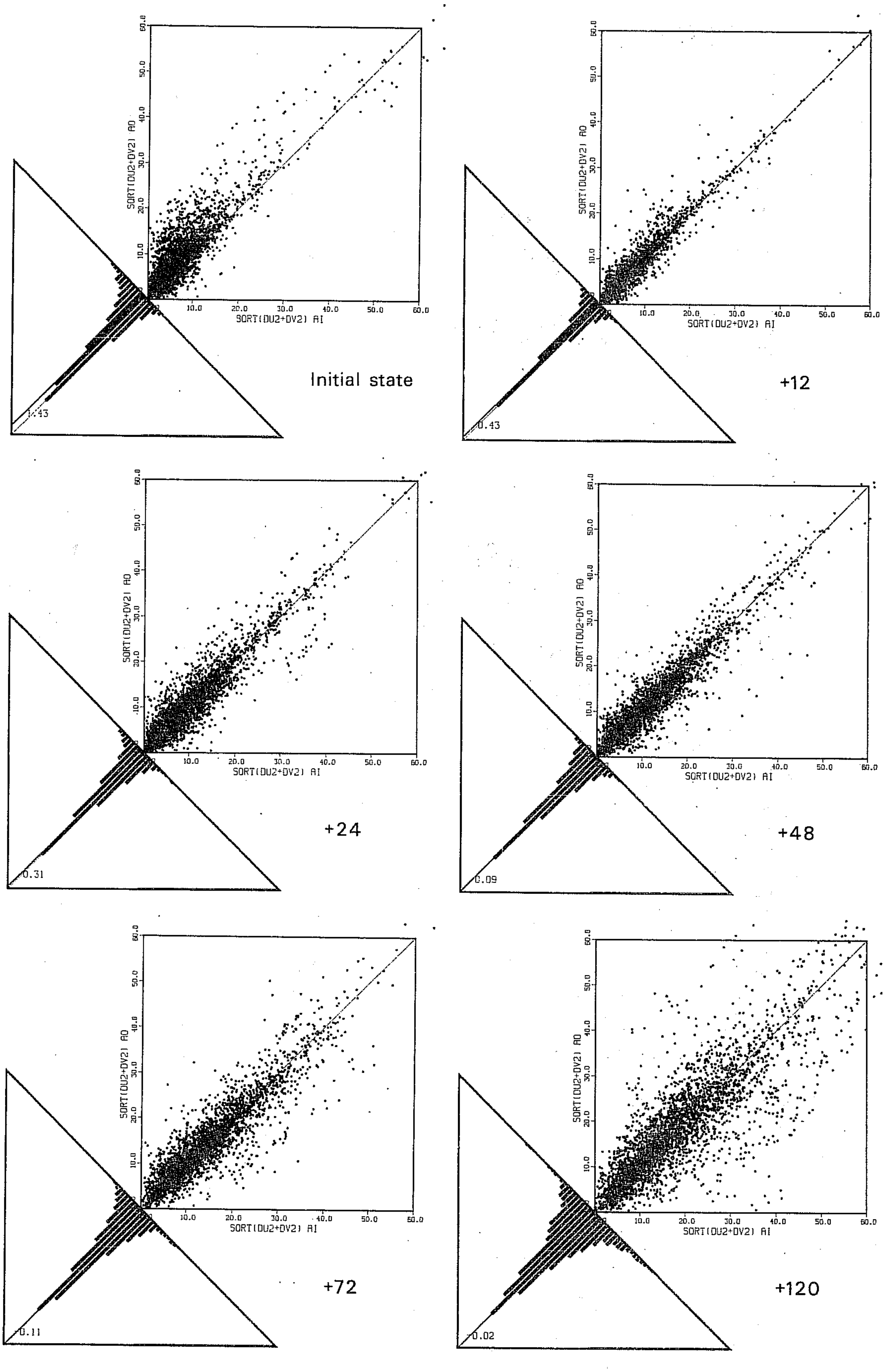

Fig. 33 Distribution of wind speed error (AI error vs. AO error) for four 10-day forecasts at 250 mb and histograms of the eviation of each point from the diagonal. Results for the initial state T, T+12, T+24, T+48, T+72 and T+120.

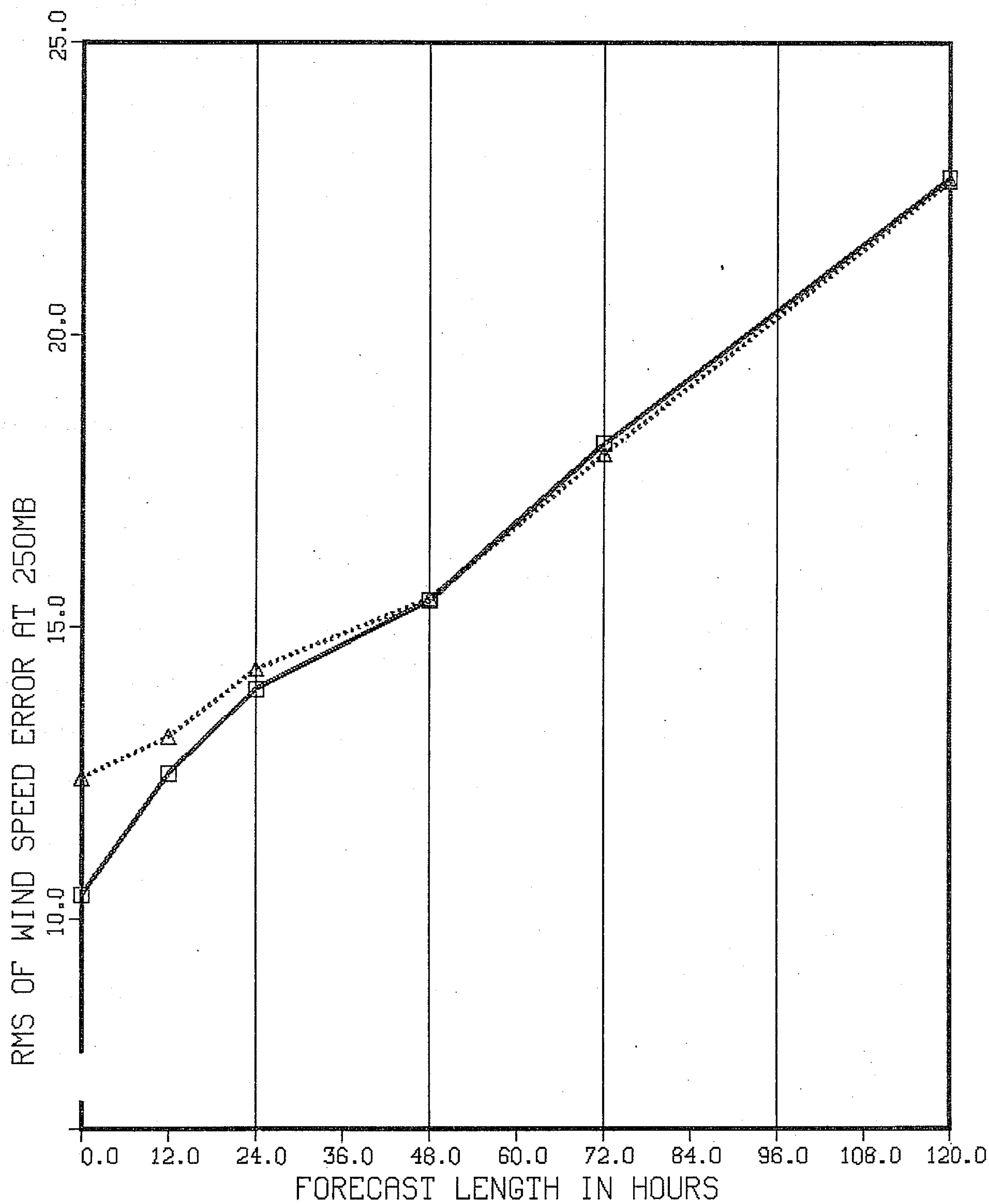

Fig. 34 The evolution of the mean RMS wind speed error at 250 mb for the AI forecasts (full line) and AO forecasts (dotted line).

(b) Verification of model forecasts up to 5 days

We have verified the wind forecasts up to 120 hour against all aircraft wind observations. The results are summarized in Fig. 33 which shows scatter diagrams of vector wind errors of forecasts AI against AO, together with histograms of deviations from the diagonal. The vector wind error is defined as

$$E = \sqrt{(u_o-u_f)^2 + (v_o-v_f)^2}$$

where u_o and v_o are the observed wind speed components, and u_f and v_f are the forecast wind speed components interpolated bilinearly to the position of the observation. Note that all winds, including those rejected by the assimilation scheme, are used in the scatter diagrams. A small bias in favour of the inclusion of aircraft data may be noticed up to 48 hours. This is also apparent from the associated histograms. It is interesting to note that even the +120 hour histogram shows some skewness in favour of AI. Essentially the same picture emerges from a verification against only ASDAR data which were set apart for independent verification purposes.

An alternative way of representing these results is to plot the mean RMS vector wind error at 250 mb for the forecasts AI and AO as a function of forecast length. This is done in Fig. 34 for the verification against all aircraft winds and again it reveals the positive impact of aircraft data during the early part of the forecasts.

4.8 Résumé

In this Section 4 we have shown that aircraft data can contribute to an improvement of the forecast of the structure and phase of synoptic features in the northern hemisphere, not only at 250 mb but also sometimes near the surface. A synoptic evaluation of forecast differences shows that the impact is local and generally small compared to forecast errors, and is often associated with synoptic systems originating from data-void areas. A comparison with forecasts from analyses without satellite data reveals that satellite data partially mask the impact of aircraft data. A detailed evaluation of the 11 November case, however, shows very convincingly that it is the aircraft data that make AI and SX the superior forecasts. Moreover one should realize that the amount of satellite data was exceptionally high during this FGGE period, certainly higher than normal. Therefore under operational, non-FGGE, conditions a larger impact may certainly be expected.

An objective assesment of the impact was obtained in two ways. Initially the forecasts were verified against analyses by computing bulk scores such as the anomaly correlation and RMS differences. Then, in order to eliminate the large influence of the model error on all forecasts, they were compared with the AI forecasts. Both methods reveal the high information content of aircraft data in the northern hemisphere. These data alone seem to be able to define the structure of the large scale flow. In this study this was masked to a large extent by the overwhelming presence of satellite data, although even under these circumstances a small positive impact was found.

A comparison of our results for the 11 November case with those by Barwell and Lorenc, reveals the importance of the design of the assimilation scheme for the extraction of information from single level data. It seems that a multi-variate three-dimensional analysis scheme with a properly adjusted vertical structure function is advantageous for a successful assimilation of single level wind data. It should be noted, however, that in our experiment the structure functions have not been specially adjusted at all, a fact that may be responsible for the slightly negative impact we found in the stratosphere.

Finally a verification of our forecasts against aircraft wind observations shows a systematic and significant positive impact of aircraft data on short range forecasts along flight tracks at flight level. This may be important for flight-planning purposes.

5. SUMMARY AND CONCLUSIONS

In order to assess the impact of aircraft wind data on numerical analyses and forecasts, two data assimilations with and without such data were performed on FGGE data for the period 8-19 November 1979; two corresponding assimilations were also run with SATOBs and SATEMs excluded. From four pairs of the analyses parallel sets of 10-day forecasts were produced. Both the assimilations and the forecasts were carried out using the operational ECMWF system with some minor changes.

The analyses and forecasts were evaluated both subjectively and objectively. Moreover results were available from a similar experiment carried out at the UKMO. The following conclusions were drawn:

- The quality of the automatically transmitted aircraft data is high. In the AIREP data set, however, many apparent errors were spotted. This in itself is a strong argument in favour of automatic aircraft data transmission.
- In the upper troposphere considerable differences between both sets of analyses were found, locally over 20 m/s at 250 mb, particularly over the tropical and northern hemisphere oceans. Further evaluation showed that the impact is beneficial and most significant in the tropics. Outside the tropics there seems to exist a considerable redundancy between aircraft and satellite data, a redundancy that could be smaller, however, under operational non-FGGE conditions or under different synoptic conditions.
- Most of the impact is retained after normal mode initialization, indicating that the impact is concentrated in the external and lower internal Rossby modes.

- A careful comparison of 2-5 day forecast results revealed a small but significant impact on some synoptic features at 250 mb and sometimes near the surface.
- A comparison of forecast 250 mb height anomaly correlations showed only a slight impact, indicating that the impact was indeed small.
- In an attempt to explain why the impact was small it was found that it was much larger when satellite data were excluded from the assimilation. The conclusion must be that the ECMWF system is able to extract valuable information from the aircraft data, but that during this period this information was partly masked by satellite data.
- A comparison of 6 hour forecasts with corresponding wind measurements revealed a positive impact of aircraft data on short range tropospheric wind forecasts. A slightly negative impact was found in the stratosphere, possibly due to the assumed structure of the vertical prediction error correlation.
- A systematic improvement of short range wind forecasts was found, in particular along flight tracks. Mean improvements of up to 0.9 m/s per wind component were found but in individual cases the impact may be considerably larger.
- A comparison of our results with those of Barwell and Lorenc at the UKMO showed that a multi-variate scheme with proper vertical coupling is advantageous in extracting information from single level data. Further work to investigate and optimize this is desirable.
- From this assessment it is concluded that a future system of aircraft data, based on ASDAR, will be a valuable contribution to the global system.

Several times in this report reference was made to the fact that a certain amount of redundancy seems to exist between aircraft and satellite data. It is worth summarizing some different aspects of this phenomenon.

First of all it should be realized that a certain degree of redundancy is necessary in the Global Observing System. Without it data checking and quality control would be impossible. The quality and reliability of the automatically transmitted aircraft data are important assets in this respect.

However the fairly high degree of redundancy between aircraft and satellite data found in this study may have been caused by a number of fortuitous circumstances; the satellite data density was larger than under non-FGGE conditions. Moreover it appears that also the degree of redundancy may strongly depend on the synoptic situation. In particular the importance of a data system may depend on the geographical position of the main areas of activity. As we have seen, during the period studied in this report, the main activity took place over the northern oceans, in particular the mid-Pacific. This favours the impact of satellite data and therefore their mutual redundancy with aircraft data. Also the availability of data may depend on the synoptic situation. SATOB and aircraft data may be redundant when sufficient high level clouds are available but may complement each other in cloud-free areas. Similarly SATEMs and aircraft data may be redundant in mid-latitudes in cloudless conditions but may complement each other when in cloudy areas microwave retrievals are poor.

A final aspect that should be mentioned is the present assimilation technique. Aircraft data are high quality single level data. The assimilation techniques available at present may not take sufficient advantage of such data. The apparent redundancy, therefore, may partly be due to the way we extract information from single level data systems. Techniques may be found that make better use of the unique features of aircraft data and thus reduce the redundancy.

In summary we can conclude that redundancy is a desirable feature of the Global Observing System but that, before conclusions are drawn concerning the degree of redundancy between two observing systems, attention should be paid to specific circumstances, such as the synoptic situation and the assimilation technique.

Acknowledgements

We thank Dr. H.M. de Jong and Mr. G.W. Brouwer of KNMI for carefully checking the AIDS and ASDAR data for the selected period and making these data available to us. Thanks are also due to Ms. Birgit Kok for her excellent typing.

REFERENCES

Arpe, K., A. Hollingsworth, A. Lorenc, M.S. Tracton, S. Uppala and P.Kållberg (1985): The response of numerical weather prediction systems to FGGE II-B data, Part II: Forecast verifications and implications for predictability. Quart. J. Roy. Meteor. Soc. 111, 67-102.

Baede, A.P.M., S. Uppala, and P. Kållberg (1983): Impact of aircraft wind data on ECMWF analyses and forecasts during the FGGE period, 8-19 November 1979.
Preprints, 6th conference on Numerical Weather Prediction, Omaha, Nebr., 1983. Published by the American Meteorological Society.

Barwell, B.R. (1982): The impact of aircraft data on the Meteorological Office FGGE data assimilation suite: February 1979 case study. Met O 20 Technical Note No. II/189.

Barwell, B.R. and A.C. Lorenc (1985): A study of the impact of aircraft wind observations on a large scale analysis and numerical weather prediction system. Quart. J. R. Met. Soc. 111, 103-129.

Bengtsson, L. (1982): Problems in Observing System Experiments. Report of the J.S.C. Study Conference on observing systems experiments, Exeter, 19-22 April 1982.
GARP WCRP WGNE Report No. 4, Appendix C, paper No. 16.

Bengtsson, L., M. Kanamitsu, P. Kållberg and S. Uppala (1982): FGGE 4-Dimensional Data Assimilation at ECMWF. Bull. Am. Met. Soc., 63, 29-43.

Bjørheim, K., P. Julian, M. Kanamitsu, P. Kållberg , P. Price, S. Tracton and S. Uppala (1982): FGGE III-B Daily Global Analysis, part 4, September 1979 - November 1979, published by ECMWF.

Cats, G.J. (1982): Current problems in medium range forecasting at ECMWF, Data Assimilation Scheme.
Proceedings of ECMWF Seminar 1981, Reading, U.K., 14-18 September 1981, published by ECMWF.

Cats, G.J. and O. Åkesson (1983): An investigation into a Marked difference Between Two Successive ECMWF Forecasts of September 1982. Contr. Atm. Phys. 56, 440-451.

Daley, R. (1980): On the Optimal Specification of the Initial State for Deterministic Forecasting. Mon. Wea. Rev., 108, 1719-1735.

Gilchrist, A. (1982): J.S.C. Study Conference on observing systems experiments. Exeter, 19-22 April 1982. GARP WCRP WGNE Report No. 4.

Hollingsworth, A., A.C. Lorenc, M.S. Tracton, K. Arpe, G. Cats, S. Uppala and P. Kållberg (1985): The response of numerical weather prediction systems to FGGE II-b data, Part I: Analyses. Quart. J. Roy. Meteor. Soc. 111, 1-66.

De Jong, H.M. (1981): Planning and implementation aspects of specific observing systems and techniques.
Document CBS/WGGOS-III/Doc. 11 for 3rd session of the Working Group on the Global Observing System. Geneva, 2-13 November 1981.

De Jong, H.M. (1982): Observing features of the ASDAR system as derived from an analysis of three ADIS/ASDAR data bases.
Document ICAP/DOC. 7 and Addendum, Interim Committee of Asdar Participants, Geneva, 6-10 December 1982.

Lorenc, A.C. (1982): The impact of aircraft data on the Meteorological Office FGGE data assimilation suite: November 1979 case study.
Met. 0 20 Technical Note No. II/190.

Nitta, T., T. Hiraki and Y. Nagano (1979): Review and formulation of the basic requirements to be placed on a future WWW System.
Document CBS/IPM-ISS/Doc. 12 for Informal Planning Meeting on Integrated WWW System Study, Geneva, 8-12 October 1979.

Saker, N.J. (1975): An 11-layer general circulation model. Met 0 20 Technical Note No. II/30.

Simmons, A. and B. Hoskins (1979): The downstream and upstream developments of unstable baroclinic waves. J. Atmos. Sci., 36, 1239-1254.

Sumi, A. (1982): On an Evaluation of Data Impact with respect to Rossby Mode in the Data Assimilation Cycle.
Part 1. On the Impact of Satellite Wind Data with Special Emphasis on the Vertical Prediction Error Correlation.
J. Met. Soc. Japan, 60, 917-930.