MONITORING OF OBSERVATION AND ANALYSIS QUALITY BY A DATA ASSIMILATION SYSTEM

A. Hollingsworth, D.B. Shaw, P. Lönnberg,
L. Illari, K. Arpe, A.J. Simmons
European Centre for Medium Range Weather Forecasts
Reading, U.K.

ABSTRACT

The purpose of this paper is to demonstrate the ability of a modern data assimilation system to provide long-term diagnostic facilities to monitor the performance of the observational network. Operational data assimilation systems use short range forecasts to provide the background, or first-guess, field for the analysis. We make a detailed study of the apparent or perceived error of these forecasts when they are verified against radiosondes. Assuming that the observational error of the radiosondes is horizontally uncorrelated, the perceived forecast error can be partitioned into prediction error, which is horizontally correlated, and observation error, which is not. The calculations show that in areas where there is adequate radiosonde coverage, the prediction error is comparable with the observation error.

This statement is discussed from a number of viewpoints. We demonstrate in the Northern Hemisphere mid-latitudes for example, that the forecasts account for most of the evolution of the atmospheric state from one analysis to the next, so that the analysis algorithm needs to make only a small correction to an accurate first-guess field. If the doubling time for small errors is two days, then analysis error will amplify by less than 10% in 6 hours.

This being the case, the statistics of the forecast/observation differences have a simple statistical structure. Large variations of the statistics from

station to station, or large biases, are indicative of problems in the data or in the assimilation system. Case studies demonstrate the ability of simple statistical tools to identify systematically erroneous radiosonde wind data in data sparse as well as in data rich areas, errors which would have been difficult to detect in any other way. The statistical tools are equally effective in diagnosing the performance of the assimilation system.

The results suggest that it is possible to provide regular feed-back on the quality of observations of winds and heights to the operators of radiosonde networks, and other observational systems. This capability has become available over the last decade through improvements in the techniques of numerical weather analysis and prediction.

1. INTRODUCTION

The purpose of this paper is to demonstrate the ability of a modern data assimilation system to provide long-term diagnostic facilities to monitor the performance of the observational network. Operational data assimilation systems use short range forecasts to provide the background, or first-guess, field for the analysis. Recent studies (Hollingsworth and Lönnberg 1985, Lönnberg and Hollingsworth 1985 - hereafter called I and II) have shown that over North America the 6 to 12 hour forecasts used as a background for the operational analyses at ECMWF have an accuracy comparable to that of the radiosonde observations. In this paper we extend the earlier calculations to Europe and the Tropics using data from a more recent period. The relative accuracy of the forecasts can provide the basis of a monitoring capability.

The plan of the paper is as follows. Following a brief outline of the assimilation system in Section 2, Section 3 provides an introductory overview of a variety of statistics concerning the differences between the observations and the three gridded fields produced by the assimilation system, namely the forecast, analysis and initialised analysis fields. The main aim of the section is to show that the assimilation works in a reasonable way by verifying all three fields against observations, and comparing the verification statistics. We also show that the differences between observations and fields are approximately Gaussian at any point, and that in mid-latitudes the magnitudes of these differences are considerably smaller than the climatological variability. By contrast the tropics are data sparse, and the perceived forecast errors are not as small relative to the climatological variability.

Section 4 uses the methods described in papers I and II to partition the perceived forecast errors over Europe, the Tropics, and North America into prediction error and observation error (which includes both instrumental errors and errors of unrepresentativeness, sometimes called sampling errors). For Europe and North America the calculations are in agreement on the fact that observational and short-range forecast error for heights and winds are comparable. Even in the Tropics this statement is true for the height field at most levels and for the wind field in the lower and middle troposphere. Since this statement has not been made before it is discussed from a variety of viewpoints in order to check that it is reasonable.

The results of section 4 are supported by the discussion in Section 5 where it is shown that if F, A, I are measures of the changes (increments) made by the Forecast, Analysis, and Initialisation steps of the assimilation, then F >A >I in Northern Hemisphere mid-latitudes. This would be impossible if the forecasts were substantially less accurate than the observations, assuming of course that the analysis is reasonably faithful to the observations. The forecast model describes most of the evolution of the atmosphere from one synoptic hour to the next; the analysis needs to make only small changes to an accurate background field.

A data assimilation system combines current observational information with earlier observational information coming from the first-guess. It is therefore reasonable to suppose that in areas with even moderate data coverage the accuracy of the analysis should be comparable with the accuracy of the observations. The doubling time for small analysis errors in mid-latitudes is generally estimated to be about 2 days, although it may be as short as 1.5 days for the shortest baroclinic scales (Arpe et al. 1985). A 2 day doubling time implies a growth rate of 9.05% per six hours. Even with pessimistic assumptions about the model performance, it is unlikely that the forecast error would grow to, say, twice the analysis error in twelve hours.

The general accuracy of the forecasts provides a new set of tools to monitor the performance of both the Global Observing System (GOS) and the assimilation system. These tools are demonstrated in Sections 6 and 7, where case studies use the statistics of the perceived forecast errors to identify systematically erroneous observations, and indeed systematically erroneous aspects of the assimilation system. Section 6 deals with wind observations while Section 7 discusses pressure observations. Although the examples mainly concern radiosonde data and Synop data, the methods are applicable to all kinds of data (Delsol 1985). The examples suggest that it is possible to provide useful information on instrumental performance to the data producers on a regular basis, as discussed in section 8.

Comparison with earlier studies of system performance suggest that this capability has become available through improvements in the techniques of numerical weather analysis and prediction over the last decade.

2. THE ASSIMILATION SYSTEM

The ECMWF assimilation system is an intermittent insertion system consisting of three main steps: analysis, initialisation and forecast to provide the next first-guess field. The analysis system is described by Lorenc(1981) and is an application of the optimal interpolation technique discussed by Gandin (1963), Rutherford (1972), Schlatter (1975) and Bergman (1979). Similar methods are used in several operational centres (Gustavsson, 1981). The distinctive feature of the ECMWF implementation is that the analysis is performed for a large number of grid-points and variables simultaneously, which requires the selection of a large quantity of data for each analysis volume. The demands on computer power are correspondingly large; a typical analysis requires the inversion of several thousand matrices with orders between 100 and 200.

The initialisation scheme is an application of the non-linear normal mode scheme proposed by Machenhauer (1977), and described by Temperton and Williamson (1981) and Williamson and Temperton (1981), and modified to include diabatic effects by Wergen (1982, pers.comm.).

The model used to produce the 6-hour forecast is the operational ECMWF forecast model. Up to mid-April 1983 this was a finite difference model (Burridge and Haseler, 1977), and since then it has been a spectral model as described by Simmons and Jarraud (1984). The physical parameterisation package has been described by Tiedtke et al. (1979).

The ECMWF assimilation system has been used to produce global IIIb analyses for the FGGE year (Bjorheim et al. 1982, Bengtsson et al. 1982). The response of this and other assimilation systems to the FGGE IIIb data has been compared in Hollingsworth et al. (1985) and Arpe et al. (1985).

3. STATISTICS OF THE PERCEIVED FORECAST ERROR

The goals of this section are (i) to show that the assimilation works in a reasonable way, by verifying the forecast, analysis, and initialised analysis fields against observations and comparing the verifications, (ii) to indicate the value of these data in quality control studies, (iii) to show that the statistics of the observation minus field differences are approximately Gaussian at any point and (iv) to show that in mid-latitudes the magnitudes of these differences are considerably smaller than the climatological variability. By contrast the tropics are data sparse, and the perceived forecast errors are not as small compared to the climatological variability. Finally, the problems of estimating instrumental and model bias are also touched on.

Two sets of observational data are used in this section, the unscreened set, and a subset called the screened set. The unscreened set consists of all data which was presented to the analysis. This data has passed telecommunications checks on the integrity of the message, as well as simple checks on the departure of the reported values from climatology, but it has not been screened by the analysis check. The histograms of this data are discussed in 3.2. The screened subset consists of all data which were accepted as correct or probably correct by the analysis system (Lorenc 1981). The screened data is used in the verification of the data assimilation fields in 3.1.

We verify the first-guess, analysed and initialised fields against 12Z radiosonde observations in January-March 1984 for three regions: North America (30-60°N, 50-140°W), Europe(35-65°N, 12W-25°E) and the Tropics (20°S-20°N). Only screened observational data were used in the verifications, and only stations that reported frequently enough (minimum of 40 reports in the Tropics and 60 reports elsewhere) were used. Certain tropical networks have very noisy height reports but good quality wind reports; such networks have been excluded from the present calculations. The North American, European and Tropical regions cover 4.5%, 1.7% and 34.2% of the area of the globe. The number of 12Z radiosonde reports at 200 mb were 6847, 6046 and 4066 respectively, from 83, 72 and 47 stations. We shall refer to the observation-forecast, observation-analysis and observation-initialised analysis statistics as the O-F, O-A and O-I statistics respectively.

3.1 The fit to observations of the assimilation fields

(a) Overview of the Assimilation Cycle

Figs. 1 and 2 show the verifications of the screened v-component and height data against each of the three gridded fields, for the three areas mentioned above. The first-guess is furthest from the observational data, the analysed field is closest to it, while the initialised field is intermediate between the other two. The fit of the uninitialised and initialised analyses to the observations is very similar in mid-latitudes. This indicates that the uninitialised analysis is sufficiently well balanced that most of the observational information is communicated to the slow modes of the forecast model. This is important in ensuring an accurate first guess for the next analysis.

In the Tropics on the other hand, the analysis changes to the wind field survive the initialisation procedure much more readily than the height information, which is largely lost during the initialisation procedure. This effect comes mainly from changes in surface pressure, rather than changes in thickness, because the largest effect of the initialisation occurs in the external mode.

The mean effect of the initialisation on the meridional component of the wind in the upper Tropical troposphere is to reinstate a bias which occurs in the first-guess and is largely corrected by the analysis. This feature is possibly sensitive to the number of modes used in the initialisation (Hollingsworth and Cats 1980), although the current initialisation differs from the 1980 version through the incorporation of a smoothed representation of the model's diabatic forcing (Wergen 1982, pers.comm.)

(b) The wind forecasts over North America and Europe

The mid-latitude first-guess wind fields Fig. 1(a,b) are furthest from the observations at the jet level, with the perceived forecast error being about 5 m/s rms per component. In mid-latitudes the winter rms climatological variability for v increases the surface value of 5 m/s to about 15 m/s at jet level and then decreases again in the stratosphere. (Oort 1983, Figs. A44-A47 for December to February). The rms O-F differences for wind show the same pattern of variation with height. The perceived forecast error of the short range wind forecast is about one third the climatological variability at the level of maximum error in the troposphere; however the perceived forecast error has a substantial component arising from the observational error.

As demanded by the theory of optimal interpolation, the analyses fit the observations to within the specified observation error, but do not fit the observations exactly. To do so would be unreasonable since the analysis is

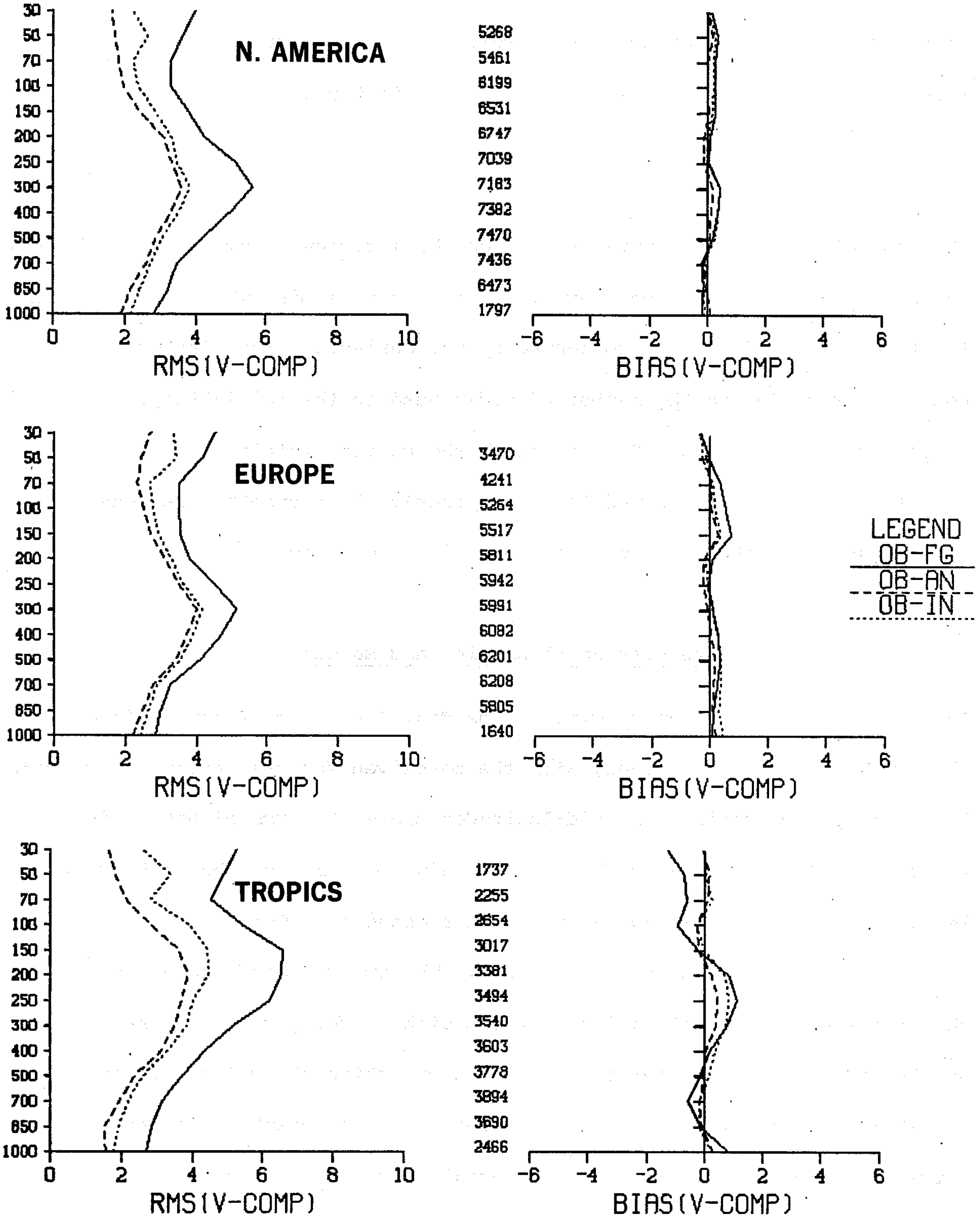


Fig.1 Root mean square (left) and bias (right) statistics on the fit of the gridded assimilation fields to screened rawinsonde reports of the northward (v) component of the wind over North America (top), Europe (centre) and the Tropics (bottom). The fit of the first-guess, analysis and initialised analysis is indicated by the solid, dashed and dotted lines respectively. For the bias, the result shown is for observation minus field.

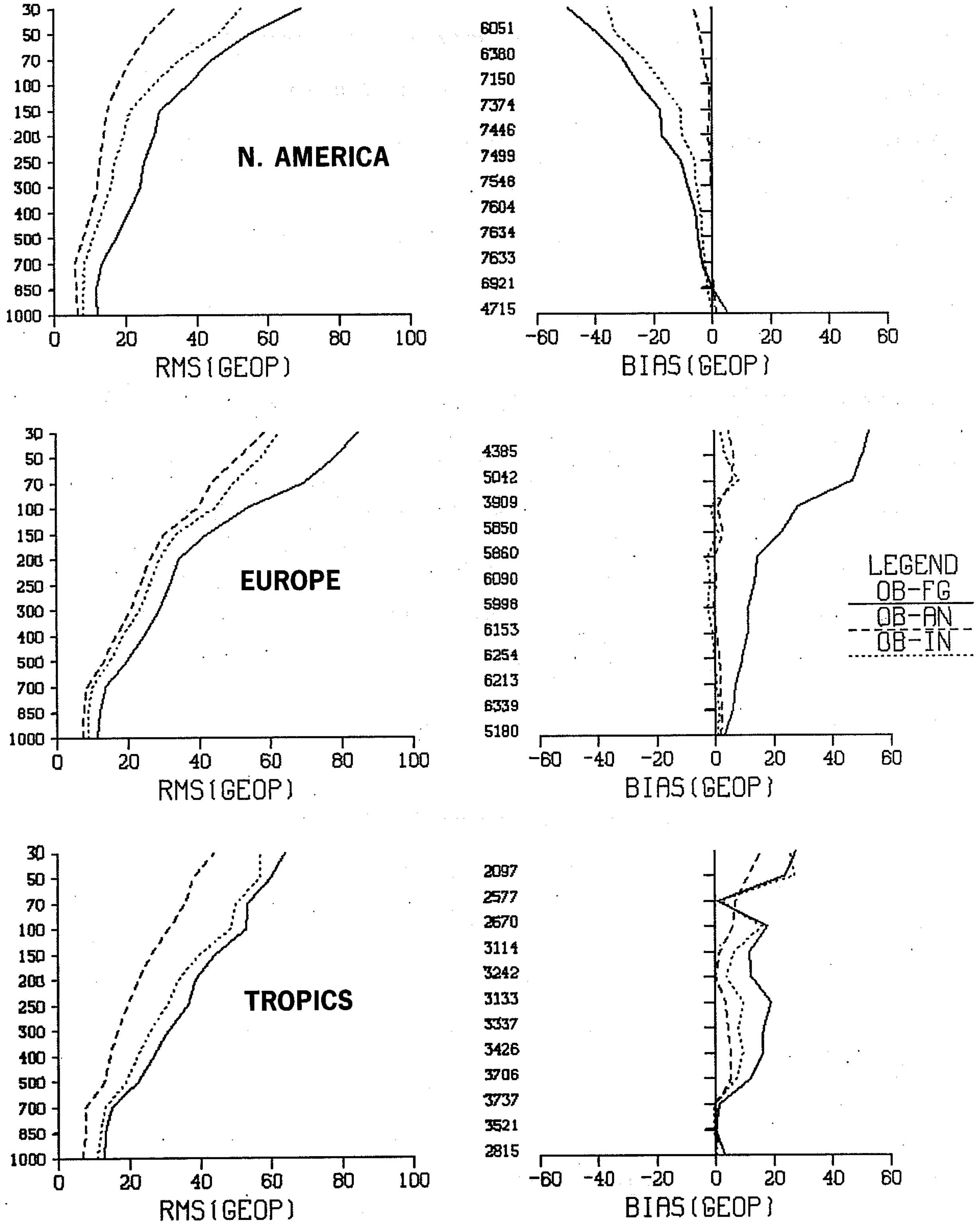


Fig.2 As Fig. 1 for the height reports. The unit is meter.

seeking to combine the observational data and the forecast data to produce an estimate that is more accurate than either on its own. If one supposed, for the sake of discussion, that the analyses were error free, then O-A would represent an estimate of the observational error. Because of correlation between the observation error and the analysis error, the O-A differences probably underestimate the observation error.

(c) The height forecasts over North America and Europe

The first-guess height differs from the observed height by about 28 m rms at 200 mb over North America and by about 34 m over Europe, Fig. 2(a,b). As will be shown later this is mainly because the European network is noisier. Both the American and European O-F rms height differences have substantial contributions from bias effects which will be discussed Section 3.4. The biases mainly affect the stratospheric values.

The rms O-F height values increase monotonically with height. Oort's data indicates that the climatological standard deviation for daily variations in these latitudes is about 75 m at the surface increasing to about 150 m at the tropopause; there is then a decrease to about 125 m at 100 mb with a further increase higher in the stratosphere. Thus the tropospheric values for the perceived forecast errors show substantial skill relative to climatology.

(d) The Wind Forecasts over the Tropics

Fig. 1(c) shows the plots of the rms O-F, O-A, O-I differences for the tropical wind field between 20°N and 20°S. The perceived forecast errors per component increase from about 3 m/s at the surface to about 6.5 m/s in the upper troposphere. Oort's estimates of the climatological variability are also about 3 m/s at the surface and about 7 m/s in the upper troposphere.

The difference between the rms O-F values and the rms climatological values is small. This does not mean that the forecasts are useless, because the O-F differences contain an important contribution from observational error.

At and below 300 mb the rms O-F values for the v-component in the Tropics are no larger than over Europe or North America. The large differences in the climatological variabilities of the two regions implies that the mid-latitude forecasts are much more skilful. This may be due to many factors, not least of which must be the differences in data density. Although other types of data beside rawinsondes are received from the Tropics, notably Cloud Track Vectors and satellite thicknesses, their information content for the wind field is limited compared to the radiosondes, as they sample the wind field at only two levels, or not at all.

(e) The height forecast over the Tropics

Oort's figures for the variability of daily variations in the tropics indicate values of about 15 m near the surface, increasing slowly to about 40 m at 200 mb, and then more rapidly to about 70 m at 50 mb. The results in Fig. 2(c) show very similar values for the O-F differences.

The fact that the O-F values for the Tropical heights are so close to the climatological values is of no less concern than for the Tropical winds. Although height data is sometimes said not to be of importance for tropical analysis, there are theoretical grounds for the opposite view, particularly on the larger scales. As discussed by Cats and Wergen (1982), it is very important that observational data be correctly partitioned between Rossby modes and Kelvin modes in the tropics. To make the correct discrimination it is necessary to have an accurate estimate of the height field, because both wind components are nearly geostrophic for the Rossby modes, while only the zonal component is nearly geostrophic for the Kelvin modes.

Up to 150 mb the bias in the geopotential associated with the O-F differences is consistent with the known tendency of the model to cool in the tropics. The rapid variations in the geopotential bias near 70 mb arise from the problem of interpolating a sufficiently accurate 70 mb height from geopotentials on the model's intermediate coordinate levels at about 50 and 100 mb, in a region where the tropopause is a very sharp feature.

Using the screened data, the perceived forecast errors in both height and wind are much lower than the climatological variability in mid-latitudes than in the tropics. Forecasting for the Tropics is difficult both because of the scarcity of data and because of the difficulties of model formulation. Heckley (1985) and Kanamitsu (1985) show that there are substantial mean forecast errors in the Tropics, probably due to problems in the formulation of the physical parameterisations. As shown in Section 4, the large O-F differences for the screened tropical data occurs mainly because of model problems. Yet there may be a contribution from inadequate quality control on accepted data, as quality control is necessarily less reliable in data sparse areas.

3.2 Unscreened histograms of perceived forecast errors

We now consider some statistics of the unscreened data. As noted above this data has been subject to telecommunications, decoding, and climatological checks. It has not been subjected to the multivariate check against the other data and the first guess which is a feature of the ECMWF system (Lorenc 1981). The O-F differences may be larger for the unscreened data than for the screened data, for obvious reasons.

Provided all the stages between the taking of a measurement and its presentation to the analysis system are correctly executed, one should expect the differences between the measurement and the forecast to have a Gaussian distribution. Departures of the histograms of the unscreened data from a Gaussian shape have been useful in identifying errors in our own procedures as data users, and occasional errors in the working practices of data producers. Examples include previously undetected errors in coding or decoding procedures. Thus studies of the unscreened data, as presented to the analysis, can be as valuable as studies of the screened data.

(a) Unscreened histograms of perceived wind forecast errors

Fig. 3 shows histograms of the unscreened 200 mb u component O-F differences for North America, Europe and the Tropics for the three month period. All three histograms are roughly Gaussian in shape and all show small biases on average. The fact that the bias averaged over all stations is small does not imply that all stations have small biases. The compact mid-latitude histograms with low standard deviations indicate that there are few problems with radiosonde wind reports from North America and Europe.

There are large differences in the spread of the histograms between mid-latitudes and the Tropics; the standard deviations of the distributions are 4.7 and 4.2 m/s in North America and Europe but 6.7 m/s in the Tropics. Much of this difference is due to lower forecast skill in the Tropics.

(b) Unscreened histograms of perceived height forecast errors

Fig. 4 shows histograms of the unscreened O-F differences for the 200 mb

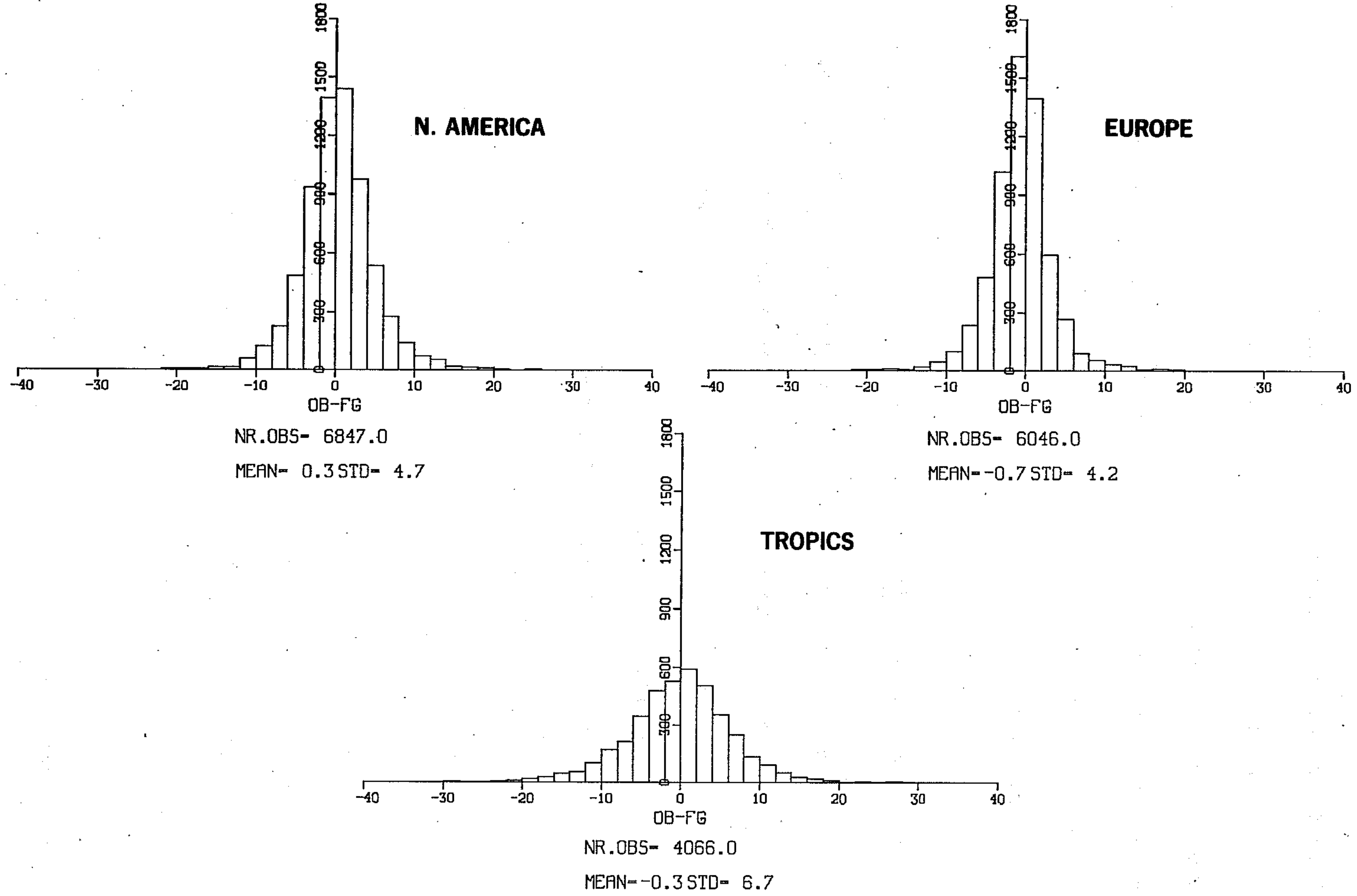


Fig.3 Histograms of the observation minus first guess differences for the unscreened 200mb u-component reports for North America (left), Europe (right), and the Tropics(bottom). The units are m/s.

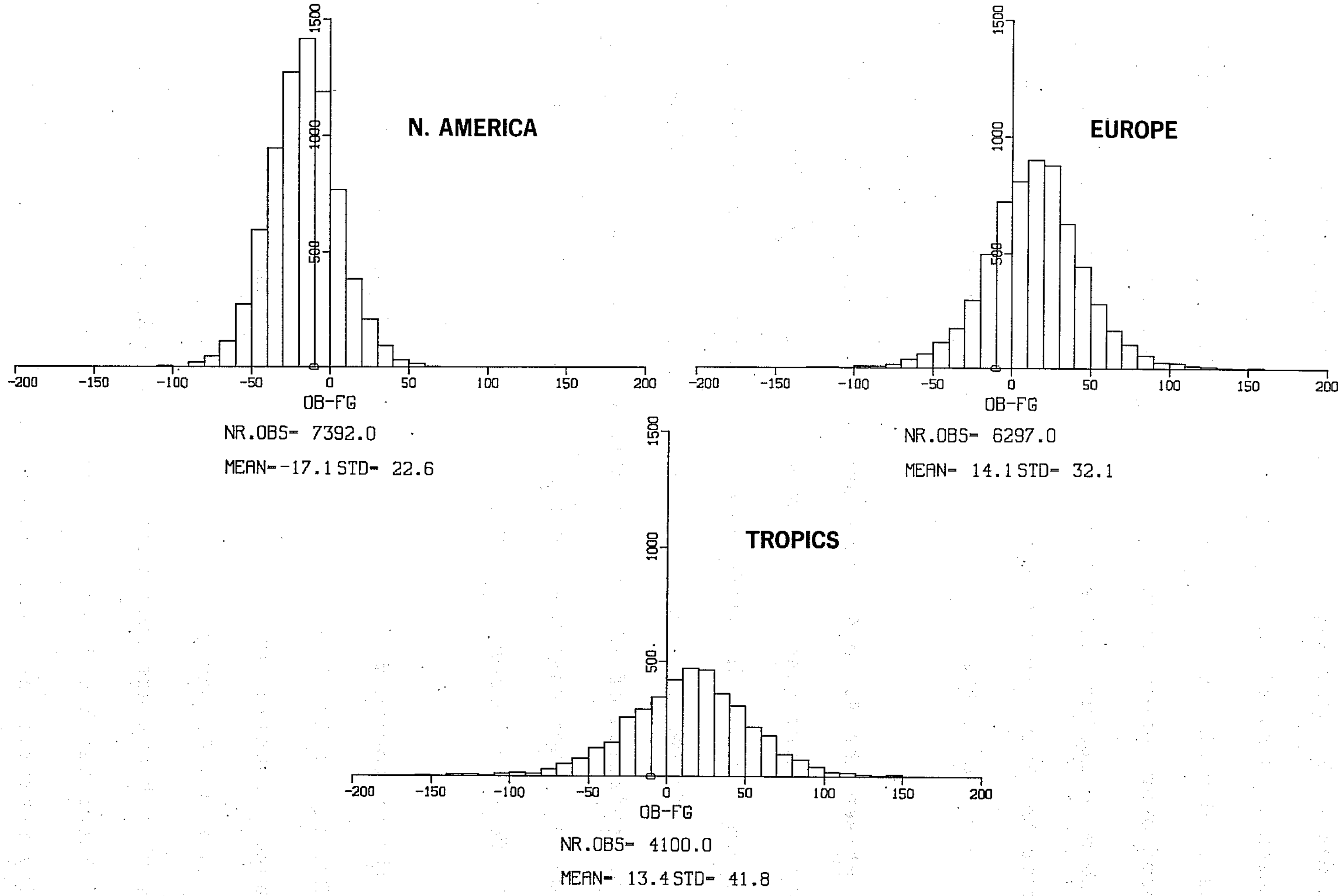


Fig.4 As Fig. 3 for the unscreened 200mb height reports. The unit is m.

height field over North America, Europe and the Tropics for the same period as Fig. 3. The North American histogram is compact, with very short tails. This indicates that there are few quality control problems with the North American height data. By contrast, the European histogram is broader (standard deviation 32.1 m, cf 22.6 m) and the tails are longer, with many more occurrences of deviations in excess of 50 m. The short range forecast quality is as good over Europe as it is over North America, if the verifications are made against the analyses or by the methods of section 4. This suggests that the European network is noisier than the American network, possibly in part because of the variety of sonde types used in Europe.

The difference between the North American and European standard deviations may not seem large at first glance. The proper way to compare them is to calculate the square root of the difference of the squares. This calculation indicates that the difference in the noise level between the two networks has itself a standard deviation of 22.8 m. In effect, the European network has a noise level which is higher than the American noise level by an amount equivalent to the forecast errors over North America.

Fig. 4(c) shows the histogram for the distribution of unscreened O-F height differences at 200 mb for the Tropics for the same period. The histogram appears approximately Gaussian with a very wide spread, the standard deviation being 41.8 m. Comparison of this histogram with that for North America suggests that there are important problems with the quality of these height reports. These results do not show the full extent of the problem as certain very noisy networks have been excluded from the calculation. The average bias is small (13.4 m), but we have not examined the variations of the bias on a station by station basis. Such an investigation is needed to understand the large differences between the screened and unscreened data.

3.3 Separation of radiosonde and forecast bias in mid-latitudes

Detailed examination of the mid-latitude statistics shows substantial height biases in the observation minus forecast differences at individual stations. It is important to be able to separate the role of instrumental bias from model bias if one is to attempt to correct the observational bias; Nash (1984) discusses radiosonde height biases.

Mean forecast errors in the free atmosphere above the boundary layer may be expected to show an approximately geostrophic relationship between the height and wind biases. If no such relationship is seen between the biases then the likelihood is that the biases are instrumental in origin.

Figs. 5 and 6 show the 200 mb height and wind observation minus forecast biases at radiosonde stations in North America and Europe for June 1983. Despite some noise in the height biases over North America, a pattern is nevertheless discernible there, with a clear tendency for the forecast to have too large heights over the western part of the continent. The wind biases show a pattern which appears, at least qualitatively, to have a geostrophic relationship with the height biases. Wallace et al (1983) discuss forecast errors in this region.

Over Europe, no such association can be seen. The European height biases are generally larger and much more random on the largest scales. On small scales some international boundaries can be identified from changes in sign of the biases. There is little apparent association between the height biases and the wind biases, which suggests that the height biases are dominated by instrumental bias. Thus the case for correcting the individual European

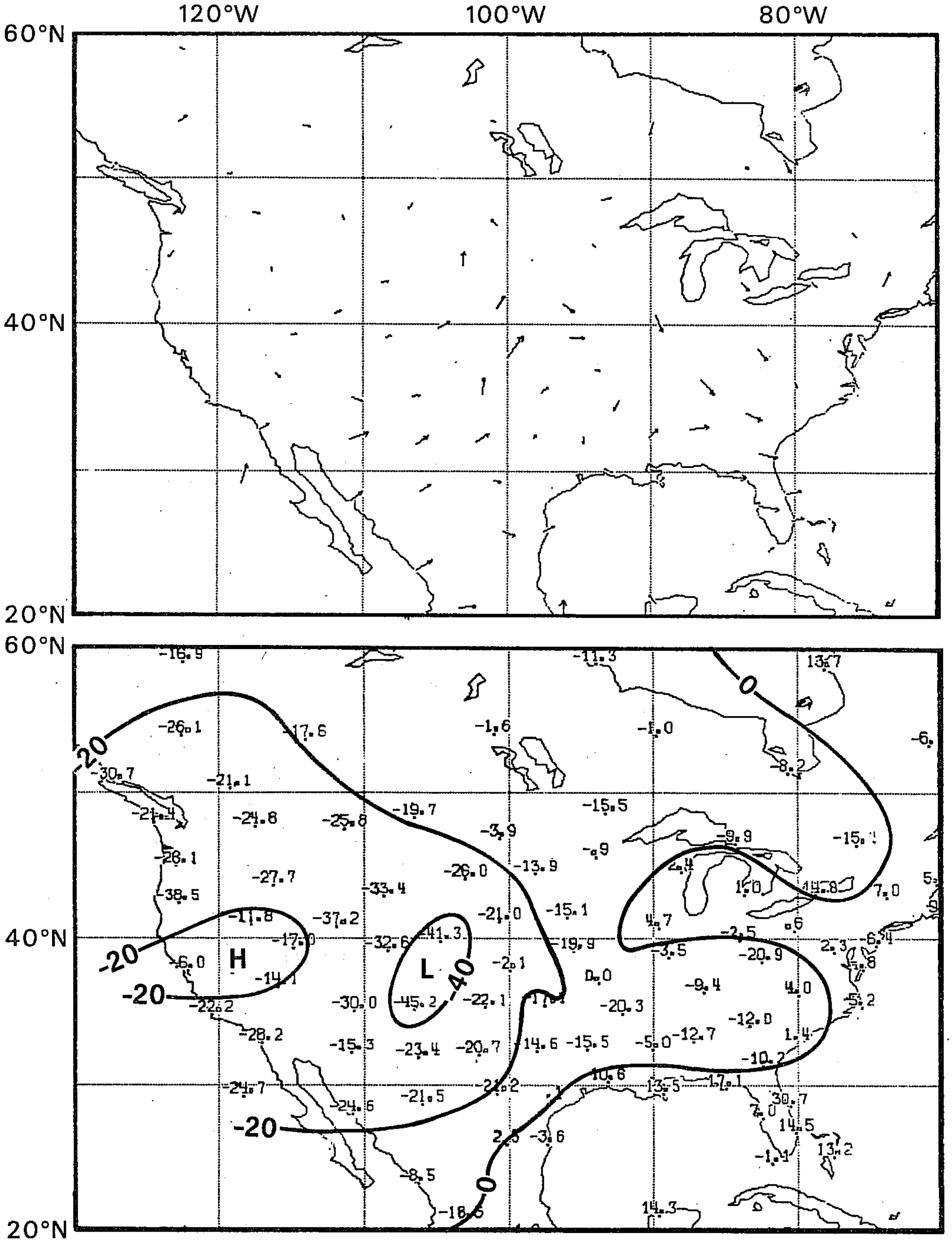


Fig.5 Bias (observation minus first guess) between the reported and forecast (top) 200mb wind field over North America, (bottom) 200mb height field over North America, for June 1983, at 1200 GMT. The contours on the height bias field were hand-drawn.

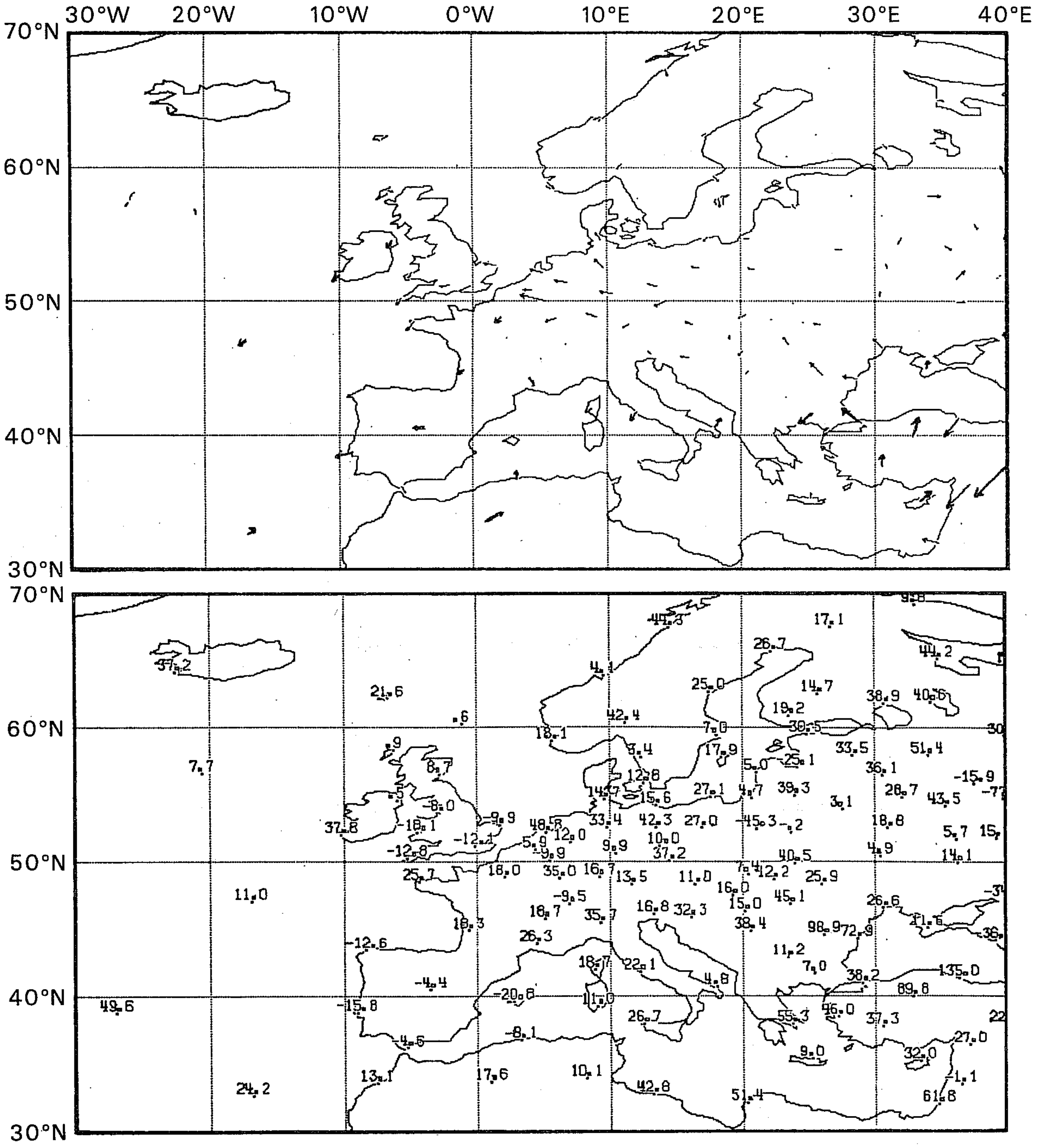


Fig.6 As Fig. 5 for Europe. No smooth contours could be drawn over Europe.

radiosondes is stronger than that for the individual North American sondes. On the other hand, studies by Nash (op.cit.) and Lange (1985, pers.comm.) suggest that the North American sondes have a uniform radiation error which should be corrected for.

The gradients in the sonde biases over Europe are large. There are numerous instances where the gradient in the instrumental bias exceeds 50 m in 500km. In a multivariate analysis system this could lead, in the absence of wind data, to randomly distributed geostrophic wind biases in excess of 10 m/s.

3.4 Discussion

In the mid-latitude areas from which we receive adequate data, the perceived error of the short-range forecasts is quite low, indicating considerable short-range forecast skill. The quality of the data is generally good, but the European height data is clearly noisier than the North American data. The analysis draws for the observational data to within the expected tolerances. Most of the information is retained during the initialisation, and is communicated to the slow modes of the forecast model.

In the Tropics, an area from which we receive relatively little radiosonde data, the forecasts are clearly of poorer quality than in the other areas studied. The quality of the wind data is shown to be generally good in the next section. The unscreened tropical height data appears to be rather noisy, even though some networks with very noisy height data have been excluded. It is shown in the next section that the screened Tropical height data is of similar quality to the European height data. The wind data largely survives the initialisation procedure in the Tropics, but the geopotential data is largely rejected, mainly in surface pressure.

The perceived short range forecast errors over Europe have a large contribution from instrumental bias in the height field. The biases in the height reports over Europe are so large that in the absence of wind data they could easily cause randomly distributed geostrophic wind biases in excess of 10 m/s over the area.

4. THE ACCURACY OF THE SHORT RANGE FORECASTS

In this section we use the methods of papers I and II to partition the perceived forecast errors over Europe, North America and the Tropics into prediction error and observation error; the working assumption is that the observational errors are horizontally uncorrelated. Such a calculation gives a more precise estimate of the accuracy of the short-range forecasts than is possible from the verifications discussed in the last section. The term observation error includes both instrumental error and sampling errors. The screened dataset is used for the calculations.

The correlation of the perceived forecast error is determined for all possible station-pairs, and the isotropic component is determined by compositing the results, and fitting a function defined by a truncated series expansion. When extrapolated to the origin this function provides a partitioning of the perceived forecast error into the desired component contributions. As discussed in papers I and II the estimate of the observation error depends on the truncation used in the fitting of the data. In the light of the results of those papers, we decided to limit the truncation to 10 terms in a Bessel series, with a search radius of 3000km; the truncation is equivalent to a truncation at planetary wavenumber 68. The expansions of the series are terminated at a lower truncation if the coefficients of the series become negative.

The calculations in this section treat the biases in the data differently from the last section. The mean difference between forecast and observation is removed separately for each station, and the average variance of the resulting time series is defined as the ensemble mean of the resulting station variances. In this way the effect of variations in station bias is excluded from the calculation.

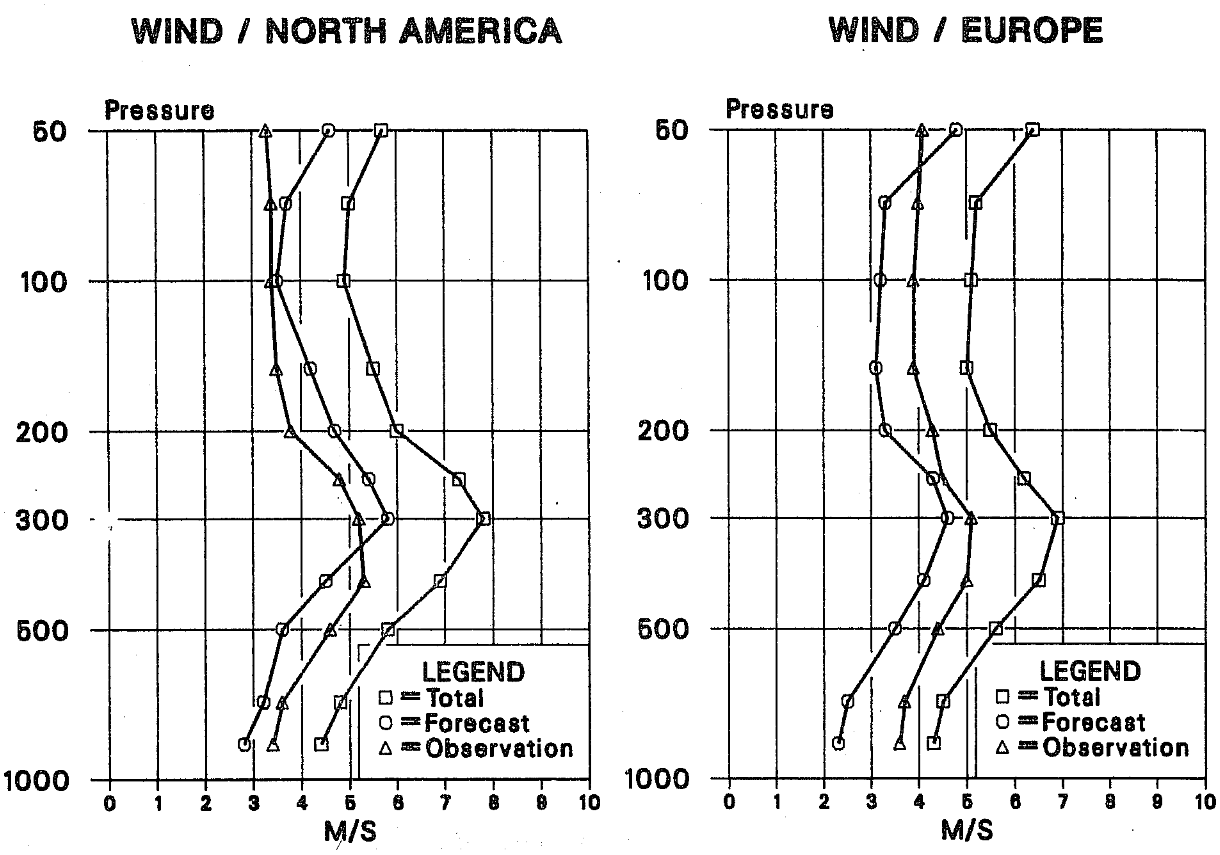
Fig. 7 shows the rms perceived forecast error, together with the observational error and prediction error components, for the vector wind, for each of the three areas for the first quarter of 1984. In comparing these results with those of the last section, it is important to note that here we discuss rms vector wind errors, while there we discussed rms wind component errors; in addition the effect of variations in station bias have been removed.

There is excellent consistency between the three estimates of observational error in the troposphere, with peak values of 5 m/s for the observational error in all three areas. The values for the observation error agree very well with those estimated by WMO (1983). The very low estimate of observational error around 100 mb in the Tropics is suspect, because the volume of available data decreases rapidly at these levels.

There is little doubt that the prediction error for the vector wind in the upper tropical troposphere is substantially larger than the corresponding values in mid-latitudes, 7.5 m/s compared to 5 m/s. As anticipated in the last section the vector wind prediction error in the Tropics is rather less than the climatological variability of 10m/s. The North American and European vector wind prediction errors have similar vertical profiles, with magnitudes of 2 to 3 m/s at 850 mb, 4.5 to 5.5 m/s at 300 mb and 3.5 to 4 m/s above 200 mb. The mid-latitude climatological variability for the vector wind is in excess of 20 m/s near the tropopause.

Fig. 8 shows the corresponding results for the height forecasts. The estimated prediction error for height is very similar over North America and Europe (about 15m rms at 200mb), which agrees with verifications of the short range forecasts against gridded analyses. The difference in perceived forecast error in the two areas when the forecasts are verified against radiosondes therefore arises from the higher level of observational error in the European network (about 25m rms compared to 15m, at 200mb). The analysis system removes this noise, so that the verifications against gridded analyses show lower levels of error.

It is noteworthy that the screened tropical sondes used in our calculations have very similar height observational errors to the European network. This result is not true for all tropical height measurements, as some networks with very high noise levels have not been used in the calculations. It is also encouraging to note that the prediction error for height in the upper tropical troposphere is substantially lower than the climatological variances. The height forecasts in this region are definitely more accurate relative to climatology than the wind forecasts.



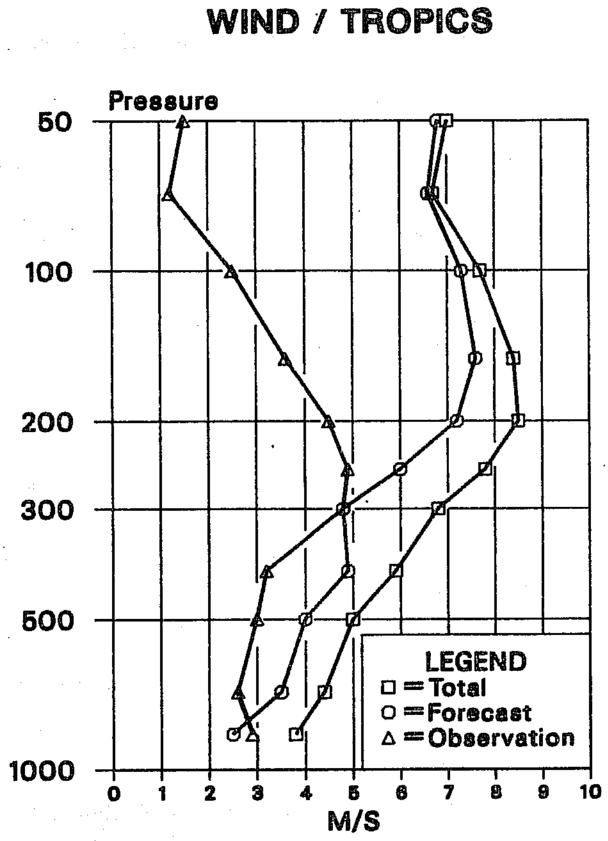


Fig. 7 For each of the three areas North America(top left), Europe(top right), and the Tropics (bottom), the plots show the total perceived forecast error for the vector wind using screened data, and the calculated prediction error and observation error as indicated in the legend. The unit is m/s.

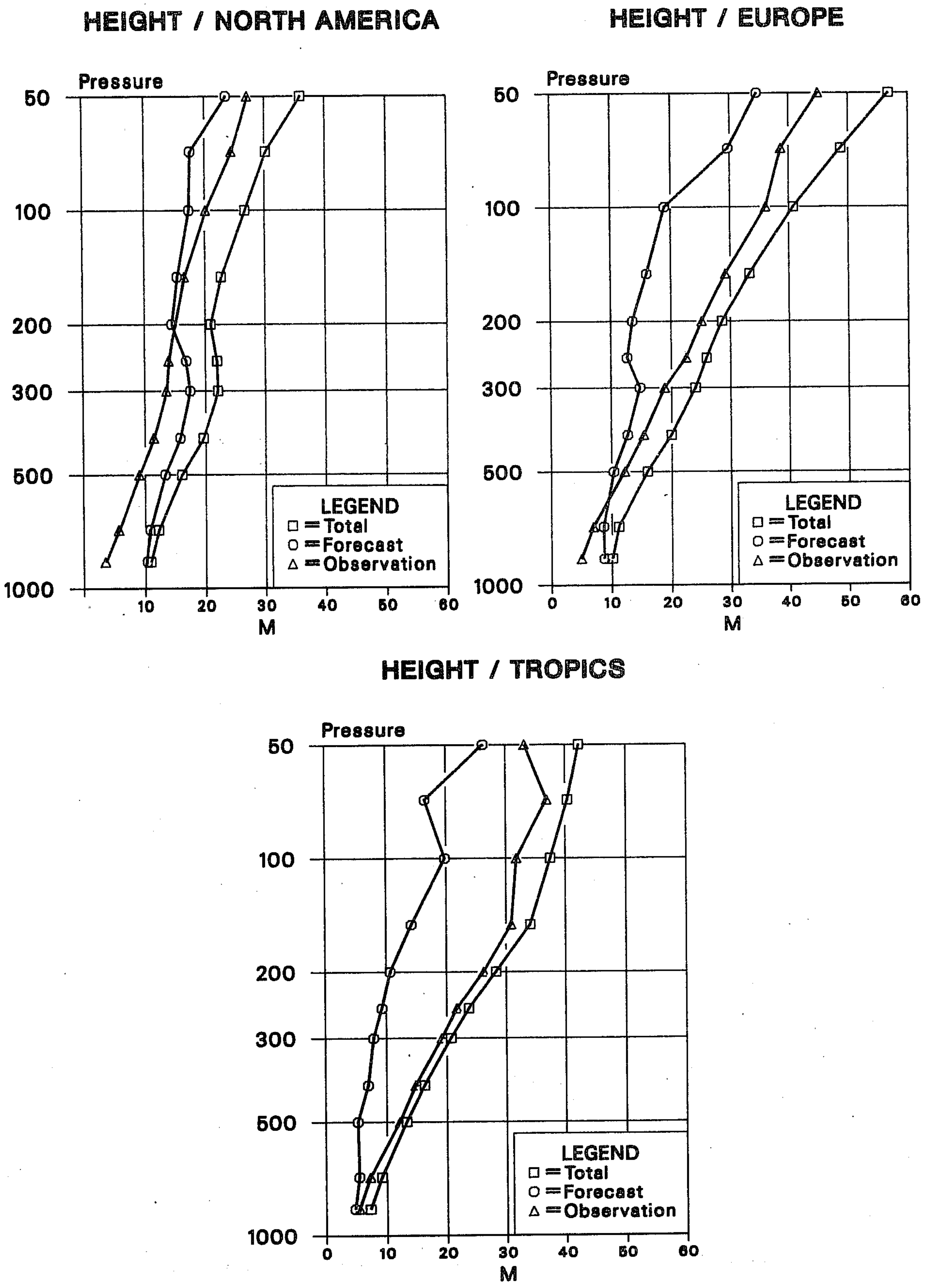


Fig.8 As Fig. 7 for the height in m.

The prediction errors for height and wind over Europe and North America are clearly comparable with the observation errors, and over Europe are smaller than the observation errors. In the Tropics by contrast, the wind forecast errors are somewhat larger than the observation errors in the middle and lower troposphere, and substantially larger than the observation errors in the upper troposphere.

4.1 Changes in forecast skill over the last decade

The mid-latitude results are further examined in the next section through a detailed investigation of the magnitudes of the changes made by each of the steps of the analysis procedure. Here we make a simple check on the validity of our main contention by a comparison of present levels of forecast skill with those of a state of the art assimilation system of about a decade ago. Fig. 9 compares the total perceived errors of the short-range forecasts over North America, with those discussed by Hollett (1975) for similar forecasts for essentially the same area, and verified in the same way. Hollett used the Canadian five-level primitive equation model and his data pertains to the period August to October 1974. The ECMWF data is shown for the periods January-March 1984 and June-August 1984, to indicate the seasonal variations.

Given that there has been little change in radiosonde accuracy or coverage, it is clear that the forecast accuracy has improved substantially over the last decade. Many factors may have contributed to this improvement, including higher model resolution and better analysis and initialisation techniques.

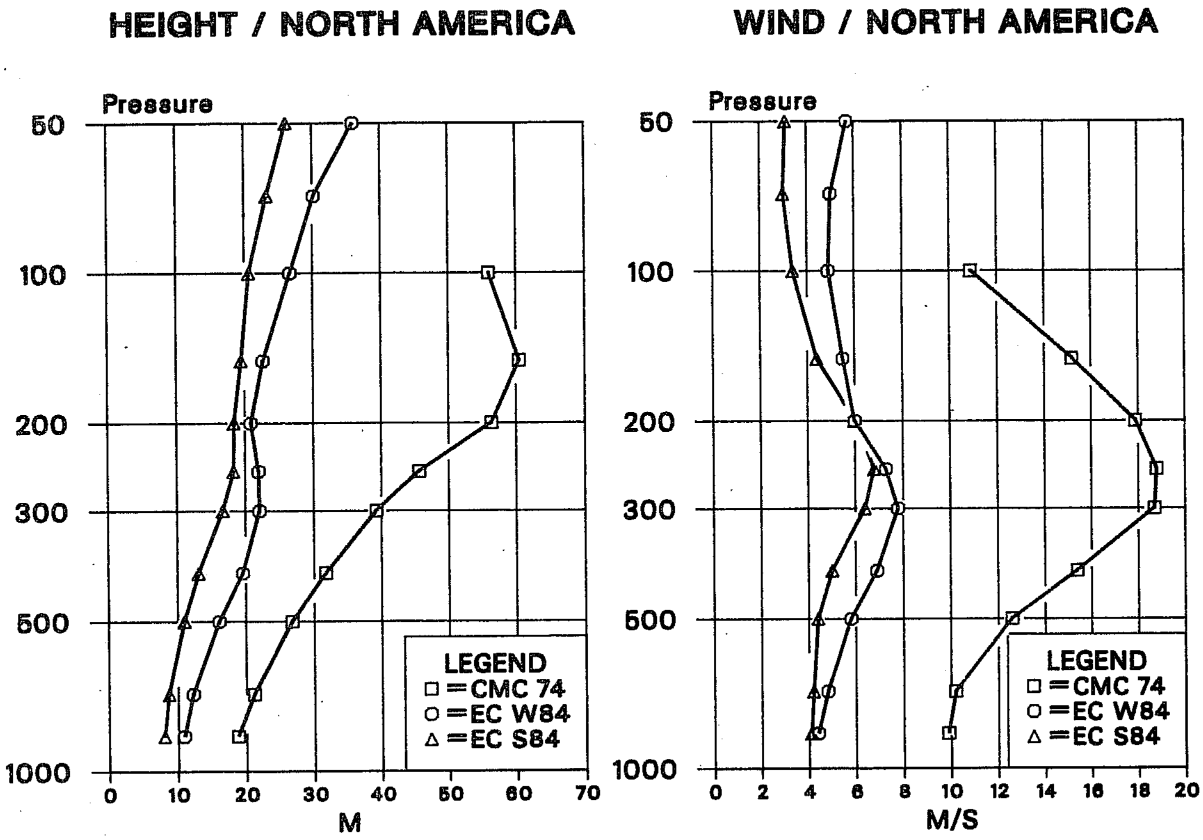


Fig.9 The plots show the perceived forecast errors for height (left, in m), and wind (right, in m/s), from verifications against North American radiosonde data. The curves are for the results of Hollett (1975) indicated by CMC74, and for the ECMWF forecasts for January to March 1984 (indicated by EC W84) and June-August 1984 (indicated by EC S84).

5. INTERNAL CONSISTENCY OF THE ASSIMILATION SYSTEM

Each of the three main steps in the assimilation (forecast, analysis, initialisation) 'increments' the output of the previous step to provide the input to the next step. The statistics of these increments provide a useful measure of the changes made by each step. If the short-range forecast accuracy is comparable to the observation accuracy, then the forecast must capture most of the evolution of the atmosphere from one synoptic hour to the next. In such a case one should expect that

$$F > A > I$$

where F, A and I are measures of the magnitudes of the forecast, analysis and initialisation increments; the measure used is the rms height increment at 500 mb.

A necessary requirement for this result to be true is that the spatial distribution of the temporal variance of each of the three time series of forecasts, analyses, and initialised analyses should be similar; in other words it is necessary that neither the forecast nor the initialisation should have a significant damping effect on the variance of their input fields. Hollingsworth and Arpe (1982) examined this question in a pilot study for the present paper. They found that the variances of the first-guess fields, the uninitialised analyses and the initialised analyses were very similar in mid-latitudes. This necessary condition is also satisfied in the present study, but no results will be shown.

5.1 Northern Hemisphere

Fig. 10 shows Northern Hemisphere maps of the rms 500 mb height increments due to the forecast, analysis and initialisation steps, for operational 12Z analyses in October 1983. In the main storm tracks over the oceans it is clear that F > A, since the changes made by the 6-hour forecast are substantially larger than the analysis changes. Over the continental areas this is also true, but not to such a marked degree, mainly because the forecast changes are smaller. The overall spatial homogeneity of the analysis increments is striking. The area mean values for F and A are 20.9 and 12.6 m respectively.

In most places it is also clear that A > I, as the initialisation increments are definitely smaller than the analysis increments. The area mean values of A and I are 12.6 and 6.3 m. In general the effect of the initialisation is to reject some of the change made by the analysis, and to bring the analysis closer to the first-guess. The results at 500 mb are typical for the height fields; the effect of initialisation on the height field is generally larger than on the wind field.

5.2 Southern Hemisphere

Fig. 11 shows the corresponding results for the Southern Hemisphere. The magnitudes of the forecast increments in the main oceanic storm tracks are rather similar to, though somewhat larger than, those in the Northern hemisphere. The magnitudes of the analysis increments in the Indian Ocean and Atlantic Ocean sectors are less than the forecast increments.

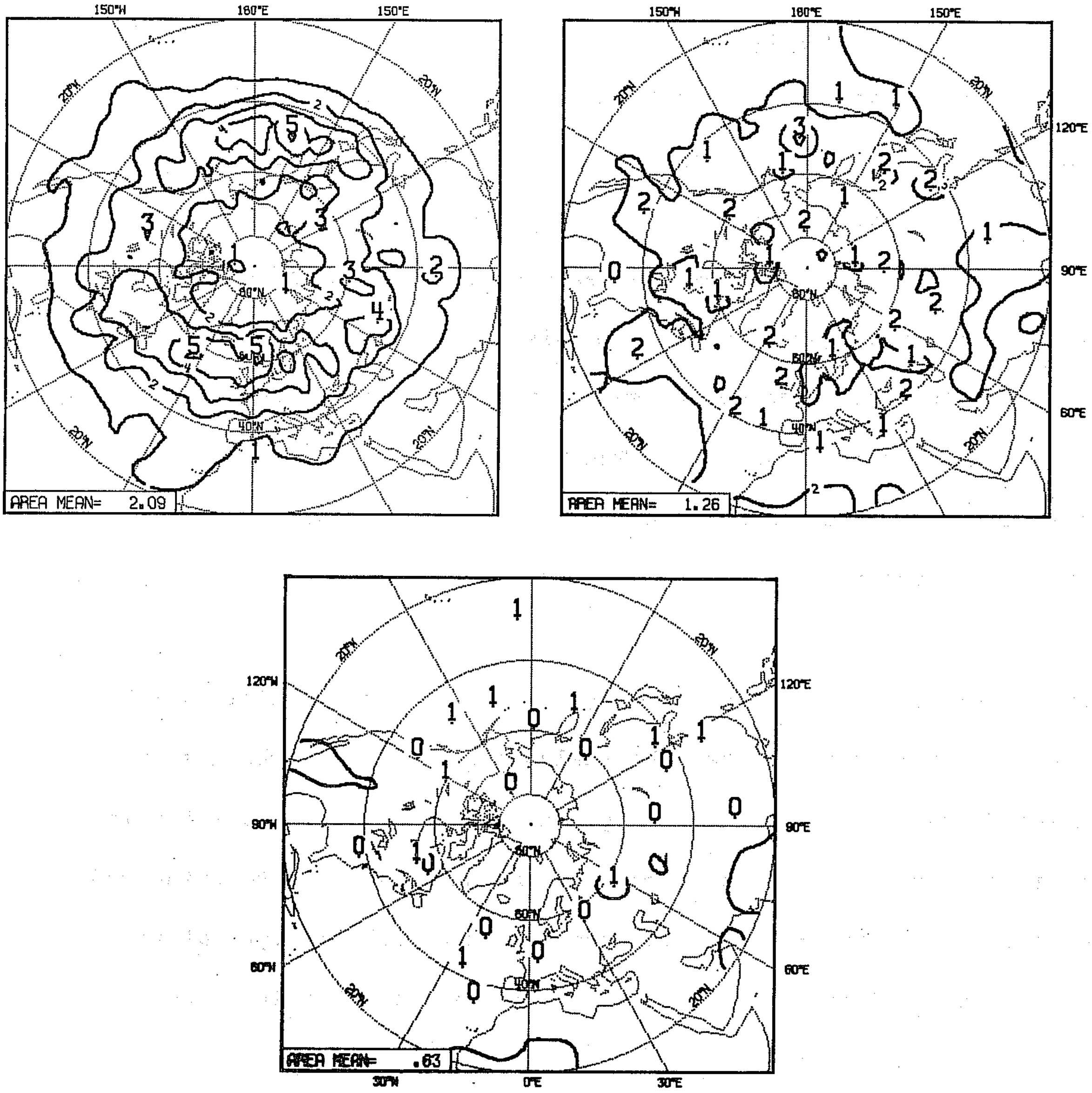


Fig.10 Root mean square values of the 500mb geopotential increments made by the 6-hour forecast (top right), analysis (top left) and initialisation (bottom) for October 1983 in the northern hemisphere.

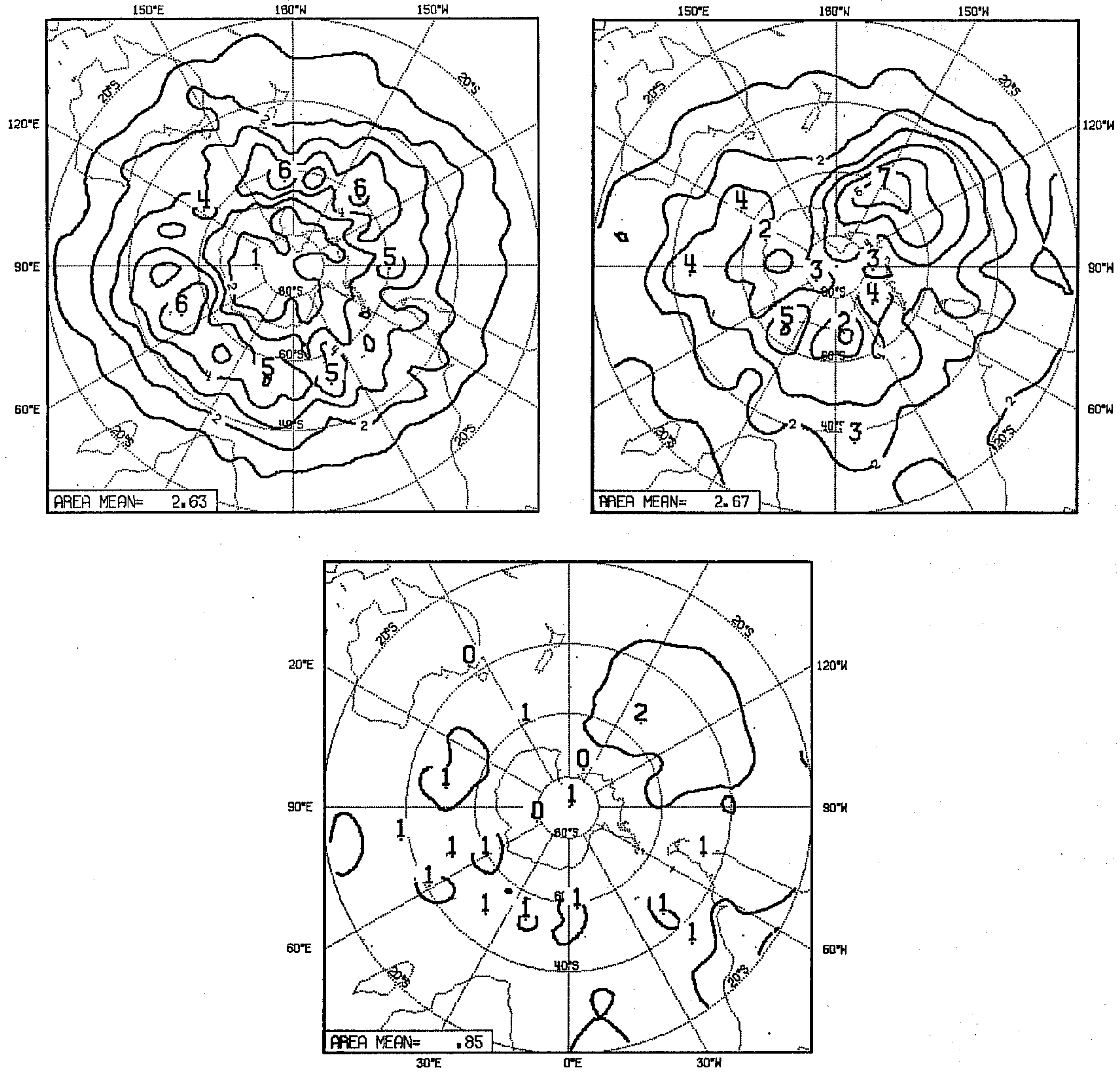


Fig.11 As Fig. 10 for the southern hemisphere.

There is a substantial analysis problem in the South Pacific sector at latitude 60°S, to the south-east of New Zealand. This is perhaps the most data void area in the world. The main data sources are the polar-orbiter SATEM retrievals, which provide thickness data, and the Australian PAOBs which provide estimates of surface pressure derived from interpretation of satellite imagery (Guymer 1978). The problem in the South Pacific sector arises from the absence of independent surface pressure data to guide the producers of the PAOB data (Guymer 1985, pers. comm.). There are also indications that the error level assigned to the PAOBS in our system is too low. A recent study (Shaw et al. 1984) indicates that the time continuity of the analyses can be improved by giving less weight to the PAOBs.

The initialisation increments in the height field are substantially smaller than the analysis increments. This indicates that even with large differences between observation and first-guess, the multivariate analysis system derives a well balanced analysis.

5.3 Discussion

Taken together, these results indicate that, particularly in the mid-latitudes of the Northern Hemisphere, much of the evolution of the observed state of the atmosphere is described by the forecast model with a good degree of accuracy. This is true to some extent even in the Southern Oceans, in the Atlantic and Indian Ocean sectors. Moreover the initialisation changes in the extratropics of both hemispheres are small compared with the analysis changes. Thus it is fair to claim, in areas where data is regularly available in reasonable quantities, that

$$F > A > I$$

It is therefore consistent with the internal behaviour of the assimilation system to claim that in well-observed areas, the observations and forecasts have similar accuracies.

6. THE ASSIMILATION SYSTEM AS A MONITOR OF WIND DATA

A variety of techniques are in use to monitor the long-term quality of radiosonde height reports (Nash 1984, McInturff et al 1979). No reliable technique has been available to monitor the long-term quality of wind reports. In this section and the next we use case studies to illustrate the value of simple statistics of the assimilation system for long-term monitoring of observational data on both height and wind.

The first case study demonstrates how bias information permitted the identification of a serious error in the reported wind direction from one of the most isolated radiosonde stations in the world. The second case study exploits the smooth variation of forecast quality in space to compare the quality of wind reports from adjacent radiosondes, by removing the large synoptic variations. The third example demonstrates that the statistics are also a powerful tool to diagnose the use of wind data in the assimilation itself.

6.1 Wind direction error at an isolated radiosonde station

The radiosonde station at (45°S, 38°E), in the Southern Ocean, is one of the most isolated stations in the world. Our first statistical investigations showed that there were large mean differences between the wind reports and

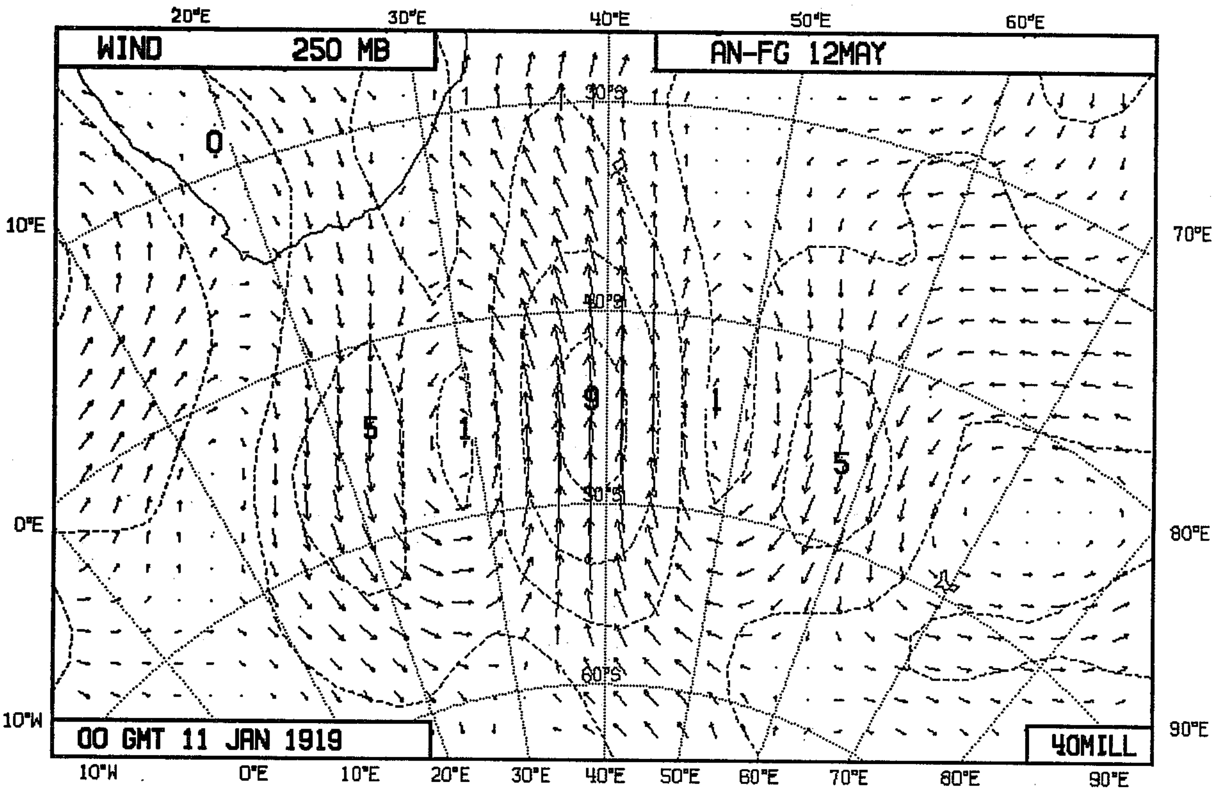


Fig.12 The mean analysis increment in the 250mb wind field near 40E,45S during May 1981. The isotach interval is 2m/s.

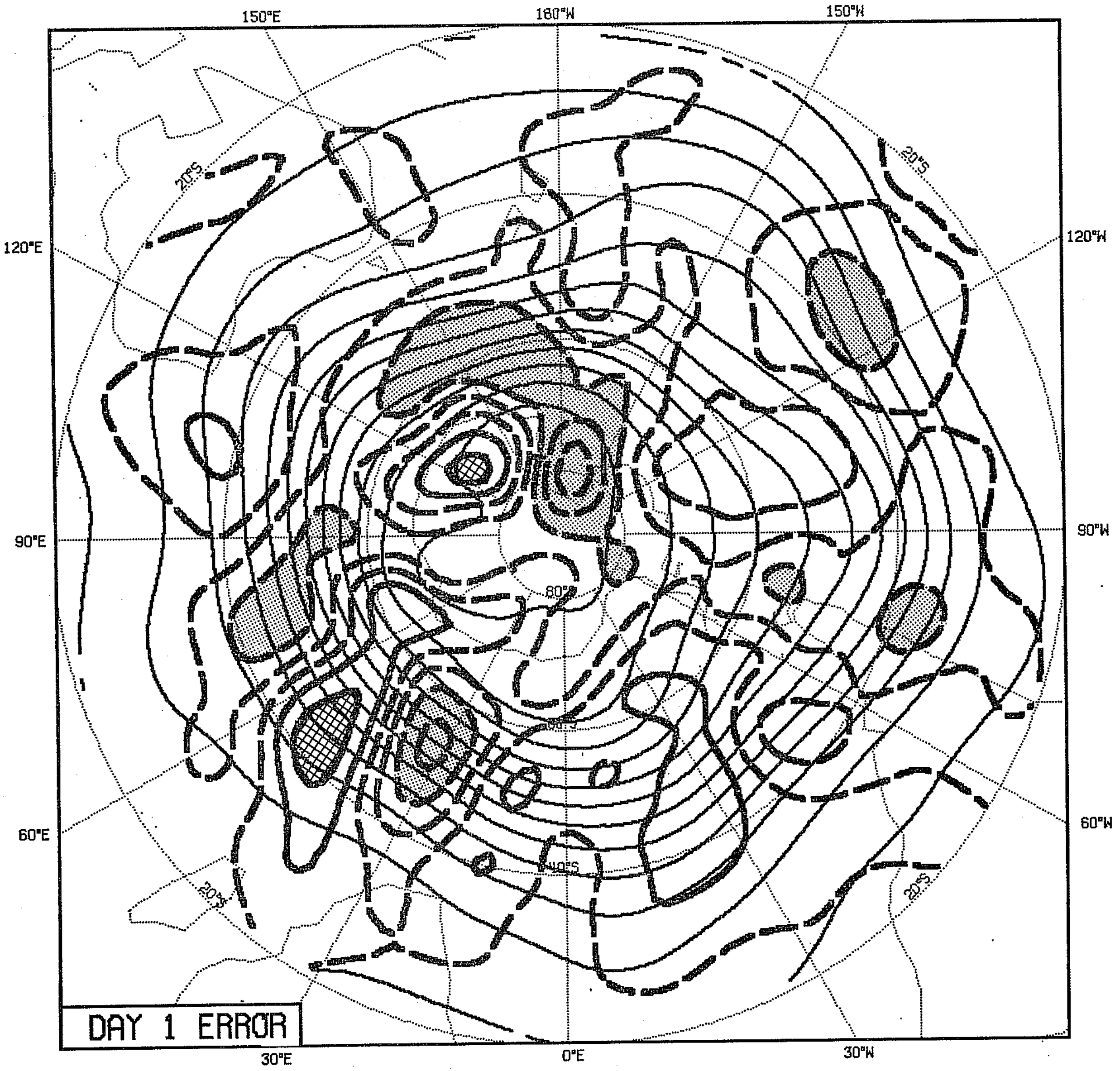


Fig.13 The thin contours show the mean 500mb analysis in the Southern Hemisphere in May 1981. The heavier lines show the mean error in the day 1 forecast. Errors higher than 15 m are hatched, errors lower than -25m are stippled.

wind forecasts at the station. No close stations were available for comparison. To pursue the investigation it was found useful to investigate the mean difference between the analysed and forecast fields.

Fig. 12 shows the mean analysis increment in the 250 mb wind in the vicinity of the station for the winter month of May 1981. The mean increment shows the typical signature of the non-divergent wind structure function used by our multivariate analysis system in the presence of an isolated wind observation. Over the station itself the increment is a southerly wind of 9 m/s. The mean increment over a period of a month ought to be close to zero. There are no other increments of comparable magnitude in the southern hemisphere. Investigation revealed that a similar pattern was found in the mean increment at the station during a randomly chosen period of the FGGE year, and again in 1982.

Fig. 13 shows the mean 1-day forecast error in the Southern Hemisphere height field at 500 mb for the month of May 1981. There is a clearly marked dipole error in the height field in the vicinity of the station. This forecast error indicates that much of the information introduced by the data at the station disagrees with the information introduced by data upstream.

These findings prompted a study of the wind-rose for the station, and for similarly isolated stations upstream and downstream. The upstream and downstream stations both showed a mean wind from 270°, with a reasonable degree of dispersion about this value. The station in question showed a similar dispersion about the direction 258° from true North. This suggested an error in the reported direction of some 12°.

As a further check, the reported mean southerly component was compared with the reported East-West gradients of geopotential, using mean geopotential data from the upstream and downstream stations. This was necessarily a crude calculation, but it did not support the existence of the reported mean southerly component. The evidence therefore pointed to an error in the reported wind direction, with the implication that the analysed fields were contaminated by the error.

To investigate this suggestion, the dynamical consistency of the analyses was tested in the following way. Using the monthly mean winds at 300 mb, we calculated the forcing Q necessary to satisfy the steady state barotropic vorticity equation

$$J(\psi, \nabla^2\psi + f) = Q$$

where ψ is the monthly mean stream-function. Fig. 14 shows the monthly mean value of ψ at 300 mb for the southern and northern hemispheres. The trough in the vicinity of the station in the southern hemisphere is clearly evident. Fig. 15 shows the values of Q for the southern and northern hemispheres. Over most of the globe the structure of Q is dominated by very large scales, except in the vicinity of the isolated station where there is a strong dipole in Q. This we interpreted as evidence that the monthly mean analysis was not dynamically consistent in the area in question.

These results led to a local investigation of the instrumentation, which indicated that the reported wind direction had been in error by somewhat more than the 12° suggested. The necessary corrective action was taken.

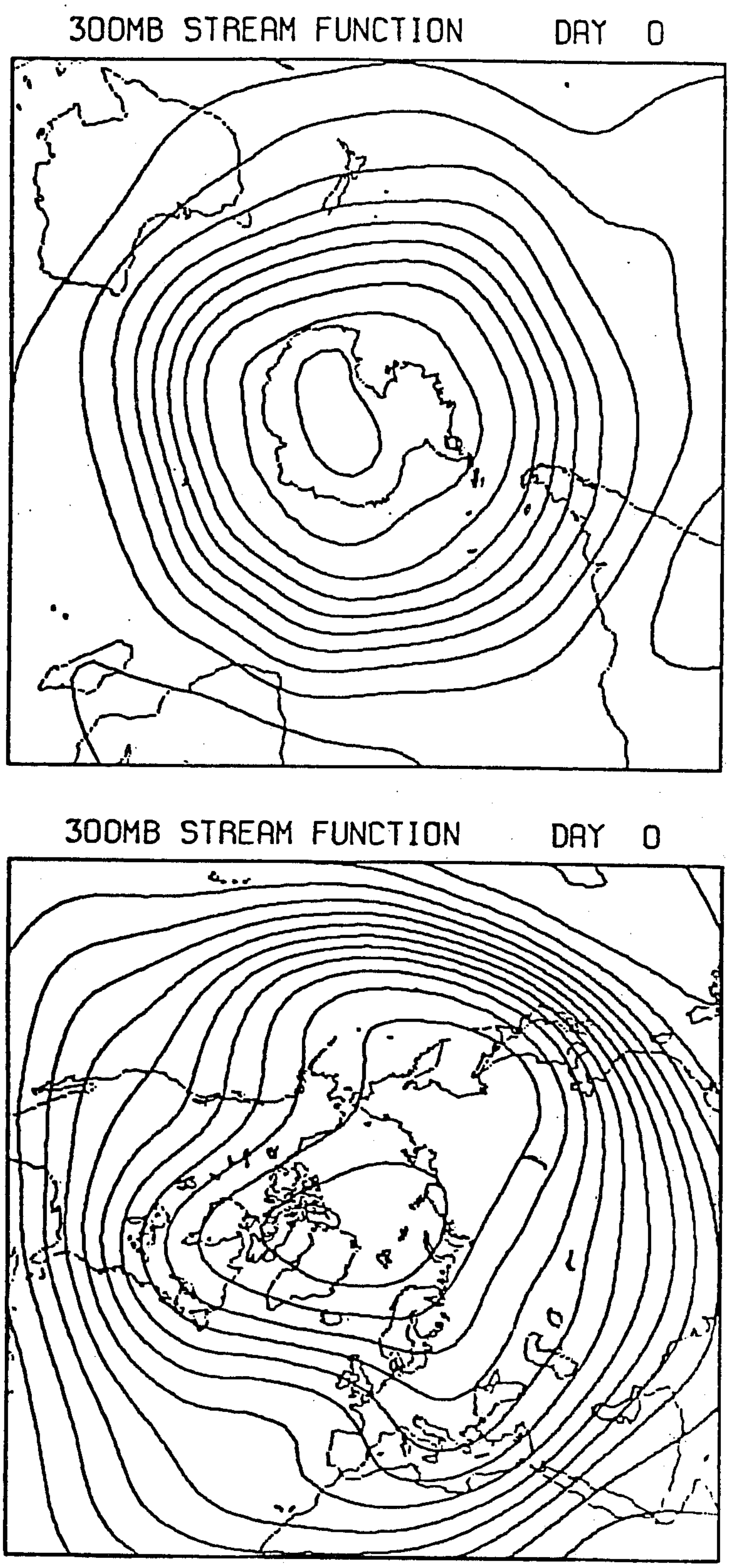
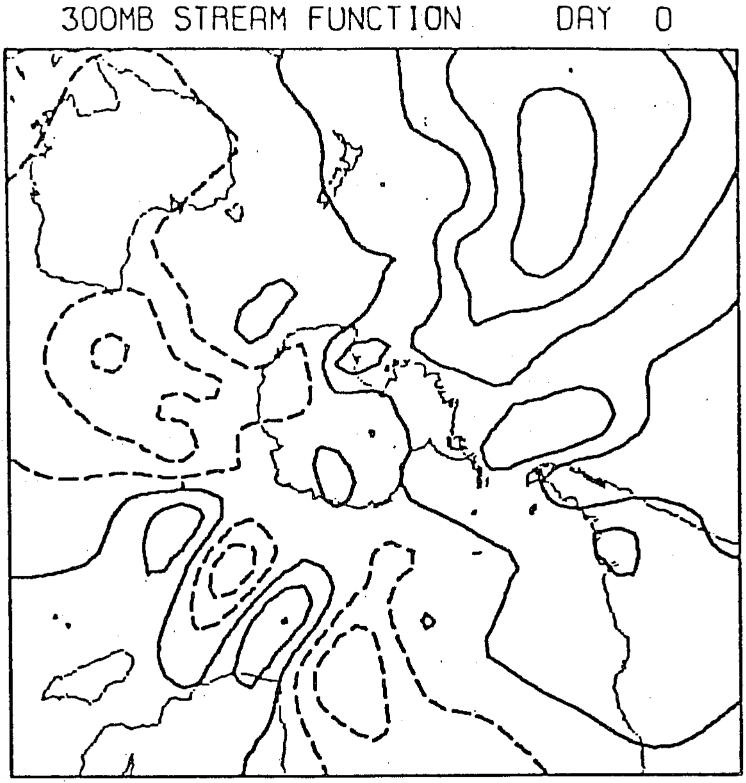


Fig. 14 300mb stream function for the Northern and Southern hemispheres based on 100 days of data from December 1 1980.



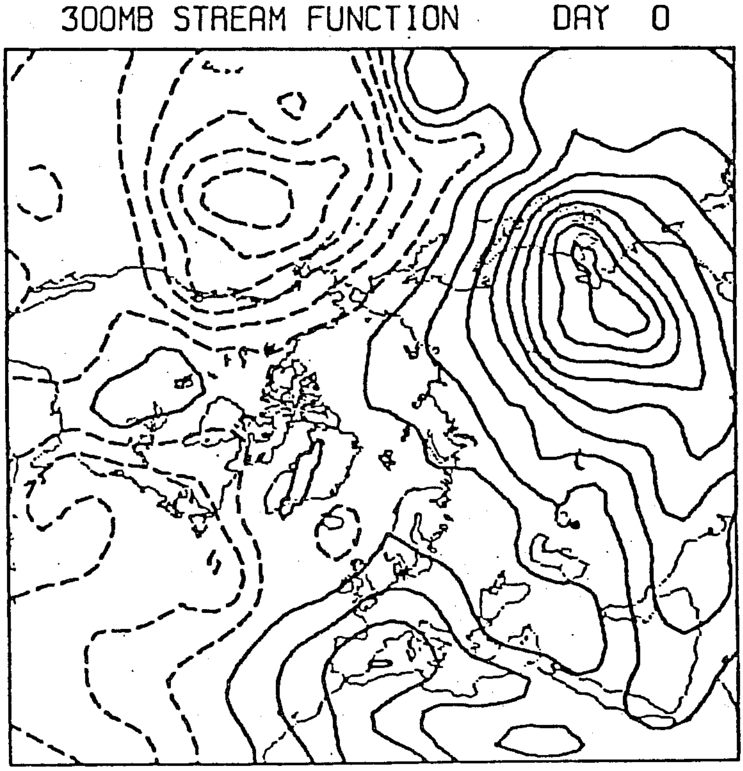


Fig.15 Calculations of the forcing Q necessary to satisfy the steady state barotropic vorticity equation when the steady state stream-function is given by Fig. 14.

The case-study demonstrates the ability of the Data Assimilation system to combine data from disparate sources - SHIPS, SATEMS, and PAOBS - so as to provide a long-term check on a set of radiosonde wind reports which perhaps could not have been checked in any other way.

6.2 Quality control of winds in a data rich area

The next example shows how a study of station to station variations of the rms O-F differences led to the detection of a problem in the wind reports from a station in a fairly data dense region. Since the forecast accuracy varies smoothly in space, the occurrence of a much larger rms O-F value at one station relative to its neighbours is indicative of problems with the instrumentation there.

Table 1 shows the standard deviations of the O-F differences for November 1982 for the 250 mb wind at two radiosonde stations A, B, which are part of a Northern Hemisphere national network. The stations are separated by about 700 km and are the only stations in the network using a particular type of wind-finding equipment. The value of about 4 m/s at station A was typical for most of the radiosondes in the area. Station B, showing a value in excess of 8 m/s, was therefore suspected of having technical problems. Local investigation confirmed this suspicion, and the problems were quickly corrected. As a result, the statistics for the doubtful station in January 1983, two months later, were considerably better, with both stations providing high quality reports, as shown in the table.

Table 1 Standard deviation of forecast minus observed V(250)

Station	A	B
Nov 1982	4.2	8.4
Jan 1983	4.4	4.3

This example, though not as dramatic as the earlier one, demonstrates how rapid communication between data users and data producers can improve the performance of both.

6.3 Identification of problems in the assimilation system

Our final wind example is just one of numerous cases where study of the statistics of the forecast-observation differences have helped identify (and eventually cure) problems in the assimilation system.

Fig. 16 shows the mean difference between the observed SYNOP reports, and the 10 m wind estimated from the forecast model's boundary layer formulation over the North East Atlantic. The statistics show a marked over-estimate of the mean west-south-westerly winds by the procedure which generates the estimated 10 m wind from the forecast. The problem was most serious on exposed western coasts, and was not apparent at Ocean Weather Ships, nor in the interior of continents.

This had important consequences for the analyses. The analysis algorithms analyse the differences between the forecast and the observation. In many cases where the surface pressure observations indicated an under-forecast of

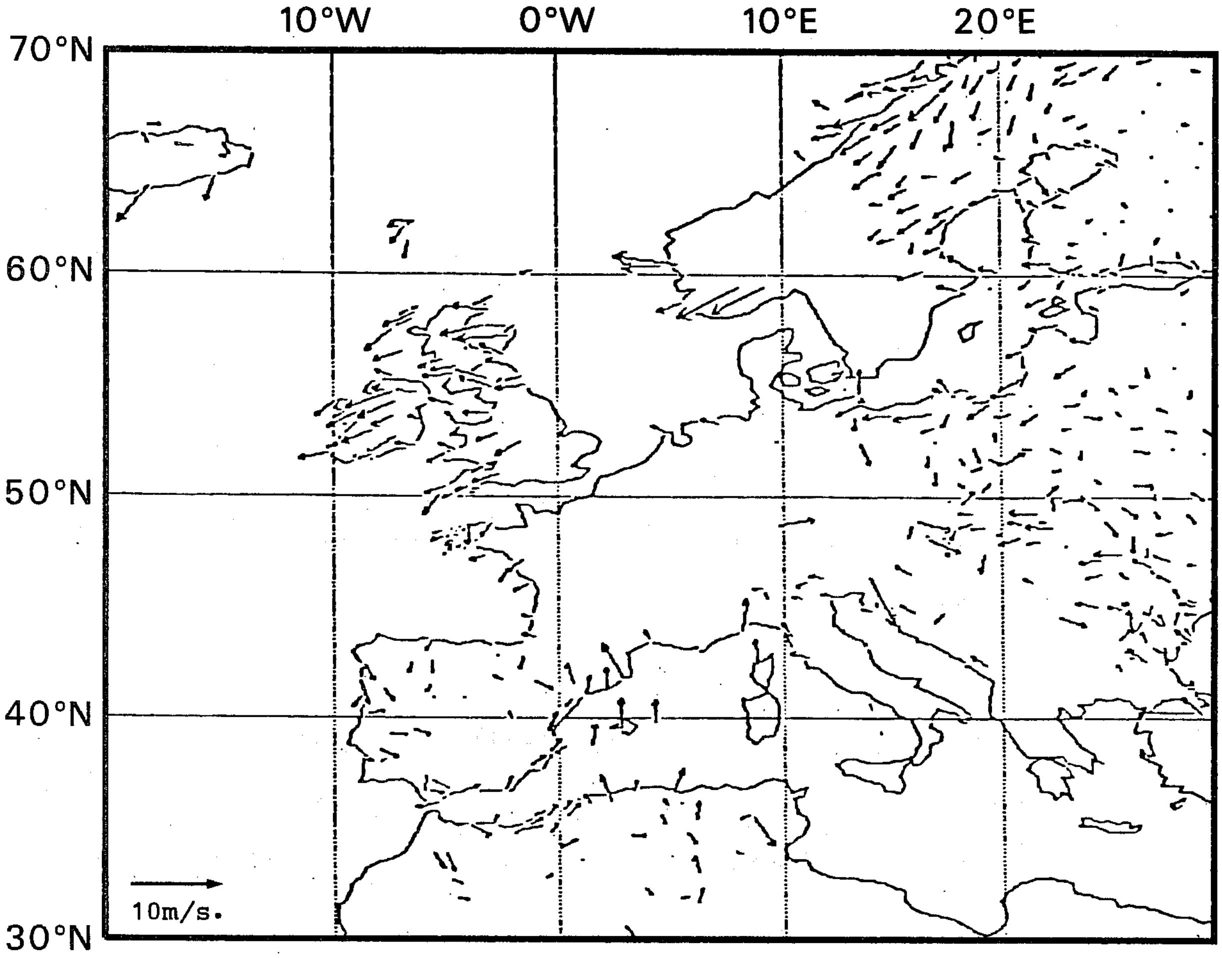


Fig.16 The bias (observation minus first guess) of the forecast wind at 10m against the reported Synop wind at selected stations for 12Z in January 1983.

a deepening low, the wind observations were interpreted by the algorithm as showing the opposite, because of the over-strong low level winds extracted from the model. The system's compromise between the apparently conflicting sets of data resulted in a degraded analysis.

Part of the problem was traced to excessive smoothing of the surface roughness parameter, which has a sharp transition between land and sea; the extent of the error is expected to be reduced by use of substantially less smoothing of the roughness parameter in a higher resolution version of the forecasting model.

6.4 Discussion

The need for intercomparisons of meteorological instrumentation has long been recognised. Intercomparisons of radiosonde behaviour have been of two types. Short-term expensive experiments involving the near-simultaneous launch of two or more sonde types for a period of some weeks are reported in Richner and Phillips (1984), and several of the references cited therein. Longer term monitoring of the behaviour of the instruments has mainly focussed on the need for radiation corrections to the height data (Nash (op.cit.) and McInturff et al. 1979, Lange 1985).

Reliable methods for long-term monitoring of wind-finding systems have not been available, because of the large synoptic variations in the wind field on short time and space scales. The case studies just presented show that it is now possible to monitor the long term performance of wind-finding systems, since most of the synoptic variations can be eliminated through study of the statistics of the difference between the observed value and the forecast for that observation.

7. THE ASSIMILATION SYSTEM AS A MONITOR OF PRESSURE DATA

Surface pressure observations play an important role in the assimilations, particularly in those regions and at those times when comprehensive upper air data is lacking. In the mass and wind analysis, the two surface elements of greatest concern are the reported pressure and wind. The reported pressure will, most commonly, be mean sea level pressure, but some hundreds of stations report alternatives such as station level pressure or the geopotential at a near surface pressure level; some stations report both a mean sea-level pressure and a station level pressure.

In common with other assimilation systems, our analysis algorithms seek to exploit such data in the three-dimensional calculation. This application of all reported pressures is not found in manual practice- a human analyst seeking to analyse a mean sea-level pressure field will not exploit a report of station-level pressure directly, though he will use the reported pressure tendency. The objective analysis uses the station-level pressure directly, and in so doing relies on the three dimensional spatial consistency of such data. To utilise a report of station level pressure a reference datum is required namely the station height, available in the index of station data (WMO 1984, Vol.A). Errors in such data will be just as serious as a systematic error in the reported pressure. The station height information is also required to use the station-level pressure datum found in radiosonde observations.

The following example is just one of many to suggest that there are numerous errors in such station height data, and that such errors have had a detrimental impact on the analyses.

7.1 Station pressure reports in the Himalayas

Fig. 17 shows biases of (observed minus first-guess) SYNOP pressure values, based on 1200 GMT data for January 1983 in eastern Asia. Some of the stations reported mean sea-level pressures; stations in the Himalayas generally report a station-level pressure. Only data accepted by the analysis system have been plotted. All biases are expressed in geopotential meters. Over peninsular India there is reasonable spatial consistency, with an average bias of approximately -25 m. This is indicative of a first-guess error and is a consequence of the absence of a diurnal cycle in the radiation scheme of the model, so that the forecast lacked the diurnal and semi-diurnal tides (an omission remedied in May 1984). Of greater concern in the present study is the lack of spatial consistency near the Himalayas, particularly in the region delineated by the dashed lines in the figure. The biases range from 59.1 m in the centre of the region to -48.1 m at a station in the south east. Discrepancies of such magnitude, with such a modest spatial separation are unlikely to be caused by errors in the first-guess field; they are judged to be errors stemming from incorrect station height data. (Some of the station heights quoted in the WMO Vol.A index are indeed designated as approximate , and their use in the assimilation system is, with hindsight, naive).

The consequences for the analysis are serious. Used in isolation the data in the centre of the region in question could be expected to yield a substantial

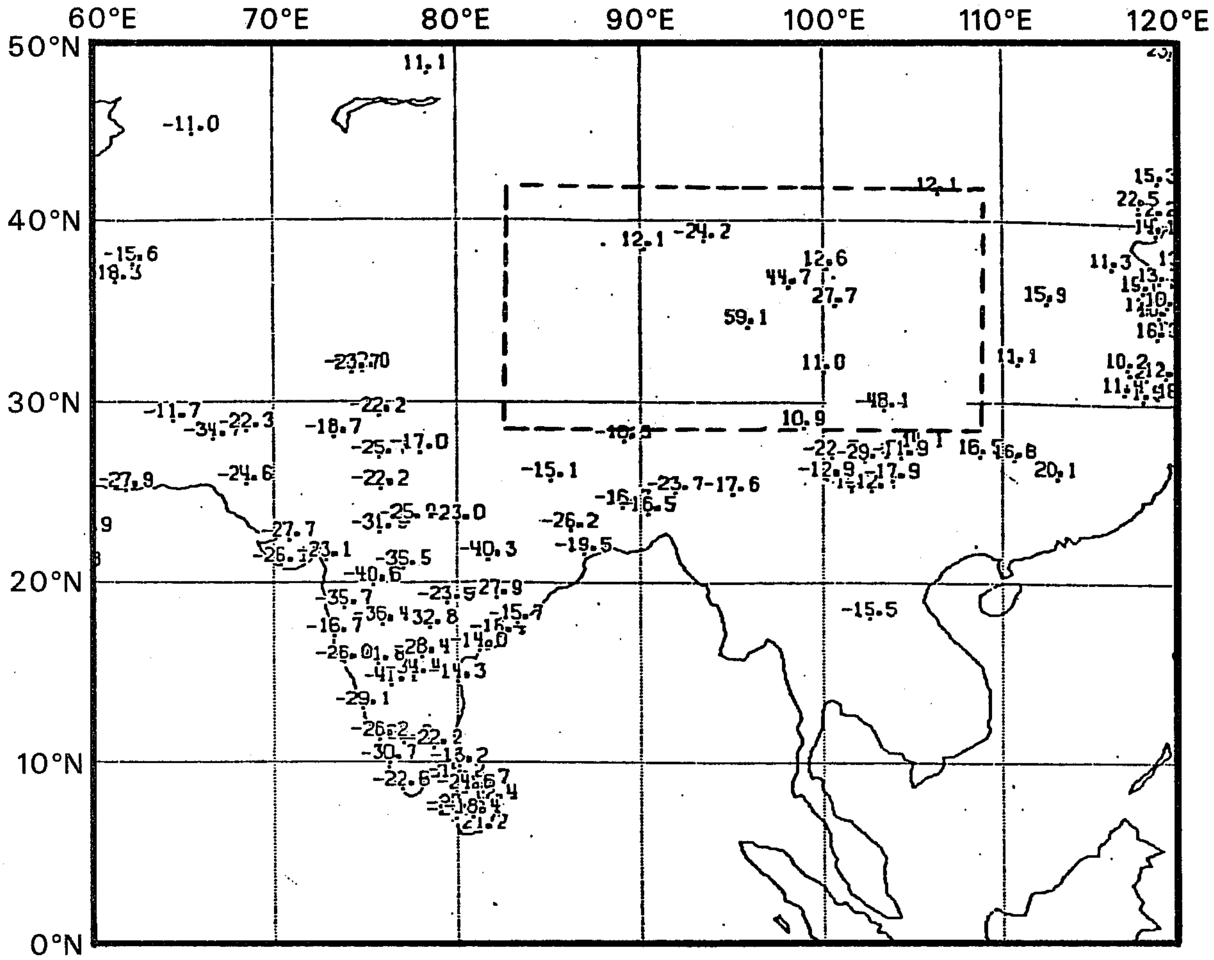


Fig. 17 The bias (observation minus first guess) between the pressure report used by the analysis (mean sea-level or station pressure) and the corresponding first-guess in south east Asia for 1200GMT data for January 1983.

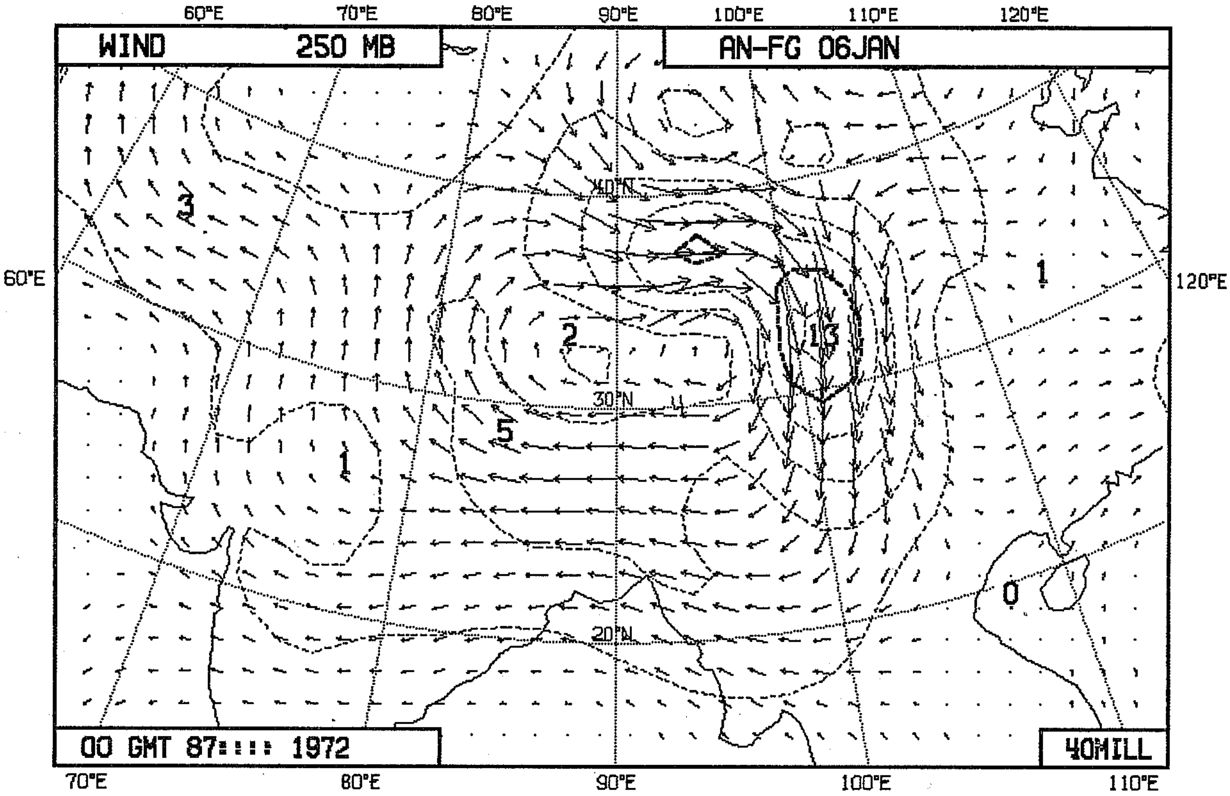


Fig.18 Mean analysis increment in the 250mb wind field over the Himalayas for 0600 GMT analyses in January 1983.

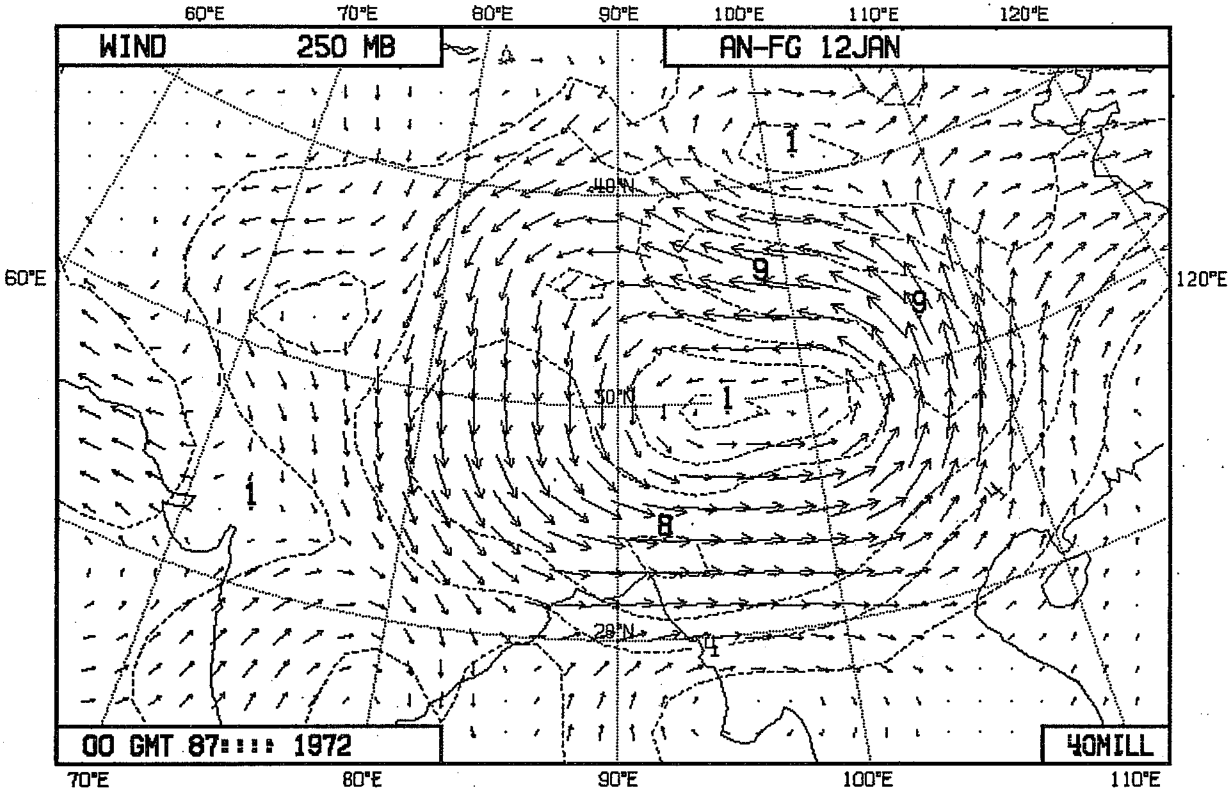


Fig.19 As Fig. 18 for the 1200GMT analyses.

increase in the first-guess geopotential field which, because of the geostrophic coupling should generate a anticyclonic wind increment. The times at which the data are used in isolation are at 0600 an 1800 GMT, when few other data are available. Fig. 18 shows the mean analysed wind increment for the 0600 GMT analyses for January 1983 for the 250 mb level, with the consequence of the incorrect data clearly evident. Because the increment is geostrophically balanced, it will survive the initialisation, and will be retained to some extent in the subsequent 6-hour forecast to 1200 GMT. At 1200 GMT, a similar set of incorrect data will be presented to the analysis, but now in competition with appreciable volumes of rawinsonde data at the standard levels in the vertical. The effect of this wind data is to outweigh the surface pressure data, and reverse the incorrect anticyclonic circulation of the first-guess field. This process is seen in Fig. 19 which shows the mean analysed wind increment at 1200 GMT. Noting that the analysis does not fully eradicate errors in the first-guess field, one can conclude that the quality of the final analysis in this region is adversely affected by the use of the surface data.

The use of surface data in our assimilation scheme is dependent, in part, on the orography of the model in the sense that observations are not used if to do so required excessive extrapolations of the prediction model in order to create an appropriate first-guess value. A substantial change to the model orography was introduced in April 1983 (Wallace et al 1983), and so there was a changed signal in the features of Figs. 17, 18 and 19 arising from the model change. Changes will also occur from time to time as a consequence of new observing practices in certain regions, from changes in receipt of data via the GTS, and indeed from changes to the WMO Volume A index, which is updated at six monthly intervals.

It could be argued that erroneous information of the type described should be identified and rejected by the real-time quality control checks within the analysis. Much of it is of course; what are depicted in the diagrams are the discrepancies which are appreciable in terms of their impact on the analysis, but which are nevertheless sufficiently modest to escape detection by the quality control algorithms currently used. The conventional checks failed to detect these errors, and while more sophisticated checks would be partially successful, it is clearly desirable to remove the cause of the errors.

8. DISCUSSION

The purpose of this paper has been to demonstrate the ability of a modern data assimilation system to provide long-term diagnostic facilities to monitor its own performance, and the performance of the observational network. The capability has been developed through improvements in computer power and the techniques of numerical weather analysis and prediction over the last decade. The diagnostic ability exists because:

(i) The analysis algorithm uses both the observations and the first-guess for extensive quality control of the observations. It then combines the information in the first-guess with the information in the diverse observational database, to produce an estimate of the state of the atmosphere that is more accurate than either of its inputs.

(ii) The analysis is sufficiently well balanced to permit most of this information to be communicated to the slow modes of the forecast model, ensuring an accurate and noise free first-guess for the next cycle.

The usefulness of the monitoring tools in comparing the long-term consistency of adjacent stations, or of different data types, stems from the fact that the accurate forecasts enable one to remove the large synoptic variations which would otherwise render the comparisons impossible.

The examples we have discussed are only a small sample of the useful results of this work. The diagnosis of an anomaly in the statistics requires investigations at the analysis centre to ensure that the anomaly is not due to the assimilation system. Collaboration between data analysis centres can be valuable in cross-checking the findings. Regular exchange of data monitoring results among analysis centres, and between these centres and the data producers, would be beneficial.

In this report we have only studied two elements of the Global Observing System, namely land SYNOP reports and radiosonde reports. The techniques have been extended to all aspects of the Global Observing System (Delsol, 1985). Systematic exploitation of these techniques will certainly improve the quality of the initial data for forecasts.

REFERENCES

Arpe, K., A. Hollingsworth, A.C. 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.Met.Soc., 111, 67-102.

Bengtsson, L., M. Kanamitsu, P. Kållberg, S. Uppala 1982: FGGE 4-Dimensional Data Assimilation at ECMWF. Bull. Amer. Meteor. Soc. 63, 29-43

Bergman, K.H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. Mon.Wea.Rev., 107, 1423-1444.

Bjorheim, K., P. Julian, M. Kanamitsu, P. Kållberg and P. Price, 1981: FGGE III-b Daily Global Analyses. Available from ECMWF.

Burridge, D.M. and J. Haseler, 1977: A model for medium range weather prediction -adiabatic formulation. ECMWF Tech.Rept.No.4. Available from ECMWF.

Cats, C.G. and W. Wergen, 1982: Analysis of large scale normal modes by the ECMWF analysis scheme. in Proceedings of ECMWF workshop on Current Problems in Data Assimilation, 343-372. Available from ECMWF.

Delsol, F., 1985: Monitoring the availability and the quality of observations at ECMWF. In Proceedings of ECMWF Workshop on the Use and Quality Control of Meteorological Observations, shortly available from ECMWF.

Gandin, L.S., 1963: Objective analysis of meteorological fields. Translated from Russian by the Israeli Program for Scientific Translations (1965).

Gustavsson, N., 1981: A review of methods for objective analysis, Dynamic meteorology - data assimilation methods, edited by L. Bengtsson, M. Ghil and E. Kallen, published by Springer-Verlag.

Guymer L.B.,1978: Operational application of satellite imagery to synoptic analysis in the southern hemisphere. Tech Rept 29. Bureau of Meteorology, Australia.

Heckley, W.A., 1985: Systematic errors of the ECMWF operational forecasting model in tropical regions. Submitted for publication.

Hollett, S.R., 1975: 3 dimensional spatial correlations of PE forecast errors. M.Sc. thesis. Department of meteorology, McGill University, Montreal, Canada, 73 pp.

Hollingsworth, A. and C.G. Cats, 1980: Initialisation in the Tropics. In Proceedings of ECMWF workshop on 'Tropical Meteorology and its effects on medium range weather prediction at middle latitudes', 105-142; available from ECMWF.

Hollingsworth, A., K. Arpe, 1982: Biases in the ECMWF assimilation system, ECMWF Tech.Memo.No.46, available from ECMWF.

Hollingsworth, A., A.C. Lorenc, M.S. Tracton, K. Arpe, C.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.Met.Soc., 111, 1-66.

Hollingsworth, A. and P. Lönnberg, 1985: The statistical structure of short range forecast errors as determined from radiosonde data. Part I The Wind Field. To appear in this volume

Kanamitsu, M. 1985: A study of predictability in the tropics based on the ECMWF operational forecast model. To be published in Proceedings of FGGE Sigtuna Workshop on Atmospheric Modelling by WMO/ICSU

Lange, A. 1985: The quality of TEMP data. To appear in proceedings of ECMWF workshop on The Quality and Use of Meteorological Observations, shortly available from ECMWF.

Lönnberg, P. and A. Hollingsworth, 1985: The statistical structure of short range forecast errors as determined from radiosonde data. Part II Covariance of height and wind errors. To appear in this volume.

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

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model, with application to normal mode initialisation. Contrib.Atmos.Phys., 50, 253-271.

McInturff, R.M., F.G. Finger, K.W. Johnson and J.D. Laver, 1979: Day-Night differences in radiosonde observations of the stratosphere and troposphere. NOAA Tech. Memo. NWS/NMC/63 published NOAA, Washington.

Nash J., 1984: Compatability of radiosonde measurements in the upper troposphere and lower stratosphere for the period 1/11/1981 to 31/10/1982. O.S.M.24 Meteorological Office, Bracknell.

Oort, A, 1983: Global Atmospheric Circulation Statistics 1958-1973. NOAA professional paper 14, NOAA, Rockville, Md., pp.180.

Richner, H. and P.D. Phillips, 1984: The radiosonde intercomparison SONDEX. Published Birkhausr Verlag Basel, pp.324

Riehl H., 1979: Climate and Weather in the Tropics. Academic Press. pp.454.

Rutherford, I.D., 1972: Data assimilation by statistical interpolation of forecast error fields. J.Atmos.Sci., 29, 809-815.

Schlatter, T.W. 1975 Some experiments with a multivariate statistical objective analysis scheme. Mon.Wea.Rev., 103, 246-257.

Shaw, D.B., P. Lönnberg and A. Hollingsworth, 1984: The 1984 revision of the ECMWF analysis system. ECMWF Tech.Memo. 92, available from ECMWF.

Simmons A.J., M.Jarraud 1984: The design and performance of the new ECMWF operational model. in proceedings of ECMWF 1983 seminar on Numerical Methods for Weather Prediction. Available from ECMWF.

Temperton, C. and D.L. Williamson, 1981 : Normal mode initialisation for a multilevel grid-point model Part I : Linear aspects, Mon.Wea.Rev., 109, 729-743.

Tiedtke, M., J-F. Geleyn, A. Hollingsworth, J-F. Louis, 1979: ECMWF model parameterisation of sub grid scale processes. ECMWF Tech.Rep. 10.

Wallace, J.M., S. Tibaldi, A.J. Simmons, 1983: Reduction of systematic forecast errors in the ECMWF model through the introduction of an envelope orography. Quart. J. Roy. Meteor. Soc., 109, 683-717

Williamson, D.L. and C. Temperton, 1981: Normal mode initialisation for a multi-level grid-point model, Part II: Nonlinear aspects. Mon.Wea.Rev., 109, 744-757.

WMO 1983: Guide to meteorological instruments and methods of observations. 5th Edition, WMO pubn. No. 8., Geneva.

WMO 1984: Weather Reporting, Volume A: Observing Stations; pubn. No.9, WMO, Geneva.