CENTRAL FORECASTING MONITORING NOTE NO. 7

THE USE OF OUTPUT FROM A NUMERICAL MODEL TO MONITOR THE QUALITY OF RADIOSONDE OBSERVATIONS

by C D Hall

February 1992

Central Forecasting Division
Meteorological Office
London Road
Bracknell, Berkshire
RG12 2SZ

THE USE OF OUTPUT FROM A NUMERICAL MODEL TO MONITOR THE QUALITY OF RADIOSONDE OBSERVATIONS

by C D Hall

SUMMARY

To make optimum use of meteorological observations it is essential that regular monitoring is performed to identify those of poor quality. Output from numerical forecast models has proved to be very valuable for this purpose; short-period forecasts or background fields provide accurate global reference values against which observations may be compared. This paper presents some recent results of the monitoring of radiosonde observations, and describes a number of different methods that may be used to identify cases where errors of observation, over a period of a month or more, are significantly larger than normal.

1. INTRODUCTION

The paper by Hall, Ashcroft and Wright which appeared in the August edition of this publication, described some of the ways in which output from a numerical model, in particular the short-term forecast or background values, may be used to provide valuable information on the quality of meteorological observations. Some general principles were outlined, and examples demonstrated how observations of pressure and wind from ships and buoys could be monitored. In this paper it is shown how the quality of observations from radiosondes may be assessed using similar monitoring methods, and some revealing characteristics of the errors are identified.

Radiosondes are the cornerstone of the meteorological observing network, providing in most cases detailed vertical profiles of wind, temperature and humidity of high accuracy. The importance of monitoring radiosonde performance has long been recognised, and this has been achieved at the international level through intercomparisons, sponsored by WMO, where different sondes have been carried on a single balloon. Following the first two phases of the intercomparisons in 1984 and 1985 (Nash and Schmidlin, 1987), and a later phase in 1990, systematic differences between many of the sondes in regular use have been identified. The results set a standard which is obtainable under the best operating conditions; in actual practice performance may not be the same as in a trial, as routine monitoring of the daily observations, received in real time over the GTS, readily reveals. Such monitoring may be performed in various ways. For instance, attention is frequently focused on the reported geopotential height at 100 hPa as this value usually reflects the integrated effects of errors in the measured temperature at lower levels. Differences from the observed 100 hPa height at neighbouring stations, or between observations made in night-time and daylight conditions are useful idicators of quality. In the absence of nearby stations, comparison is best performed against some reference values, and numerical models, which provide global fields of high quality, have often been used for this purpose (eg Kitchin, 1989a). Much work in this field has also been performed at ECMWF and results are given in Hollingsworth et al (1986) and Radford (1987). This paper will summarise what can be achieved using output from the UK 15-level model which was operational up to June 1991. Observations of temperature, geopotential and wind will be considered, but not of humidity.

2. MONITORING METHODS

Central to the monitoring methods described here are differences between the observed value and the value of the model background field interpolated to the observation position (referred to throughout as O-B). The background fields, derived from cycles of data assimilation, reflect the information contained in past observations as well as information relating to the structure of the atmosphere provided by the numerical model. Great advances have been made in numerical modelling in the past two decades, and today global fields are available at high resolution. Their quality is sufficiently high for them to have an important role in observation monitoring. Where the values of O-B relate to observations from one source over a long period of time, the long-term performance of the observing system may be assessed. For instance, a time sequence of values of O-B for radiosonde observations from a given station may reveal changes during the period, of larger magnitude than known errors in the background field, which can only be attributed to changes in the characteristics of the observations. Background, rather than analysis values are used for the monitoring of observation quality because it is assumed that, being derived prior to the observation time, they are independent of the observation itself. This is probably not always strictly true; persistent systematic observation errors are not always filtered out by the data assimilation system and may influence the background field. A second basic assumption is made: namely that both the systematic and random background errors, averaged over periods of a month or more, vary only smoothly in space. This is probably true in the free atmosphere away from steep orography and the model's upper and lower boundaries. In contrast, errors arising from inaccurate measurements may vary greatly from station to station or between national groupings of stations. Differences from background which are larger than the local average can in most instances be attributed to larger than average errors at the observing station. Errors in the background fields, which are largest in data sparse areas, are a limiting factor in their use for observation monitoring, and indeed it is essential that all monitoring results are set in the context of estimates of model errors.

The difference between an observed value (O) and the value of the background interpolated to the observation position (B) may be expressed as

$$O-B = (O-T_o) - (B-T_b) - (T_b-T_o)$$

T_o is the true value of the observation. If the observation is a spot value, T_o is the true spot value, while if the observation represents some time or space average, T_o is the true value averaged over time or space. Likewise T_b is the true value on the scale that the model can resolve, which in the case of the global model results presented here is approximately a 150km x 150km x 80hPa grid-box average. $O-T_o$ will be referred to as the measurement error, $B-T_b$ as the background error, and T_b-T_o as the representativeness error. Squaring and taking an average over many observations the following is obtained

$$\overline{(O-B)^2} = \overline{(O-T_o)^2} + \overline{(B-T_b)^2} + \overline{(T_b-T_o)^2}$$

$$= E_m^2 + E_b^2 + E_r^2$$

It has been assumed that the various cross-product terms are zero or can be neglected. E_m, E_b and E_r are respectively the rms measurement, background and representativeness errors.

Several factors contribute to the measurement error (E_m): there are errors due to the malfunctioning of the instrumentation; there are errors arising from the wrong estimate of the pressure level; and finally, there are errors introduced on encoding, either due to truncation (the upper air code only allows for the direction to be reported to the nearest 5 degrees) or inaccurate ground procedures. The background error (E_b) represents numerical forecast errors on the scale that the model can resolve. There may be additional background errors if account is not taken of the time for the balloon to make its ascent and its horizontal displacement from the release point in strong winds. The representativeness error (E_r) is a measure of the sub-grid scale detail measured by the sonde but beyond the model resolution. There will be a contribution to E_r from fine structure in the vertical (eg temperature changes across an inversion or strong vertical wind shear through a jet) as well as from mesoscale features with horizontal scales less than 150 km.

Kitchen (1989b) has estimated values for many of the components of O-B listed above using observations from the UK operational radiosonde network. He finds that E_m is the smallest of the three components of O-B (0.6-1.5 ms^{-1} for wind and 0.06-0.16° for temperature) while E_r is typically 2.5-3.0 ms^{-1} and 0.6-0.8° for wind and temperature respectively. He shows that a failure to interpolate the background field in space and time to the actual balloon position only leads to large errors in the relatively uncommon cases where the observations are valid 3 hours from the validity time of the background field or the sonde is 100 km downwind of the point of release.

In the monitoring results presented in the next section the background values are taken from the UK operational global model valid at the main synoptic hour (00, 06, 12 or 18 GMT) nearest to the observation time. Interpolation in time has not been performed between model fields, nor has the downwind displacement of the sonde been taken into account. This will lead to errors, as discussed above, but they are not thought to be large on average; in the case of operational radiosondes where most ascents start at, or one hour prior to, one of the main synoptic hours and take perhaps 60-90 minutes before balloon burst, the difference between the actual observation time and the validity time of the model field is seldom more than 1 hour.

For most radiosonde observations the vertical profile obtained from the full TEMP report contains far more detail than is available from the 15 levels of the numerical model. To achieve the best match between observation and background, the reported profile has been averaged across each of the model layers to give a layer-mean value. By smoothing the data in this way its vertical resolution is reduced to exactly that of the model and the contribution to E_r from fine structure in the vertical is eliminated.

3. MONITORING RESULTS

Figure 1 shows vertical profiles of the mean and rms O-B differences in the 3-month period October to December 1990 at Hemsby (53°N, 2°E). Plots such as these are used routinely to monitor the quality of radiosonde observations, and in all cases it is essential to assess what part of O-B may be attributed to model error and what part to observation error. This question is usually best answered by comparing the values with those obtained at nearby stations, and in the data rich area round Hemsby, stations over the UK or other parts of Northern Europe may be used. It turns out that these stations show similar values of O-B at all levels. The rms of O-B for temperature is a little more than 1°C and for wind it is around 3-4 ms^{-1} rising to 6 ms^{-1} at jet-stream levels. Both values are considerably larger than the reproducibility of good quality sonde and wind-finding systems (typically 0.2°C and 1.0 ms^{-1} respectively) and the major contributions must come from the background error (E_b) and the representativeness error (E_r). From the estimates of E_r noted in the previous section it can be seen that in an area such as Northern Europe E_b

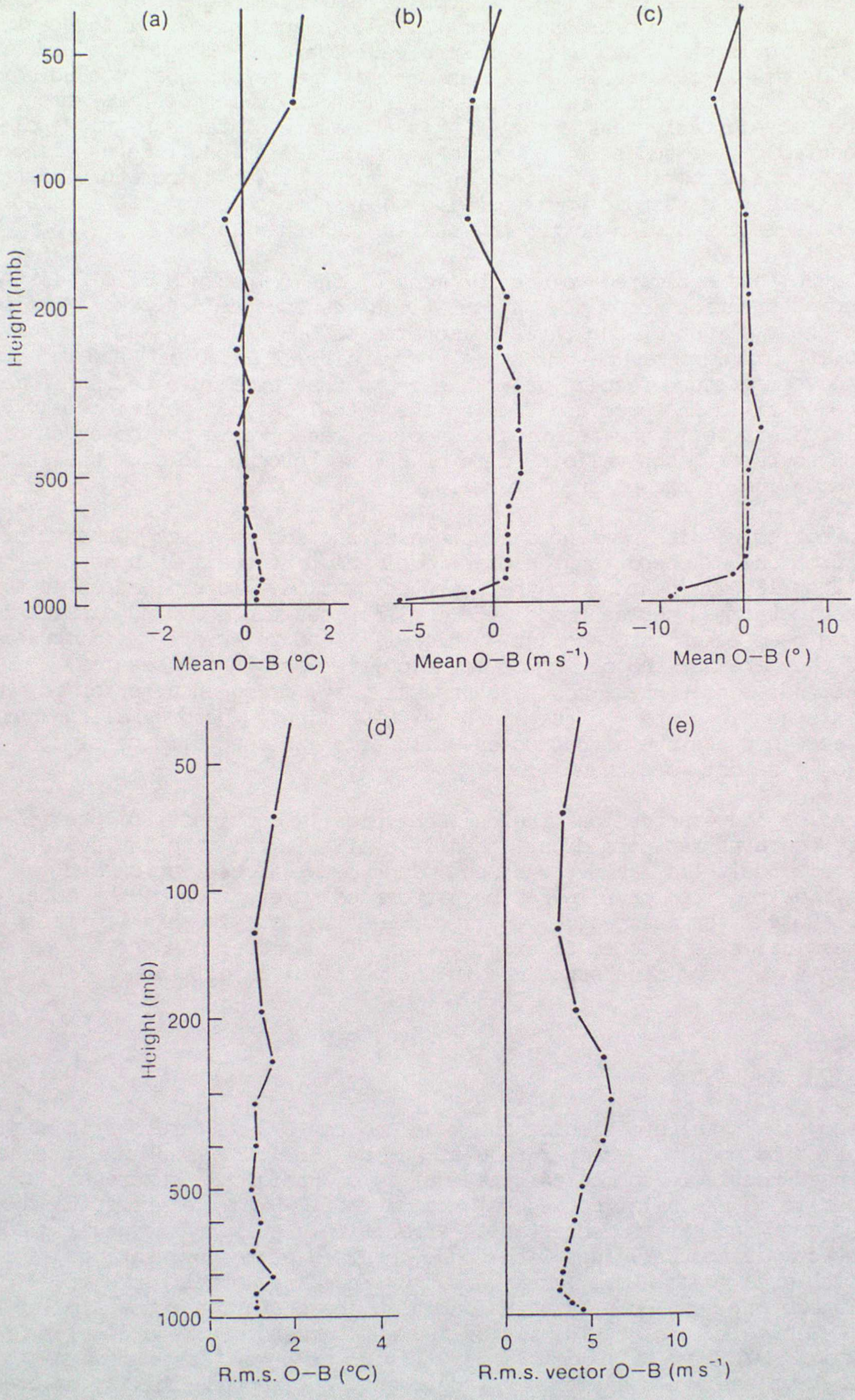


Figure 1. Vertical profiles of O-B for radiosonde observations from Hemsby (53°N 2°E) in the period October-December 1990; (a) mean temperature differences, (b) mean wind speed differences, (c) mean wind direction differences, (d) root mean square temperature differences, and (e) root mean square vector wind differences.

is a little larger, but not by much; typical values are around 3-5 ms^{-1} for wind and 0.8-1.2° for temperature. The bias of O-B is mostly very small, and where there are significant departures from zero, values similar to those at Hemsby are found at all neighbouring stations. Consistent biases such as these point to regional systematic errors in the background values. There are positive O-B temperature biases of around 1°C above 100 hPa showing that the model atmosphere is too cold at these levels. There are negative speed and direction biases close to the surface which is a characteristic found at most land stations and probably reflects inadequacies of the surface processes in the model. The positive speed biases, which are largest in the upper troposphere, is another characteristic found at many middle-latitude stations and shows up most noticeably as a tendency of the model to underestimate the strength of jets. Apart from the biases in O-B noted above due to systematic errors of the model, the mean differences from background are small; less than 0.2°C for temperature, 0.5ms^{-1} for speed and 2° for direction. The largest values of the rms vector wind differences are at around 300 hPa and are associated with large random model errors within the jet stream and with the strong horizontal wind shears often observed at this level which are beyond the resolution of the model. Consequently the height of this maximum varies with latitude and season in the same way as the level of the jet stream. Where comparisons are required between stations at different latitudes or between statistics in different seasons, it is usually advisable to average the rms vertically through a deep layer of the upper atmosphere. In this way dependence on the height of the jet stream maximum is largely avoided.

Time sequences of values of O-B from a single station provide a sensitive test of quality as the paper by Hall, Ashcroft and Wright showed for marine observations. Figure 2 shows a sequence of monthly mean values of O-B for 100hPa geopotential height at Hemsby over the period September 1989 to August 1990. Observations at 00 GMT only have been selected to avoid the complicating effects of solar radiation. For comparison monitoring values using the ECMWF model are also shown. The backgrounds from both models indicate that a change of bias occurred after January 1990, conninciding closely with the replacement of the Mark 3 sonde by the Vaisala RS80 at that station on the 23rd of the month. The known tedency for the Mark 3 to measure too cold is clearly evident as is a 15-20m systematic difference between the background values from the two models.

The RS80 has now become the most widely-used sonde over Western Europe and this allows an intercomparison of O-B statistics for a common instrument to be made over a large region. In Figure 3 the mean and standard deviation of O-B temperature differences at all stations where it is operational are plotted for the period October to December 1990 using 00 GMT observations only. The values, in tenths °C, represent vertical averages performed over a deep layer of the atmosphere from 850hPa to 100hPa. To avoid individual observations, differring from background by a very large amount, distorting the sample characteristics, values of O-B have only been included for those observations passing the automatic quality-control checks. In practice very few observations are excluded as quality-control flags are generally raised on less than 1 percent of the occasions. Mean O-B lies between 0.0°C and +0.3°C at most stations, but the are exceptions principally over Spain and Italy where the mean differences are larger. The larger positive values can almost certainly be attributed to the different radiation correction schemes in operational use. Most stations (indicated by the closed circles at the station position in Figure 3) use Vaisala "1986" corrections based on an evaluation by Vaisala of results of the WMO International Radiosonde Intercomparisons. A few stations still use earlier "1982" corrections (indicated by open circles) and in almost all cases they are the ones showing the larger mean temperature differences. The sign and magnitude of the difference agrees closely with the difference between the correction schemes at zero solar elevation (Kitchin

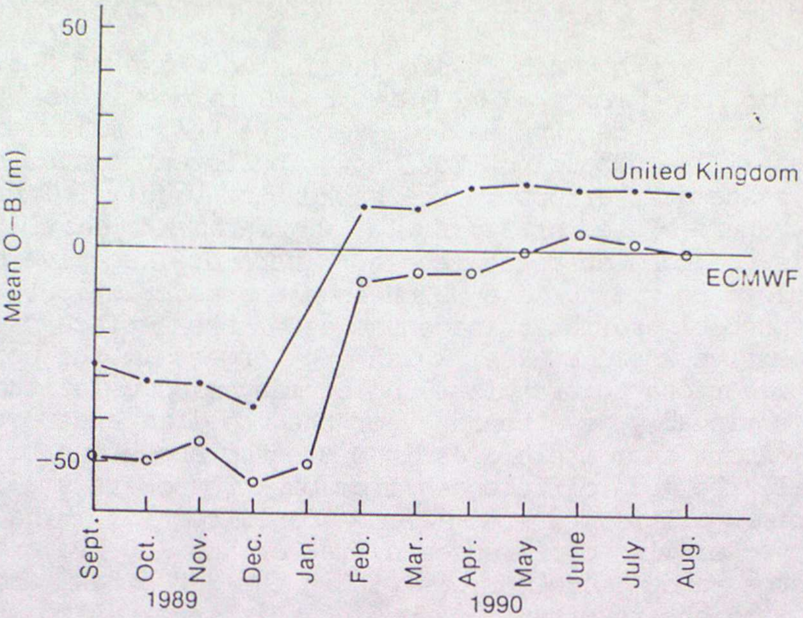


Figure 2. Monthly mean O-B differences for 100 hPa geopotential height at Hemsby using the UK and ECMWF operational models. 00 GMT data only.

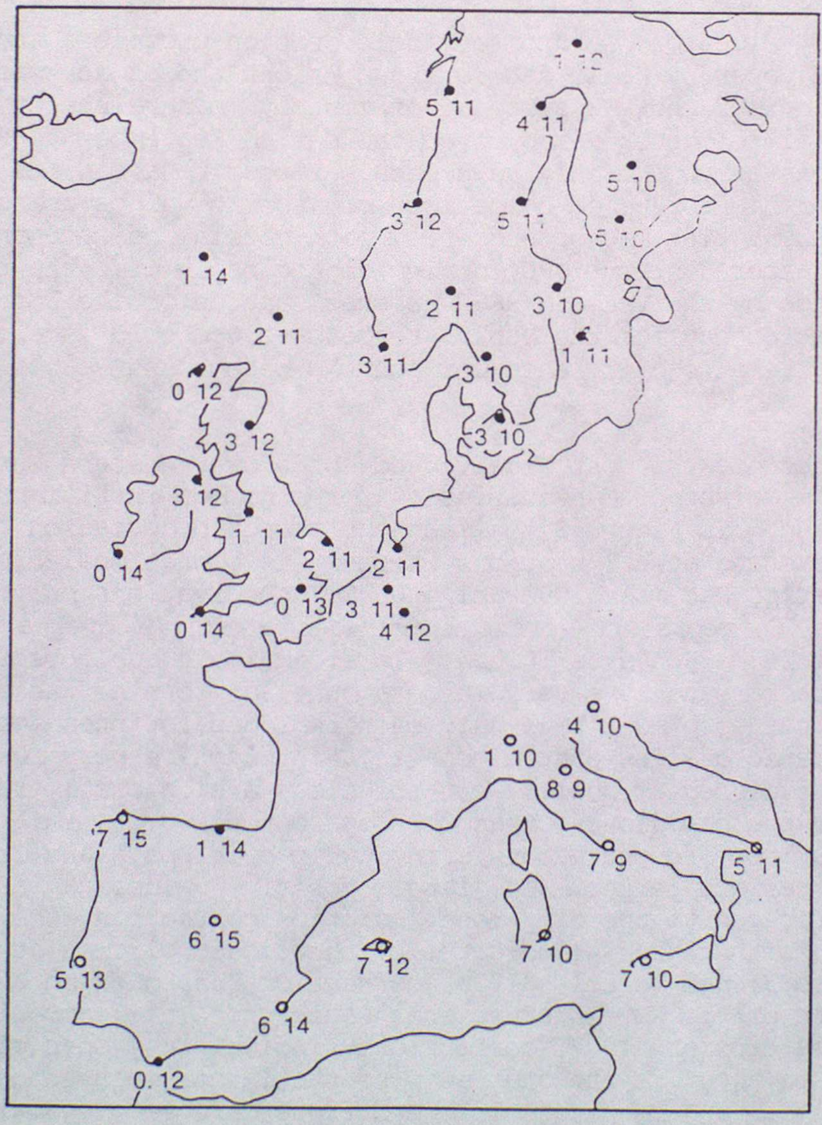


Figure 3. Mean (left) and standard deviation (right) of O-B temperature differences at radiosonde stations using the Vaisala RS80 sonde. Values in tenths °C for the period October-December 1990 have been averaged over the layer 850-100 hPa. Stations applying the "1982" corrections are marked by an open circle.

1989a). The second set of values in Figure 3 gives the standard deviation of O-B temperature averaged over the layer and it can be seen that it varies smoothly over the region, confirming the uniform pattern from station to station. As noted earlier, the values, lying between 1.1-1.5°C, principally represent the contributions from E_r and E_b. They are a little larger in the west than in the east, but this is to be expected as background errors are likely to show a similar regional variation rising to a maximum over the data sparse Atlantic.

Where there is a good coverage of stations providing reasonably accurate observations, as was the case above, values of rms O-B are found to vary smoothly over the whole region. Stations with a large observation error stand out as having values which are larger than at neighbouring stations. Figure 4 shows vertically-averaged values of rms O-B differences of the vector wind in units of tenths ms^{-1} in a data-rich region for the 12-month period January to December 1988. The vertical averaging has been performed over the layer 400-150hPa in order to obtain a representative value through the depth of the jet stream. As in the case above there is an underlying smooth variation over the region and values lie within the range 4-5 ms^{-1}, but this time three stations, identified by the letters A, B and C, stand out with values which are considerably larger than the local average. Large observation errors at these stations are the only reasonable explanation for the large differences from background.

The methods described above are valuable for identifying unreliable stations, but they do not provide much information on the nature of the problems. A more detailed study of O-B differences can reveal much more useful information, especially if it is based on a knowledge of the likely sources of error in the instrumental system in use. Three such examples are presented in sections below.

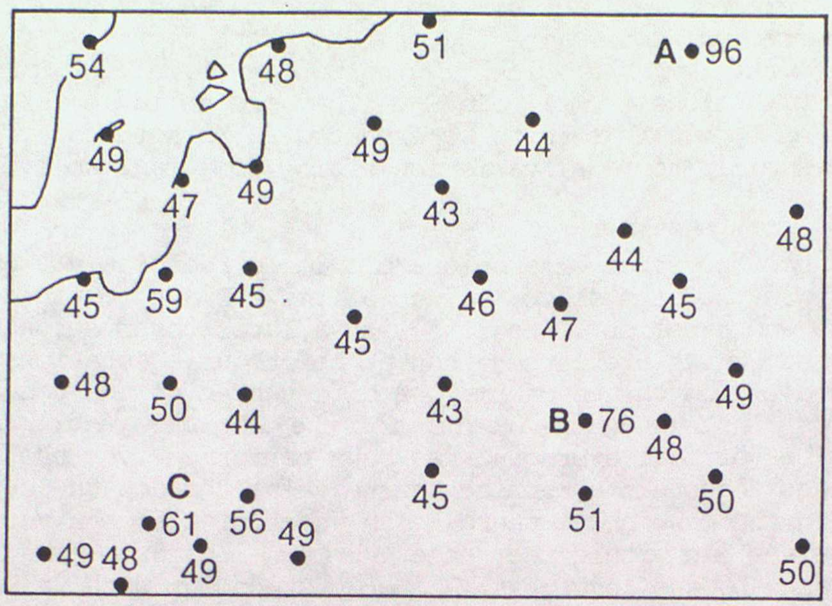
Wind direction errors

One type of wind error that is easiest to identify comes from a misalignment of the direction of true north, and it shows up as a bias in wind direction relative to background and nearby stations which is constant with height. Two of the stations (A and C) identified in Figure 4 as having abnormally large observation error are found to have a clear direction bias as Figure 5 demonstrates. In each case the vertical profile of O-B direction differences have been plotted for the station in question and its nearest neighbour. In both cases there is a systematic difference between the pair of profiles: relative to the local average the reported directions are backed by 17° at station A and by 10° at station C. It is interesting to note that at all stations in the region there are negative direction biases in the boundary layer similar to the bias noted at Hembsby (Figure 1). There are around 20 stations worldwide having O-B direction biases in excess of 10° which can confidently be attributed to observation error, and at least another 20 where the O-B bias is smaller and observation error is considered probable.

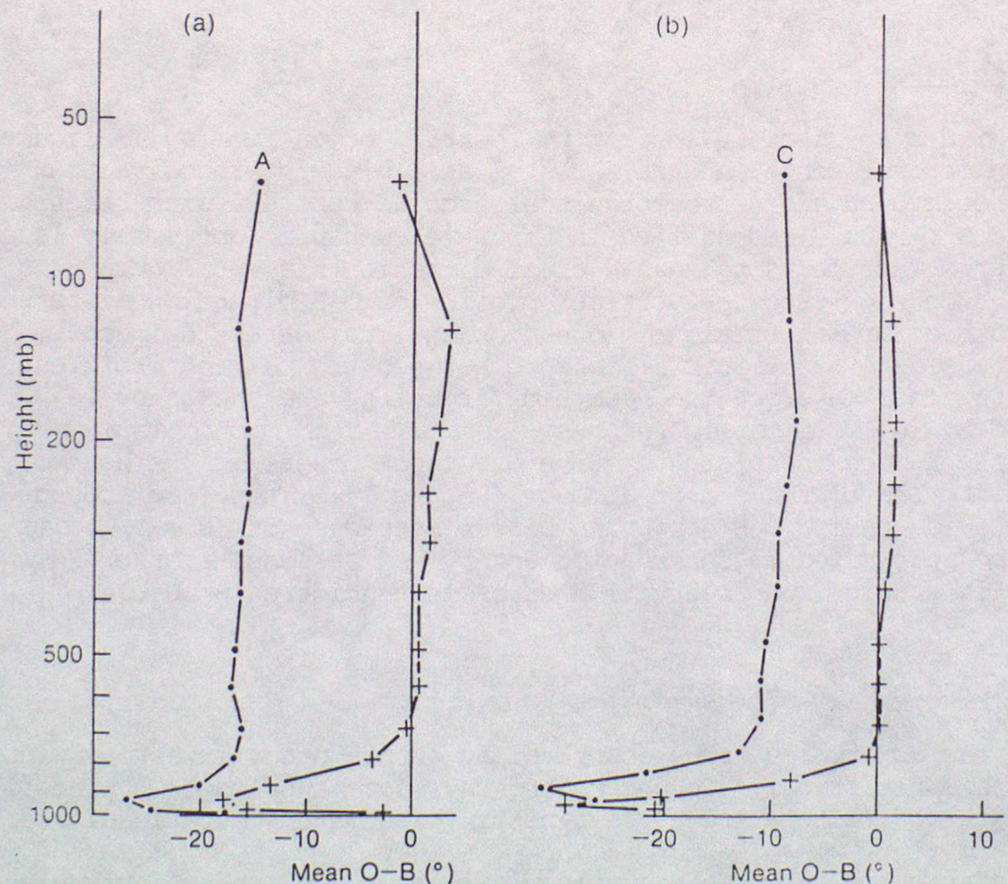
Wind error dependence on balloon elevation

At some stations abnormally large observations errors occur in strong winds at the level of the jet stream. To understand why, some knowledge of wind-finding systems is required. There are three types in widespread operational use:

1. Primary radar which measures elevation, azimuth and slant range provides in general the most accurate wind finding. Tests made at Beaufort Park, where a balloon was tracked by two independent radars (Edge et al 1986), have demonstrated that the reproducibility of wind measurements from the UK



<u>Figure 4</u>. Root mean square vector wind O-B differences for radiosonde stations in the period January-December 1988. Values in tenth ms^{-1} have been averaged over the layer 400-150 hPa.



<u>Figure 5</u>. Mean O-B direction differences in the period January-December 1988 at (a) station A in Figure 4 and its nearest neighbour, and (b) station C and its nearest neighbour.

operational radar using 1-minute averaging was better than 1 ms^{-1} rms vector error at slant ranges less than 60 km, and about 1.5 ms^{-1} rms vector error at slant ranges of 90 km.

2. NAVAID (navigation aids) is the general term applied to systems for determining horizontal location at any point on the globe through the use of electromagnetic waves in the radio frequencies. Synchronised signals are transmitted from a number of well-spaced stations, and differences in the time of receipt at a sensor enable its position to be determined. Omega is the NAVAID system in most widespread use and achieves an accuracy of 1-2 ms^{-1} for 2-minute averages on most occasions. Loran systems are in use at some UK stations which achieve a somewhat greater accuracy.

3. Radiotheodolite is the most common wind-finding system in use today. Radio signals from the sonde are tracked by direction-finding antenna at the ground station enabling azimuth and elevation to be measured. The height is usually determined by integrating the hydrostatic equation using the measurements of temperature, humidity and pressure in the same way as in NAVAID systems. At high balloon elevations and short slant ranges the accuracy obtained from radiotheodolites is comparable to the accuracy of NAVAID winds. However, at low balloon elevations, which are frequently encountered in the strong jet streams in middle latitudes, the reported wind is much more sensitive to errors in the measured elevation. At some stations the operational practice is that winds are not reported where the elevation falls below some critical value. At other stations secondary radar or transponder systems are used to provide direct measurements of the slant range used in the wind finding, eliminating the dependence of the derived wind on measurements of balloon elevation. Both practices lead to a reduction in the largest errors associated with radiotheodolite systems.

The problems of wind measurement at low balloon elevations at stations using radiotheodolite equipment is a major source of error in the global radiosonde network. Where there is no means of measuring the slant range, winds are calculated from measured values of balloon azimuth, elevation, and a value of the height derived from the pressure/temperature profile from the sonde. If the balloon encounters a strong jet it may be carried 100km or more downwind and its elevation at the ground station will be less than 10°. Accurate wind measurements require an accurate measurement of the balloon elevation which is critically dependent on the precise alignment of the receiving antenna. Where misalignment occurs, errors are likely to be much larger in the component of wind along the line of sight to the balloon than in the component perpendicular to the line of sight.

The characteristics of radiotheodolite systems at different balloon elevations can be investigated using model background values. It is necessary to work in wind components which lie along the line connecting the balloon and the station (a-component) and perpendicular to that line (p-component). Differences from background for each of these components can be calculated at various balloon elevations. The balloon downwind range can be estimated from the observed wind profile given in the radiosonde report, and the height can be calculated assuming a constant rate of ascent (taken to be 5 ms^{-1} here). Of course it is not known how model errors in the a and p-components of wind differ; for small elevations they are both likely to be larger than average as low balloon elevations result from strong winds at jet-stream level where it is known that model and representativeness errors are large. In addition the magnitude of model errors at low balloon elevation may depend on location; low elevation implies the existence of a strong (usually westerly) jet, which in turn implies the rapid propogation of errors. In such cases, model errors on the western coasts of continents, just downwind from data-sparse oceans, are likely to be larger than at sites further inland. These model characteristics are impossible to quantify without working from observational results, and as before background plus representativeness errors will be estimated by reference to wind-finding systems of known high quality.

Figures 6a-d show the dependence on the balloon elevation of the mean and rms O-B differences for the a and p-components of wind. The closed circles represent values from a wide selection of stations in Europe providing observations of good quality from either radar or NAVAID wind-finding systems. All observations have been included with the exception of those making an exceptionally large contribution to the variance of O-B. These outliers have been identified using standard statistical techniques on each sample of observations having values of the elevation within a specified range. In practice far fewer observations are eliminated than have flags raised by the routine quality control checks. All reports (TEMP and PILOT) received in 1988 have been used and vertical averages have been performed over the band 400 hPa to 150 hPa which includes the jet-stream maxima in most latitudes and seasons. As anticipated there is indeed an increase in O-B differences with smaller values of elevation, and the increase is a little greater in the a-component than in the p-component. These values provide a standard against which other stations may be compared. The crosses in Figure 6 are for station B which in Figure 4 had rms differences from background considerably larger than at neighbouring stations. It is immediately clear that the suspected observation error is contained in the a-component; the rms O-B of this component becomes very large at low balloon elevations, while that of the p-component differs little from the standard. At some stations observations cease where the elevation falls below some critical level, no doubt as a result of the local observing practice. Where observations continue at elevations below 10° massive rms O-B differences may be found as the third example in Figure 6 shows (indicated by the open circles) which is for a station in Asia.

The only reasonable explanation of the characteristics shown in these examples is an error in the radiotheodolite wind-finding systems in use at these two stations. In both cases the mean O-B of the a-component is also large and accounts for much of the variance. Quite possibly there is some misalignment of the antenna at these stations resulting in a constant bias in the measured elevation.

Errors in the assignment of height

The level assigned to a radiosonde observation, reported as a pressure in a TEMP report, may be derived in a number of different ways depending on the instrumentation: the pressure sensor on a sonde gives a direct measurement of the pressure level; alternatively the height in metres, derived from the slant range and balloon elevation, may be converted to a pressure level by applying the hydrostatic equation to the virtual temperature profile measured by the sonde. For systems with range-finding radar and a sonde with a pressure sensor, these two independent estimates of the height may be obtained and cross checked, providing probably the most accurate values of observation level. For systems with a sonde providing pressure and temperature but with no range finding, for example NAVAID and radiotheodolites without secondary radar, the pressure level assigned to the observations is simply the value measured by the sonde.

Some systems have no pressure sensor and rely on the range and elevation provided by the wind finding, and the virtual temperature profile provided by the sonde to give the pressure level. In some cases the elevation is measured by radiotheodolite and, as in the case of wind observations, errors can arise through misalignment of the instrument.

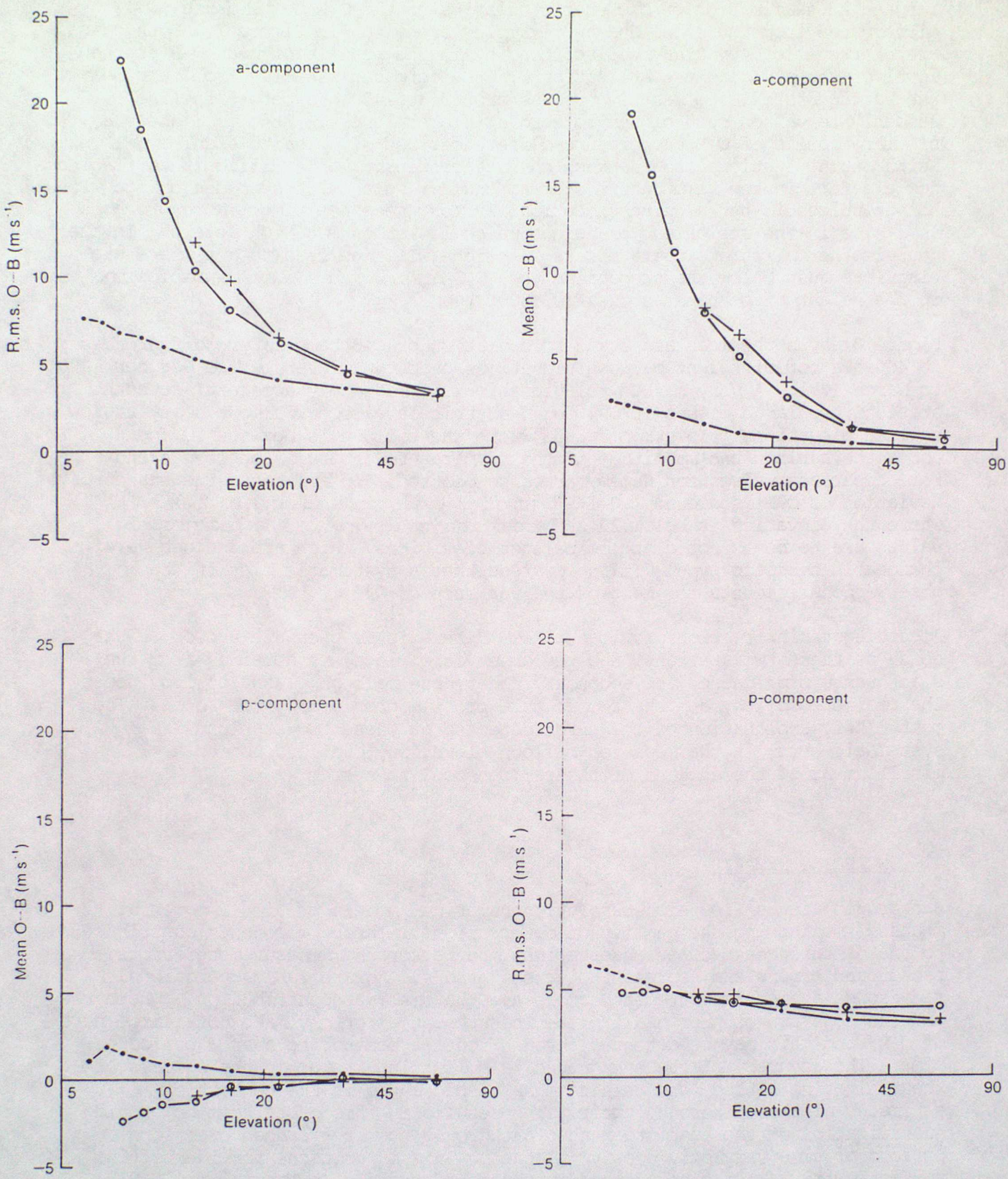


Figure 6. Mean and root mean square O-B differences for the wind components across (a-component) and perpendicular to (p-component) the line of sight to the balloon at various balloon elevations. Values have been averaged over the layer 400-150 hPa for the period January-December 1988. The closed circles are for radiosonde stations in Europe using NAVAID or radar wind-finding systems, the crosses for station B in Figure 4, and open circles for another station appearing to have large wind errors.

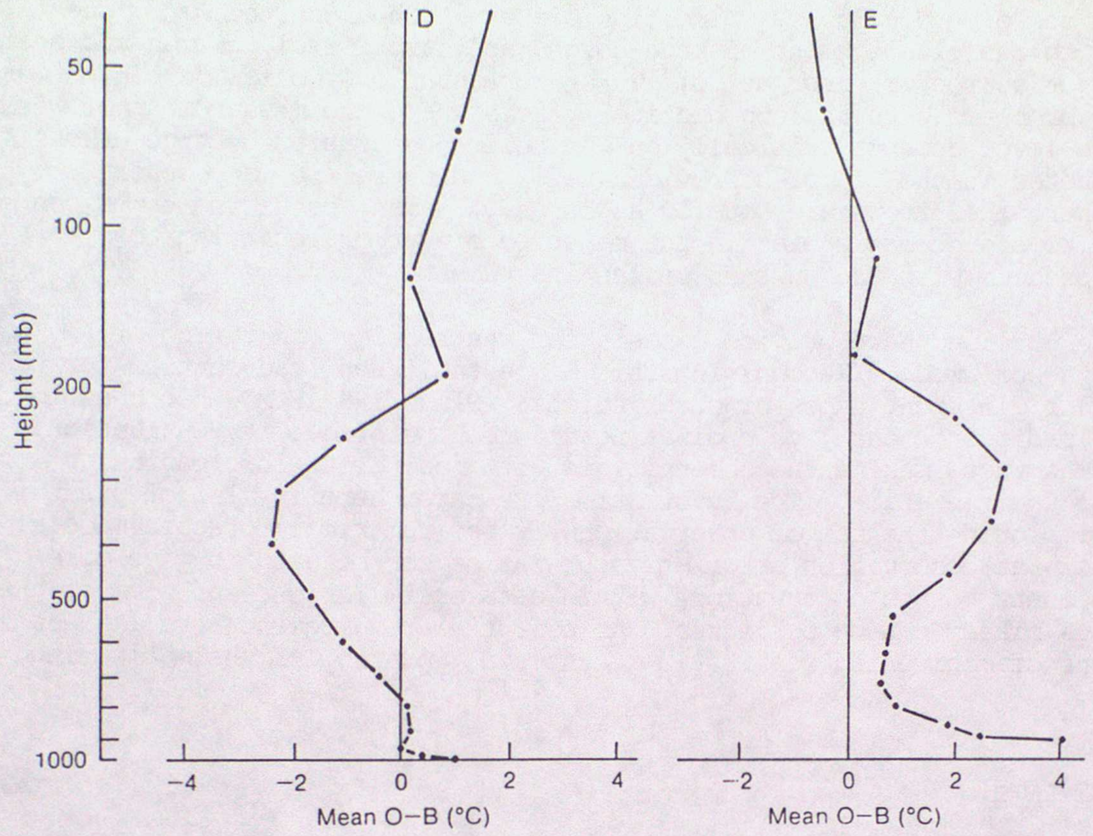
A possible way of detecting systematic biases in the height assignment is through an examination of the characteristics of the O-B temperature differences from the sonde and two examples are given in Figure 7. Both cases are characterised by a sharp discontinuity in the profile of mean O-B at the level of the tropopause. At station D, O-B increases steadily from zero near the surface to a large negative values around 300-400 hPa before falling suddenly to values much closer to zero at higher levels. Station E shows a similar profile of O-B temperature differences but of opposite sign. Neighbouring stations show no such characteristics. It is difficult to imagine how a defect in the temperature element could result in this sudden change with height unless it has a quite exceptionally long response time. More probable is a bias in the assignment of height which is largest at high levels. In the near-isothermal stratosphere errors in height will not lead to a temperature bias, but just below the tropopause the +3°C bias implies a systematic error in the height of perhaps as much as +500m.

Figure 8 shows the mean and standard deviation of O-B temperature differences at 400 hPa for different balloon elevations over the winter period October 1990 to March 1991. A very marked relationship is immediately apparent; at both stations the large biases in O-B found at this level occur almost soley at low balloon elevations. In strong winds where the balloon is between 10 and 20° above the horizon the magnitude of the temperature errors is between 4 and 6°C. For comparison values for Hemsby are also plotted. In all cases the standard deviation of O-B is around 1.0-1.5°C at high evelations rising to 2.0-3.0°C where the elevation is below 20°. Larger random errors in the background values are to be expected at low balloon elevations, which are indicative of a changeable synoptic type. It is apparent that a systematic bias in the observations accounts for most of the variance of O-B.

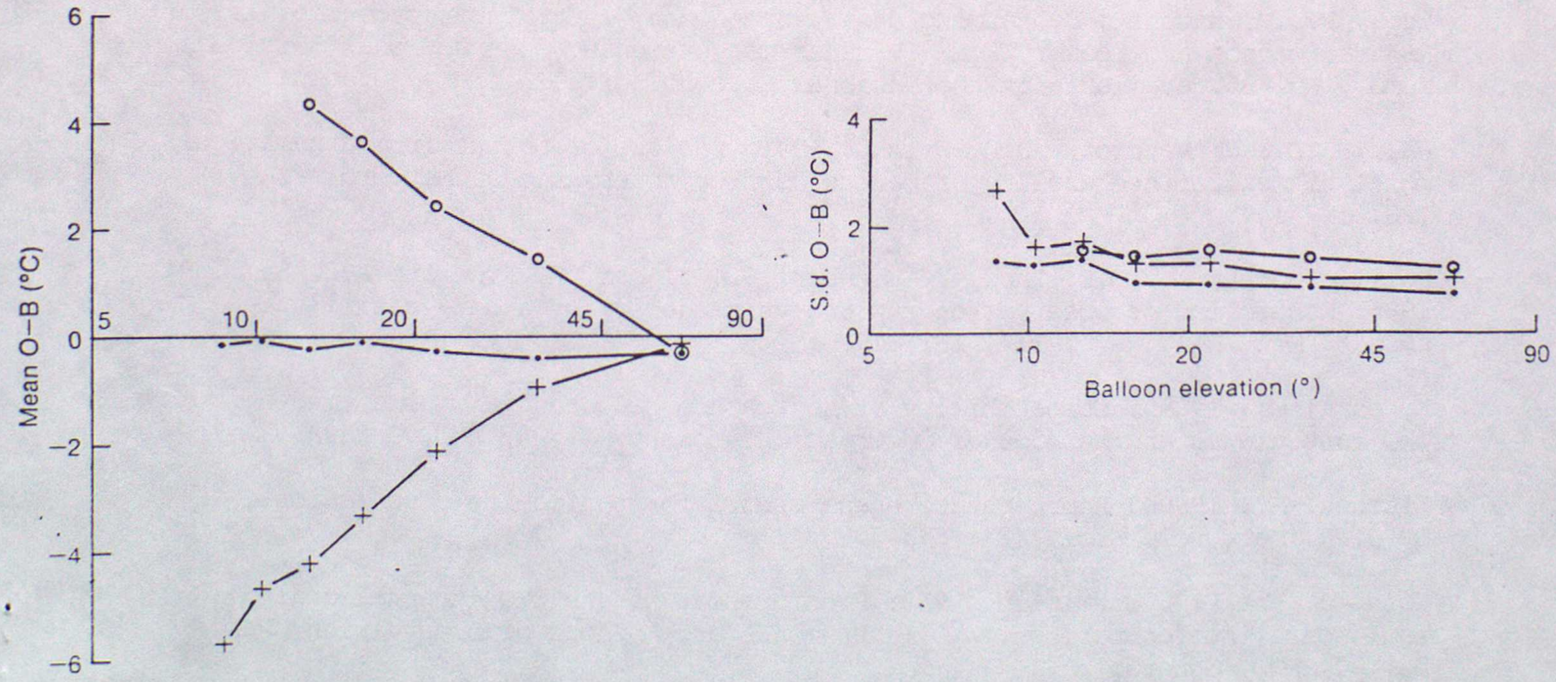
According to information provided to WMO, there is no pressure sensor on the sonde at these two stations, and the observation level is obtained from the slant range provided by the secondary radar, the balloon elevation provided by the radiotheodolite, and the temperature profile provided by the sonde. The most likely explanation of the error detected at these two stations is a systematic error in the measured balloon elevation, due no doubt to a misalignment of the antenna of the radiotheodolite system.

4. CONCLUDING REMARKS

Background values from high-resolution numerical models provide a powerful means of monitoring the quality of observations. Three components contribute to the differences between observations and background: measurement errors, background errors and representativeness errors. For reliable operational radiosonde systems measurement errors are the smallest contribution. Where the observations are of poor quality the measurement error may make up a large part of O-B, and this may be detected through routine monitoring over a period of time. The accuracy of the background is a limiting factor in the success of the monitoring method, and results must be presented in the context of estimates of the background error. The results in this paper seem to provide justification of the assumption that background errors averaged over long periods of time vary only slowly in space. This is critical for identifying stations with larger than average measurement errors. In general rms O-B values have uniformly low values highlighting the reliability of the observations, but a few stations stand out with values significantly larger than others in the immediate neighbourhood. Where this is the case observation error is the prime suspect.



<u>Figure 7</u>. Vertical profiles of the mean O-B temperature differences at two radiosonde stations for the period October 1990 - March 1991.



<u>Figure 8</u>. Mean and standard deviation of O-B 400 mb temperature differences at various balloon elevations for the period October 1990 - March 1991. The closed circles are for Hemsby, and the crosses and open circles for stations D and E of Figure 7.

Background values also provide a useful tool for investigating the cause of some of the errors. A number of techniques have been outlined here and no doubt more could be developed. Direction biases in the reported wind, due presumably to the misalignment of true north, are surprisingly common and seem to be a major source of error at some 20 or more stations worldwide. Even more common are problems with radiotheodolite systems where the measurement of wind or pressure level depends critically on the balloon elevation. Large errors have been noted in the examples provided here, and there are many other stations where the errors are equally as large. In many cases the error can be attributed to a systematic bias in the measured elevation, pointing to a levelling problem with the antenna of the instrument.

The methods outlined here provide a basis for the regular monitoring of observations worldwide. Recognising this WMO established lead centres for the monitoring of different types of observations. Since 1987 ECMWF has been lead centre for radiosonde data, co-ordinating all results of quality evaluation, and providing those making the observations with monitoring information relating to their station. The information presented here is also of great value for improving the use of observations by the numerical forecast models; estimates of the observation error at each station can be used to give more reliable weights to the observations within data assimilation, and some of the more obvious biases can be corrected. An essential requirement of all these applications of monitoring results is a continual updating of the monitoring information.

ACKNOWLEDGEMENTS

The author would like to thank John Ashcroft and Jonathan Wright for their help in developing the computer programs needed to provide the results presented here, Alan Radford for providing the ECMWF monitoring results for Hemsby, and John Nash for his helpful comments on the text.

REFERENCES

Edge, P., M. Kitchen, J. Harding and J. Stancombe, 1986: The reproducibility of RS3 radiosonde and Cossar WF MK IV radar measurements. Observational Services Memo., No. 35, available in the National Meteorological Library, Bracknell, UK.

Hall, C.D., J. Ashcroft, J.D. Wright, 1991: The use of output from a numerical model to monitor the quality of marine surface observations. Meteorol. Mag., 120, 137-149.

Hollingsworth, A., D.B. Shaw, P. Lonnberg, L. Illari, K. Arpe and A.J. Simmons, 1986: Monitoring of observation and analysis quality by a data assimilation system. Mon. Wea. Rev., 114, 861-879.

Kitchen, M., 1989a: Compatibility of radiosonde geopotential measurements. WMO Instruments and Methods of Observation, Report No. 36, WMO, Geneva.

Kitchen, M., 1989b: Representativeness errors for radiosonde observations. Q. J. R. Meteorol. Soc., 487, 673-700.

Nash, J. and F.J. Schmidlin, 1987: Final report of the WMO international radiosonde intercomparison. WMO Instruments and Methods of Observation, Report No. 30, WMO, Geneva.

Radford, A.M., 1987: ECMWF radiosonde monitoring results. Proceedings of ECMWF/WMO Workshop on radiosonde data quality monitoring, 95-114.