Met O 11 Technical Note No. 261

MODIFICATIONS TO THE AUTOMATIC QUALITY CONTROL OF SHIP DATA AND AN ASSESSMENT USING CASE STUDIES

by

B R Barwell
and
C A Parrett

Met O 11 (Forecasting Research)
Meteorological Office
London Road
Bracknell
Berkshire RG12 2SZ

1. Introduction

Analysis schemes for numerical models generally assume that observations and background fields contain randomly distributed errors. However, observations in particular may also contain either systematic or gross errors from various sources (see Fig.1) which may be large enough to seriously degrade analysed fields. It is the job of a quality control scheme to detect and flag these data so that they can be excluded from the analysis.

In the Meteorological Office, a quality control scheme using Bayesian techniques has been developed for surface data from ship and buoy observations extracted from the Synoptic Data Bank (SDB) (Lorenc and Hammon, 1986). In this scheme, an initial estimate of the probability of gross error is made for each item of data based on observation type and SDB flags. This probability is updated by comparing the data first with a background field and then with nearby observations ('buddy checks'). Data with a final probability of gross error greater than a given threshold (e.g. 50%) can then be flagged for exclusion from the analysis.

Output from the new quality control program has not yet been used operationally but has accumulated in an Observation Processing Database (OPD) since the program started running regularly in early March 1986. The first seven or eight months of data in this OPD forms the basis for the statistical calculations described in this report.

The OPD stores details of differences between observations and background values but partition into separate observation and background errors is not possible without additional information. Indeed, since errors of unrepresentativeness (see Fig.1) can be considered either as part of the observation or background error, such a partition needs further definition. Objective methods for estimating the observational error have been proposed (Hollingsworth *et al.* 1986) but the continual movement of ships, their non-uniform geographical distribution and the large variations in instrumentation from one ship to another complicate the application of these methods to surface data over the oceans. In our work we determine background errors by subtracting pre-existing fixed estimates of observational errors from the OPD statistics.

Lorenc and Hammon also reported on the results of case studies and an operational trial of the new scheme. These highlighted two particular areas where performance could be improved. First, if an observation was flagged by the SDB, there was very little probability that it would be accepted by the quality control however well it agreed with the background or with surrounding observations. Second, the background error fields used were fixed climatological fields with 5° resolution and did not depend on the local synoptic situation. Modifications to the program made to overcome these problems are described in sections 2 to 4 and an assessment of the performance of the modified program using case studies is made in section 5. Conclusions are presented in section 6.

2. New initial estimates of error statistics

Statistics from a previous version of the program can be fed back to improve the performance of a new version. In this way, the OPD was used to calculate mean probabilities of gross error for reported position, pressure, wind and temperature for each observation type and level of SDB flagging. The results were used as the basis of new initial estimates of these probabilities.

In most cases there was little change from the previous estimates. However, the statistics showed that for observations which had failed either the SDB check for consistency between a ship's reported position and movement, or that between the reported pressure and tendency, the initial probability estimate had been set too high. This provides objective evidence supporting the conclusion from the operational trial mentioned above. Initial estimates of the probability of gross error for these cases were therefore reduced from 90-95% to 70-75% depending on observation type, giving the data a better chance of being accepted for the analysis.

As more recent data on operational error levels was available (Bell 1985, Annex 8), the observational errors assumed by the program were updated. This involved a reduction in wind errors and a small reduction in temperature errors, but no change was judged necessary for pressure errors. Table 1 gives the new values of error levels and initial estimates of probabilities of gross error for various observation types and levels of SDB flagging.

Histograms of observation-background differences for data flagged by the quality control (i.e. final probability >50%) were plotted to check the validity of the rectangular ('top hat') distribution assumed for the representation of gross errors. Some disagreement was found for temperature errors as a result of which the assumed distribution was broadened and reduced in magnitude. Changes for pressure and wind data were judged not to be necessary.

3. Variation of background errors with synoptic situation

Climatological error fields can be used to represent the geographical variation of background errors but do not show variations with time. Background errors can be expected to vary with time in a way closely related to the ever-changing synoptic situation. The time variation may be important in the more active situations where accurate forecasting is more difficult and observations therefore all the more valuable.

	OWS	SHIP	SHIP (AUTO)	RIG/ PLAT	R/P (AUTO)	MOORED BUOY	DRFT. BUOY	SYNOP	SYNOP (AUTO)
Probability of gross error in location (%)									
Observation failed SDB sequence test	75	75	--	--	--	--	--	--	--
Observation failed other SDB test	90	90	90	90	90	90	90	90	90
Observation passed SDB sequence test	0.4	2	--	--	--	--	--	--	--
No SDB test done or flag set	1	4	4	2	1	1	7	0.8	1
Probability of gross error in pressure (%)									
Observation failed SDB tendency test	75	75	75	75	75	70	75	70	70
Observation failed other SDB test	95	95	95	95	95	95	95	95	95
Observation passed SDB tendency test	0.4	4	4	--	--	1	--	0.6	1.5
No SDB test done or flag set	0.8	6	6	4	1.5	1.5	5	2	2
Probability of gross error in temperature (%)									
Observation flagged by SDB	95	95	95	95	95	95	95	95	95
Observation not flagged by SDB	0.5	7	5	4	3	2	8	1.5	1.5
Probability of gross error in wind (%)									
Observation flagged by SDB	95	95	95	95	95	95	95	95	95
Observation not flagged by SDB	0.5	7	5	4	2	2	8	2	2
Observation error variances									
Mean sea level pressure (mb^2)	1.0	1.0	1.0	1.0	1.0	1.0	2.56	1.0	1.0
Temperature ($^oC^2$)	2.25	4.00	3.24	3.24	3.24	4.41	8.00	4.41	4.41
Vector wind ($(m/s)^2$)	9.68	12.50	10.58	10.58	7.22	7.22	18.00	11.52	8.82

Table 1. Initial estimates of probabilities of gross error for different observation types and levels of SDB flags. Also tabulated are observational errors used by the quality control program for each observation type.

For example, an explosively developing depression may be poorly represented in background fields due to inadequate data coverage at previous times, or because the resolution of the model was insufficient to determine its intensity adequately and predict its evolution. Any subsequent observations in the area are then particularly important for forecasting developments accurately, but they may differ significantly from background fields. Without some prior knowledge that larger than usual errors in the background field are likely, these valuable observations may well be rejected by the quality control.

On the other hand, except in very data-sparse regions, the background fields within a large quasi-stationary anticyclone will usually be of good quality because the slow evolution and weak gradients of pressure, wind and temperature will minimise errors due to poor timing of developments or incorrect geographical location. Subjectively, it may then be possible to identify a reported pressure from a ship as being in error even if it differs from the background field by only a few millibars, but an automatic scheme relying only on climatological statistics would be most unlikely to reject such data.

In an attempt to incorporate a background error field dependent on synoptic situation into the automatic ship quality control program, an investigation of the dependence of observation-minus-background differences was carried out for pressure and wind observations. The OPD generated by the old version of the program formed the database for this work. Background fields of mean sea level pressure were extracted from the WGDOS printfile archive (Lowther 1986) but corresponding surface wind and temperature fields were not available.

3.1 Background error in sea level pressure

Using the background fields and OPD data, an expression of the following form was fitted to sea level pressure differences between observation and background;

$$\text{Mean square pressure error} = A + B|\nabla p|^2 + C(\partial p/\partial t)^2. \qquad (3.1)$$

The unknown coefficients A, B and C are expected to vary geographically but are taken to be independent of time and local conditions. In Eq.(3.1), the first term on the right allows for climatological variation of the mean variance, the second is significant in regions where timing or positional errors in the background fields are most likely to be important, and the last identifies areas where errors due to incorrect development are most likely to occur.

Coefficients A, B and C were determined for each 10-degree latitude band using data from the OPD. All pressure observations within a band which were given a final probability of gross error less than 50% were used. About 260,000 pressure observations were involved covering the period from early March to mid-October 1986. Each observation was assigned to a 'box' depending on the background value of $|\nabla p|$ and $\partial p/\partial t$. $|\nabla p|$ was computed from the model background field by finite differences within the grid square containing the observation. $\partial p/\partial t$ was estimated

from the change of pressure at the observation location during the 6-hour forecast which generated the background field, i.e.,

$$\partial p / \partial t = p_{t+6} - p_{t+0} \quad \text{mb/6hrs} \qquad (3.2)$$

The resolution of each box was 2×10^{-3} mb/km for $|\nabla p|$ and 1 mb/6hrs for $\partial p/\partial t$. Fig.2 shows a typical distribution of observations on a chart of ∇p against $\partial p/\partial t$. Observations with large pressure tendency generally occur in regions of significant pressure gradient but the reverse is not necessarily true: observations in a large pressure gradient need not have a large tendency.

The mean square pressure difference between observation and background was computed for each box containing at least two observations. Equation (3.1) was fitted to the data using a least squares technique with each box weighted by the number of observations. This prevented unusually high or low values in boxes with only a few observations from dominating the results.

As an example of the calculation, Fig.3 shows the details for the latitude band 50-60°N. The mean square values of the observation-minus-background pressure differences are plotted together with contours of the fitted function (3.1). Although computed values differ from fitted values in some boxes, the function captures much of the large scale variation especially in the region where most of the observations are concentrated.

The symbols in Fig.4 show the variation of A, B and C for 10° latitude bands (except for the 80-90°S band for which there were no observations). Values for each 10° latitude circle were derived by averaging values for adjacent bands; this also filtered out some of the small scale variation. Some extrapolation was necessary in the polar regions. The full lines in Fig.4 show the resulting latitudinal dependence of the coefficients.

'A' remains roughly constant over much of the northern hemisphere and tropics but increases rapidly below 30°S reflecting the greater background errors in the extra-tropical southern hemisphere where the data coverage is poor. Since B and C do not show the same increase, they do not appear to be influenced by differences in data coverage. The distribution of B is bell-shaped, being a maximum in the tropics where coupling between mass and wind fields is weakest. The values of C show some variation from band to band but remain much the same magnitude over the whole globe. (Values of B and C in the tropics are difficult to determine with confidence because of the small range of ∇p and $\partial p/\partial t$, but because the range is so small values are not required to the same accuracy.) The weak dependence of C on latitude and the broadly symmetrical shape of B increase confidence that different values for different seasons are not required. Increases in background errors in seasons with greater activity will be reflected in increases in ∇p and $\partial p/\partial t$. In the absence of a full year of statistics, it is difficult to determine the seasonal dependence of A. However, since A represents the pressure error for the case where there is no synoptic activity ($|\nabla p|=0$

and $\partial p/\partial t=0$), any changes in A due to seasonal variations in synoptic activity are probably no more significant than those in B and C.

As a check on the degree of confidence in the computation of the regression coefficients, the calculation was repeated for 5° latitude bands. Results were very similar to those of Fig.4 for the large scale features but showed more small scale 'noise' especially in the tropics. Other runs in which coefficients were computed for 'boxes' rather than complete latitude bands gave unreliable results because the number of observations in some boxes was insufficient to permit meaningful calculations of mean square pressure differences. The use of 10° latitude bands therefore suffices to take account of the most important latitudinal variation of pressure errors without introducing numerical errors due to non-uniform data distribution.

To represent the background error in mean sea level pressure, we use a formula similar to (3.1);

Mean square pressure error in background =

$$A' + B|\nabla p|^2 + C(\partial p/\partial t)^2. \qquad (3.3)$$

where A' is calculated by subtracting 1 mb^2 from A, this being the observational error variance assumed for reports from ships and land stations by the operational analysis scheme (Bell 1985). The final values of A', B and C are given in Table 2. (The last column of Table 2 will be explained in section 4.)

Latitude	A' (mb^2)	B (m^2)	C (6 hrs)^2	Mean square error (mb^2)
90°N	1.62	1.1×10^9	0.125	2.75
80°N	1.29	1.1×10^9	0.093	2.13
70°N	1.32	1.1×10^9	0.056	2.14
60°N	1.26	1.4×10^9	0.080	2.42
50°N	1.10	2.4×10^9	0.106	2.71
40°N	1.35	2.8×10^9	0.102	2.64
30°N	1.25	5.5×10^9	0.119	2.04
20°N	1.13	8.4×10^9	0.134	1.88
10°N	1.60	6.4×10^9	0.075	1.98
0°	1.80	9.9×10^9	0.076	2.09
10°S	1.64	14.3×10^9	0.111	2.20
20°S	1.59	8.2×10^9	0.146	2.27
30°S	1.65	3.3×10^9	0.157	2.56
40°S	2.73	3.0×10^9	0.091	5.00
50°S	5.60	3.2×10^9	0.096	9.55
60°S	7.28	3.0×10^9	0.115	11.91
70°S	6.20	2.0×10^9	0.079	9.12
80°S	6.20	2.0×10^9	0.093	9.12
90°S	6.20	2.0×10^9	0.125	9.12

Table 2. Regression coefficients A', B and C for background pressure errors (see Eq.(3.3)) and mean square error.

3.2 Background error in vector wind

Wind observations can differ from background wind fields for a variety of reasons. As well as observation errors, there may be errors of representativeness from such sources as non-standard height of anemometer or the influence of local topography or man-made structures. These errors are likely to be local, non-random and dependent on wind direction.

A major source of wind errors is unrepresentativeness due to differences in scale (Fig.1). A wind observation is typically a 10-second mean value and can therefore be affected by fluctuations on scales greater than a few hundred metres even in a strong wind. In contrast, the grid length of a numerical model is measured in tens of kilometres. Unrepresentativeness due to wind fluctuations on intermediate scales can make a substantial contribution to differences between observation and background.

The magnitude of fluctuations in the observed wind is strongly dependent on the mean wind speed. We therefore expect the wind speed to be the dominant atmospheric variable on which wind errors will depend. Working in 20° latitude bands and using data from the OPD for the same period as in section 3.1, wind observations were grouped in 50 boxes corresponding to their reported wind speed in knots. (Ideally, the boxes should be based on background wind speed but the appropriate background fields were not available.) As in section 3.1, only observations whose probability of gross error was less than 50% were used; about 248,000 altogether. (The few reports of winds > 50kt were not used. Reports of zero knots were also rejected because of contamination by corrupt but unflagged drifting buoy reports.)

Mean square wind errors were computed for each box and plotted against wind speed. As an example, the graph for the 50-70°N band is shown in Fig.5. The points define a fairly smooth curve though with some scatter at high wind speeds where there are fewer observations per box. In less data-rich bands there is more scatter but the general shape of the curve is the same. The variation suggests that the errors can be approximated by a parabolic function of the form

$$\text{Mean square wind error} = D + E(\text{speed})^2. \qquad (3.4)$$

Values of D and E for each latitude band were determined by fitting this function to the data using a least squares technique. The appropriate curve for the 50-70°N band is shown in Fig.5. The symbols in Fig.6 show the latitudinal variation of D and E for each band except 70-90°S for which there was insufficient data. Like A in Fig.4, D increases sharply in the southern hemisphere where the data coverage is poor. E, however, shows no such increase but has similar values in each hemisphere giving confidence that seasonal variations are not significant.

Coefficients for each 20° latitude circle were generated by averaging values for neighbouring bands (full lines in Fig.6). Some adjustment and rounding was also done for D; in particular, the value

for 50-70°S was reduced because the small sample size for this band made accurate determination impossible.

In the absence of additional information on the characteristics of observational errors in ship winds, a constant 35 kt^2 was assumed corresponding approximately to the error used by the operational model for a surface wind report from a ship. This value was subtracted from D and the results rounded to generate a value D' for use in the equation,

$$\text{Mean square wind error in background} = D' + E(\text{speed})^2. \qquad (3.5)$$

Final values of D' and E are given in Table 3.

Latitude	D' (kt^2)	E	Mean square temperature error $(°C)^2$
90°N	20.0	0.174	4.0
70°N	15.0	0.148	4.3
50°N	28.0	0.122	3.5
30°N	23.0	0.140	3.4
10°N	10.0	0.169	1.3
10°S	10.0	0.161	1.7
30°S	25.0	0.131	3.7
50°S	50.0	0.120	9.6
70°S	65.0	0.148	22.0
90°S	40.0	0.174	32.0

Table 3. Regression coefficients D' and E for background wind errors (see Eq.(3.5)), and mean background temperature error.

The mean bias between observation and background was found to be small (~1 knot) compared with the rms value and was neglected in the analysis. No noticeable difference was found between statistics of mean square errors for 'u' and 'v' wind components separately, but the possibility of differences in error distribution for components parallel and perpendicular to the reported wind direction remains.

3.3 Background error in temperature

A detailed study of the variation of temperature errors as a function of synoptic situation was not attempted since temperatures are not used operationally by the assimilation. Mean square temperature differences between observations and background were determined from the ship quality control OPD for 20° latitude bands in the same way as for the wind coefficients. Values for adjacent bands were averaged to derive values for latitude circles. A constant value of $4.0°C^2$ was subtracted to allow for the observational error and some adjustment was made in the polar regions using older estimates of background error (Bell 1985)

because the small number of observations in these regions made derived values unreliable. The resulting background errors are given in the final column of Table 3.

4. Operational implementation

In the derivation of the coefficients in Tables 2 and 3, an allowance was made for a fixed observational error; no attempt was made to make this error variable with synoptic situation like the background error. For observations of pressure or temperature there may not be much variation, but for winds the observation error is likely to be greater the stronger the wind, especially on a ship at sea.

Eq. (3.3) is therefore likely to overestimate slightly the dependence of background error on synoptic situation. A parameter k_p was introduced to allow adjustment of the dependence as follows:

$$\text{Mean square pressure error} = k_p\epsilon_{3.3} + (1-k_p)\epsilon_p \qquad (4.1)$$

i.e. a linear combination of $\epsilon_{3.3}$, the mean square error computed from Eq. (3.3), and ϵ_p, a mean value dependent on latitude only. Values of ϵ_p (last column of Table 2) were computed from the OPD in the same way as the A' coefficient of Eq. (3.3). As the constant k_p increases from 0 to 1, the mean square pressure error changes from a simple latitude-dependent value to the full synoptic scale dependence. As suggested by the argument in section 1, a suitable value of k_p can only be set subjectively: a value of 0.9 was found to be satisfactory.

For wind errors, a similar equation was used;

$$\text{Mean square vector wind error} = k_w\epsilon_{3.5} + (1-k_w)\epsilon_w \qquad (4.2)$$

where $\epsilon_{3.5}$ is the value computed from Eq. (3.5) and ϵ_w is a latitude-dependent mean value. OPD statistics showed that ϵ_w could be estimated by using Eq. (3.5) with a (wind speed)^2 of 200 kt^2. The constant k_w was taken as 0.9.

It should be stressed that all coefficients in Tables 2 and 3 were computed using ocean data only and do not necessarily apply over land.

5. Case studies

Three case studies have been undertaken to assess the performance of the new ship quality control scheme described in sections 2-4. The scheme was run using surface marine data extracted from the Synoptic Data Bank (SDB), and used T+6 coarse-mesh forecasts (run from archived operational analyses) for the background fields. The probabilities of gross error, etc. were written to an OPD and compared with the flags obtained from the operational OPD. The dates chosen for the cases were:

(a) 00 GMT 15th December 1986,
(b) 12 GMT 12th February 1987,
(c) 12 GMT 12th June 1987.

As with previous case studies (Lorenc and Hammon 1986) observations with a final probability of gross error [P(G)] $\geq$ 50% in the new scheme were taken as being flagged, and those with 30% $\leq$ P(G) $<$ 50% as being 'suspect'. All the results shown are with a buddy check search radius of 300km, although runs were carried out with values of 150km, 400km and 500km (Lorenc and Hammon used 150km). As in the 1985 trial, coastal land SYNOPS below 100m were used to buddy check the marine observations.

An overall assessment of the three case studies is given next with a few brief examples, followed by more detailed examples taken from areas of different data density.

5.1 Overall assessment of case studies

Figures 7a to 7c give a breakdown of the flagging of Pmsl and wind observations by the new scheme and the operational scheme for each of the dates studied. (In these figures, underlined numbers refer to Pmsl observations, other numbers to wind observations.) There were about 800 surface marine observations for each data time (about 50 drifting buoys, 30 moored buoys, the remainder being ships and platforms), with similar numbers of observations being rejected by each scheme - approximately 100-130 Pmsl and around 90 wind observations. The branches of particular interest in Fig.7 are the two central ones at the third level down where different quality control decisions were made by the two schemes (involving 65-100 Pmsl and slightly more wind observations in each case). We will look firstly at those observations rejected operationally and then at those accepted operationally, concentrating on Pmsl.

(i) Observations rejected operationally :

The majority of these were rejected by the Central Forecasting Office (CFO) during intervention; the total number ranging from 73 to 92 for Pmsl. Approximately half of these were also rejected by the new scheme (which had no access to CFO flags), but the remaining half were accepted. We will concentrate on this latter half, more than 90% of which were judged subjectively to be 'good' observations (i.e. could be considered as within the random noise level of the observations). About a quarter were rejected because of errors in their wind reports - it being easier for CFO to reject the whole ship report rather than just one element. (Similarly, most of the wind reports rejected by CFO but accepted by the new scheme were from observations actually rejected for their pressure.) Thus the new scheme performs better here simply by treating Pmsl and wind separately. Some Pmsl observations were rejected by CFO because they differed from the background, but were accepted by the new scheme if they were supported by a few neighbours - often it was difficult to judge whether they were correct or not. Figure 8 shows an example from 15th December 1986 where the T+6 forecast has moved a frontal system too rapidly eastwards towards the Bay of Biscay and the UK. Two apparently good ship reports, UGRA (Pmsl=1015.1mb) and EWEN

(Pmsl=1023.6mb), were rejected by CFO but accepted by the new scheme. A few Pmsl observations that were rejected by CFO agreed with the background, but if the background was in error then the new scheme needed several good buddies to check them against in order to flag them correctly. Virtually all of the Pmsl observations rejected by both CFO and the new scheme were clearly incorrect, e.g. ship FNFD in Fig.8. Thus, overall, the new automatic scheme shows a level of skill comparable with or better than that shown by manual intervention in CFO.

Observations given a location flag by the SDB are not used by the operational system. The new quality control scheme initially gives them a P(G)=75%, but after checking against the background field and any buddies this can be reduced below 50% and the observation reinstated. One such example is shown in Fig. 9, where ship ZRCS failed the position sequence test in the SDB, but its pressure and wind were sufficiently close to the background values to be accepted (P(G)=49% for Pmsl). (The clearly incorrect report from ship UNCD, was used for buddy checking but had a negligible effect.) The ability to reinstate such observations on the background check alone is related to the new initial estimates of gross error (see section 2) and represents an improvement over the previous version of the scheme. About 2/3 of the observations given SDB position flags but accepted by the new scheme seemed to be correct, with either an error in their previous reported position or in their reported movement.

Of the few observations flagged automatically by the operational analysis but accepted by the new scheme, most were such that it was difficult to decide whether they were correct or not. Figure 10 shows an example from 15th December 1986 where there were four conflicting observations of Pmsl in the W. Pacific, S.E. of Japan. Two were flagged by the operational system, one of which (WHRN) was just accepted by the new scheme (P(G)=48%); and two were corrected by CFO, whereas one of these two (the uncorrected version, which is plotted) was just accepted by the new scheme (XCCS). In this sort of situation an automatic blacklist of ships would be very helpful in deciding which observations are reliable (and which to correct if they have a known bias). It is planned to incorporate an automatic blacklist into the new scheme in the near future.

Most of the observations that were accepted by the new scheme but classed as "not ver.(verified) by the analysis" in Fig.7 were groups of drifting buoys (mostly in the S. Hemisphere) on CFO's blacklist, many of which agreed with the background. There were also a few ships that sent in two reports - the new scheme processed both whereas the operational system was only presented with the later one.

(ii) Observations rejected by the new scheme but accepted operationally :

One can see from Fig.7 that for each case approximately 25-30 Pmsl observations that were accepted by the operational system were rejected by the new quality control scheme. Most of these (about 80%) were rejected by buddy checks, and nearly all had small observation-background (O-B) differences and may not have come to the attention of

CFO. An example is shown in Fig.11 from 12GMT 12th February 1987, where observations indicate that the background field of Pmsl was several millibars in error in the central Mediterranean, near the N. African coast. The three ships HZXN, SQJQ and TSLH were between 1½ and 4mb too low, and despite having smaller (O-B) differences, they were rejected by the new scheme at the buddy check stage (although they were accepted with a smaller search radius of 150km).

There were a total of 12 Pmsl observations rejected on the background check, most having (O-B) differences > 5mb. Of the five Pmsl observations flagged by the SDB (for pressure tendency), three were very close to the background, but were rejected by comparison with buddies; the other two were not reinstated by the background and had no buddies.

All of these observations rejected by the new scheme were looked at in detail and assessed as to whether they should have been accepted or rejected. The majority of the Pmsl reports were assessed as being in error and correctly rejected by the new scheme. Another 1/3 may have been in error - it being difficult to come to a definite decision, due to insufficient or conflicting data. The remainder (about 1/8) were assessed as correct observations and were wrongly flagged by the new scheme; most of these were in data sparse areas where the T+6 background field was in error - near deep and/or rapidly-moving depressions or near small-scale synoptic features.

A larger proportion of the wind observations were wrongly flagged by the new scheme, mainly in data dense areas and where they were buddy-checked against a number of land stations. In such cases, coastal and/or orographic influences can produce local effects on wind reports from SYNOPS making them unsuitable for comparison with ship winds.

5.2 Examples of different quality control decisions

(i) Figure 12 shows the observations and background Pmsl field for 00GMT 15th December 1986 near the west coast of N. Africa, where the observations suggest that the pressure was slightly lower than the background field. Five ship observations of Pmsl were rejected by the new scheme but accepted operationally. Only two of these were rejected with the smaller buddy check radius of 150km: EWWJ and EMXE, both about 4mb lower than the background and almost certainly wrong. With the buddy check radius at 300km three more observations were flagged (ESSU, UUCP and EOSP), due to checking against more ships (e.g. UUCP had only 3 buddies within 150Km, but it had 8 within 300km). These three ships had pressures 2-3mb lower than the background and also lower than both the land SYNOPS and the unflagged ship reports. Thus, the increase in buddy check radius to 300km has enabled a clearer identification of those observations likely to be in error. There were no further changes with the buddy check radius increased to 400km. The wind observations were all accepted by both schemes. (Ship UHJU's Pmsl was also flagged by the new scheme, being over 12mb too high. CFO 'corrected' the pressure to 1021.0mb, but the ship had a record of poor quality reports and would have benefitted from rejection by an automatic monitoring scheme.)

(ii) The new quality control scheme should reject more Pmsl observations with (O-B) differences of just a few millibars in quiet anticyclonic regions, because the assumed background error is small (see section 3). Figure 13 shows an example of an improved quality control decision in an anticyclone in the western N. Atlantic for 00GMT 15th December 1986. Ship EADF's reported Pmsl is several millibars higher than any nearby observation and was rejected by the new scheme after checks against three neighbouring observations, whereas it was accepted operationally. Ship A8TS was also accepted operationally, but just rejected by the new scheme, largely because it failed the SDB check on pressure tendency due to an incorrectly reported tendency. The wind report from ship VCRJ was also flagged by the new scheme, whereas it was accepted operationally.

(iii) Figure 14 shows another example of improved quality control decisions in the new scheme; this time in the data dense area of the North Sea, for 12GMT on 12th February 1987. The three ships ENXA, PHKS and XP3344 appear to have reported Pmsl's about 1½-2mb too high, After comparison with 23, 23 and 20 buddies respectively, all three were flagged by the new scheme; whereas they were accepted operationally. The platform reporting a pressure of 997.6mb was about 2mb too low, but was only 'suspect' after checks against fewer (14) buddies. Ship GRHJ's Pmsl was rejected by CFO (being over 4mb too low) and this was also flagged by the new scheme. The only wind observation flagged was that from ship UJAC which was flagged by the new scheme.

(iv) A situation in which the new quality control scheme performed less well is shown in Fig.15. A small-scale depression near Japan was wrongly positioned in the background field for 12GMT 12th February 1987. This produced large (O-B) differences near the position of the forecast low centre, while further to the north there were only small (O-B) differences near the cluster of ships just off the east coast of Japan. Thus, with the new scheme, ship JBOA's wind report was flagged and its Pmsl nearly flagged [P(G)=47%] after the background check. With the smaller search radius of 150km, the complete observation was reinstated by comparison with ship 8KRZ and two close land stations; but with a search radius of 300km, the cluster of ships further north were included in the buddy checking, with the result that JBOA and also 8KRZ were flagged for both Pmsl and wind. In this case a large local gradient of background error is inconsistent with the function used for the background error correlation. With small-scale synoptic features such as these, errors in the background field can lead to observations being checked against unsuitable buddies and being wrongly flagged. This problem is difficult to remedy in the absence of any method of identifying regions where the background field is likely to be wrong.

(v) Figure 16 illustrates the benefit of inserting BOGUS data to support observations in data sparse areas where the background field is in error. The isolated ship UYIR was located near to the centre of the very deep Atlantic depression of 15th December 1986 which was positioned slightly too far east in the background field. UYIR was given an SDB flag for its pressure tendency, and its Pmsl was flagged on the mode 1 (background) check but reinstated on the mode 2 (analysis) check in the operational scheme. It was supported by two CFO BOGUS observations (not

plotted in Fig.16), both positioned within ~10km of UYIR with Pmsl=920mb (one also had a wind of 030° 60kt). The new scheme flagged UYIR on the background check but was not able to reinstate it since BOGUS data were not used. Another factor influencing the performance of the new scheme was that the values of the pressure gradient and change in pressure in the background field at the precise location of UYIR were small and not representative of the general level for points in the vicinity. Consequently, the background error computed from Eq.(3.3) was unrealistically small leading to increased probability of flagging. (There were a few more occasions where isolated ships were flagged because of large errors in the background field near deep depressions, mainly in the southern oceans.) The problem could be alleviated in these cases by the use of modified 'envelope' background error fields, i.e. those with increased values where the gradient of the error field is large. This may improve the handling of ships like UYIR (and possibly JBOA in Fig.15), especially if not supported by neighbouring observations.

(vi) The final example (Fig.17) is one from the 12th June 1987. Two ship reports of Pmsl (C4VW and UWRV) near to the West Indies were accepted operationally but flagged by the new scheme at the buddy check stage (with 8 and 6 buddies, respectively). Their pressures appear to be about 3mb too low suggesting that the new scheme made the correct decision. Ship ENXH's Pmsl report was obviously wrong and was rejected by both CFO and the new scheme. All the other reports were accepted by both schemes, although ship LIZA's wind report was 'suspect' after buddy checks against nearby land stations.

6. Conclusions

Improvements to the ship quality control scheme have been made to incorporate statistical results from the previous version and to use background field errors for pressure and wind which vary with the local synoptic situation. The new scheme performs better than both the previous version and the current operational scheme, and can match the operational scheme plus manual intervention. Where observations were rejected in their entirety by CFO, the new scheme was often able to identify some elements of the report as being acceptable. Checking against nearby observations was a valuable method of providing additional information on observation quality.

The new scheme performed less well where the background field was poor, especially in regions where the low data density did not allow much buddy checking, and where there was a poor representation of small-scale features. Better performance is difficult to achieve in these cases in the absence of a reliable method of identifying the regions concerned.

Increasing the search radius for buddy checking from 150km to 300-350km improves quality control decisions. 350km has now been made the standard value in the program. Improvements in program efficiency since the original version enable this to be done without any major penalty in elapsed time.

The following conclusions highlight areas where further improvements are desirable:

(a) Many ship wind observations were rejected after comparison with winds from land stations which may not be comparable due to the different characteristics of surface wind over land and sea.

(b) There was a tendency to flag observations too readily in data dense areas (especially wind observations) due to excessive buddy checking.

(c) Surface BOGUS data are not currently used by the new scheme. These should be included, especially where they have been inserted to support ship observations.

(d) Ships and buoys with a record of poor quality observations (e.g. a consistent bias) should be detected by an automatically updated monitoring system. ('Blacklist'.)

The following changes and proposals have been made to effect these improvements.

In view of (a), data checking for a buddy check between a ship and a land station has now been confined to Pmsl: checking of wind and temperature in this case is no longer done.

The program has now been modified to use the data sets output by the data extraction program as its data source rather than the SDB. As well as saving CPU time, this will enable access to intervention flags and observations corrected by CFO. In addition, the use of station lists in data-dense regions will help to prevent the excessive buddy checking referred to in (b) above.

Program modifications are planned to access the surface BOGUS data for processing with ship data.

The above program changes should be completed before the new scheme is considered for operational implementation. Looking further into the future, it is planned to develop a system for monitoring the quality of ships and buoys so that those with biases or large rms errors can be identified. (A comprehensive monitoring system could also identify consistently reliable ships as well as consistently unreliable ones.) As well as providing additional information for automatic quality control, such records provide the potential for information on systematic errors to be sent to data providers so that corrections can be made at source.

Background error fields can be made more representative by a spatial smoothing or similar operation. Eventually, local manual modification to background error fields may be a useful alternative intervention technique.

Further analysis of the statistics of gross error using OPD data are also required. In particular, the assumption of a 'top hat' distribution

for gross errors in the magnitude of vector wind needs to be verified. Separate 'top hats' for the *u* and *v* components may be more suitable.

Successful implementation of the above modifications and rapid display of the results to forecasters should reduce the requirements for manual intervention and provide a better basis for human monitoring of the operational suite.

References

Bell, R. S. (editor), 1985: The data assimilation scheme. *Operational Numerical Weather Prediction System Documentation Paper No.3*, Version 2, October 1985.

Hollingsworth, A., Shaw, D. B., Lönnberg, P., Illari, L., Arpe, K. and Simmons, A. J., 1986: Monitoring observation and analysis quality by a data assimilation system. *Mon. Wea. Rev.*, **114**, no.5, 861-879.

Lorenc, A. C. and Hammon, O., 1986: Objective quality control of observations using Bayesian methods - Theory, and a practical implementation. *Met.O.11 Tech. Note No. 225.*

Lowther, D., 1986: WGDOS printfile archive. *Operational Numerical Weather Prediction Scheme Documentation Paper No.9.1.*

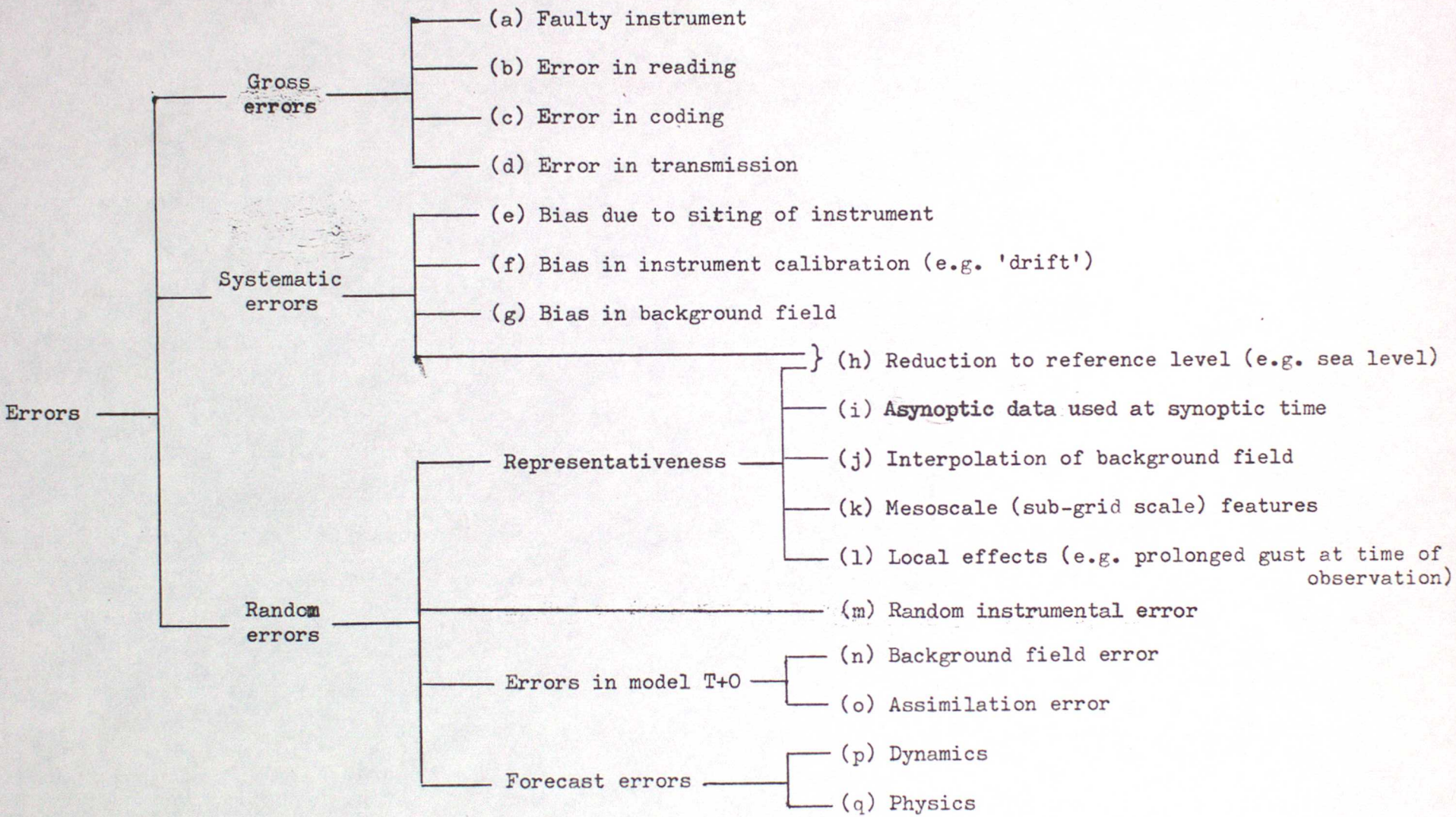


Figure 1. Sources of error in meteorological observations and numerical analyses and forecasts.

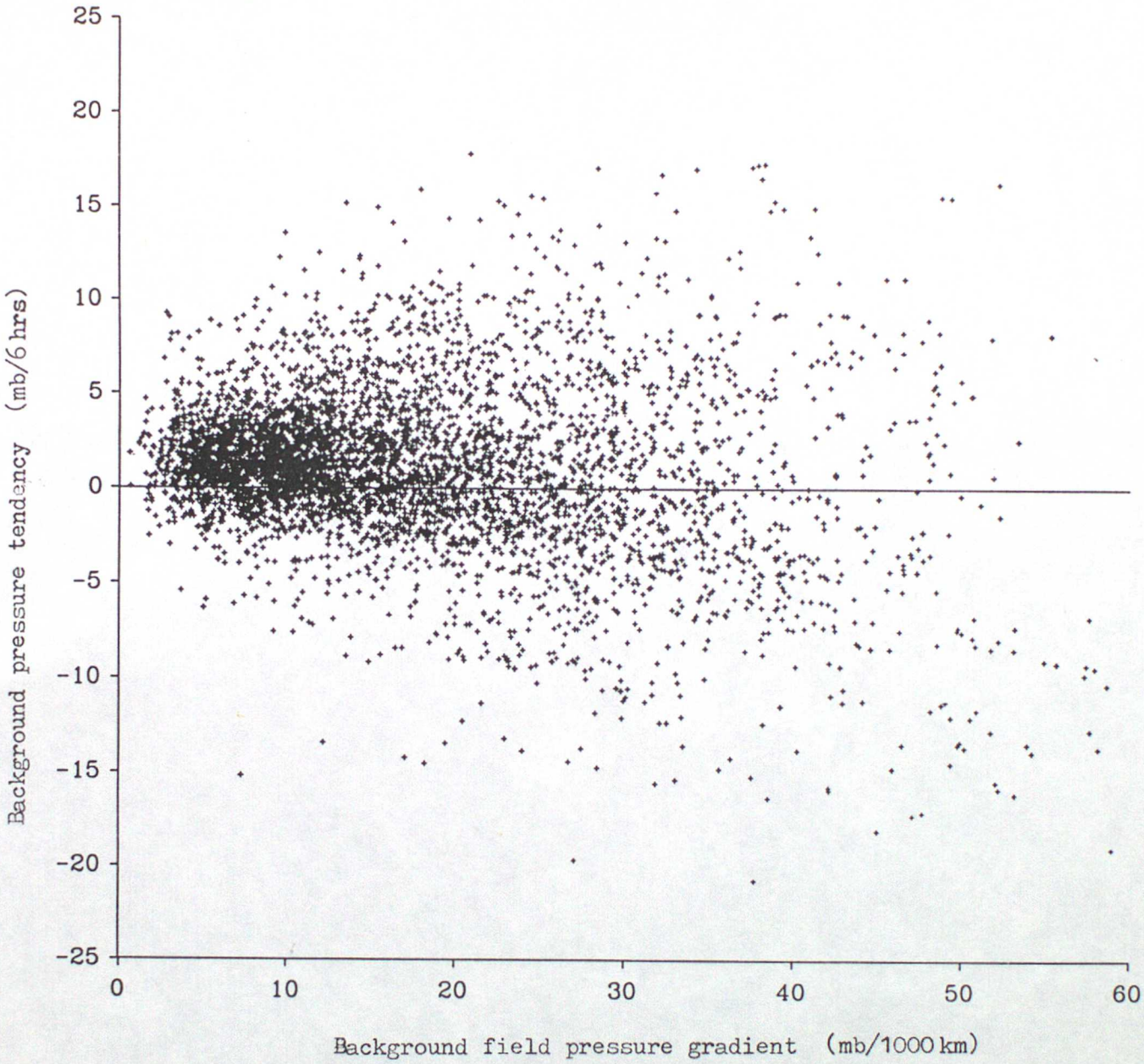


Figure 2. A typical distribution of observations plotted on a chart of background pressure gradient (horizontal axis) against background pressure tendency (vertical axis). This chart is for 5000 observations from the North Atlantic.

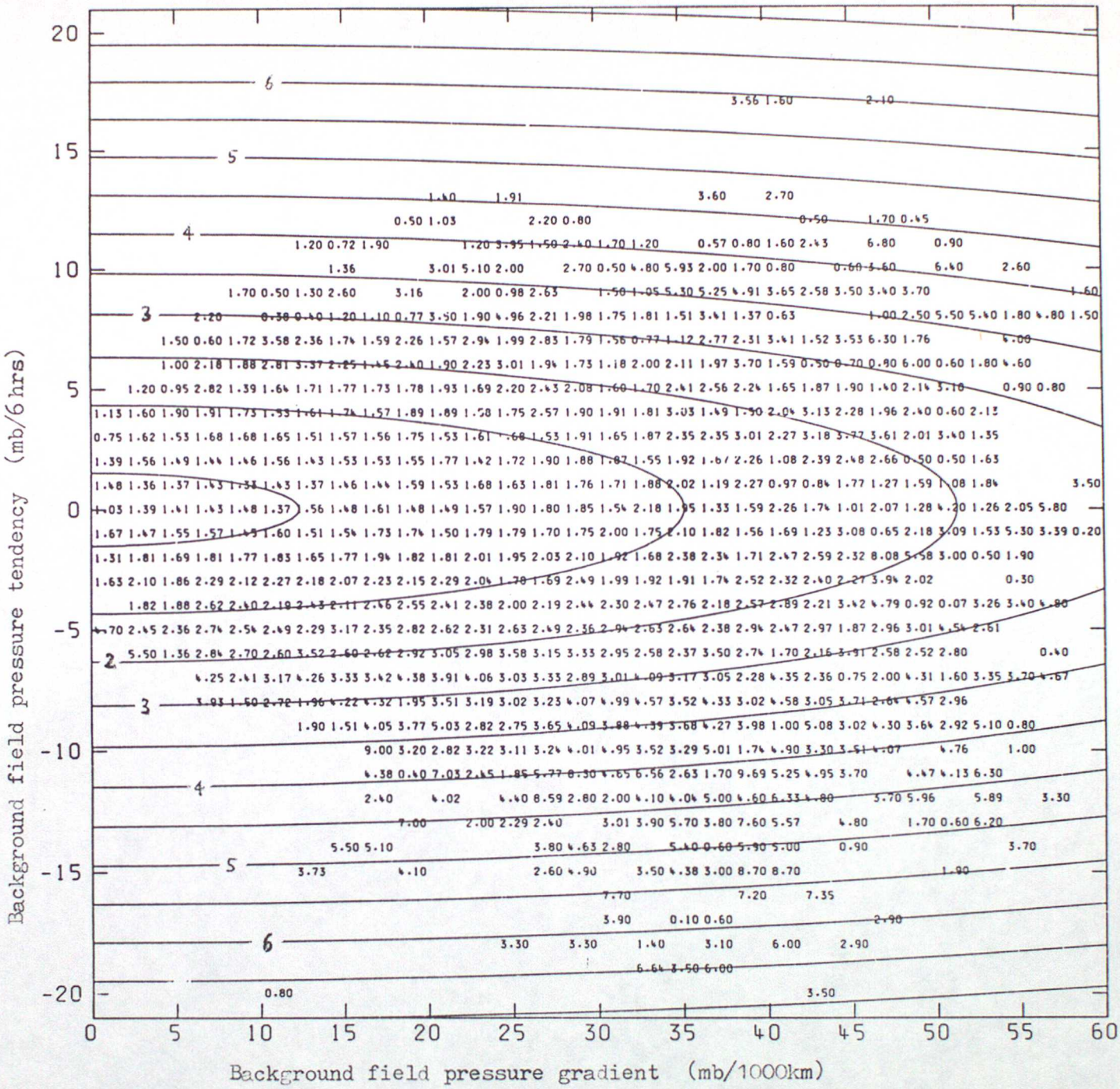


Figure 3. RMS values of observation-background pressure differences plotted on a chart of background pressure gradient (horizontal axis) against background pressure tendency (vertical axis). Numbers are 'box' values and contours represent the least squares fitted function (Eq.(3.1)) with A = 2.004 mb^2, B = 1.612×10^9 m^2 and C = 0.1057 $(6hrs)^2$. The diagram is for about 35500 observations in the 50-60N latitude band.

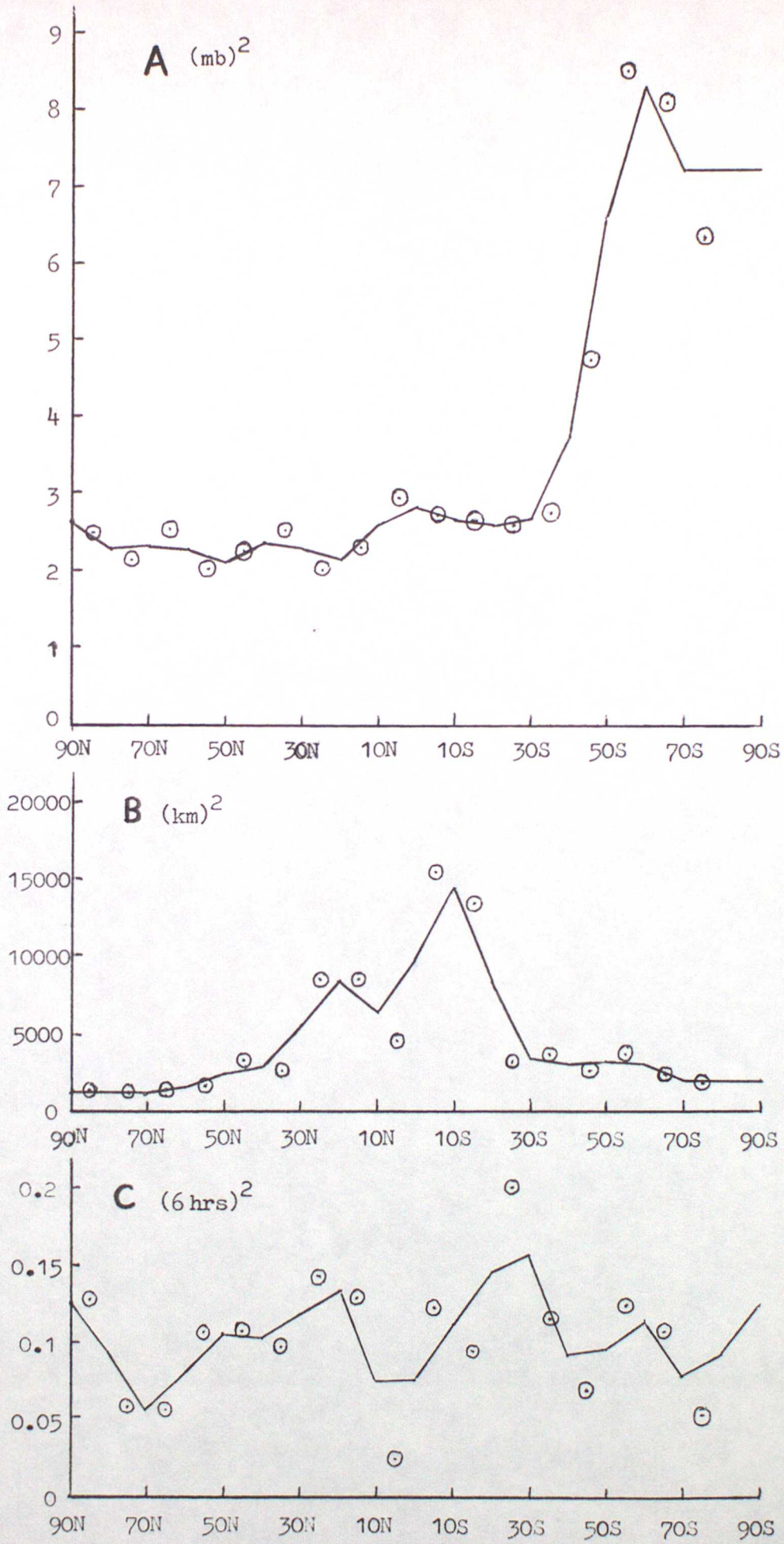


Figure 4. Latitudinal variation of coefficients A, B and C in Eq.(3.1). Symbols represent 10° latitude band values; full lines represent the function used in the quality control program.

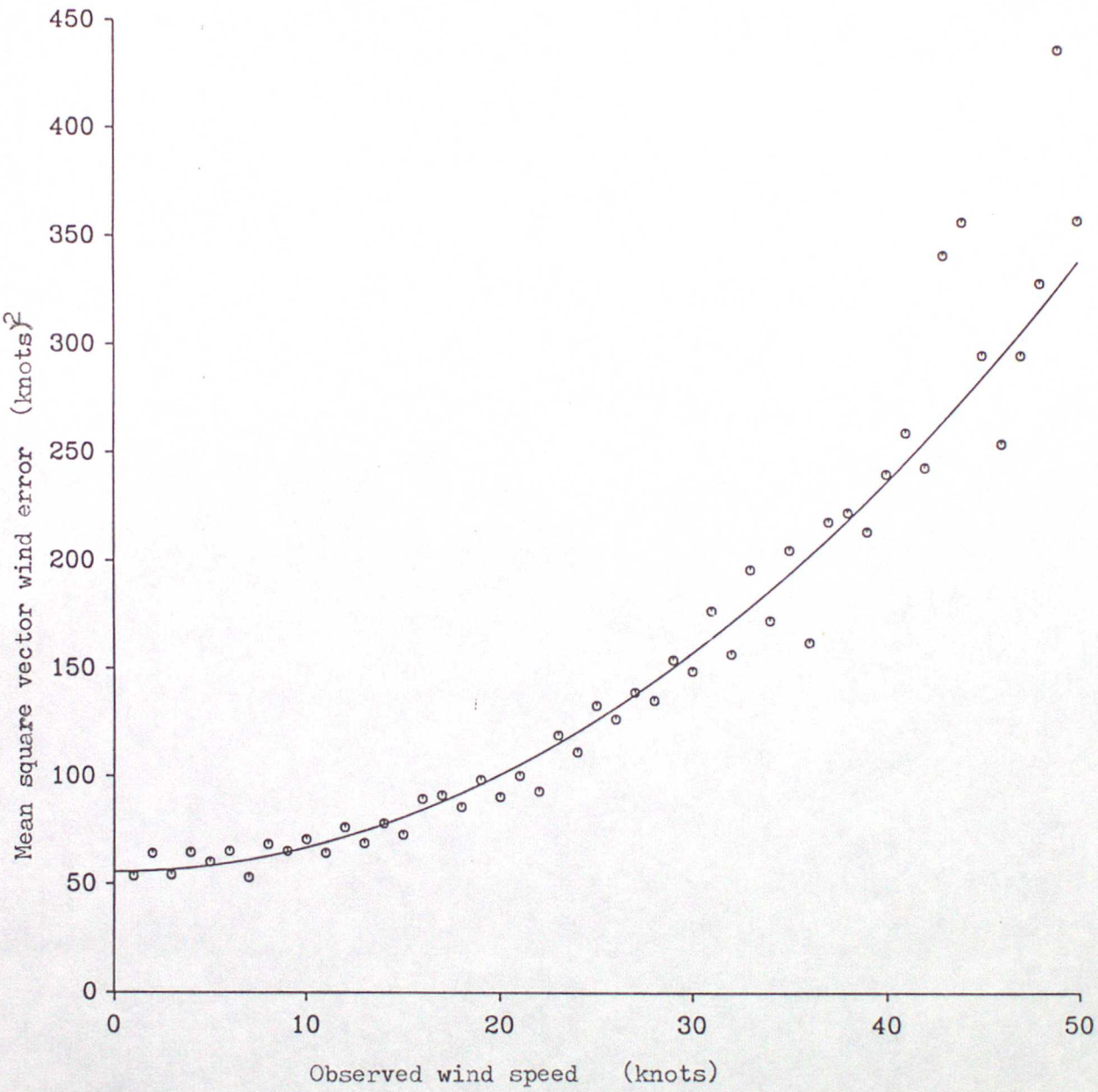


Figure 5. Mean square vector wind differences between observations and background plotted against observed wind speed. Symbols represent 'box' values for about 50000 observations in the 50-70N latitude band. The full line is Eq.(3.4) fitted to the data using a least squares technique. For this curve, $D = 55.5\ kt^2$ and $E = 0.113$.

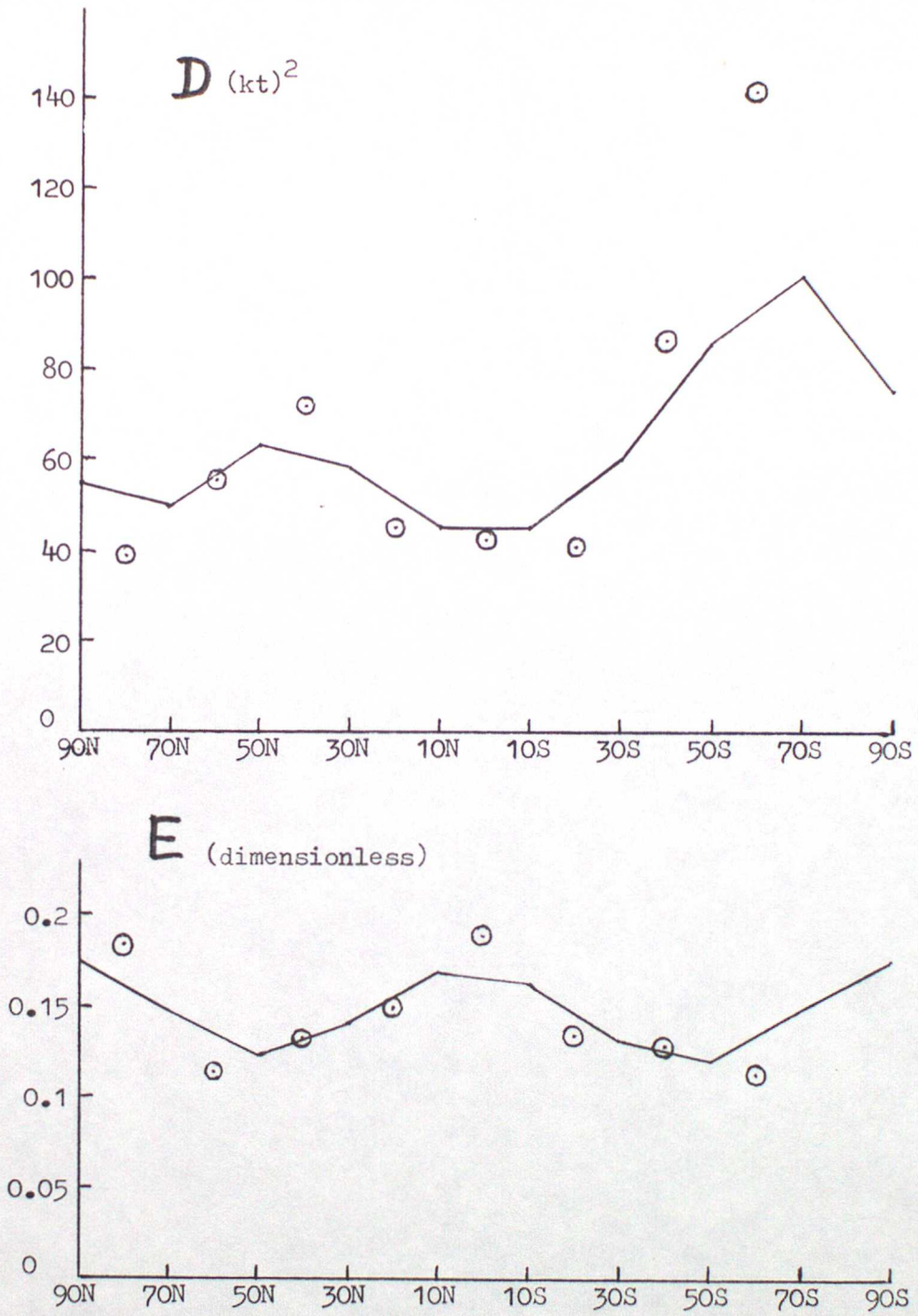


Figure 6. Latitudinal variation of coefficients D and E in Eq.(3.5). Symbols represent 20° latitude band values; full lines represent the function used in the quality control program.

A breakdown of all surface marine observations of Pmsl and wind for 00Z 15th Dec. 1986 that were quality controlled by both the current operational system and the new scheme :

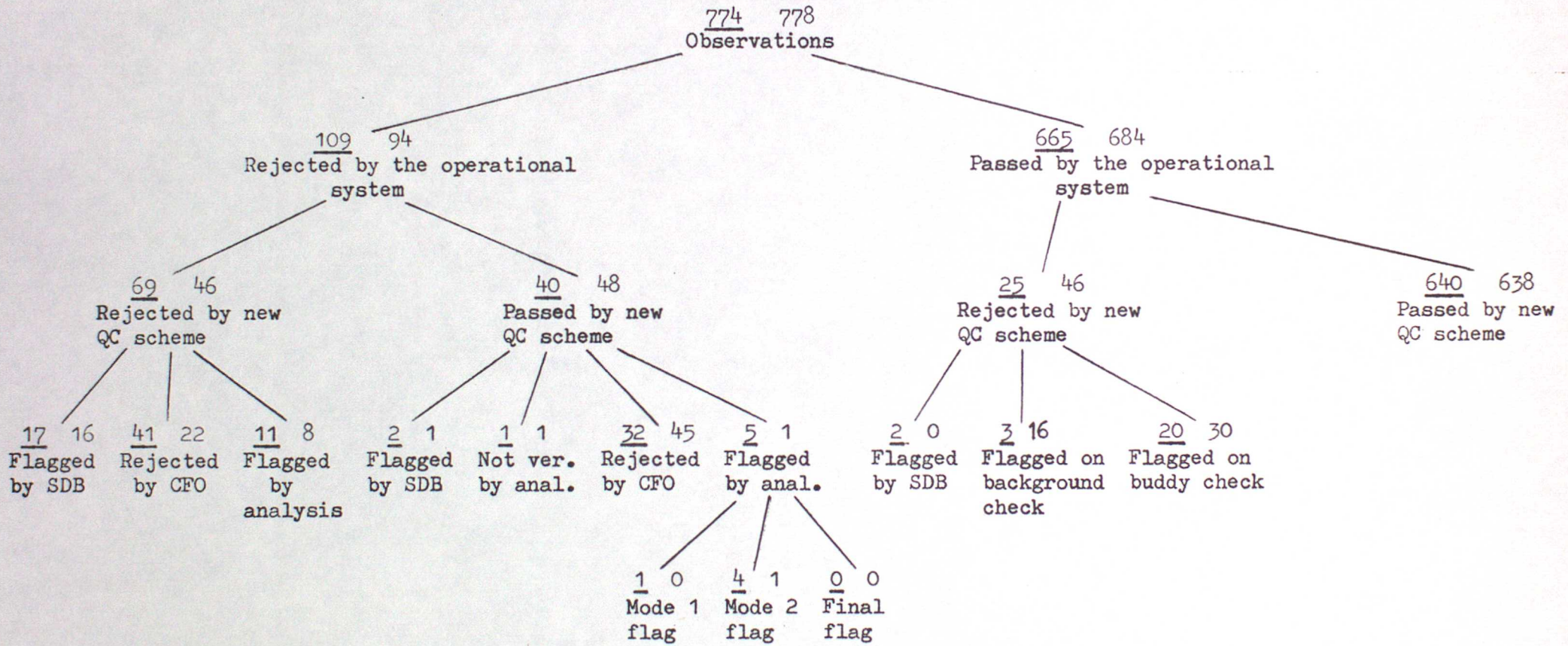


Figure 7(a).

A breakdown of all surface marine observations of Pmsl and wind for 12Z 12th Feb. 1987 that were quality controlled by both the current operational system and the new scheme :

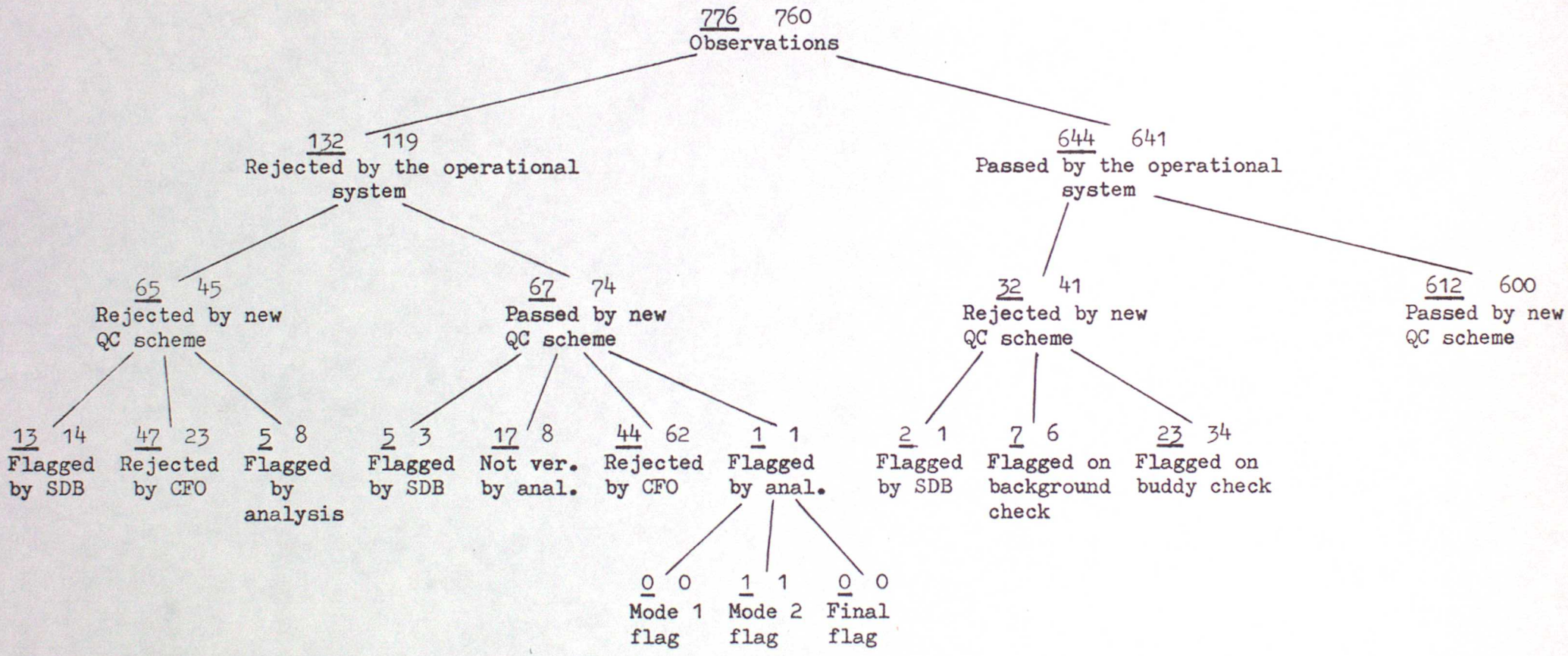


Figure 7(b).

A breakdown of all surface marine observations of Pmsl and wind for 12Z 12th June 1987 that were quality controlled by both the current operational system and the new scheme :

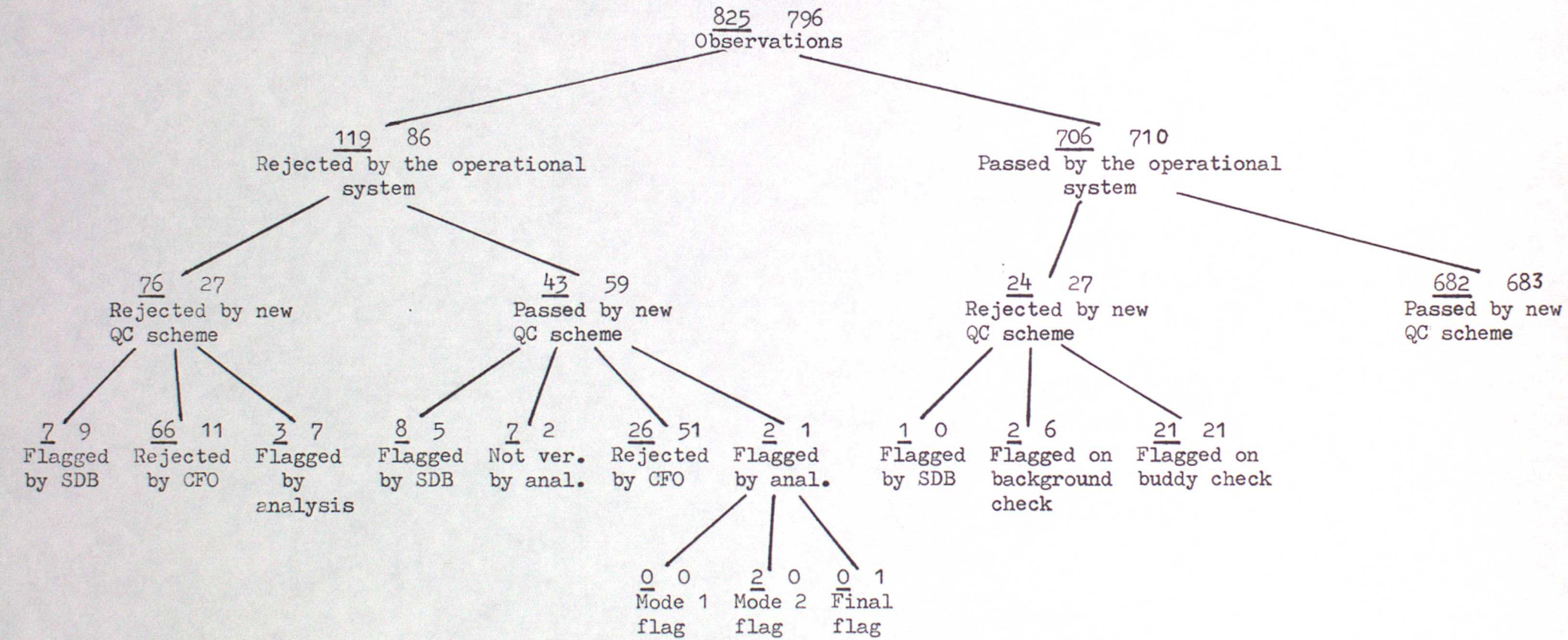


Figure 7(c).

T + 6 BACKGROUND Pmsl FIELD

and

OBSERVATIONS FOR 00Z ON 15/12/1986 LEVEL=SURFACE

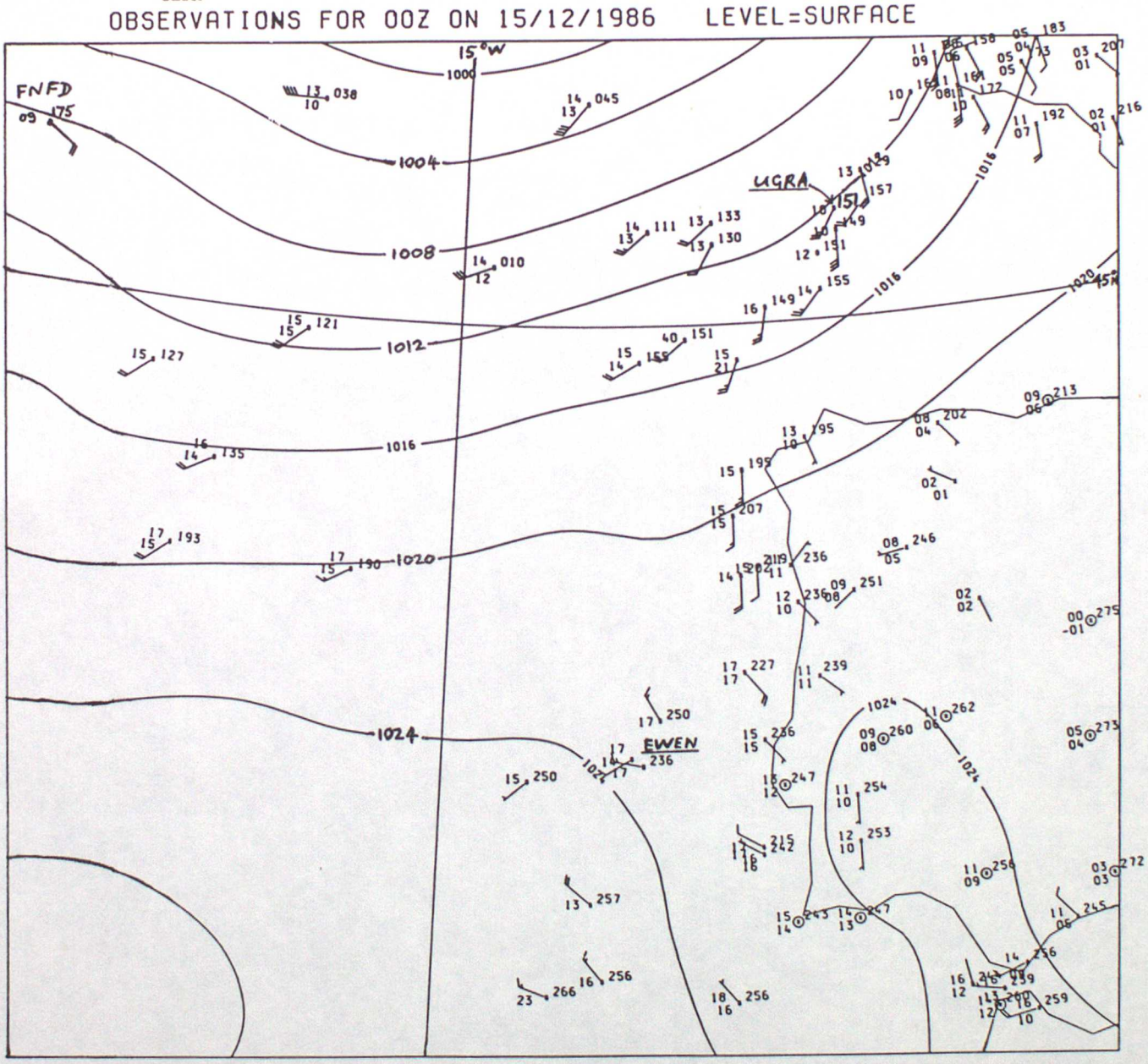


Figure 8.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G)) Pmsl	wind	No. of buddies
EWEN	Rejected by CFO	Passed (1%)	Passed (3%)	8
FNFD	Rejected by CFO	Flagged (100%)	Flagged (100%)	0
UGRA	Rejected by CFO	Passed (0%)	Passed (1%)	14

T+6 BACKGROUND Pmsl FIELD
and
OBSERVATIONS FOR 00Z ON 15/12/1986 LEVEL=SURFACE

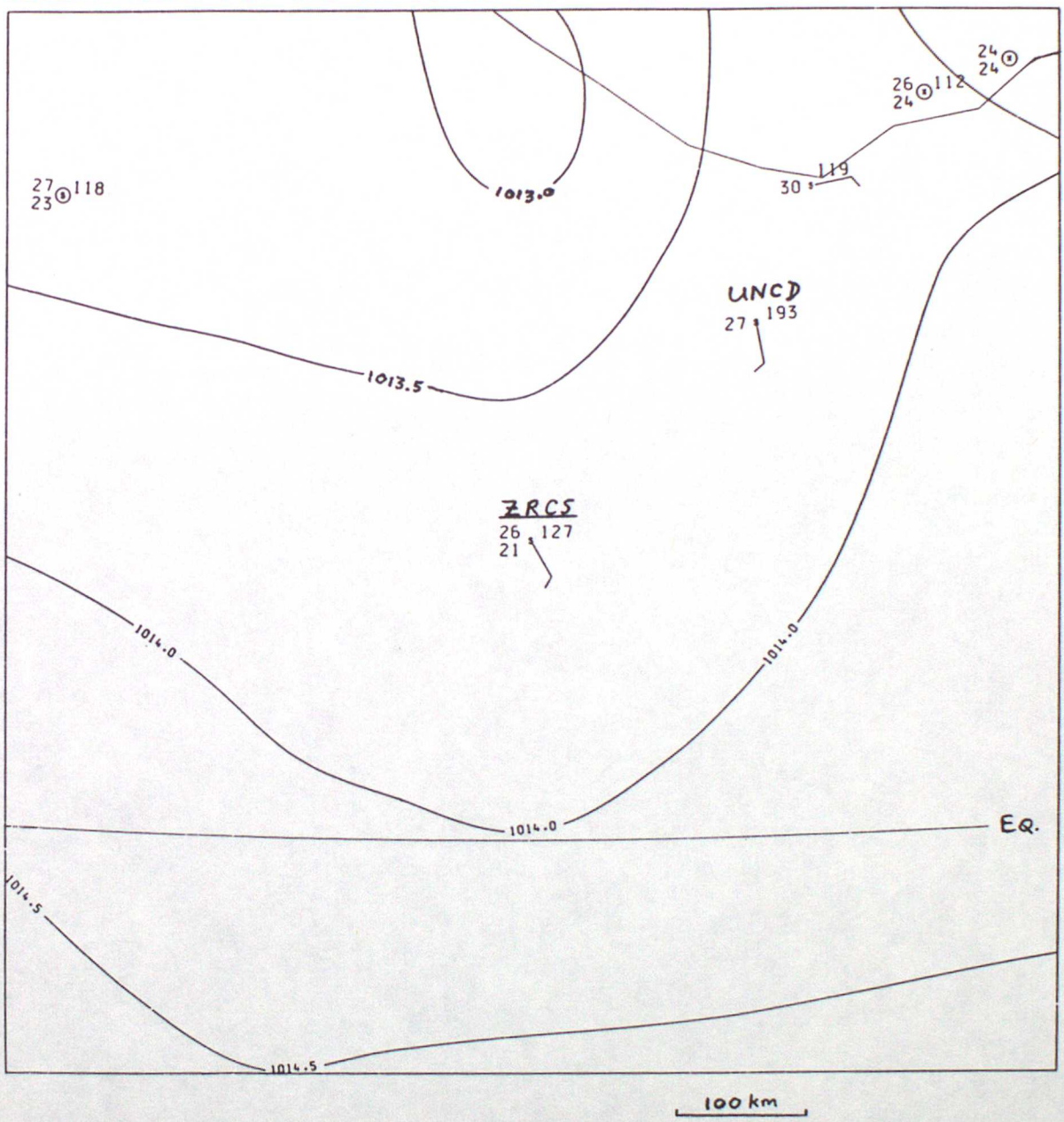


Figure 9.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G)) Pmsl	wind	No. of buddies
ZRCS	SDB position flag	Suspect (49%)	Suspect (46%)	1
UNCD	Flagged (mode 2)	Flagged (100%)	Passed (26%)	4

T + 6 BACKGROUND Pmsl FIELD
and
OBSERVATIONS FOR 00Z ON 15/12/1986 LEVEL=SURFACE

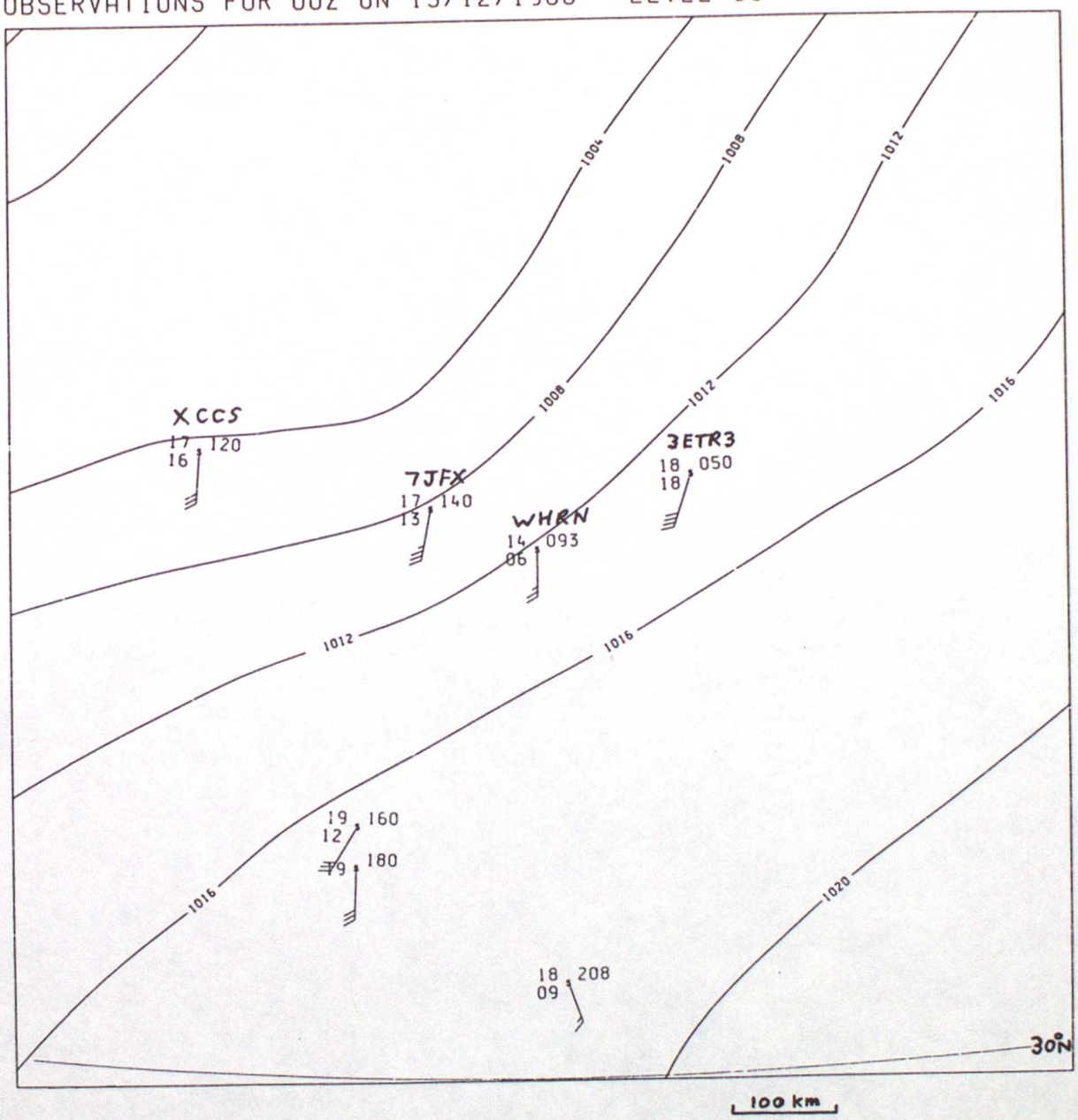


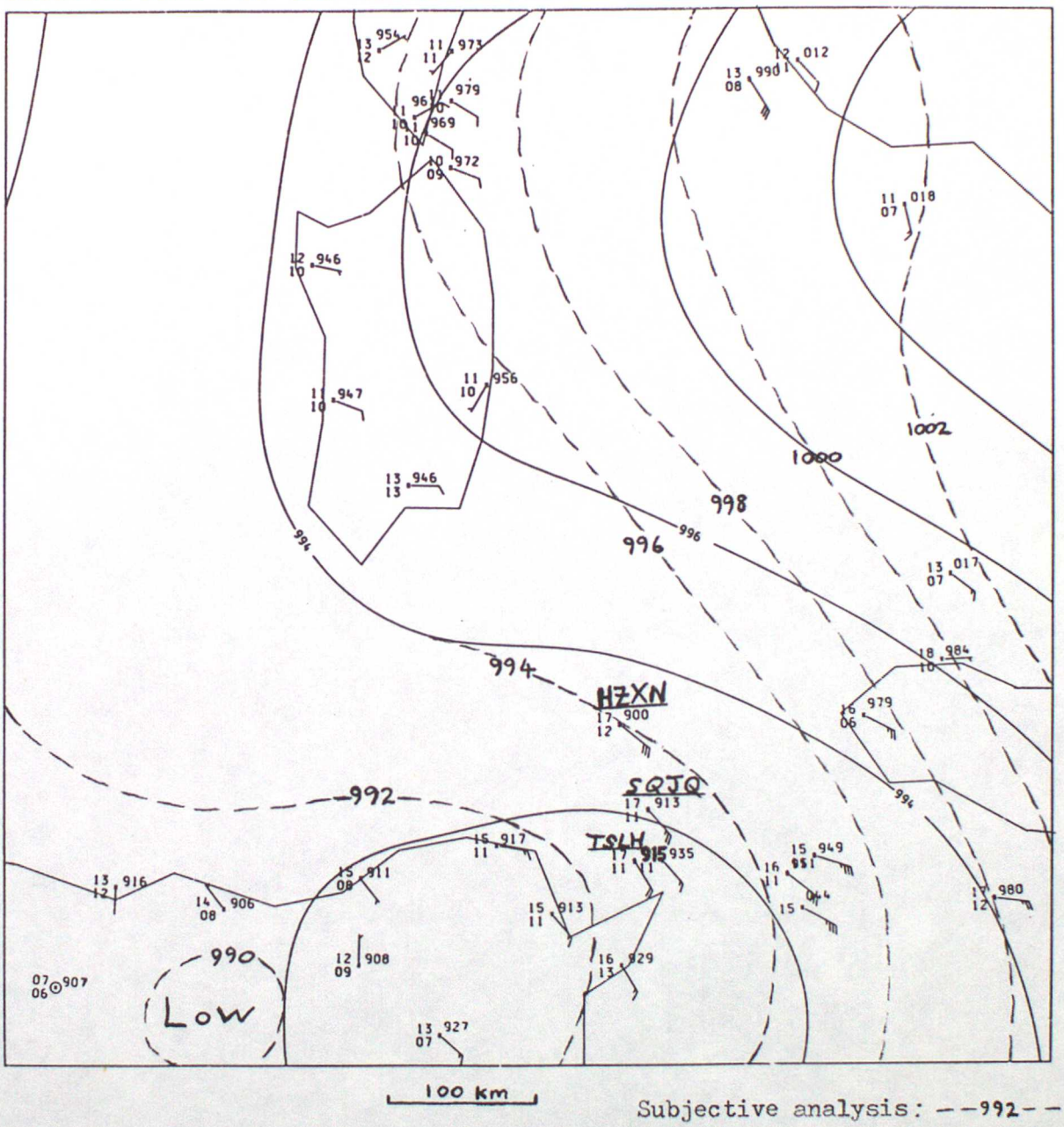
Figure 10.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G)) Pmsl	wind	No. of buddies
XCCS	Corrected by CFO	Suspect (44%)	Flagged (63%)	1
WHRN	Flagged (mode 2)	Suspect (48%)	Passed (11%)	2
3ETR3	Corrected by CFO	Flagged (86%)	Flagged (56%)	2
7JFX	Flagged (mode 2)	Flagged (58%)	Passed (8%)	3

T + 6 BACKGROUND Pmsl FIELD
and

OBSERVATIONS FOR 12Z ON 12/02/1987 LEVEL=SURFACE



Subjective analysis: --992--

Figure 11.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G)) Pmsl	wind	No. of buddies
HZXN	Passed	Flagged (100%)	Passed (2%)	17
SQJQ	Passed	Flagged (100%)	Passed (0%)	18
TSLH	Passed	Flagged (69%)	Passed (1%)	18

T + 6 BACKGROUND Pmsl FIELD
and
OBSERVATIONS FOR 00Z ON 15/12/1986 LEVEL=SURFACE

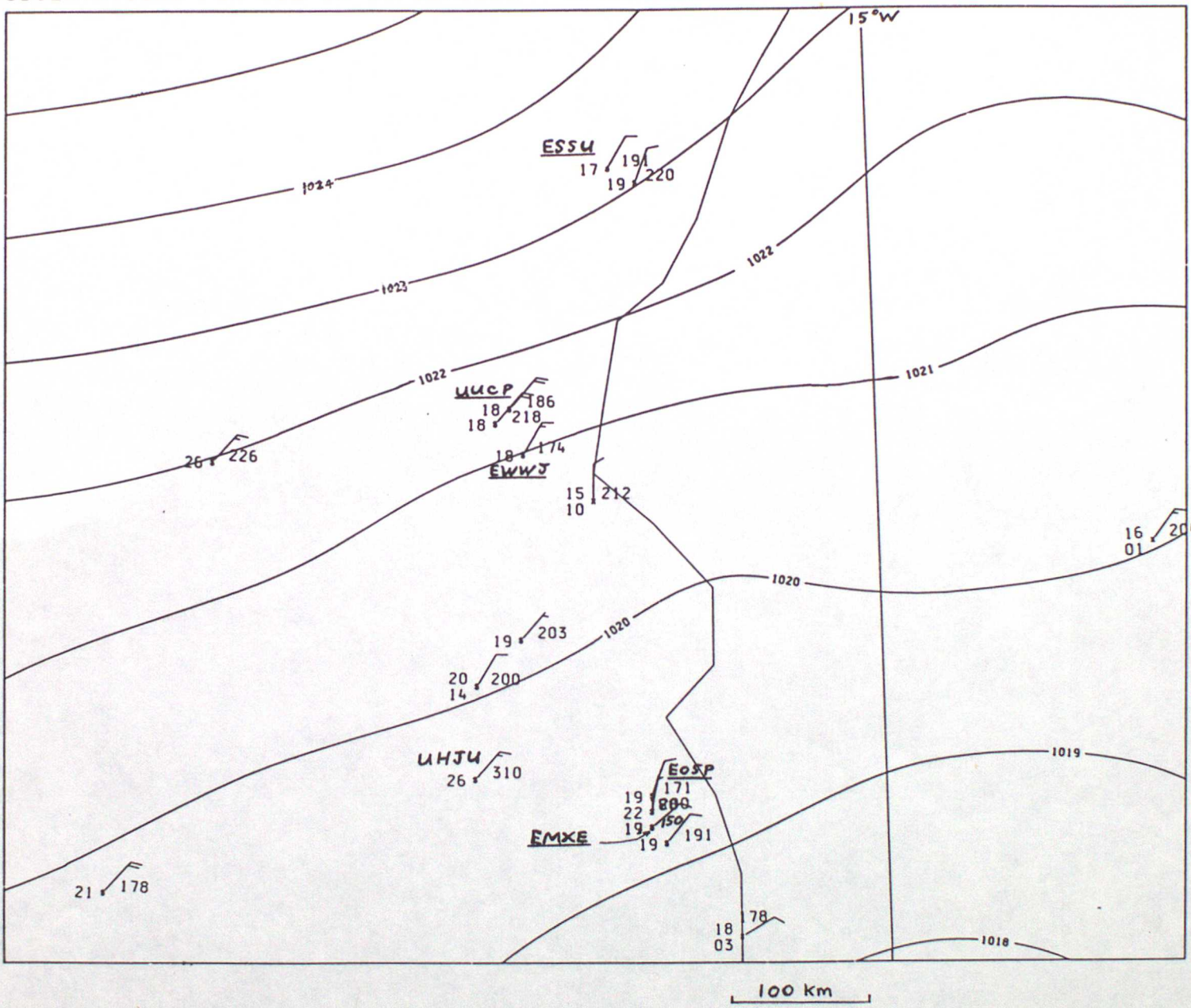


Figure 12.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G)) Pmsl	wind	No. of buddies
EMXE	Passed	Flagged (100%)	Passed (15%)	10
EOSP	Passed	Flagged (54%)	Passed (1%)	12
ESSU	Passed	Flagged (73%)	Passed (8%)	3
EWWJ	Passed	Flagged (98%)	Passed (1%)	12
UUCP	Passed	Flagged (73%)	Passed (3%)	8
UHJU	Pmsl corrected by CFO	Flagged (100%)	Suspect (34%)	13

T + 6 BACKGROUND Pmsl FIELD
and

OBSERVATIONS FOR 00Z ON 15/12/1986 LEVEL=SURFACE

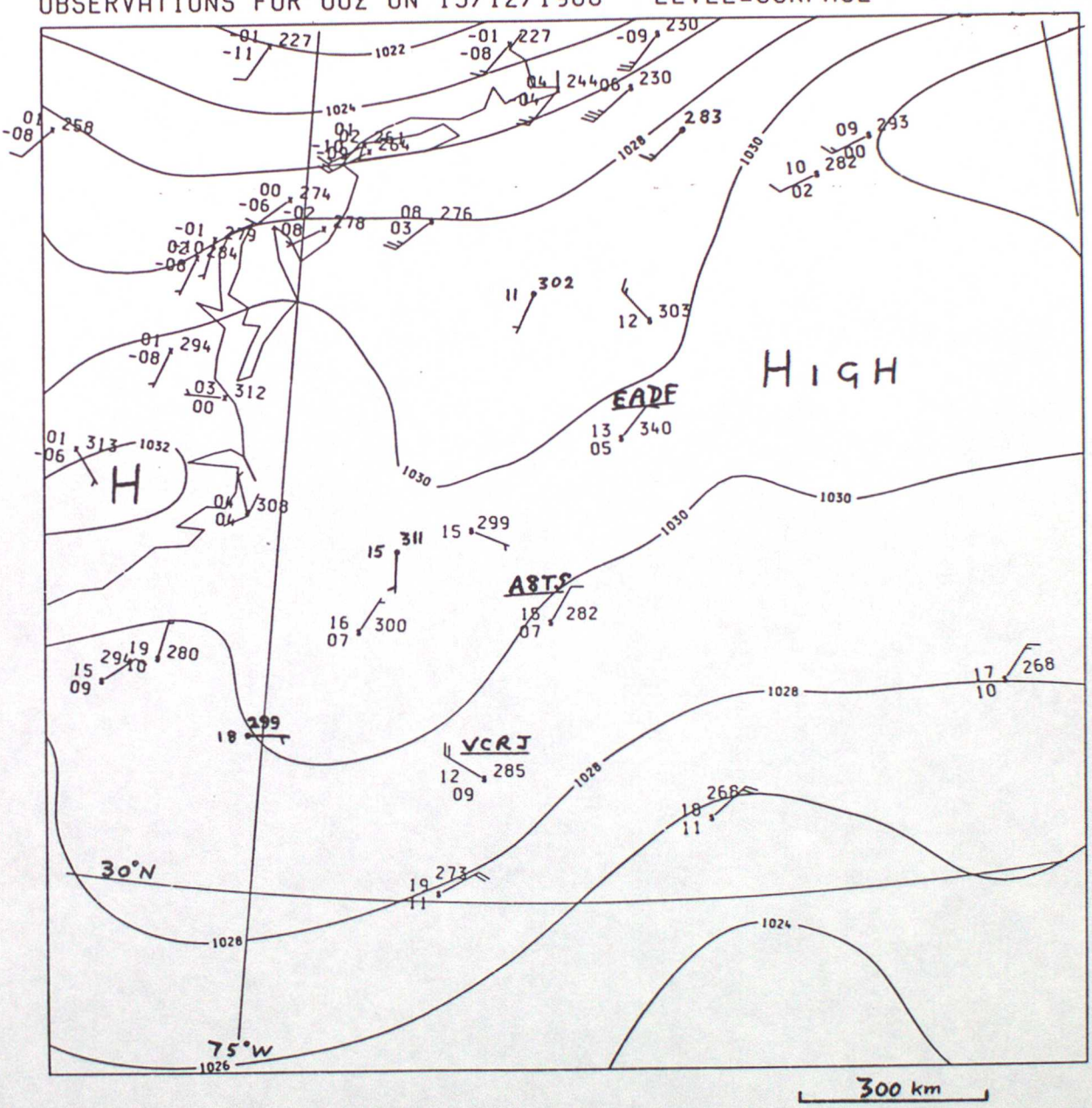


Figure 13.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G)) Pmsl	wind	No. of buddies
EADF	Passed	Flagged (64%)	Passed (12%)	3
A8TS	Passed	Flagged (50%)	Passed (5%)	3
VCRJ	Passed	Passed (11%)	Flagged (90%)	2

T + 6 BACKGROUND Pmsl FIELD
and
OBSERVATIONS FOR 12Z ON 12/02/1987 LEVEL=SURFACE

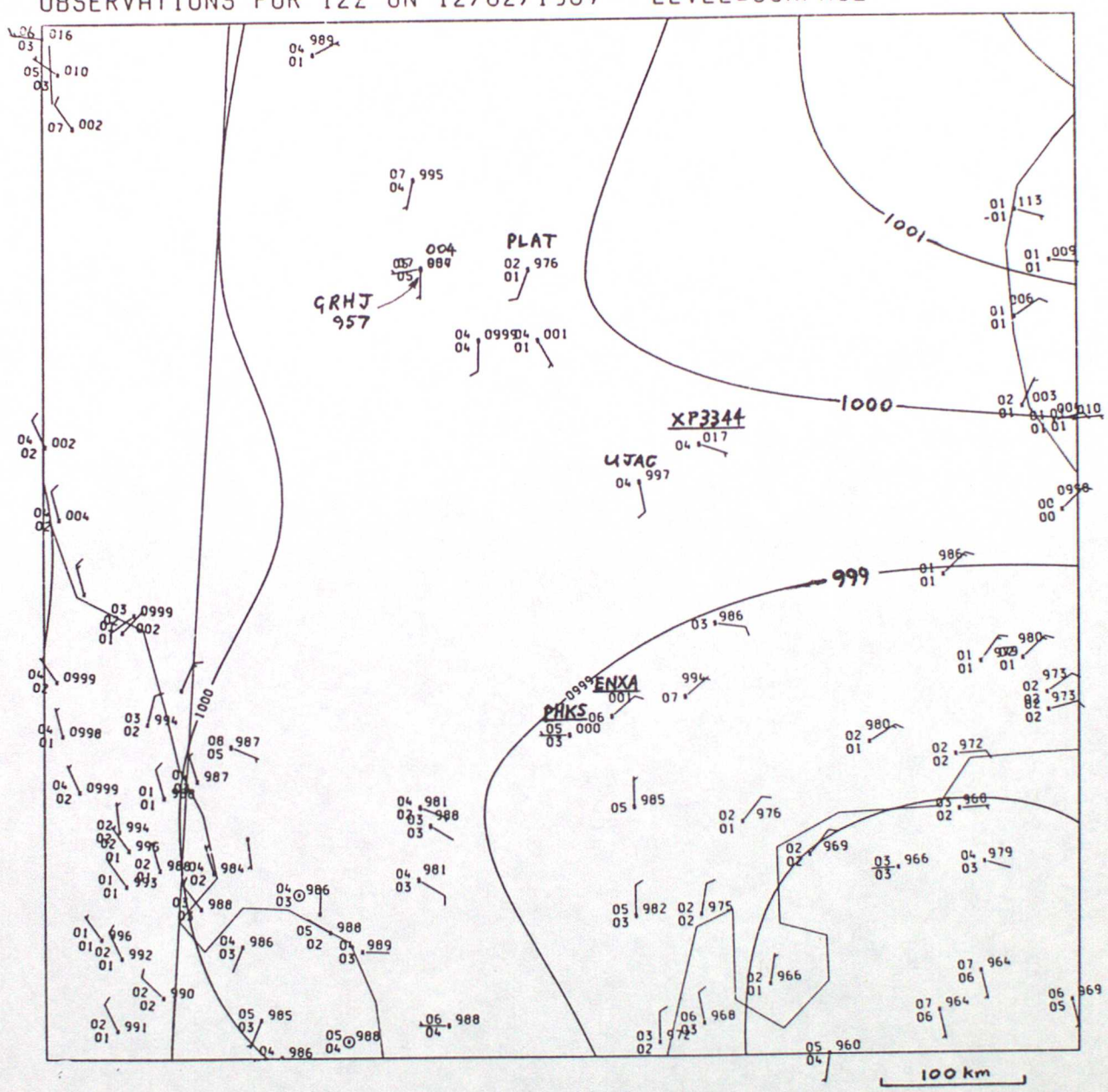
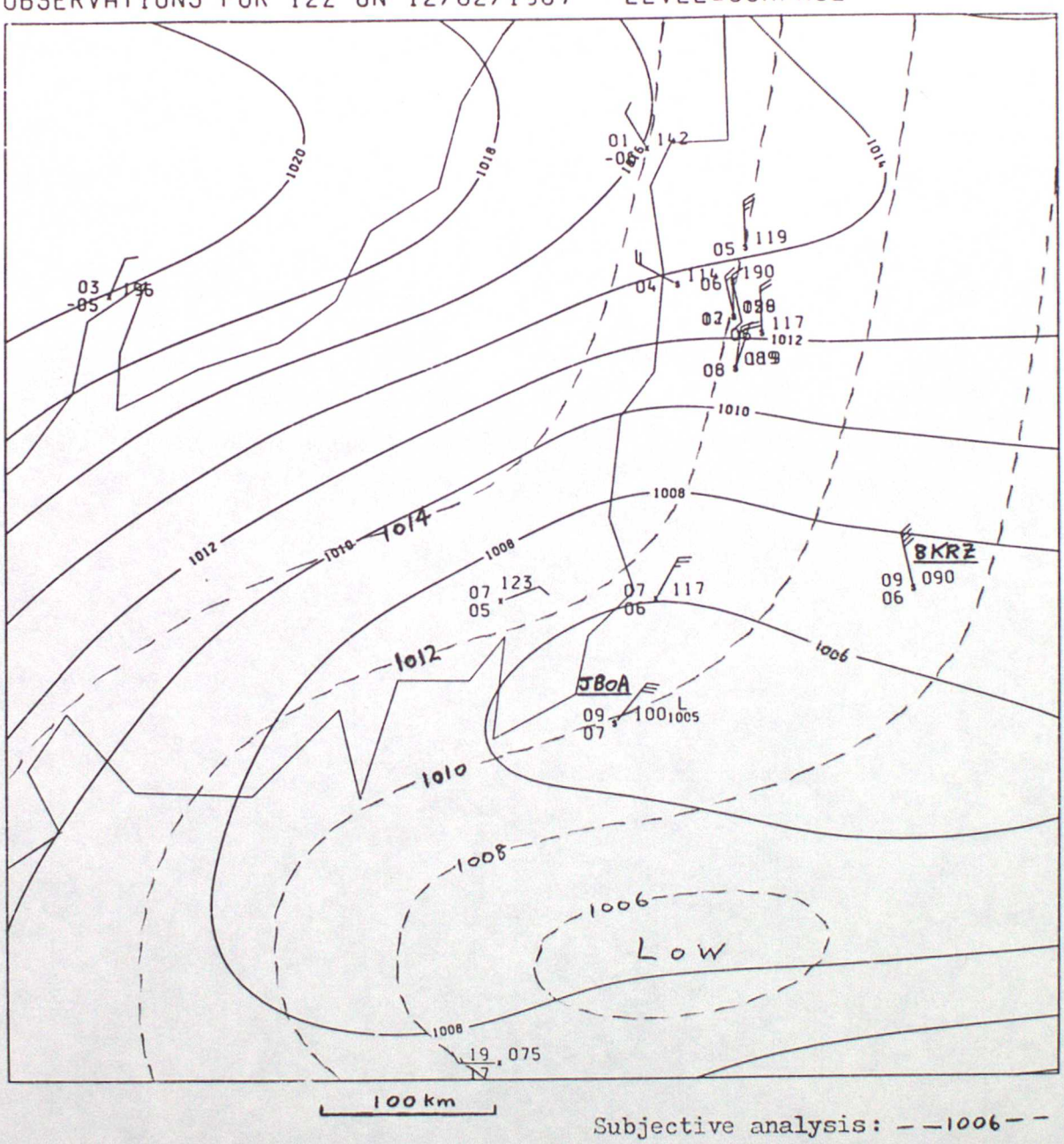


Figure 14.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G)) Pmsl	wind	No. of buddies
ENXA	Passed	Flagged (76%)	Passed (2%)	23
GRHJ	Rejected by CFO	Flagged (100%)	Passed (1%)	15
PHKS	Passed	Flagged (73%)	Passed (4%)	23
PLAT	Passed	Suspect (44%)	Passed (0%)	14
UJAC	Passed	Passed (0%)	Flagged (65%)	20
XP3344	Passed	Flagged (53%)	Passed (3%)	20

T + 6 BACKGROUND Pmsl FIELD
and
OBSERVATIONS FOR 12Z ON 12/02/1987 LEVEL=SURFACE



Subjective analysis: --1006--

Figure 15.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G)) Pmsl	wind	No. of buddies
JBOA	Passed	Flagged (100%)	Flagged (56%)	12
8KRZ	Passed	Flagged (62%)	Flagged (91%)	11

T + 6 BACKGROUND Pmsl FIELD
and
OBSERVATIONS FOR 00Z ON 15/12/1986 LEVEL=SURFACE

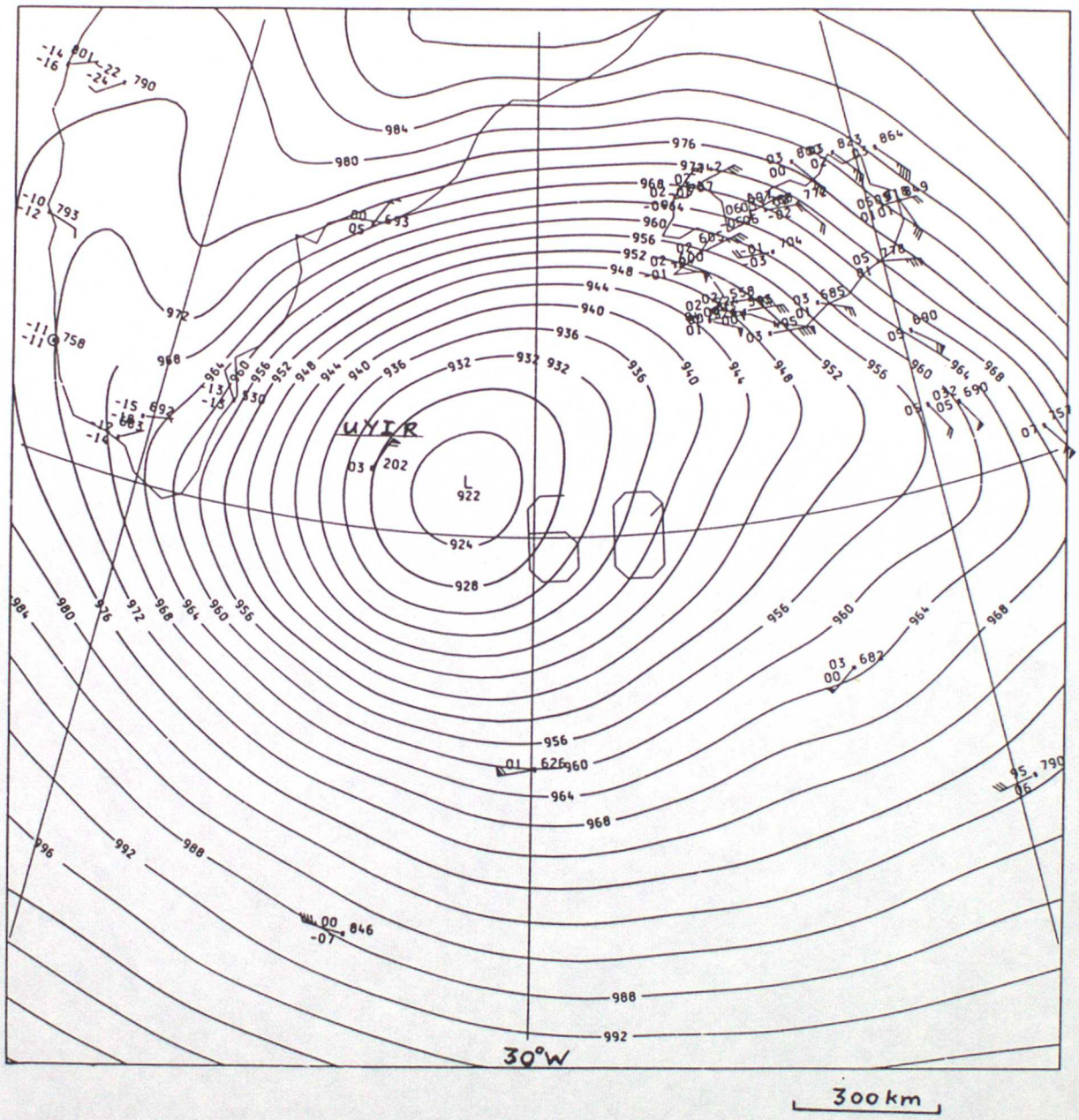


Figure 16.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G))		No. of buddies
		Pmsl	wind	
UYIR	Passed (mode 1 flag)	Flagged (97%)	Passed (18%)	0

T + 6 BACKGROUND Pmsl FIELD
and
OBSERVATIONS FOR 12Z ON 12/06/1987 LEVEL=SURFACE

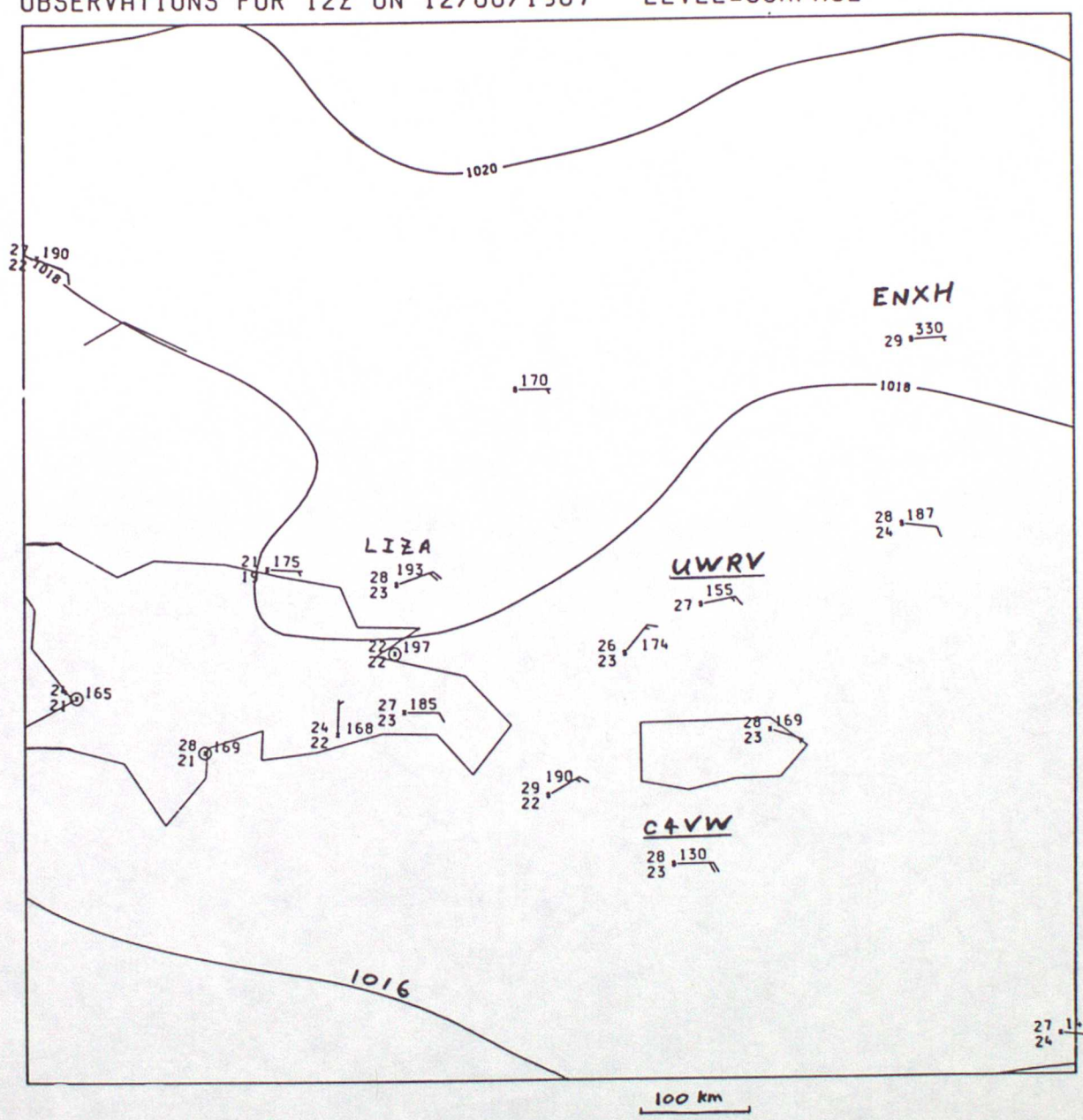


Figure 17.

QUALITY CONTROL RESULTS :

Call sign	Operational system	New Q.C. scheme (final P(G)) Pmsl	wind	No. of buddies
C4VW	Passed	Flagged (98%)	Passed (8%)	6
ENXH	Pmsl flagged, v passed	Flagged (100%)	Suspect (36%)	1
LIZA	Passed	Passed (7%)	Suspect (31%)	9
UWRV	Passed	Flagged (75%)	Passed (6%)	8