TECHNICAL MEMORANDUM

82

A method of correcting tidal biases in the data assimilation system

P. Undén

Research Department

December 1983

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

A B S T R A C T

The synoptic observations contain significant tidal components which are related to the time of day. Tidal effects are dominant in the tropics and since the operational forecast model does not reproduce atmospheric tides, the first guess fields in the data assimilation cycle will always be out of phase with the tropical observations. There is thus a systematic tidal bias in the analysis. Although the tidal pressure waves are largely barotropic in structure, they are analysed baroclinicly due to the vertical structure functions used in the analysis.

This note describes how the bias can be corrected by adding the mean analysis increment to the first guess field.

Complications arise when data from a high level station is used in the analysis of underlying levels. Therefore an attempt has been made at trying to enforce locally barotropic analysis increments by introducing bogus data below the station level.

1. INTRODUCTION

It has been known for a long time that the ECMWF forecasts suffer from excessive heating above the East African boundary layer during the first day of integration. The heating is produced by convection which often occurs in the form of "grid point storms", especially over central eastern Africa. Locally, through the effect of precipitation, this changes the climate of the model by increasing the soil moisture. Globally this physical forcing also projects on to travelling gravity and Rossby waves (Heckley, 1983).

One possible cause of the very low stability in the tropical boundary layer over Africa is the presence of strong atmospheric tides in the tropics and the way they are analysed. In Sect. 2 tidal components are introduced using the notation of optimum interpolation and the problem of analysing them in the vertical is discussed. Because these largely barotropic modes can be analysed baroclinicly when using baroclinic vertical forecast error correlation functions (sometimes referred to as just structure functions), the tidal components can be removed from the optimum interpolation procedure by adding them on to the first guess as Sect. 3 describes. Unfortunately, the tidal bias correction produces some unwanted effects when high level data are used. Tidal components cause a reverse baroclinicity below the station level and the procedure will counteract that. As a remedy for this, bogus data can be introduced at a lower level and an assimilation with such data is discussed in Sect. 4.

2. TIDAL BIASES IN THE OBSERVATIONS AND THE EFFECT ON THE DATA ASSIMILATION

The synoptic observations contain, apart from the normal "synoptic" tendencies with periods of several days, a tidal part which is related to the time of day. The tidal pressure waves are dominant in the tropics with the most significant component being the semi-diurnal tide (with a period of 12h). It has an amplitude of up to 1.4 mb near the equator (Chapman and Lindzen, 1969).

The presence of these tidal waves in the ECMWF data assimilation system was pointed out by Hollingsworth and Arpe (1982). They become apparent when examining the mean analysis increment (analysis minus first guess) for a certain analysis hour. Another example of this is shown in Fig. 1. The mean analysis increment has been calculated from mean analyses and mean first guess forecasts on σ-levels which were interpolated to pressure levels. However, some caution should be taken when interpreting the 1000 mb increments in areas of high topography. In order to filter out tendencies which are not of a semi-diurnal nature (e.g. systematic observation errors and interpolation errors) it is necessary to subtract the daily mean analysis increment from the calculated increments.

Because of the dominant semi-diurnal tide, a clear wavenumber 2 pattern appears with large negative increments over East Africa and Southern Asia at 12Z. Correspondingly large positive increments occur over Australia-New Guinea and somewhat smaller ones over the Atlantic and South America. Although the semi-diurnal tide is dominant, it is modulated by the diurnal tide which has about half the amplitude of the semi-diurnal one. This explains the higher magnitudes over East Africa and South Asia and lower magnitudes in the western hemisphere. However, over the Pacific there is also a data problem with too few regular observations.

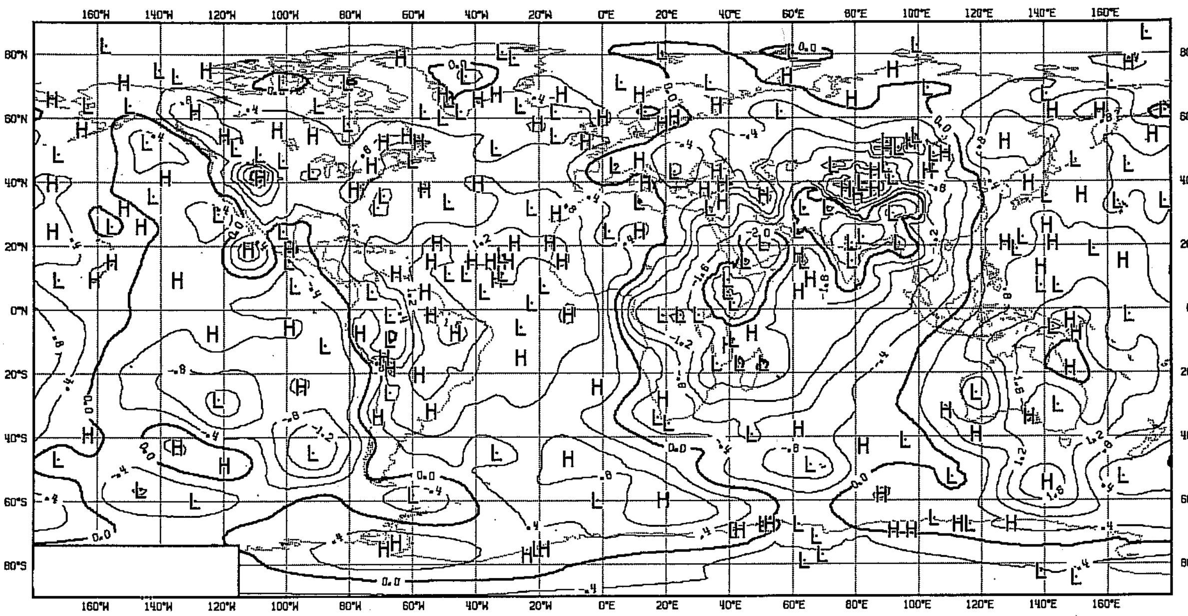

Fig. 1 Mean height analysis increments at 1000 mb for 12Z during April 1982 in Dm. The daily mean has been subtracted. Isolines for every 4m.

The amplitudes in Fig. 1 are still considerably larger than can be explained by these two tidal components. The reason for this is the way the tides are treated in the ECMWF normal mode initialisation and forecast model.

Let b and t represent the synoptic and tidal components of the observations respectively. Then the analysis of observations (b + t) can be described as

$$A(b + t) = f + O(b + t - f) \qquad (1)$$

Here f is a first guess (a 6 hour forecast from the previous analysis) and O is the linear optimum interpolation operator (a sum of weights times observational increments, see Lorenc, 1981).

Since O is a linear operator eq.(1) can be written as

$$A(b + t) = f + O(b - f) + O(t) \qquad (2)$$

This is the appropriate form for the tropics because the tidal part dominates. The ECMWF has no diurnal cycle in its radiation scheme at present so the forecast model cannot reproduce the tide (even when a diurnal cycle is included it is not clear if the tidal waves can be simulated correctly).

However, the first guess (f) will contain the Rossby mode projection of the tidal waves from the analysis 6 hours earlier.

$$f = f(b_-) + R(t_-) \qquad (3)$$

where $f(b_-)$ is the "synoptic" evolution from data 6 hours earlier (b_-) and $R(t_-)$ is the forecast of the previous tide projected on to slowly moving modes. Then eq.(2) can be rewritten as

$$A(b + t) = f(b_-) + R(t_-) + O[b-f(b_-)] + O[t-R(t_-)] \qquad (4)$$

The analysis increments consist of the last two terms in eq.(4). In the mean only $O[t-R(t_-)]$ will give a systematic contribution for a certain analysis hour. Thus the mean increments in Fig. 1 contain two components, t and $R(t_-)$ which is proportional to t_-. Since t_- is largely of the same magnitude but of the opposite sign as t (because of the dominance of the semi-diurnal tide which changes sign every 6th hour), the amplitudes in Fig. 1 can be explained.

In the vertical these mean increments first decrease in amplitude, then change sign; at higher levels they have a rather erratic structure. This sign change occurs in the tropics between 700 and 500 mb. In areas where the analysis only uses surface observation (e.g. New Guinea) this is a result of the vertical forecast error correlation function which indeed has this sign change between 1000 and 500 mb. In other areas where TEMP's are available (e.g. East Africa south of the equator) the vertical structure at 12Z and 00Z is mainly a result of the actual thermal diurnal cycle. In northern East Africa the TEMP's are very sparse and irregular. However, the existence of both sea level pressure reports and high altitude SYNOP's reporting at station level or at the nearest standard pressure level, introduces some information about the thermal stratification below the station into the analysis anyway. The vertical structure of the increments above is still determined by the analysis structure function. Since the tide itself is almost barotropic in the troposphere, it means that the analysis erroneously destabilises the boundary layer over East Africa at 12Z by using a highly baroclinic vertical forecast error correlation function.

The field shown in Fig. 1 at 1000 mb is a best estimate of a barotropic mean tidal increment $\overline{O[t-R(t_-)]}$ and will be used in the next section.

3. BIAS CORRECTIONS FOR THE TIDES

3.1 Removal of tidal biases from the analysis

In order to reduce the tidal increments in eq.(4), one could add the mean analysis increments for the analysis hour in question to the first guess and largely eliminate the negative thermal effects of the tides. Having the possible problems of initializing the tides in mind, Hollingsworth and Arpe suggested that the tidal fields should also be removed after the analysis. However, since the mean increments also contain a rest component of the previous tide (eq.(4)), this procedure would cause the system to lose "memory" of previous tides and the mean increments would not be valid anymore. This is described below.

The mean 1000 mb height increment, $\overline{t-R(t_-)}$, is added to the first guess at all levels only when forming the observational increments.

$$A(b) = f(b_-) + R(t_-) + O\left[b + t - \left(f(b_-) + R(t_-) + \overline{t - R(t_-)}\right)\right] \qquad (6)$$

Within the square brackets $t+R(t_-)$ is virtually cancelled by the mean increment $\overline{t-R(t_-)}$ so we are left with

$$A(b) = f(b_-) + R(t_-) + O\left[b - f(b_-)\right] \tag{7}$$

Instead of retaining the tide itself we carry the previous tide forward in the assimilation so that after initialization the next first guess forecast is

$$f(b) = f(b_-) + R(t_-) + S \tag{8}$$

where S is a small "synoptic" evolution given by observations b. These results will be compared with a baseline assimilation without the bias removal. That analysis and subsequent first guess forecast are

$$A^{o}(b + t) = f(b_-) + R(t_-) + O\left[b + t - f(b_-) + R(t_-)\right] \tag{9}$$

$$f^{o}(b) = f(b_-) + R(t) + S \tag{10}$$

Fig. 2 shows the difference between an analysis at 12Z 1 May 1982 with the bias removed (eq.(7)) and a baseline analysis (eq.(9)). This difference is expressed in eq.(11)) and is just the mean increment with opposite sign; this agrees quite well with Fig. 2a (compare with Fig. 1). Then taking the difference between the initialized analyses gives the Rossby mode projection of the tidal bias. As shown in Fig. 2b the pattern is the same but the amplitude seems to be damped by about a half. In order to see how the forecast model treats the tidal modes, the difference between the two subsequent 6 hour forecasts is shown in Fig. 2c. This is still the Rossby mode projection of the tidal bias as can be seen from eq.(12), which is the difference between eq.(8) and eq.(10).

$$A(b) - A^{o}(b + t) = -O\left[t - R(t_{-})\right] \qquad (11)$$

$$f(b) - f^{o}(b) = -\left[R(t) - R(t_{-})\right] \qquad (12)$$

Comparison of Fig. 2b and Fig. 2c shows that these modes appear to be stationary. That means that tidal components in the data assimilation remain stationary once they are introduced. (The term $R(t_{-})$ in eq.(8) and eq.(12) should really be $f(R(t_{-}))$ if $R(t_{-})$ were not stationary).

Carrying on the assimilation (from eq.(8) and (10)), the next pair of analyses increments would be

$$A(b_{+}) - f(b) = O\left[b_{+} + t_{+} - (f(b_{-}) + R(t_{-}) + S + \overline{t_{+} - R(t)})\right] \qquad (13)$$

$$A^{o}(b_{+} + t_{-}) - f^{o}(b) = O\left[b_{+} + t_{+} - (f(b_{-}) + R(t) + S)\right] \qquad (14)$$

The difference between the above increment is approximately $\left[-t_{+} + 2R(t) - R(t_{-})\right]$, and since the dominating semi-diurnal component changes sign every 6 hours this gives an amplitude of about twice the tide at the actual hour and an opposite sign. This is shown in Fig. 3 for a later pair of 12Z analyses. Therefore the outlined procedure replaces the tidal bias with one of larger magnitude and of opposite sign. Thus this procedure in no way solves the problem of the vertical structure, but it does give an insight into how the tidal fields are initialized and integrated by the forecast model.

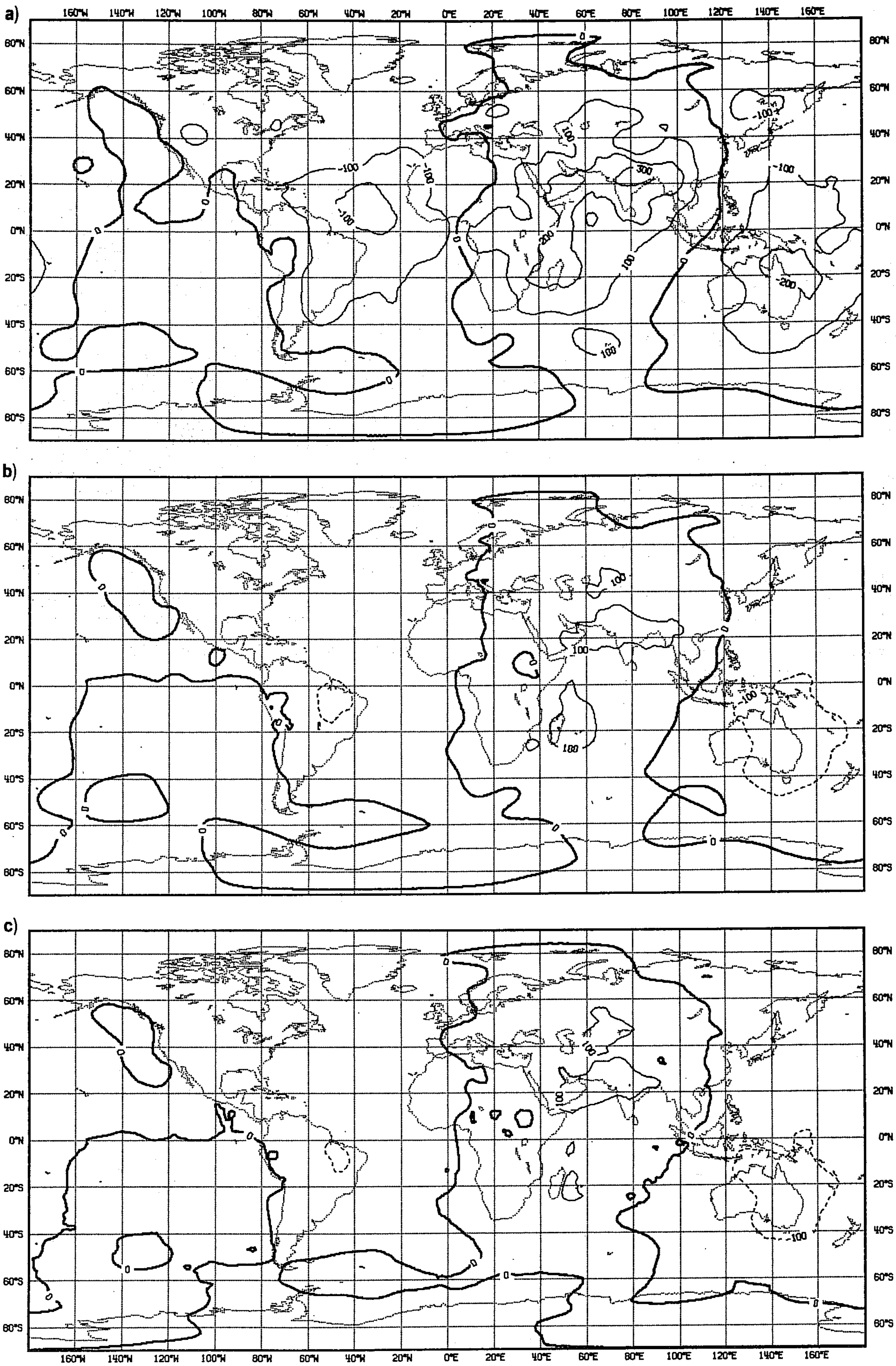

Fig. 2 Surface pressure difference for a) analyses, b) initialisations and c) 6 hour forecasts. Unit is Pa.

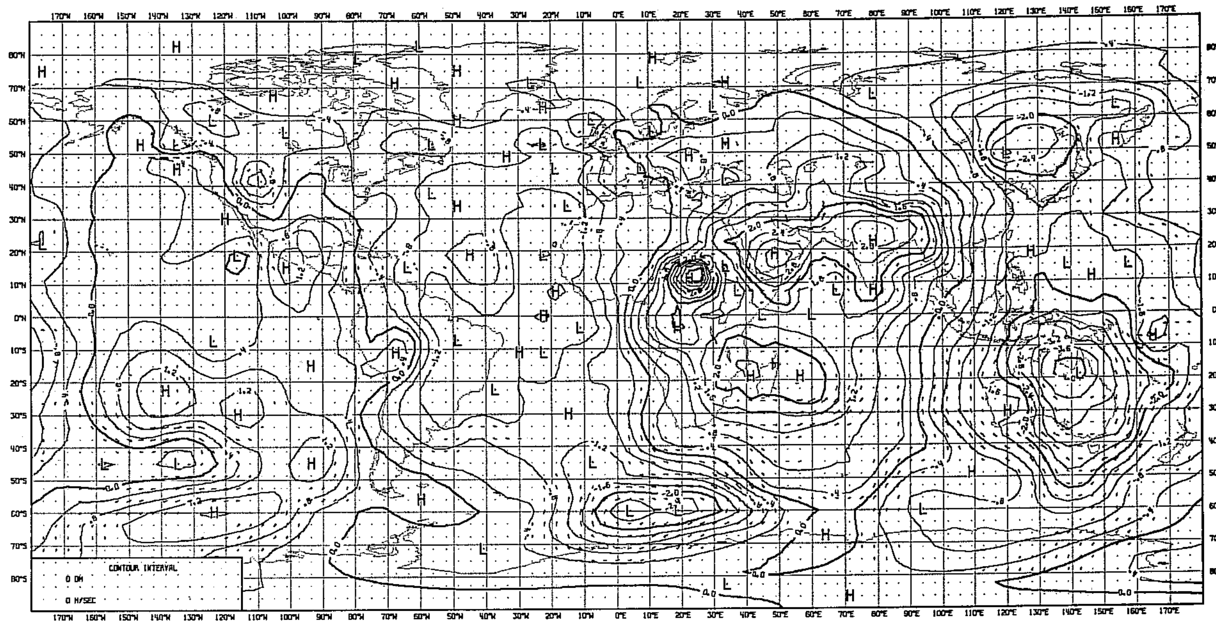

Fig. 3 Difference between analysis increments at 1000 mb for 12Z, 3 May 1982 (Dm).

3.2 Bias convection of first guess and retaining tides

An alternative way of reducing tidal biases in the data assimilation is to add the mean increments to the first guess and retain the tidal fields. In the notation used previously this would be:

$$A'(b+t) = f(b_-) + R(t_-) + \overline{t - R(t_-)}$$

$$+ O\left[b + t - (f(b_-) + R(t_-) + \overline{t - R(t_-)})\right] \qquad (15)$$

and after cancellations

$$A'(b+t) = f(b_-) + t + O\left[b-f(b_-)\right] \qquad (16)$$

$$f'(b+t) = f(b_-) + R(t) + S \qquad (17)$$

Then the next analysis in the assimilation cycle is

$$A'(b_+ + t_+) = f(b_-) + R(t) + S + \overline{t_+ - R(t)}$$

$$+ O\left[b_+ + t_+ - (f(b_-) + R(t) + S + \overline{t_+ - R(t)})\right] \qquad (18)$$

or

$$A'(b_+ + t_+) = f(b_-) + t_+ + S + O\left[b_+ - (f(b_-) + S)\right] \qquad (19)$$

It is seen from eq.(19) and eq.(16) that the increments are always "synoptic" and unaffected by tides. However, the tide is always retained and the analysis will be faithful to surface pressure observations.

The procedure was applied to 12Z 1 May 1982 with mean increments from April 1982. The performance of the subsequent analysis (18Z) is checked by taking the difference between analysis increments of eq.(19) and eq.(14). This gives $-(t_+ - R(t))$, i.e. the negative of the tidal increment at 18Z. Comparing the mean increment for 18Z in Fig. 4 and the analysis increment difference in Fig. 5, we find a good agreement, although there seems to be some "overshooting". This is probably due to the unbiased analysis drawing closer to the data which in turn causes the tides to be analysed even better. After having introduced the bias correction operationally the mean increments would thus have to be revised.

3.3 The vertical structure

Now to the crucial problem of the vertical structure in the analysis. When the observational increments (b-f) are reduced by adding the bias to the first guess, it would be expected that the negative impact on the thermal structure by the tides is reduced or eliminated. Indeed, at 18Z a difference map of analysed temperature at the lowest model σ-level (Fig. 6) shows a change in temperature in fairly good agreement with the reduced analysis increment amplitude (Fig. 5). However, this is for an analysis hour when only SYNOP's determine the tropical boundary layer temperatures. The comparison for the preceeding analysis at 12Z was not so encouraging (see Fig. 8), especially over East Africa where the reverse effect has occurred. Here heating at the lowest σ-level has taken place, contrary to what was expected, although cooling still takes place at higher levels.

This may seem puzzling, but comparing the analysis increments at 850 mb and 1000 mb in Fig. 7 shows that the magnitude increases with height over northern East Africa, whereas practically everywhere else it decreases with height as expected. Then, examination of the data coverage (Fig. 9) shows that the area is rather void of data - except along the coast. The only observation over East Africa between 0° and 25°N is a TEMP at Addis Ababa. Here the topography is high and Addis Ababa's station pressure is 771 mb. The lowest datum to enter the analysis there is the 700 mb height. Since the vertical

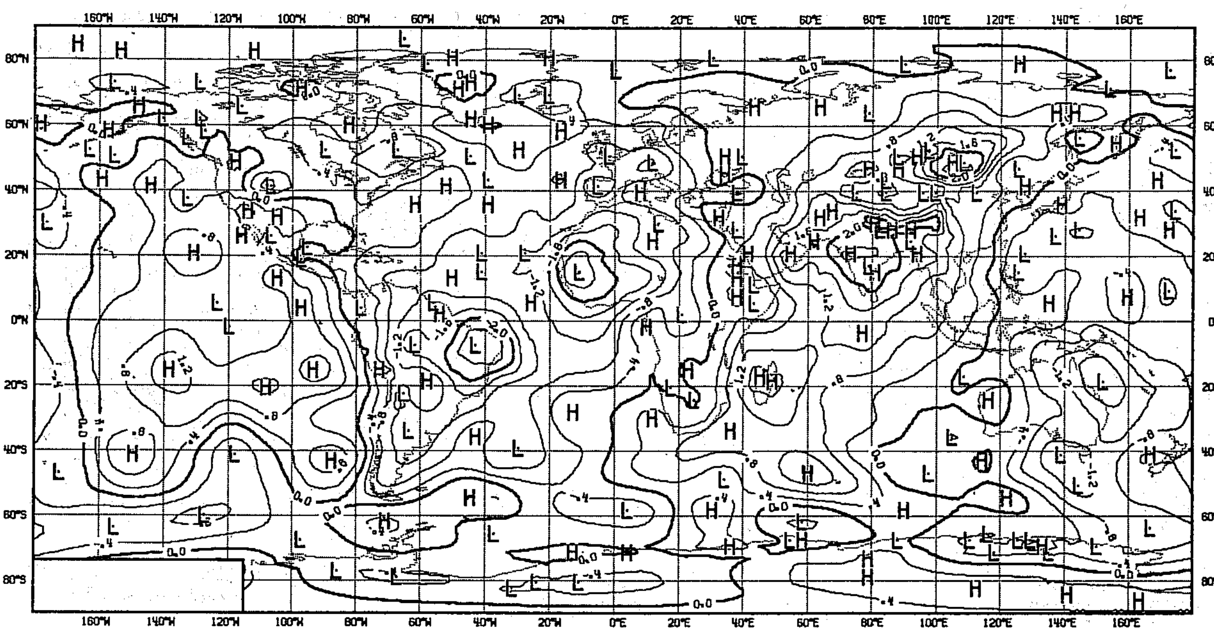

Fig. 4 Mean analysis increments at 1000 mb for 18Z. Daily mean increment has been subtracted.

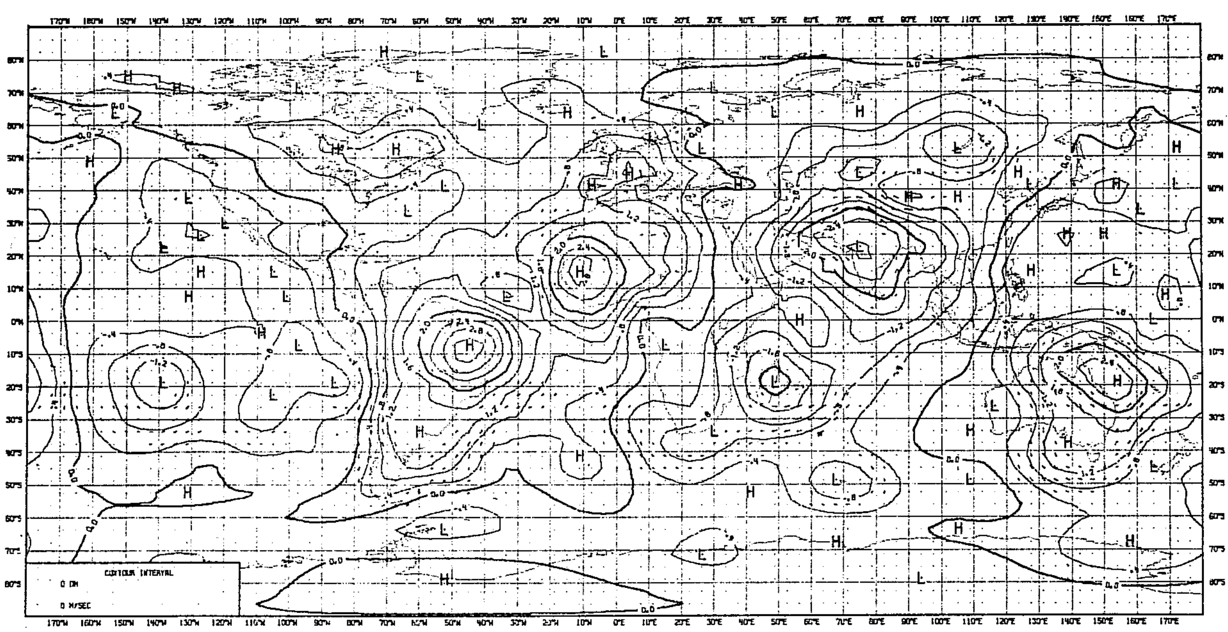

Fig. 5 Difference between analysis increments at 1000 mb when second procedure is applied.

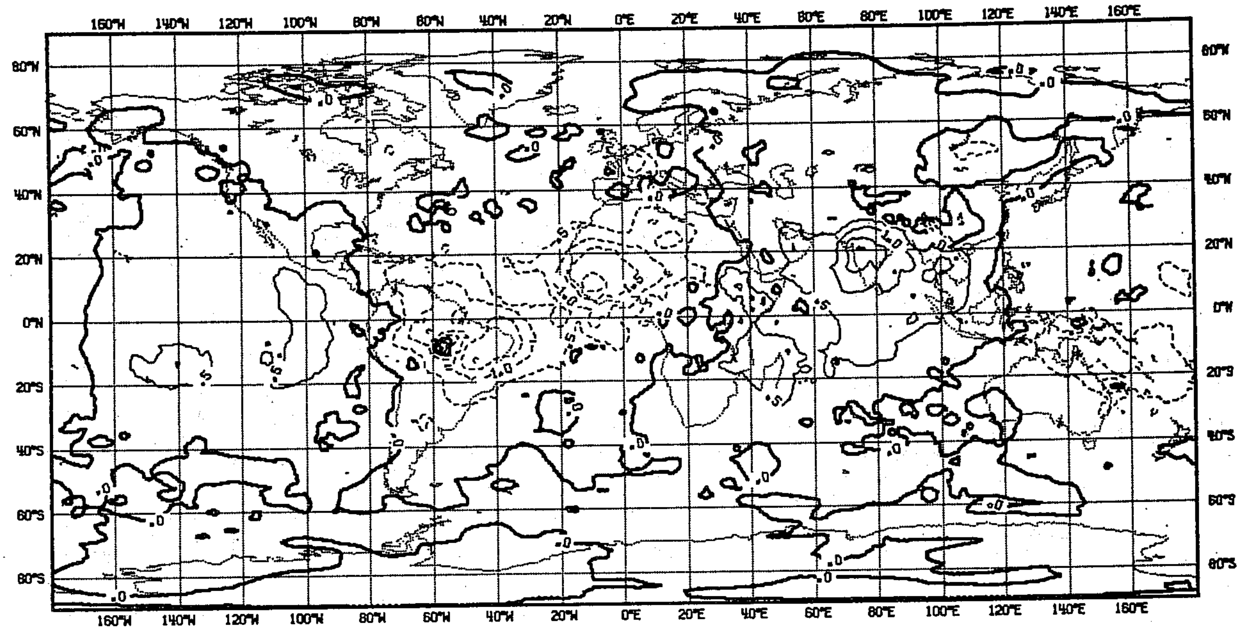

Fig. 6 Difference between analysed temperature at $\sigma = 0.996$ at second analysis cycle (18Z). Isolines energy.5K and dashed lines negative.

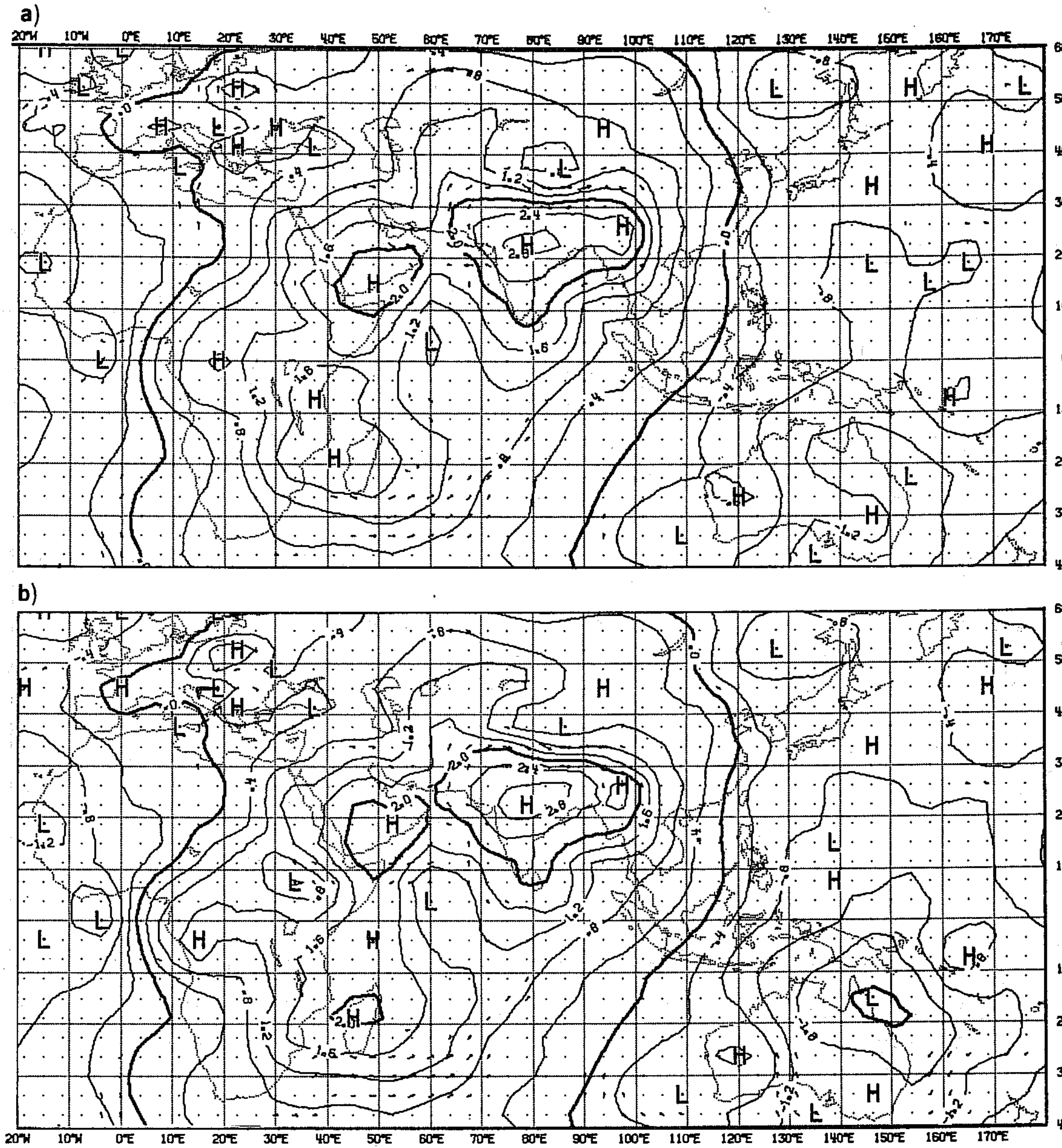

Fig. 7 Difference between analysis increments (Dm) at a) 850 mb and b) 1000 mb during the first cycle (12Z).

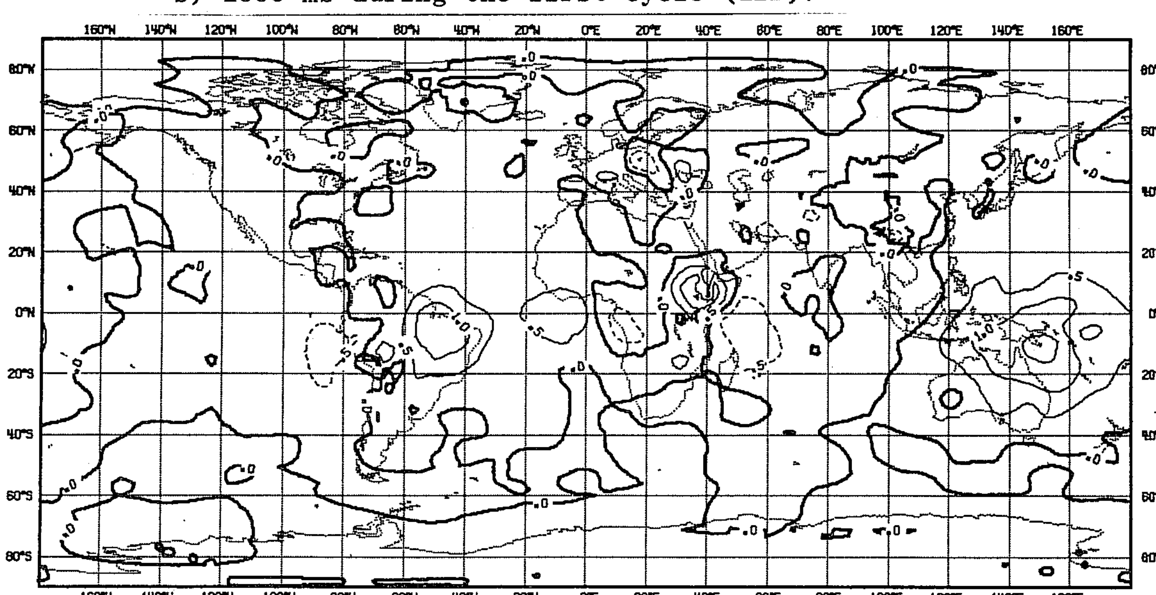

Fig. 8 Resulting temperature difference (K) between analyses at the lowest sigma level (0.996).

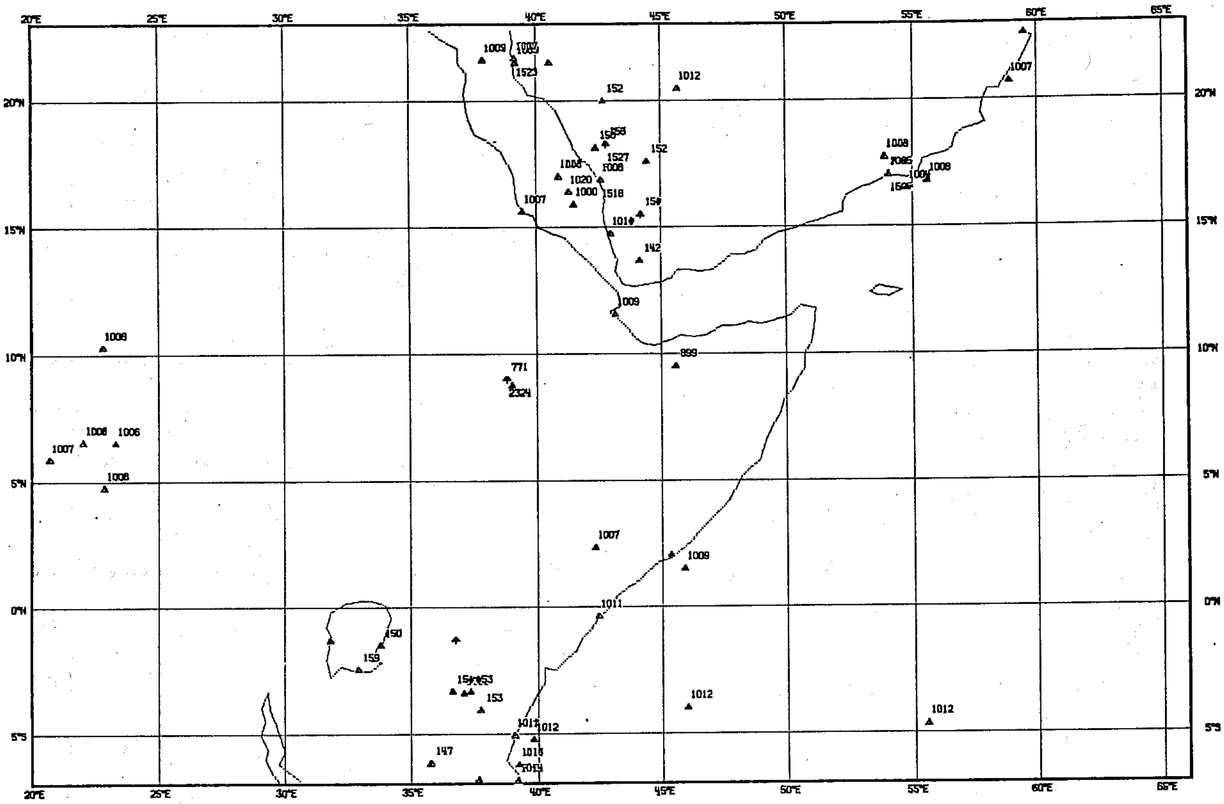

Fig. 9 Data coverage by SYNOPs and TEMPs at 12Z, 1 May 1982. TEMPs are triangles with a vertical line through the base and figures give reported pressure (mb) or height (Dm).

forecast error correlation between 700 and 1000 mb is almost zero, any bias correction at 700 mb will have very little effect on the 1000 mb analysis. Even though the model's topography reaches 2000 m at the observation point, there are areas around the mountains with topography below 400 m which are well within the radius of influence (840 km here).

The higher levels will, however, influence the 1000 mb analysis. The forecast error correlation between 500 and 150 mb is around -0.5 here. A heated extensive boundary layer will produce height increments that increase with height (since the first guess has no diurnal cycle). Adding a negative 1000 mb bias at these levels will make the increments even more positive and consequently at 1000 mb the increments will be increasingly negative. In Fig. 8 this is seen as a band of heating around the mountain in Ethiopia.

To confirm this argument the same analysis was repeated with all TEMP's removed. Then the tidal correction works beautifully as is shown in Figs. 10 and 11. A cooling has been achieved not only over East Africa, but over the whole area where the mean increments shown in Fig. 1 are negative. So, where TEMP's are available, and the data coverage is fully three-dimensional, the structure functions will not create any problems nor will a bias correction have any effect.

The TEMP coverage is shown in Fig. 12. A particular area where the bias correction will have the largest impact is around New Guinea where there are no radiosonde stations (at 12Z Australia only do soundings in the south). Consequently the heating there due to this process at 12Z is likely to be persistent.

Note also, by comparing Figs. 10 and 11, that the impact of the bias correction is larger at $\sigma = 0.845$ than near the surface. Over Ethiopia, for example, the effect is almost nil at $\sigma = 0.996$. Again checking the available SYNOP's in Fig. 9 reveals that there is one in Somalia reporting station level pressure (899 mb) which will enter the analysis as a height increment at 850 mb. Also south of the equator there are many SYNOP's reporting height at 850 mb (values around 150 Dm). In combination with adjacent low level stations reporting mean sea level pressure, they introduce to some extent a thickness between 1000-850 mb which reflects the real temperatures in the area.

To summarize, there are five basic observational configurations where the tides and bias corrections work in different ways:

a) <u>Only mean sea level pressures</u>

Tides are analysed baroclinicly irrespective of the real stratification. The bias correction eliminates the problem. A typical area is New Guinea.

(b) <u>Only high altitude surface observations entering the analysis at 850 mb (or at a higher level)</u>

Tides are analysed baroclinicly in adjacent layers with a destabilization above and a stabilization below (or vice versa depending on the phase of the tide). A 1000 mb bias correction applied at the higher level reverses the effect, at least to some degree. It will still be beneficial above the level in question and below the baroclinic effect is reduced. In practice, this configuration is not so common because low level SYNOP's usually exist within the radius of influence.

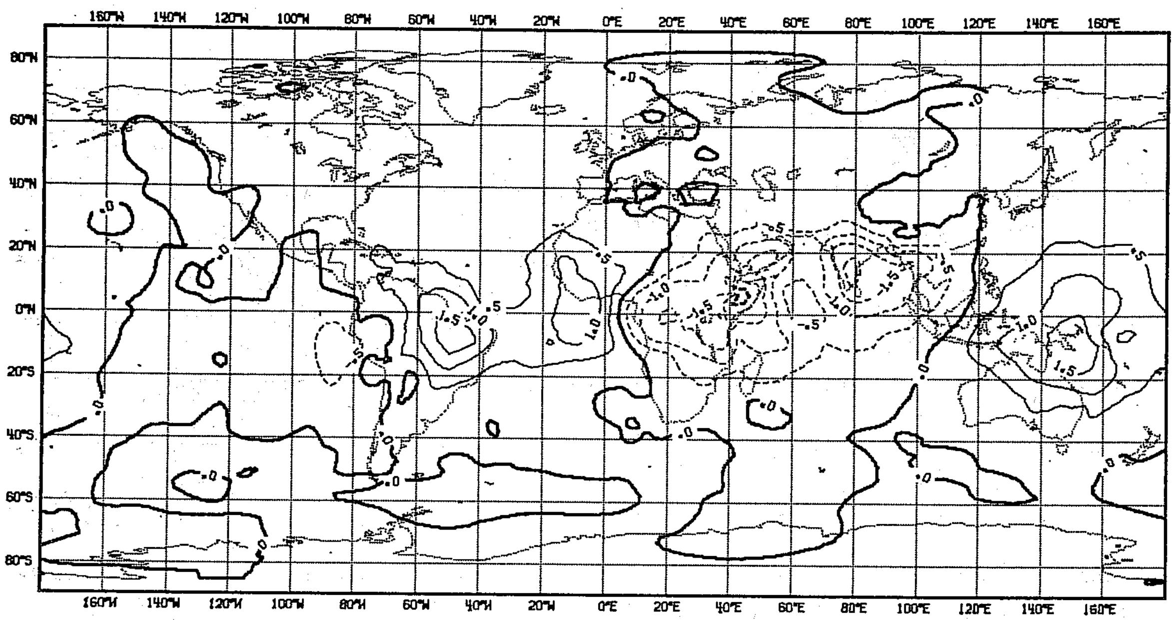

Fig. 10 Temperature differences at $\sigma = 0.845$ between a biascorrected analysis without TEMPs and a baseline analysis at 820501 12Z.

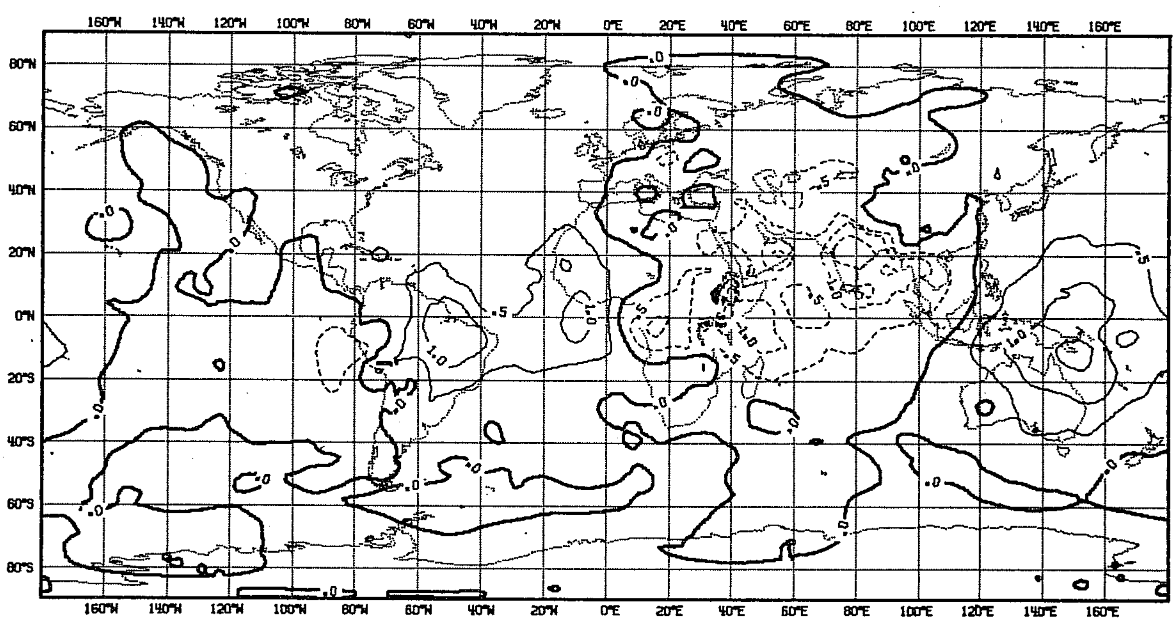

Fig. 11 As Fig. 10 but at $\sigma = 0.996$.

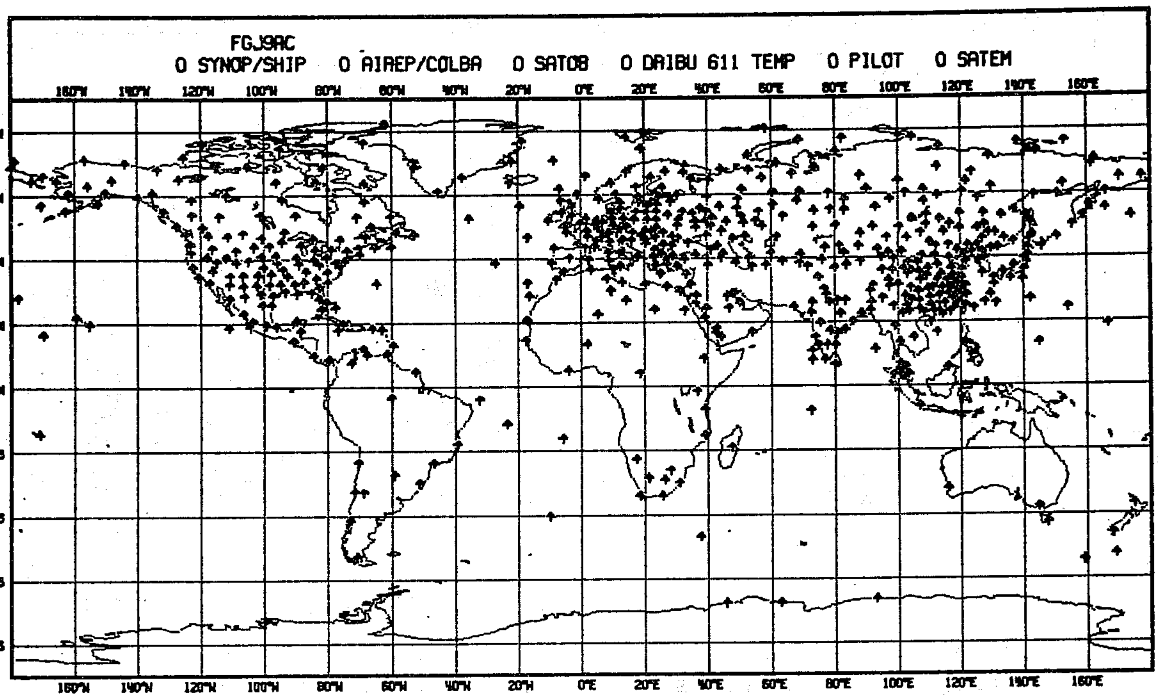

Fig. 12 TEMP reports at 820501 12Z

(c) Combination of high and low level surface observations within each others area of influence

Tides will give an erroneous baroclinic effect above the higher level, but underneath the analysis will reflect the observed temperatures. The bias correction works well above but has little benefit beneath the higher level. A typical area is East Africa at the equator and southwards.

(d) Multi-level observation (TEMP's) from a high level station

The temperature will be analysed relatively faithfully between standard levels irrespectively of the tide. The bias correction will not change anything there. Below the lowest reporting level the bias correction can have a detrimental effect due to negative correlations with high level data. This was shown earlier in the case of Addis Ababa.

(e) Multi-level observations covering all levels

Neither the tide nor the bias correction will have any impact on temperatures in this case.

4. BIAS CORRECTION AND BAROTROPIC ANALYSIS INCREMENTS BELOW LOWEST OBSERVATION LEVELS

4.1 Experiment with bogus data

To overcome the problems introduced by stations reporting at high levels there is a need to ensure that analysis increments below the lowest level are barotropic. One way of doing this is to supply bogus data at the level (or levels) below. Those data would be duplicates of the lowest observational increment from a SYNOP or TEMP (levels 850-500 mb) and the observational error and error correlation should be modified so as not to change the analysis at the originating level.

The normalized analysis increment at gridpoint k (or at a level) can be written (see Lorenc (1981)) as

$$a_k = \underline{d}^T \left(\underline{\underline{M}}^{-1} \underline{P}_k \right) \tag{21}$$

where $\underline{\underline{M}}$ is the observation correlation matrix $\left[\langle p_i p_j \rangle + \varepsilon_i^o \gamma_{ij} \varepsilon_j^o\right]$ with forecast errors p, observation errors ε^o, and observation error correlations γ for observations i and j. $\underline{P}$ is the forecast error correlation vector $\left[\langle p_i p_k \rangle\right]$ and $\underline{d}$ the normalised data vector $\left[d_i\right]$.

For one observation at a level the analysis would simply be

$$a_1 = d_1 \frac{1}{1 + \varepsilon_1^2} \tag{22}$$

The datum is then duplicated, and renormalized with the forecast errors, to give constant observational increments

$$d_2 = d_1 \cdot \frac{E_1}{E_2} \tag{23}$$

where E_1 and E_2 are the rms forecast errors.

The analysis at the original level with two data can be found by applying eq. (21):

$$a'_1 = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} p_{11} + \varepsilon_1^2 & p_{12} + \varepsilon_1 \gamma_{12} \varepsilon_2 \\ p_{21} + \varepsilon_2 \gamma_{21} \varepsilon_1 & p_{22} + \varepsilon_2^2 \end{bmatrix}^{-1} \begin{bmatrix} p_{11} \\ p_{21} \end{bmatrix} \quad (24)$$

Now, if the observational error, ε_2, was also duplicated, the analysis would be slightly changed at the higher level. This can be avoided in the two level case by substituting for ε_1 and ε_2 a common value ε'. Solving eq.(24), and requiring that the two analyses a and a' should be the same, gives

$$\varepsilon' = \left[\frac{d_1 + \mu d_2}{d_1} (1 + \varepsilon_1) - 1 - \mu\right] \frac{1}{1+\gamma} \quad (25)$$

where μ is the forecast error correlation ($p_{12} = \langle p_1 p_2 \rangle$).
The temperature difference when the analysis has been modified in this way with the bogus information is shown in Fig. 13.

The relative warming over Ethiopia occurring in the previous experiment (often due to a lack of cooling present in the uncorrected analysis) has been almost removed (Fig. 8). As discussed in Section 4 under points b) and d), a negative tidal increment at 12Z will give a cooling below the station level. Such a cooling occurs especially the next day at 12Z during the assimilation (Fig. 15). Then the bias correction together with bogus data does not achieve any cooling over the East African area even though Central Africa has been substantially cooled (or less heated) - see Figs. 14 and 16. The effect of the bogus data for this analysis hour is shown in Fig. 17.

Continuing the assimilation for another day gives a similar picture (Fig. 18) for the lowest sigma level. Note, however, that over New Guinea the correction has a consistent heating effect at 12Z (the tidal increments change sign at every 6h cycle so there is, of course, no net heating over a day). At higher levels the impact of the bias correction is much clearer (see Fig. 19). The impact in the vertical is also demonstrated by examining tephigrams from the same analysis (at sigma levels) at gridpoints near Addis Ababa and Dar-es-Salaam (Figs. 20 and 21 respectively). The boundary layer is somewhat warmer (although when compared with the actual soundings the temperatures are still far from those observed) whilst the lower troposphere has been cooled. In the higher troposphere there is some warming although changes at the highest levels are generally small.

During the third day of assimilation (after 8 and 9 analysis cycles) a reduction of vertical velocities of the "grid point storms" over Africa was actually achieved. These are best revealed in the 6 hour forecasts from the initialized analyses. Vertical velocities in the forecasts from the last pair of analyses (the 9th cycle at 12Z 3 May 1983) are shown in Figs. 22 and 23. The amplitude of the storms over Africa is almost halved. Over Bangladesh it is virtually unchanged but over New Guinea the "grid point storm" is intensified due to the heating (or less cooling effect of the bias correction).

However, when the assimilation was pursued for another day it became clear that there is no real reduction of the convective activity seen over consecutive cycles during a day. The reduction of the vertical velocities in Fig. 23 stems merely from a phase shift in time of the convection. During other 6 hour forecasts in the assimilation (e.g. during the night and in the morning) the "grid point storms" in the bias correction run can be worse. This is due to the fact that the mean increment (minus daily mean) changes

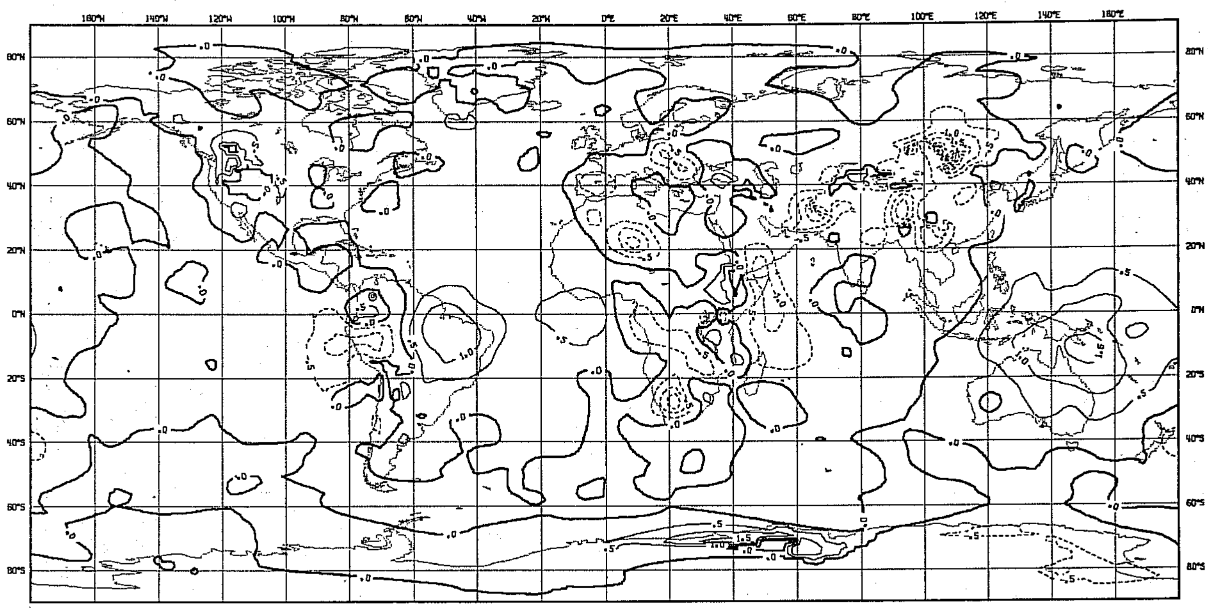

Fig. 13 Temperature difference between an analysis with bias correction and bogus data and a baseline at the lowest sigma level (0.996). For 820501 12Z as Fig. 8.

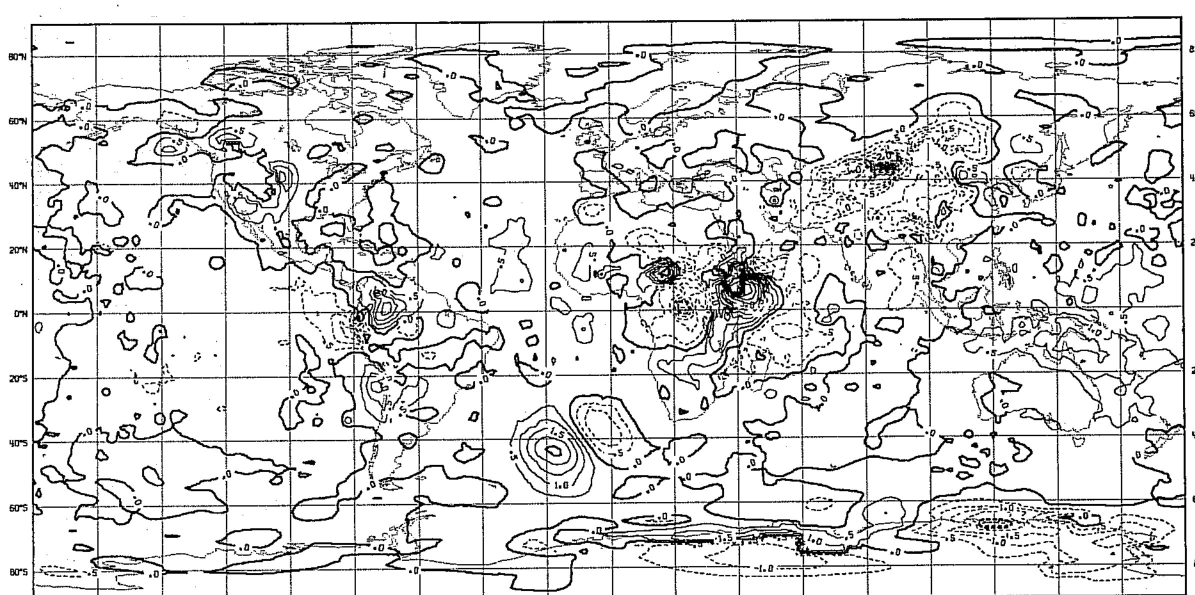

Fig. 14 As Fig. 13 but for 820502 12Z.

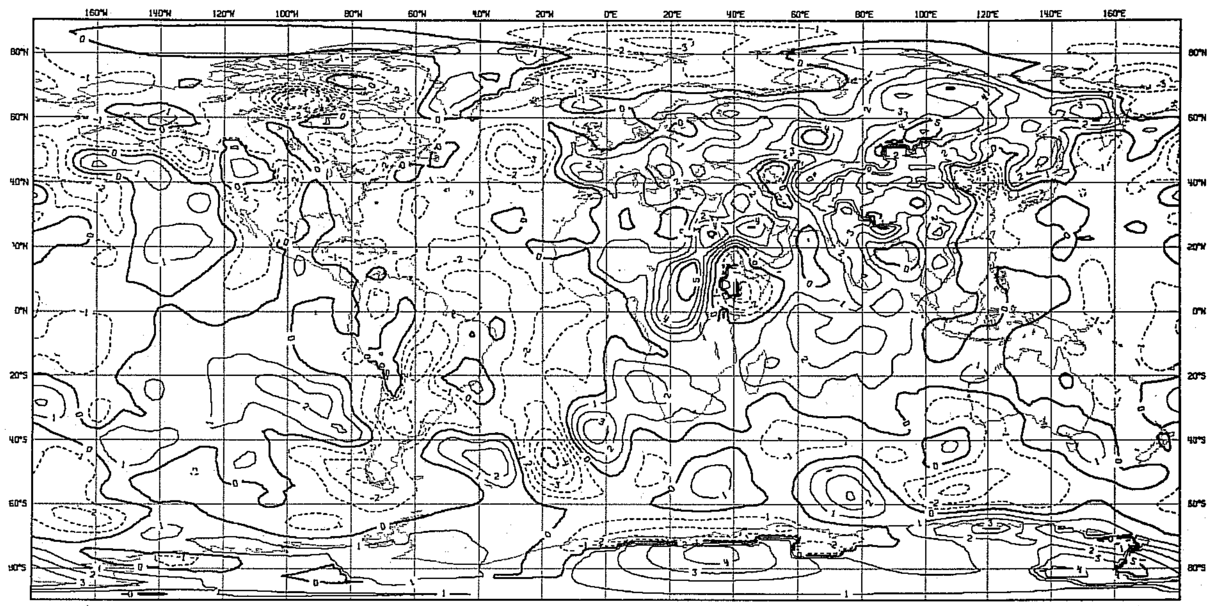

Fig. 15 Temperature analysis increments at lowest sigma level for 820502 12Z in baseline assimilation. Intervals at 1K.

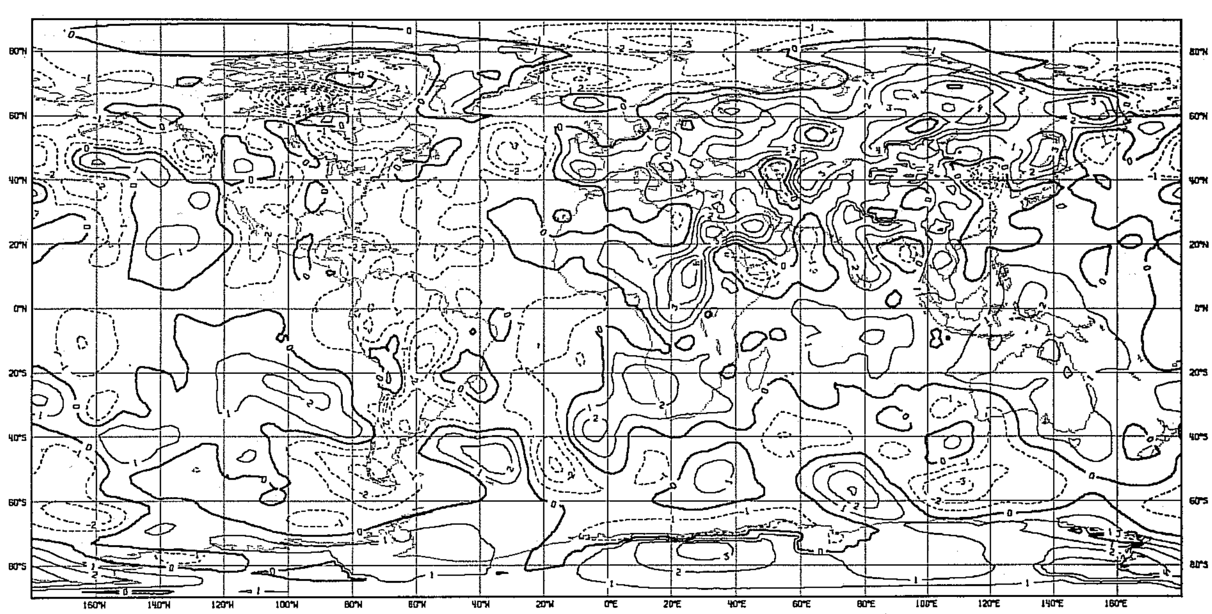

Fig. 16 As above but for experiment with bias correction and bogus data.

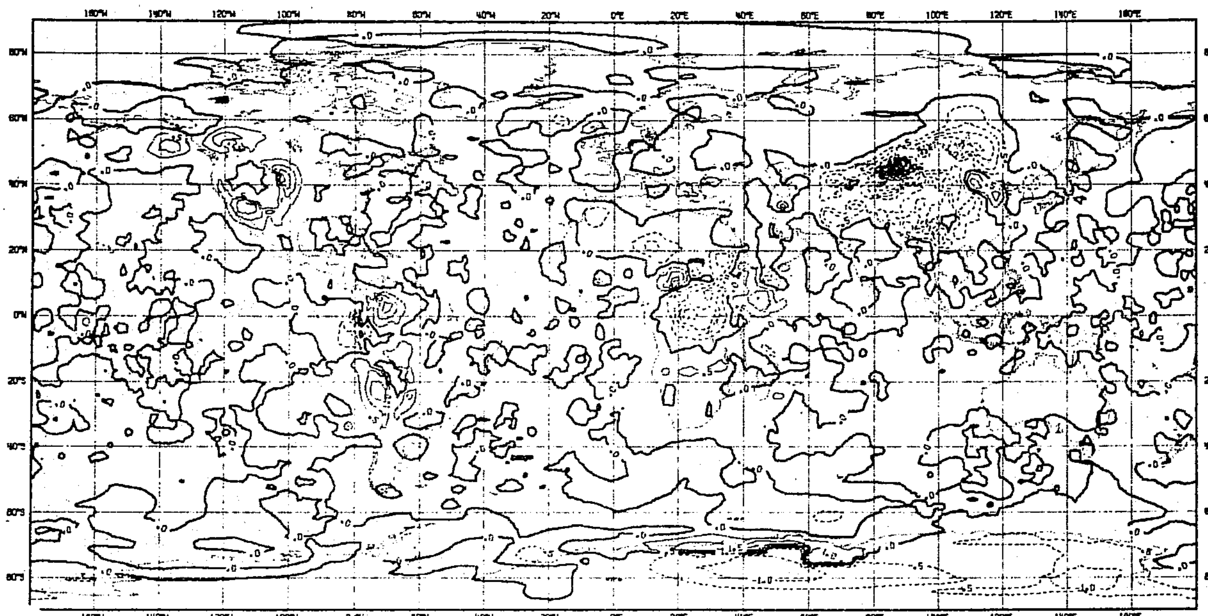

Fig. 17 The effect on temperature at lowest σ level by bogus observations demonstrated by taking difference between analyses with tidal correction and bogus data and with tidal correction only.

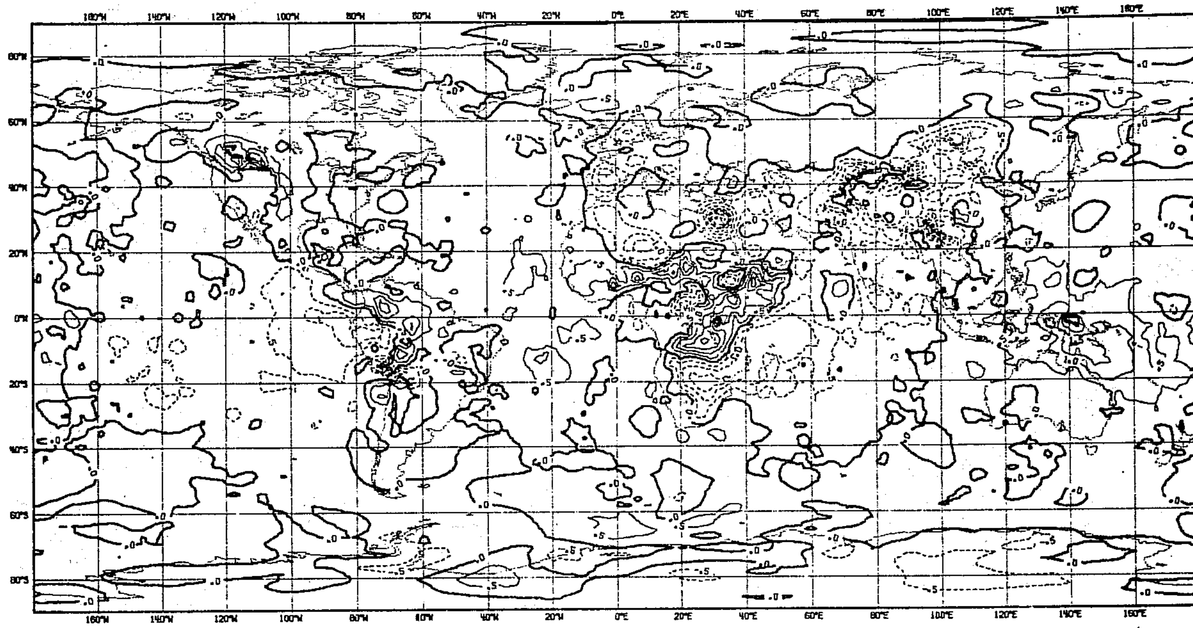

Fig. 18 Temperature difference at $\sigma = 0.996$ between analyses at 820503 12Z. Intervals at 0.5k. Analysis with tidal correction and bogus data compared with baseline.

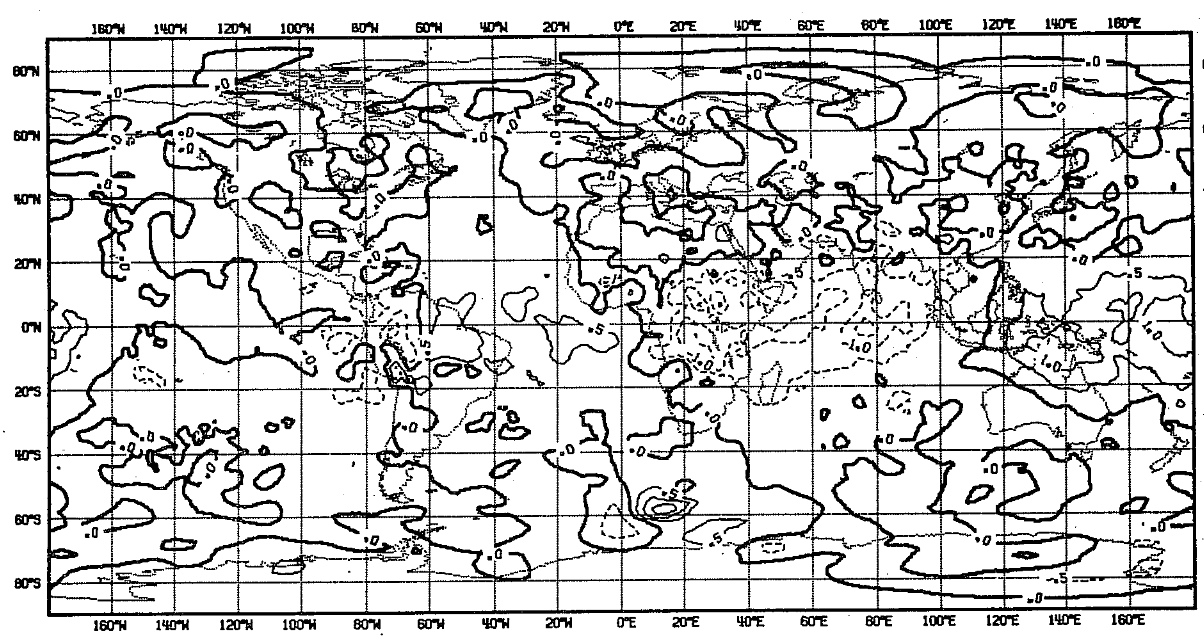

Fig. 19 Temperature difference at $\sigma = 0.589$ for 820503 12Z.

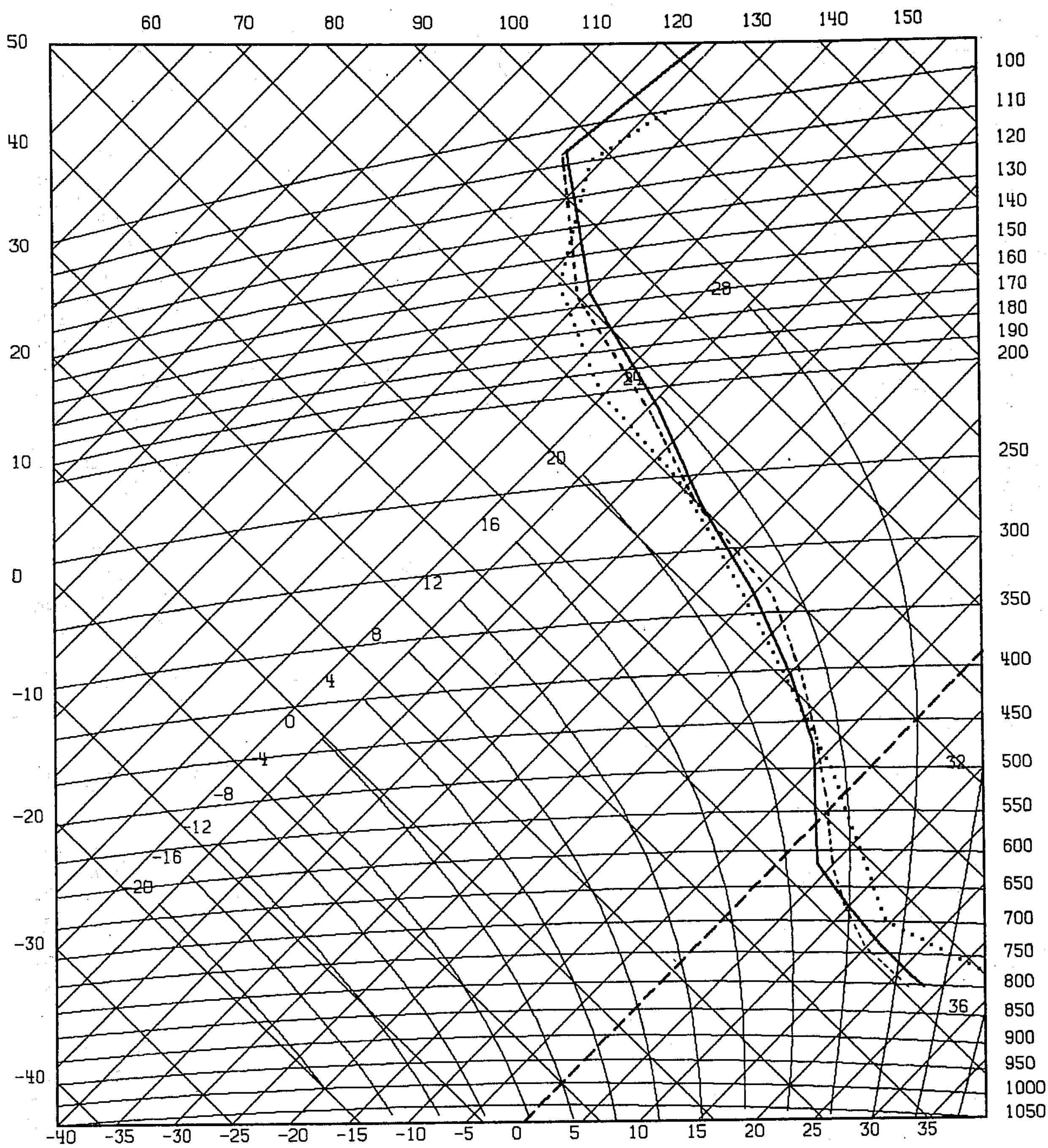

Fig. 20 Tephigrams from analysis interpolated to sigma levels at a grid point near Addis Ababa. Full line is for experiment with bias correction and bogus data. The profile from the baseline analysis is dashed. Dotted profile is the actual sounding from Addis Ababa at 820503 12Z.

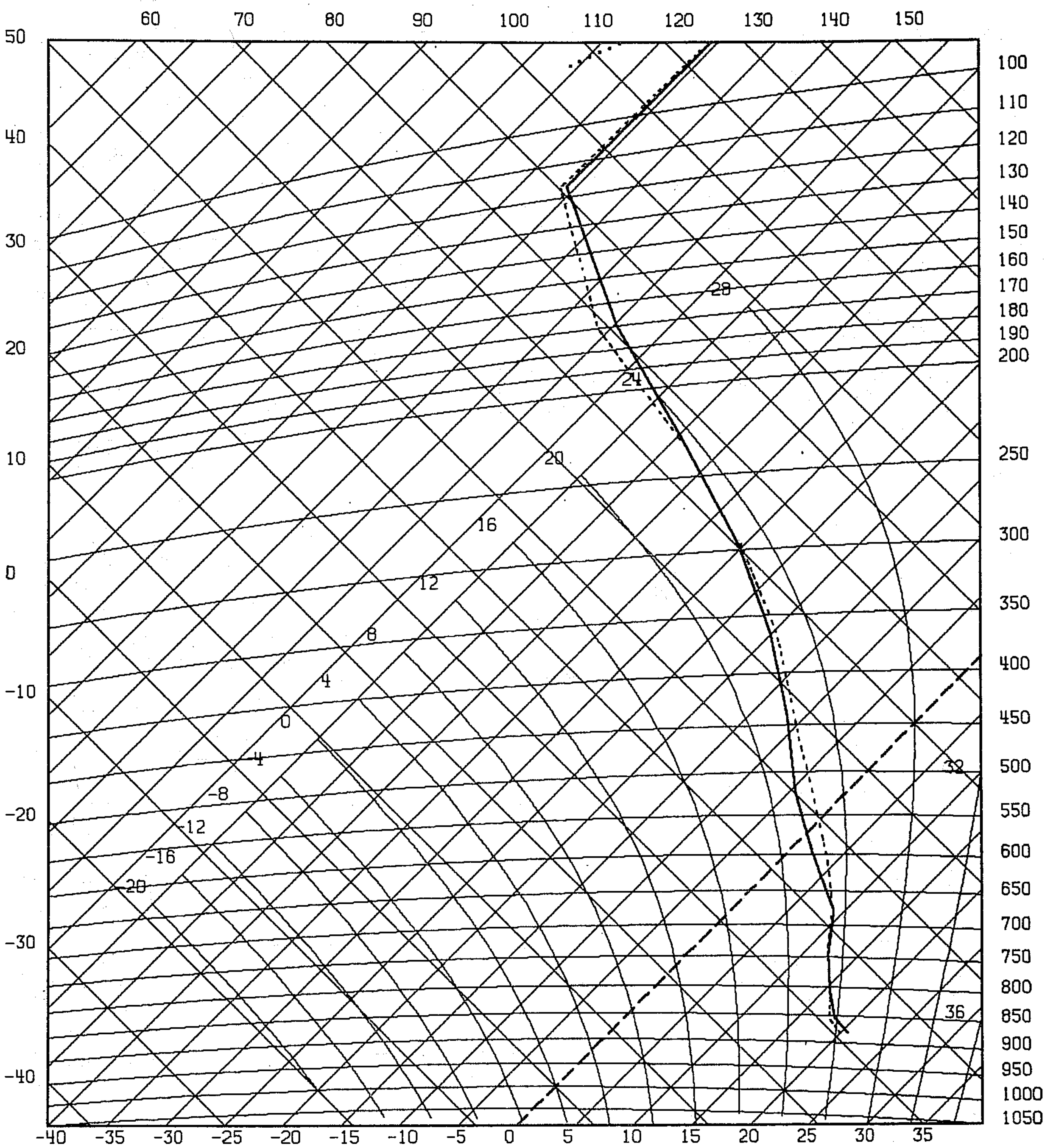

Fig. 21 As for Fig. 20 but near Dar es Salaam. Temp data was only available above 100 mb.

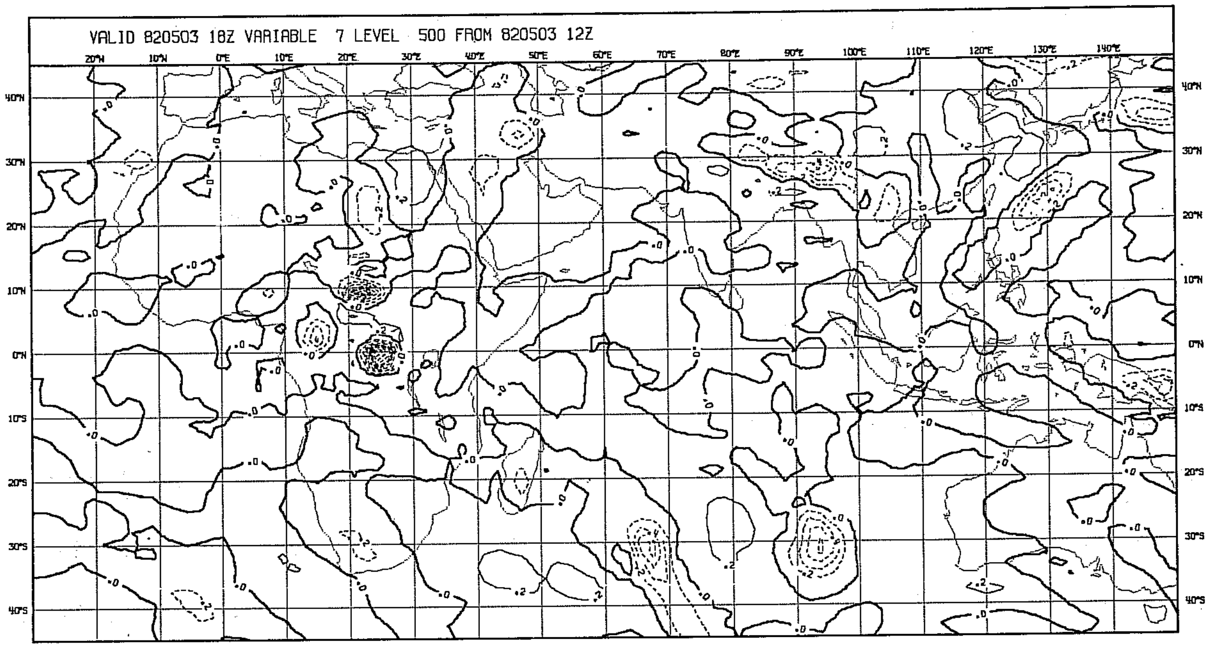

Fig. 22 Vertical velocity for a 6 hour forecast from baseline analysis at 820503 12Z. Intervals at 0.2 Pa/s and negative isolines dashed (upward notion).

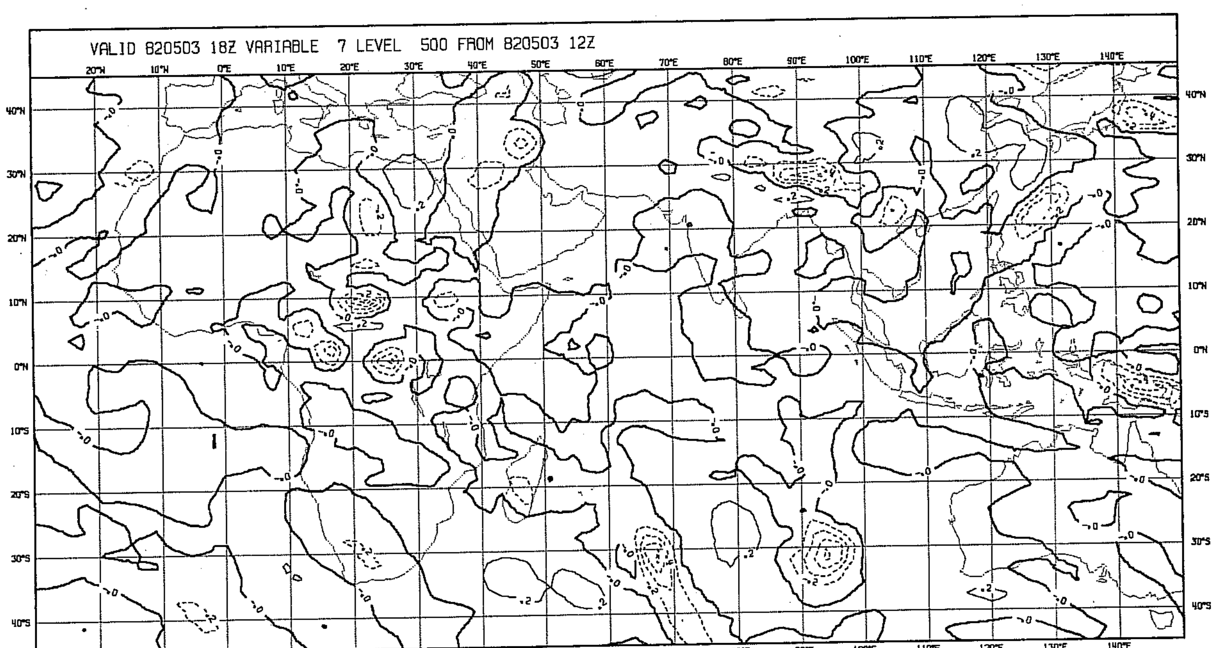

Fig. 23 Corresponding vertical velocities for the 6 hour forecast from analysis with bias correction and bogus data.

sign for each analysis hour and no net cooling or heating will occur during 4 cycles (24h) provided that the data coverage is symmetric over the same 4 cycles. This is not the case over East Africa. Soundings north of the equator are only made at 12Z there (i.e. usually one or two TEMP's from Ethiopia) whereas at other times the only mass data available are from SYNOP's. Since the mean increment over East Africa has as large positive magnitude at 06Z and an equally large negative one at 12Z, the tidal correction will prevent a spurious cooling at 06Z and do little at 12Z.

4.2 Limitations when using bogus data

In theory and practice the only property that can be obtained with the bogus data is a barotropic temperature increment in the layer between the original datum d_1 and the bogus datum, d_2, below. This is equivalent to requiring that the renormalized analysis increments (of height) should be the same at the two levels:

$$a_1' \cdot E_1 = a_2' E_2 \qquad (26)$$

where E is the rms forecast error.

a_2 can be found by replacing (p_{11}, p_{21}) by (p_{12}, p_{22}) in eq.(24). This leads to a more complicated expression for d_2 as compared to eq.(23).

$$d_2 = \frac{\mu d_1 E_2 - d_1 E_1}{\mu E_1 - E_2} \qquad (27)$$

In the experiment described in the previous section, eq.(23) was used and bogus data were inserted at all underlying levels. Even though the conditions for constant increments were not exactly satisfied, there were only small differences when a rerun with bogus data at only one level was made.

However we can only ensure that the analysis at one level (a_1') is unchanged. The analysis at any other level (above) can be found by applying eq.(24) again to give

$$a_3' = \frac{\kappa d_1 + \lambda d_2}{1 + \varepsilon' + \mu + \gamma\varepsilon'} \tag{28}$$

where κ and λ are the forecast error correlations p_{13} and p_{23}. The original analysis at that level is still

$$a_3 = \frac{d_1 \kappa}{1+\varepsilon} \tag{29}$$

Assuming that $1 + \varepsilon' + \mu \simeq 2(1+\varepsilon)$ and $d_1 \simeq d_2$, which is reasonable for two adjacent levels (e.g. 700 and 850 mb), we find that $a_3' \simeq a_3$ provided λ is of the same order as κ. This is the case for levels not too far away from the datum; but when one comes to the "tails" of the vertical correlation function, κ may be twice the value of λ (even if only 0.10 and 0.05 respectively).

Therefore in cases of a large observational increment (d_1), the bogus datum could give a relatively large change in the stratosphere if no other data is available (bearing in mind that when renormalising a_3 the rms forecast error may increase rapidly with height).

5. CONCLUDING REMARKS

It has been shown that tidal biases in the data assimilation do have a baroclinic impact in the lower troposphere (below 500 mb). This is caused by using highly baroclinic vertical structure functions in the analysis where surface data is the only source of information. By correcting the bias and adding a barotropic mean analysis increment to the first guess height fields, these artificial temperature changes can be largely reduced. The amplitude of the temperature reduction was found to be around 1.5K in areas with only surface observations at or about sea level. Complications arise in mountainous areas with high level stations. Below the station level (or rather its nearest standard level) the baroclinic effect on the analysis will be the reverse as compared to the above. In order to ensure barotropic temperature analysis increment in a layer below a station, bogus height data was produced from the lowest observed data. Then the bias correction was at least not detrimental to the analysed temperatures in the forecast model's boundary layer. However, the bogus correction may affect the analysis at other levels in an undesirable way.

In areas with a combination of high level and mean sea level observations (and especially if TEMP's are available) the boundary layer temperatures seem not to be much affected by tidal biases and a bias correction has little impact there. Some impact may, however, be seen above the boundary layer but it is generally small (0.5-1.0 K). Since East Africa is one of these areas, the bias correction (including bogus data) does not give a very clear improvement in the static stability. In some cases there may even be a destabilisation at 12Z over the Ethiopian mountains due to the bias correction cooling aloft and reducing the artificial cooling in the normal baseline analysis below the station level. However, over the rest of Central Africa the bias correction seems to have a somewhat larger cooling effect at the surface. One area where the correction has an effect is around New Guinea. At 12Z this produces a warming and destabilization of the boundary layer. Grid point storms became more severe there and, in a forecast run from these assimilations, the only consistent impact was in this area. Convection developed earlier from the bias corrected analyses. The baseline forecasts have similar cycles of intense convection but they occur later in the integration.

It is of course somewhat disappointing that so little improvement could be achieved over Africa but it can be concluded that the tidal biases are not the major cause of the low stability there (and indeed, they cause a spurious cooling at 06Z instead). The results suggest that the analysed temperatures over East Africa at 12Z are in fairly good agreement with the real ones. On the other hand, the time shift of the South East Asian convection is probably desirable but there does not seem to be any significant objective forecast improvements because of this. It is also, of course, more satisfactory to have an unbiased first guess field for the analysis (and especially for data checking), but it has to be traded off against the additional complexity of keeping and using a database of mean analysis increments.

REFERENCES

Chapman, S. and Lindzen, R.S. 1969 Atmospheric Tides. D. Reidel Pulishing Co., Dordrecht, Holland.

Heckley, W. 1983 Problems and prospects in tropical prediction. ECMWF Newsletter No. 19.

Hollingsworth, A. and Arpe, K. 1982 Biases in the ECMWF data assimilation system. ECMWF Technical Memorandum No. 46.

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