TECHNICAL MEMORANDUM

46

Biases in the ECMWF data assimilation system

A. Hollingsworth and K. Arpe

Research Department

January 1982

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

Abstract

This note presents the results of a preliminary study of the means and variances of the changes, ("increments"), made by the three stages of the assimilation procedure i.e. the analysis, the initialization and the forecast. The results indicate that there are important bias effects at all three stages, some of which can be corrected in a straightforward manner and others of which are more difficult. There are indications that over the oceans the satellite temperature measurements have little impact on the analysis because they appear to agree with the first guess.

The increments show the expected relative magnitudes, with the forecast increments being largest and the initialization increments smallest.

1. INTRODUCTION

The original purpose of this work was to study the question "Is ECMWF running a data assimilation system?", from a simple minded point of view. Ideally we should make accurate analyses in the course of an assimilation for which the following proposition is true

$$F > A > I;$$

here F, A, I are measures of the increments made by the forecast, the analysis and the initialization. As a convenient measure we use the rms of the height changes due to each of the processes in the Northern Hemisphere extratropics. In the course of this work it emerged that each of the increment fields had significant biases, some of which are understandable and which can be removed. We also found indications that in some regions the available data appears to have relatively little impact on the analysis; this merits further investigation.

In general the conclusion is that we are indeed running an assimilation system in the sense indicated but that the system certainly be improved.

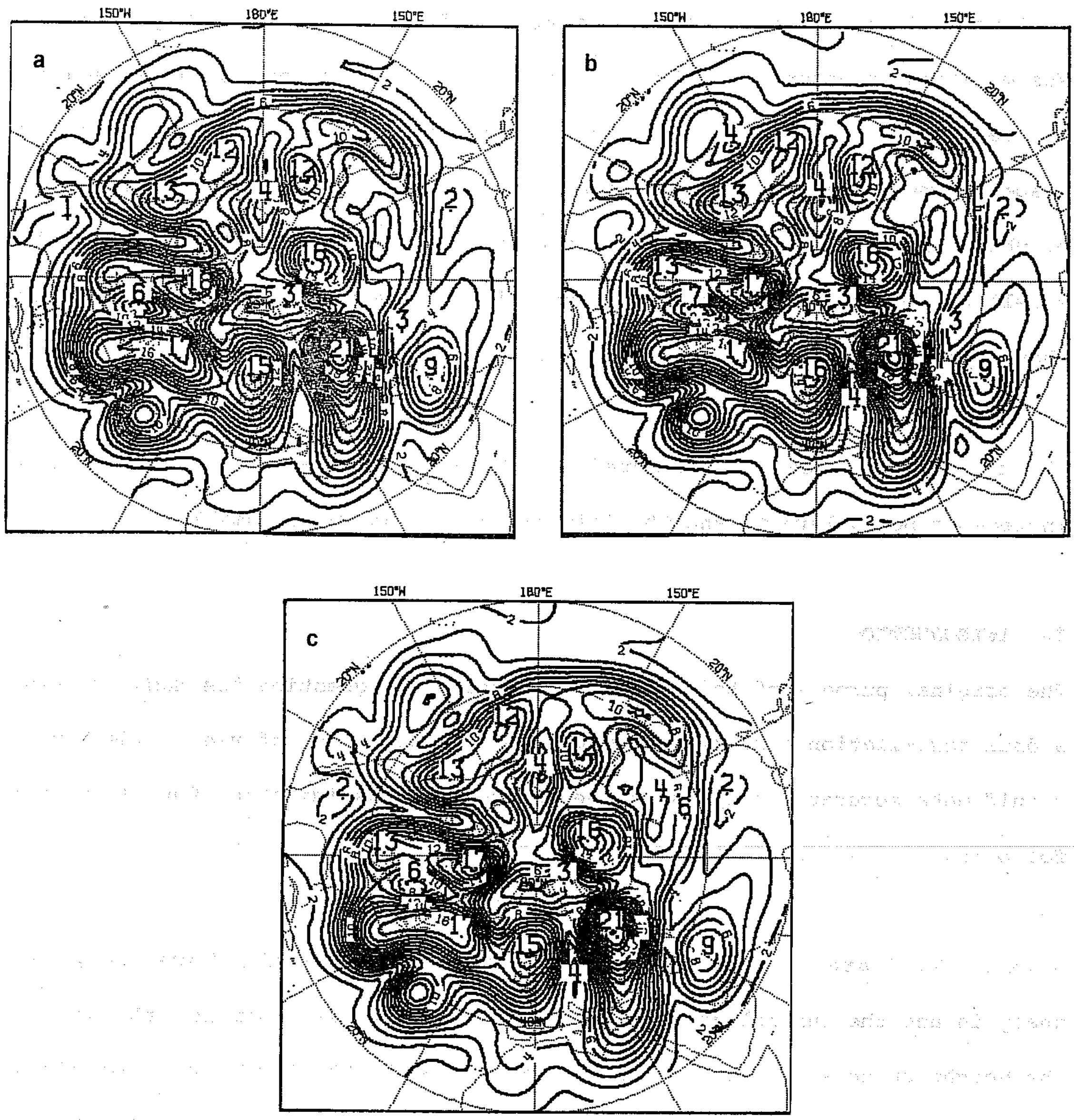

Fig. 1 Amplitude (dam) of the temporal standard deviation in a) the first guess field b) the analysed field c) the initialized field at 500 mb for the period 16-30 April 1979.

After a brief description of the data sources in Section 2, we give an overview of the increment fields in Section 3. The biases in the analysis and initialization fields are discussed in Sections 4 and 5. The impact of SATEM data is discussed in Section 6. The note concludes with a summary and suggestions for further work in Section 7.

2. DATA SOURCES

The data used in this study are archived fields from several different runs of the assimilation system using both FGGE data and operational data. The periods used are 1) the FGGE production files for the period 16-30 April 1979; 2) the archives from a re-run of the FGGE analysis for 6-19 February 1979 using the current operational system (February 1981); 3) the operational archives for March, April and May 1981.

3. OVERALL VIEW OF THE ASSIMILATION SYSTEM

As mentioned earlier,a data assimilation system should have the property that

$$F > A > I$$

where F, A, I are measures of the forecast analysis and initialization increments. A necessary condition for this to be true is that the ensembles of first guess, analysed and initialized fields should all have similar temporal variances. Otherwise two of the increments would have to be of the same order of magnitude. Suppose for example that the assimilating model were very dissipative so that the first guess fields were very much smoother than the analysed fields. Then the analysis increments would necessarily be as large as the forecast increments. It is only in the situation where all three ensembles have about the same variance that we can hope to find the desired ordering between the three increments.

Figure 1 shows the temporal variances of the first guess, analysed and initialized 500 mb fields for the April 79 period. The patterns are very similar and this similarity is typical for all levels for each of the periods studied. Thus the first necessary condition is satisfied.

We turn now to consider the variances of the increments themselves. Figure 2 presents maps of the rms value at 500 mb of the forecast increment, the analysis increment, the initialization increment, and what we shall call the net analysis increment which is the difference between the first guess and the initialized field.

At 500 mb it is clear that in the main storm tracks over the ocean we can claim that F > A, since the changes made by the 6 hour forecast are substantially larger than the analysis changes. Over the continental areas this is also true, but not to such a marked degree. We can also claim that A > I in most places, as the initialization increments are smaller in this rms sense than the analysis increments.

Finally we note that the effect of the initialization is generally to reject some of the change made by the analysis and to bring the analysis closer to the first guess.

The state of affairs outlined above is generally true at 1000 mb, away from mountains. Thus, for the lower troposphere we can claim with some confidence that our proposition is true.

4. ANALYSIS INCREMENTS

In the upper troposphere the situation is not so clear cut. Figure 3 shows the mean analysis increment at 200 and 100 mb for the April 79 period (with the FGGE production system) and also for the re-run of the February 79 period (with the Feb 81 operational system). A major difference in the two systems is the vertical interpolation from pressure to sigma. As a result of the interpolation change the mean value of the 200 mb and 100 mb height fields

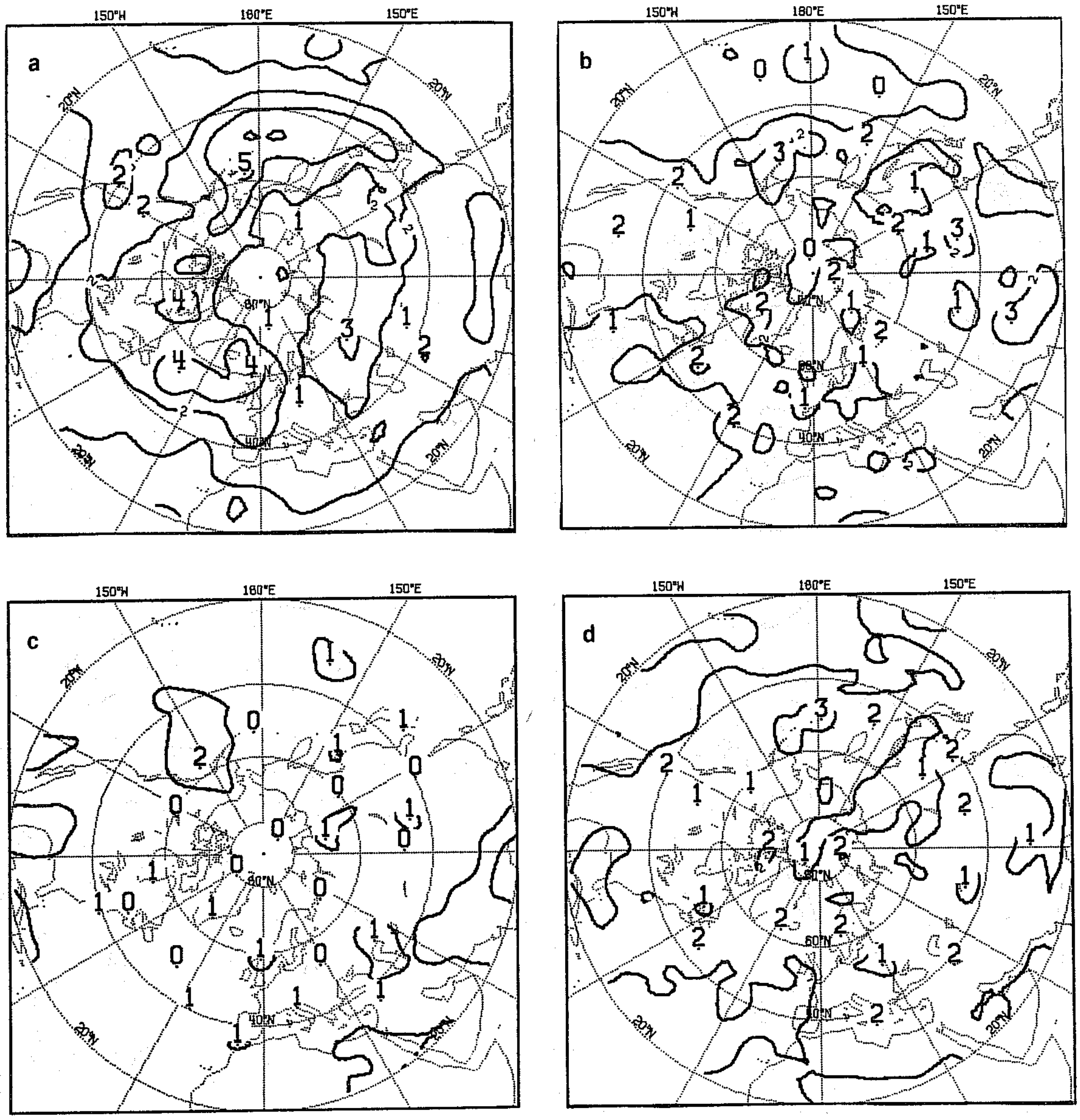

Fig. 2 RMS geopotential value at 500 mb of a) forecast increment b) analysis increment c) initialisation increment d) net analysis increment for the April 1979 period. Contour interval is 1 dam.

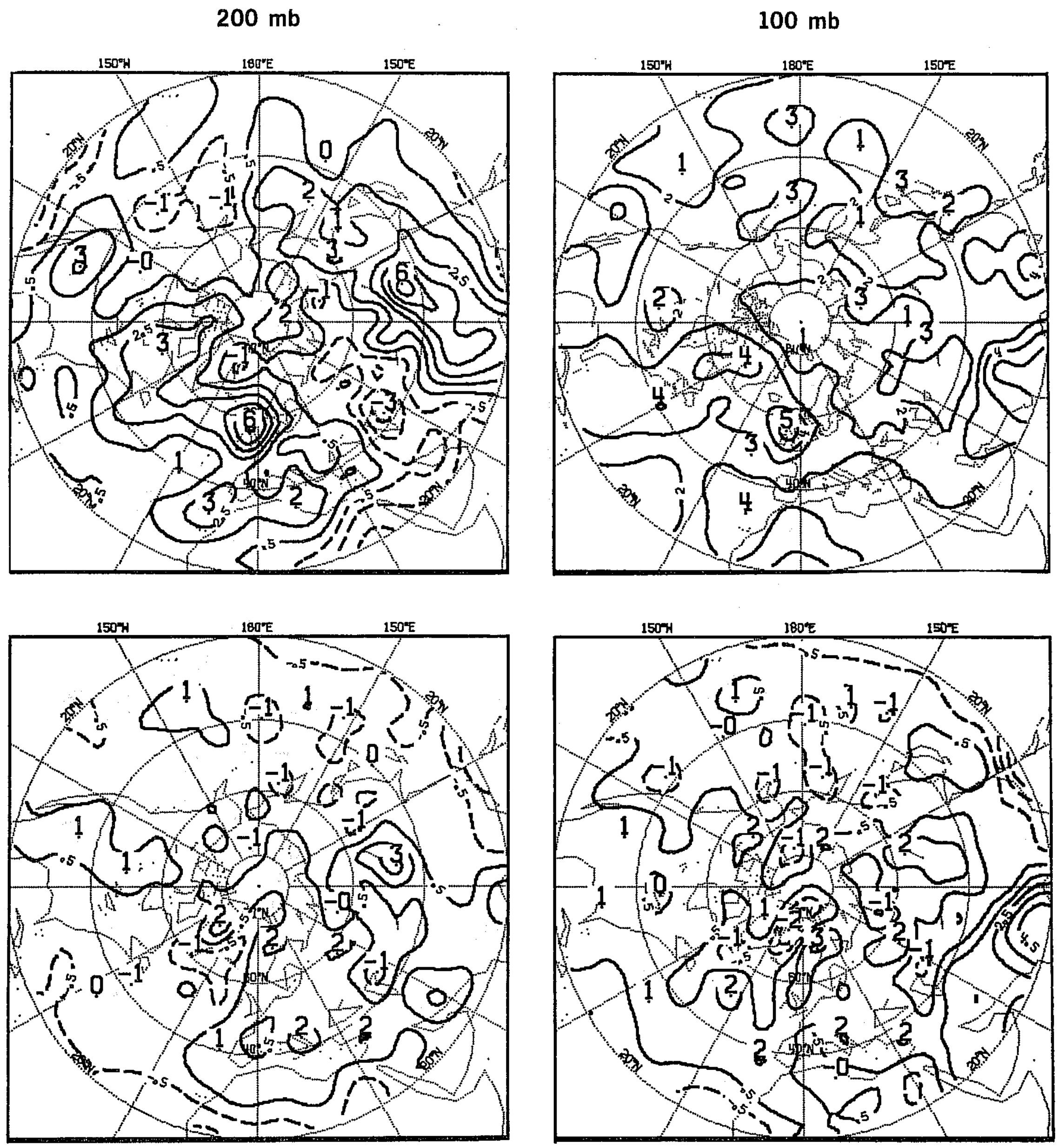

Fig. 3 Mean geopotential analysis increment at 200 and 100 mb for the April 1979 period (top) and for the re-run of the February 1979 period (bottom). Contour interval is 1 dam.

changed slightly. Despite this change both periods show clear evidence of small scale structure in the height increment field particularly over Europe. In addition there are large mean increments over the Himalayas. We believe the small scale structure is associated with radiosonde biases while the structures over the Himalayas are associated with the initialization.

4.1 Radiosonde biases

To examine the small scale structure of the 200 mb mean height increment we present in Fig. 4 the mean difference beetween the radiosonde 200 mb height reports and the first guess field for the February 79 period (current operational system). We see that there are very large variations in this field. For example all the British radiosondes show negative biases. The two different radiosondes used at the Irish stations show differences of 90 m in a distance of 400 km. Similarly the radiosonde in Iceland and on Greenland immediately to the north show differences of 70 m. Differences of this magnitude between closely adjacent radiosonde stations are also found in central Europe. Very much larger differences (up to 250 m) between adjacent stations can be found in the tropics. There is thus a clear need to introduce a method to correct each individual radiosonde in order to reduce the effect of these large differences between radiosonde types.

4.2 Diurnal cycle

The semi-diurnal and diurnal tides in the atmosphere have surface pressure amplitudes of order a few millibars with largest amplitudes in the tropics. Since the assmilating model does not yet have a diurnal cycle there will be mean contributions to the analysis increments from this source. In addition there is evidence (McInturff et al 1979) that radiosonde corrections for diurnal radiation effects are not always consistent so there will be additional biases arising from this source. Figure 5 shows the mean 12Z analysis increment in surface pressure for May 1981 (operational data). There is a clear zonal wavenumber two structure in the tropics and subtropics

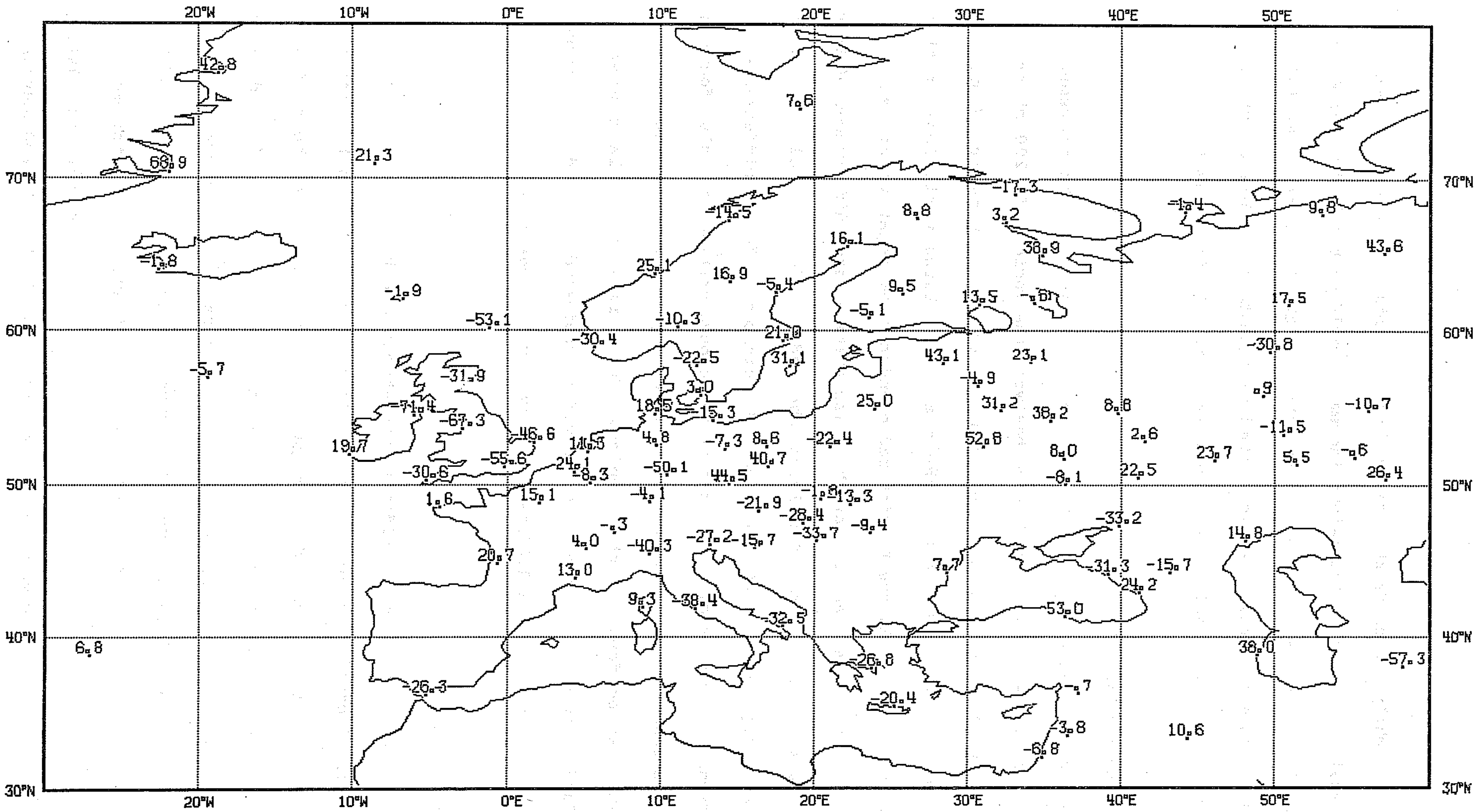

Fig. 4 Mean difference between the radiosonde 200 mb height and the first guess 200 mb height for the February 1979 period. Units in metres.

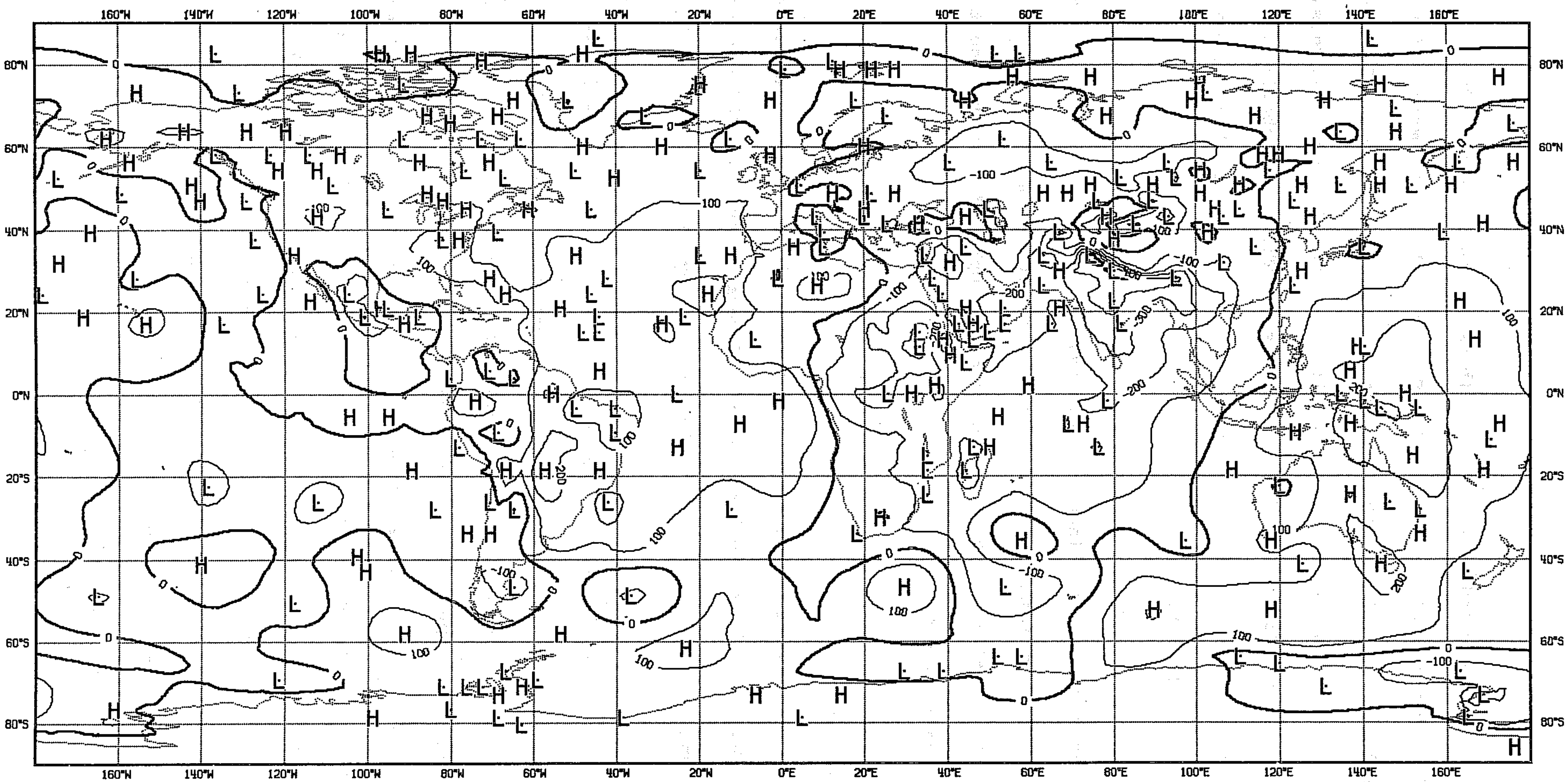

Fig. 5 Mean increment in surface pressure for operational 12 GMT analyses in May 1981.

with amplitudes of the order of a few millibars. This is clearly due to the absence of a diurnal cycle in the model. Figure 6 shows the mean analysis increments for the 0Z and 12Z 500 mb height analyses for the April 79 period. There is a clear diurnal signal present with the mean increment reversing sign between 00 and 12Z over the main land masses. This is probably a real signal since radiosonde radiation corrections are not thought to be significant below 200 mb.

A likely consequence of this lack of a diurnal cycle is an increase of the correlation length when we come to calculate first guess error structure functions. Calculations of the first guess error amplitudes must also take account of these tidal effects.

4.3 Model errors

In addition to the contributions to the mean analysis increment already alluded to there is a contribution from mean forecast errors. These cannot be removed unless the forecasts themselves can be improved.

One may well ask to what extent forecast errors in 6-hour forecasts can be identified at upper levels given the difficulties that have been indicated with some of the observing systems. Unambiguous identification of 6 hour forecast errors over Europe at upper levels is complicated by the large variations in radiosonde biases. However, over North America, the radiosonde network is much more homogeneous than over Europe. There are clear indications (Lonnberg, pers.comm.) that there are geostrophic relations between the patterns of mean observation minus forecasts fields for height and winds over North America. This geostrophic connection would indicate that these difference patterns indicate forecast errors rather than variability of radiosonde quality.

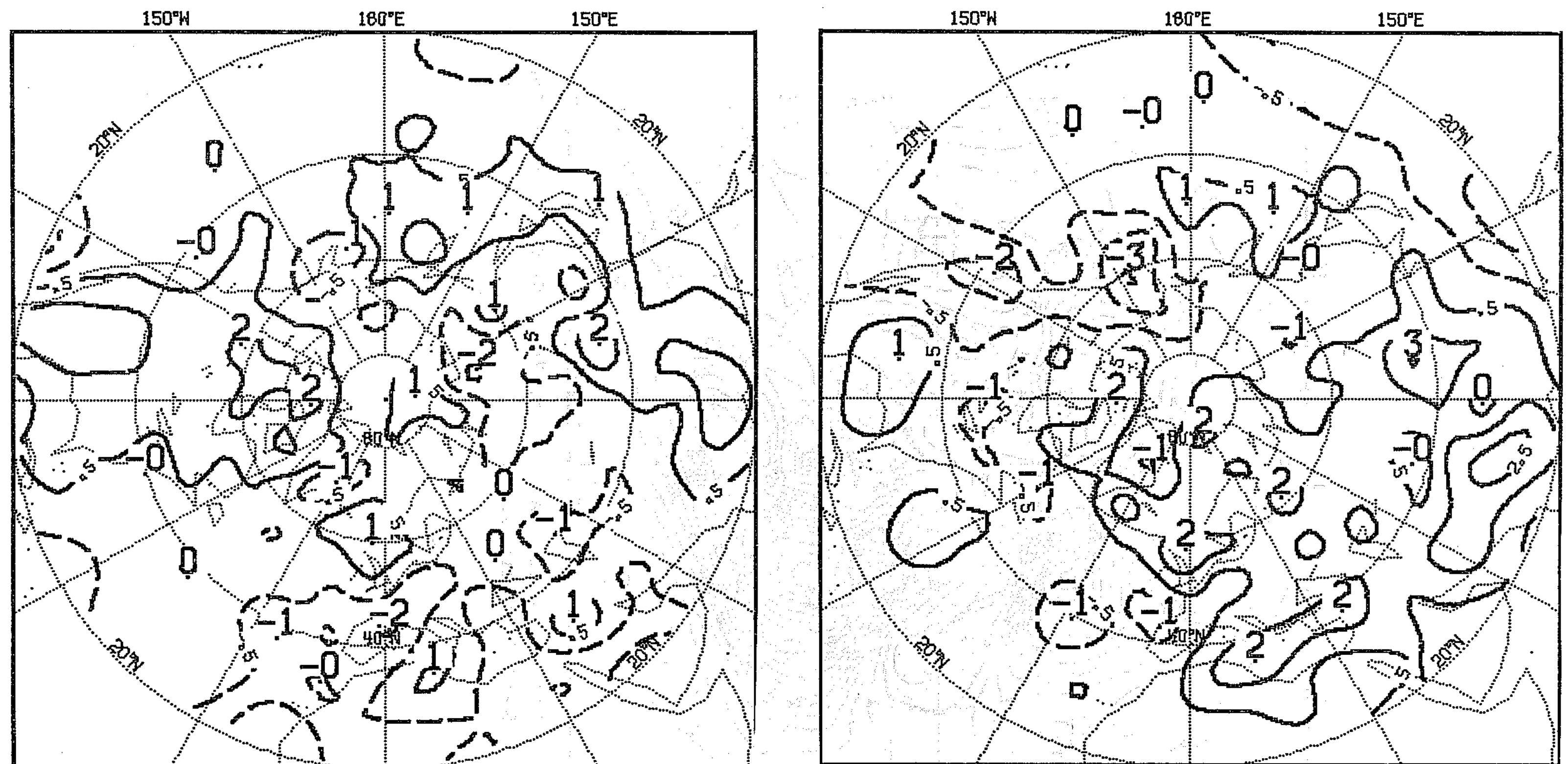

Fig. 6 Mean analysis increment in 500 mb geopotential during the April 1979 period at 12 GMT (left) and 00 GMT (right).

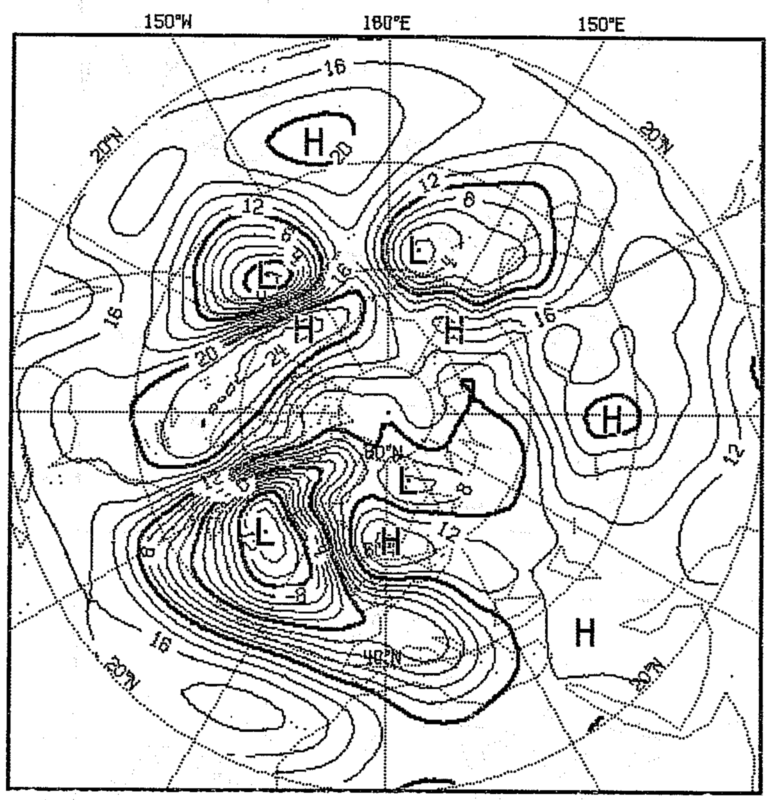

Fig. 7 Mean surface analysis at 1000 mb for the period 8-19 February 1979.

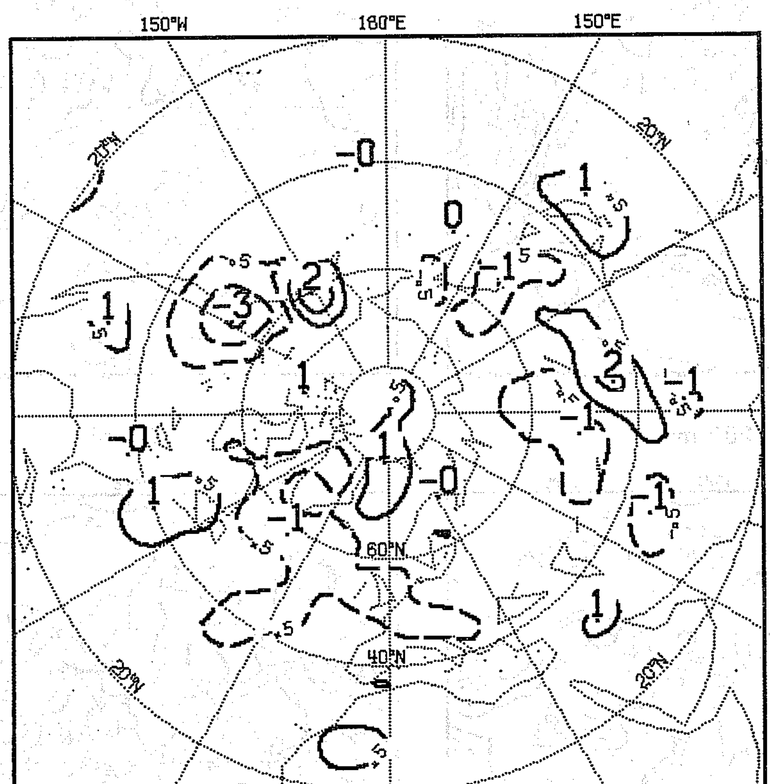

Fig. 8 Mean error (forecast-observed) in the 6 hour forecasts for the February period.

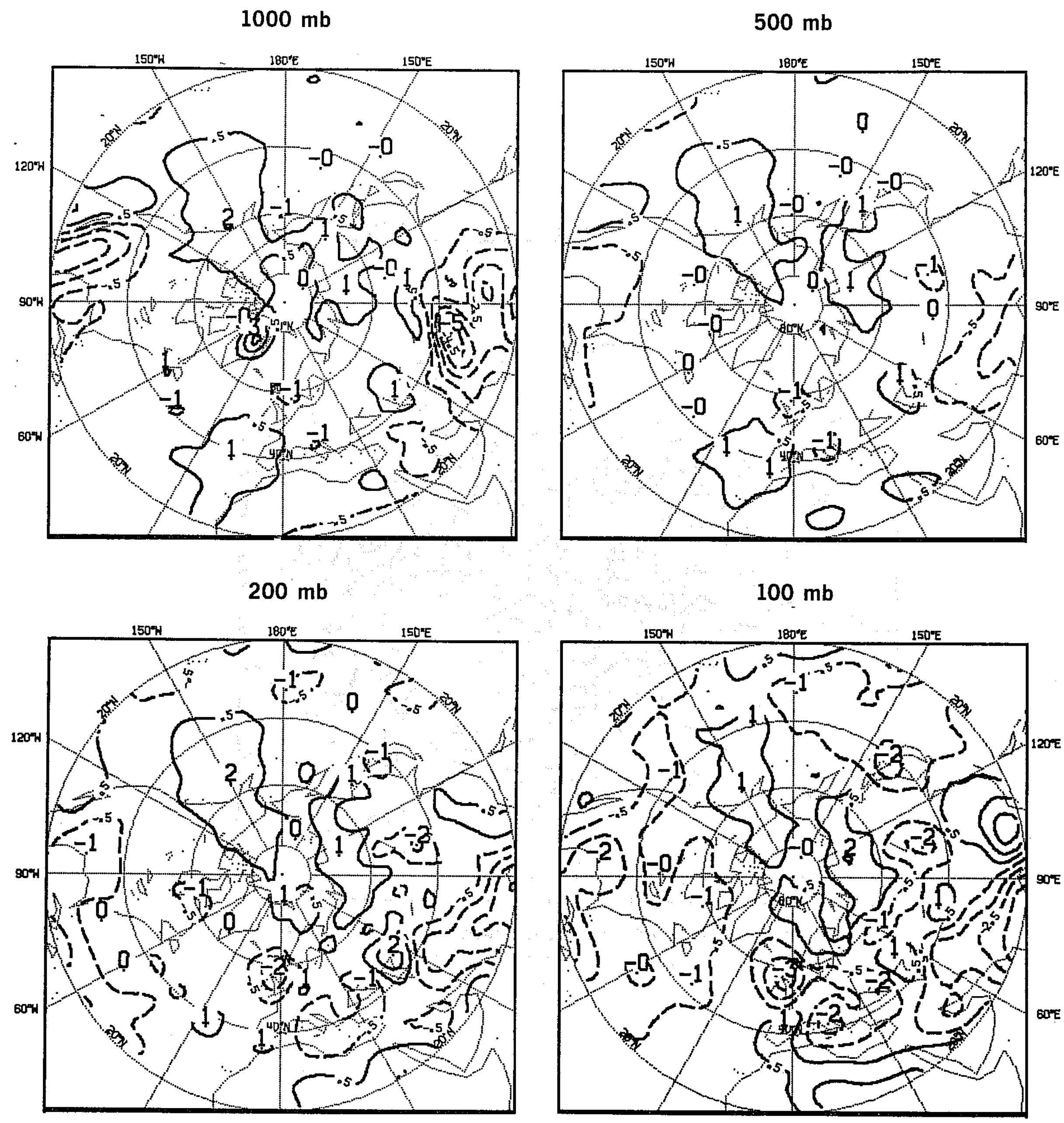

Fig. 9 Mean initialisation increment in geopotential at 1000, 500, 200 and 100 mb for the April 1979 period.

Unambiguous identification of forecast errors is also possble at the surface. Figure 7 indicates the mean 1000 mb height field for the February 79 period. Over North America there is a typical winter situation with low pressure in the Gulf of Alaska and a cold high in the lee of the Rockies. The mean 6-hour forecast error (Fig. 8) shows a clear dipole error field which indicates that the low tends to extend inshore. This error pattern, in a much more intense form is often found to some degree in the surface forecast fields for much longer ranges.

4.4 Summary

We have tried to demonstrate in this section that several different contributions to bias in the analysis increment can be identified. We have by no means exhausted the list of possible conributions. One further contribution will be tentatively identified in the next section on initialization increments.

5. MEAN INITIALIZATION INCREMENTS

In the northern hemisphere height fields the rms initialization increments are dominated by the mean increment. In the troposphere these increments have a strong equivalent barotropic component associated with mean changes in surface pressure. This equivalent barotropic structure can be seen clearly over the Eastern Pacific at 1000, 500 and 200 mb for the April period (Fig. 9). A global view of the mean surface pressure initialization change is given in Fig. 10, calculated from the operational May 81 sigma files. Mean changes in the northern hemisphere extratropics do not exceed 2 mb and are below 1 mb in most places. The mean changes reach their maximum in the subtropics.

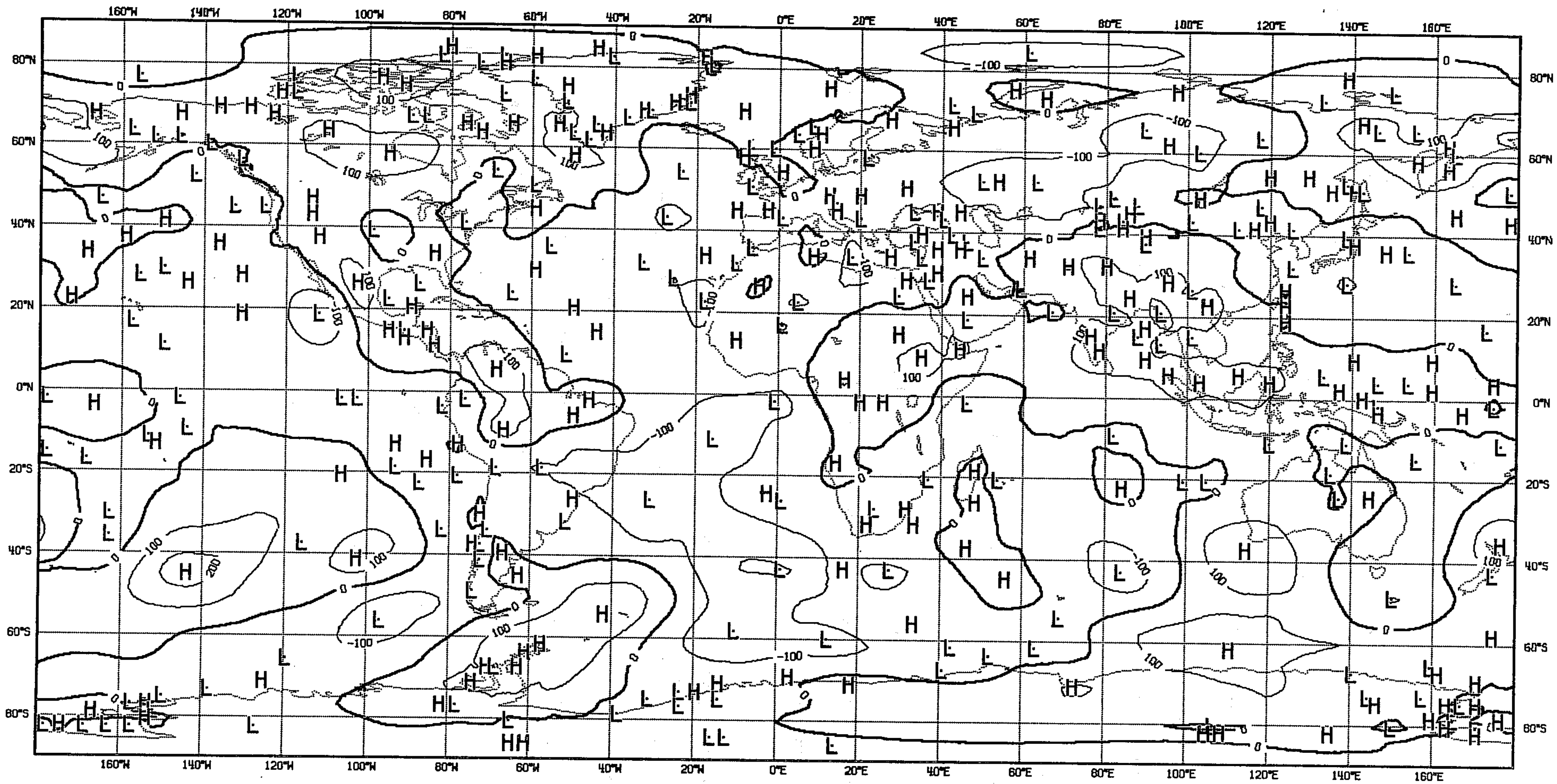

Fig. 10 Mean initialisation increment in the operational 12 GMT analyses for May 1981. Contour interval is 100 Pa (= 1 mb)

5.1 Tidal fields

The net effect of the initialization is usually to bring the analysis closer to the first guess. Figures 5 and 10 support this view particularly in the tropics and subtropics. It is tempting to speculate that the initialization is rejecting some or all of the tidal data. Thermal tides cannot occur in the model and the adiabatic initialization cannot preserve them. The spatial structure of free and forced modes are quite different. Thus forced structures such as the thermal tides will probably have projections on both the free gravity waves and the free Rossby waves. The initialization will then suppress the gravity wave projection but leave the Rossby wave projection alone.

Figures 5, 10 suggest that it would be quite interesting to extract the tides from the monthly mean sigma files and to make experiments on initializing the tide itself and initializing the analysis with the tide removed.

Such experiments could perhaps justify the following scheme to allow the diurnal effects to be treated better in the analysis system without the difficulties associated with an accurate fine tuning of the diurnal cycle in the model. One would calculate the monthly average diurnal (and semi-diurnal) fields from last year, add them to the background field, perform the analysis, remove the same tidal field and then initialize and resume the assimilation. Such a scheme would offer most of the advantages of having a better first guess in the tropics at little cost. We will not need to implement this scheme if all goes well with the testing of the diurnal cylce.

In any event it would be of interest to establish the extent to which tidal effects are manifested in the mean initialization fields in order to understand the reasons why the initialization tends to bring the analysis closer to the first guess. If such a tendency is intrinsic to the initialization then a study of its mechanism could shed light on mean forecast errors.

5.2 Stratospheric mass field

We turn now to consider the mean effect of the initialization in the stratosphere. Comparison of Figs. 3 and 9 at 100 mb show significant compensation between the mean analysis increment and the mean initialization increment over the Himalayas.

It appears as if the analysis is introducing height data which is promptly rejected by the initialization scheme. A more detailed examination indicates that it is stratospheric temperature data which is being rejected.

Figure 11a,b shows the mean initialization increment for April 81 in the temperature of the top two sigma-levels of the model (at σ = .025, .077). These pictures show a clear indication of large changes occurring in the areas of steepest topography. The corresponding maps for the analysis increment are extremely similar but with opposite sign. Sundquist (1975) has pointed out that truncation errors in the pressure gradient calculation in σ-coordinates increase rapidly as one moves from the troposphere across the tropopause into the stratosphere. Expressed in terms of equivalent geostrophic winds the errors can easily reach 5 m/s. There is good reason to believe that the temperature changes shown in Fig. 11 are associated with just such a truncation error. Phillips (pers.comm.) has found that the problem described by Sundquist does indeed arise with significant amplitude in a multilevel pe model. Simmons (pers.comm.) indicates that the use of special hybrid coordinate systems can considerably reduce the truncation error identified by Sundquist. We need to check if the use of the hybrid coordinate in the assimilation will reduce the magnitude of the temperature changes. If this proves to be the case then it will be a good indication

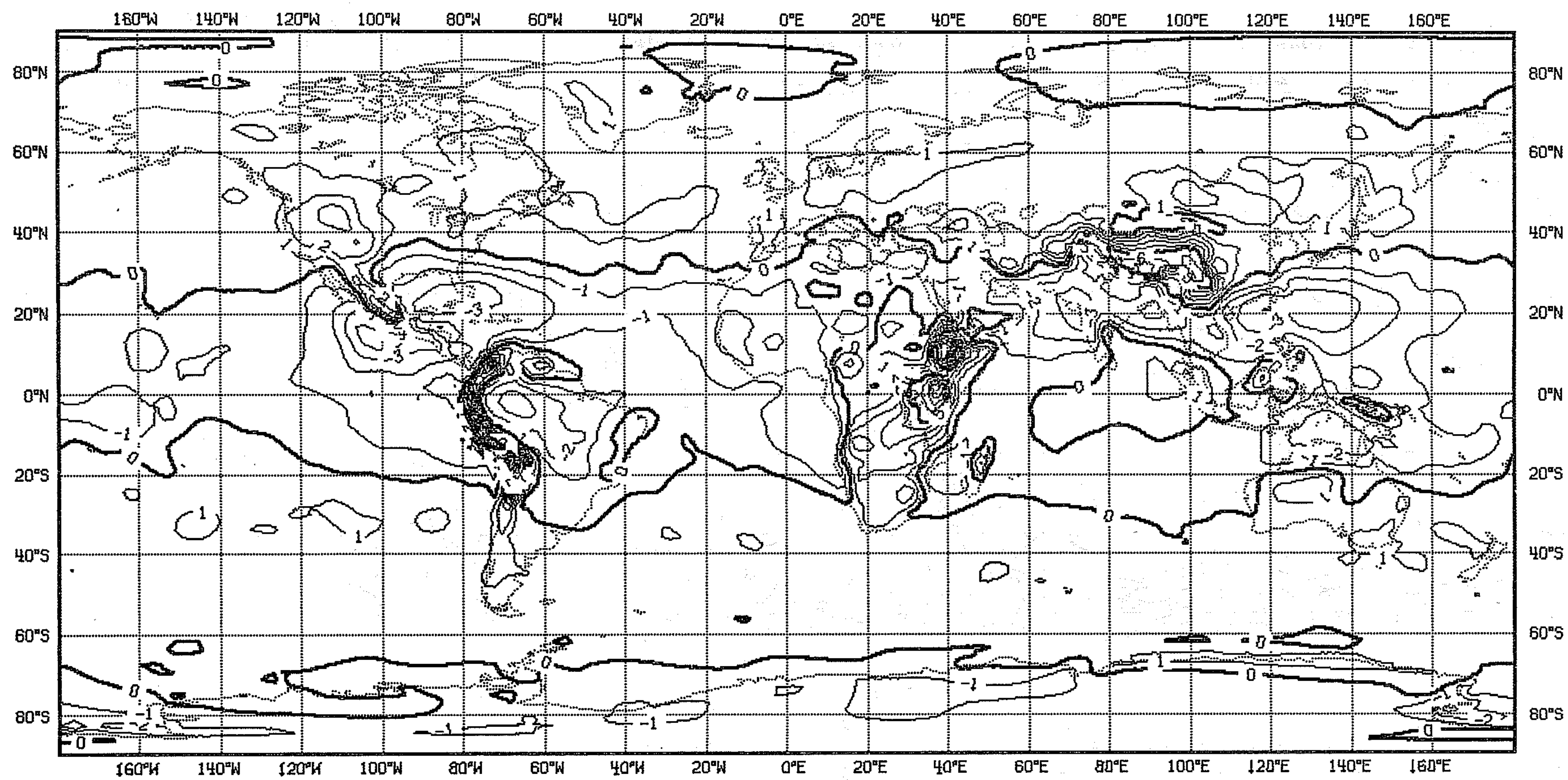

Fig. 11a Mean initialization increment in temperature at the top, $\sigma = 0.025$, sigma level for operations in April 1981.

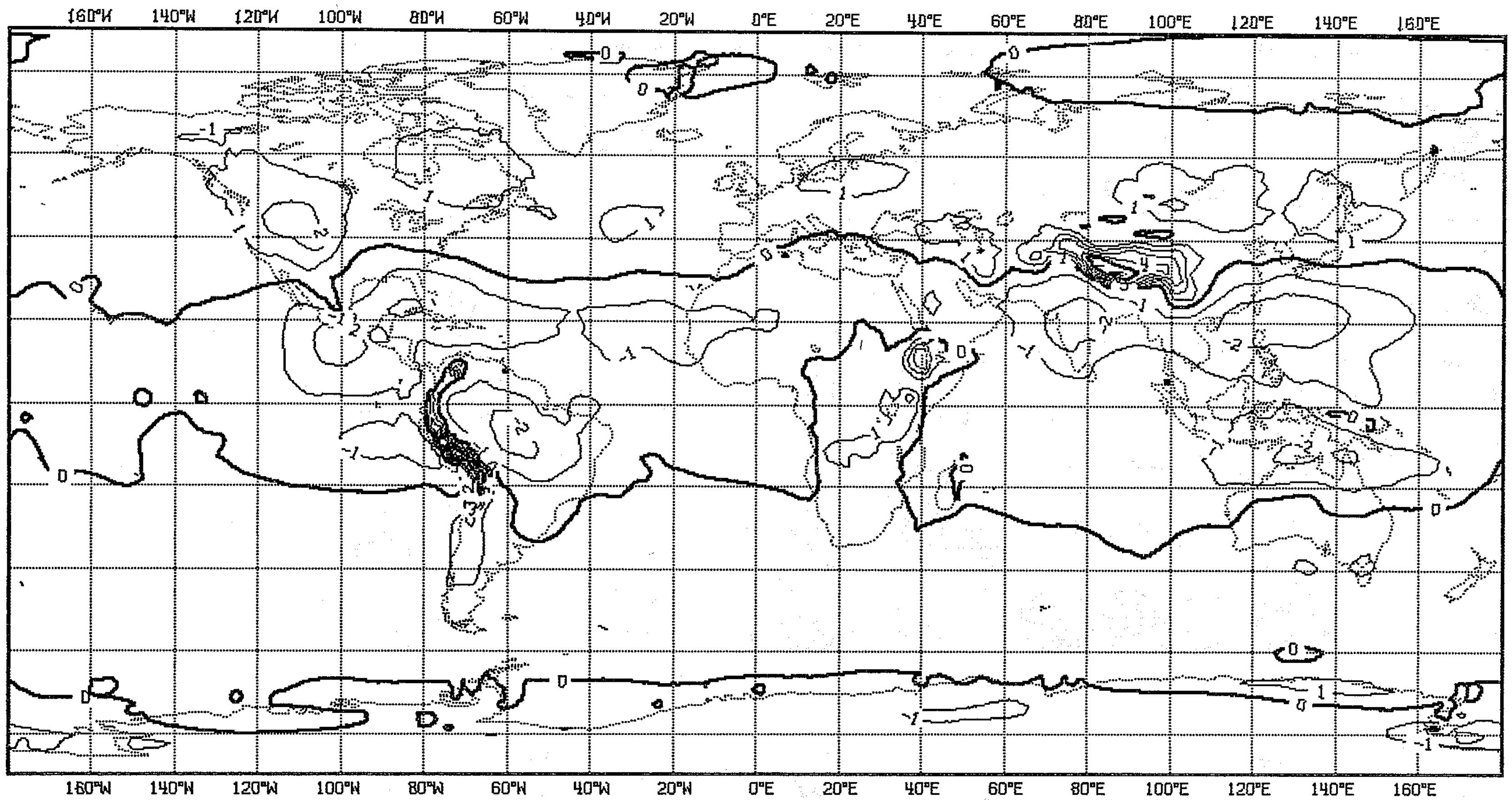

Fig. 11b Mean initialization increment in temperature at the second highest, $\sigma = 0.77$, sigma level for operations in April 1981.

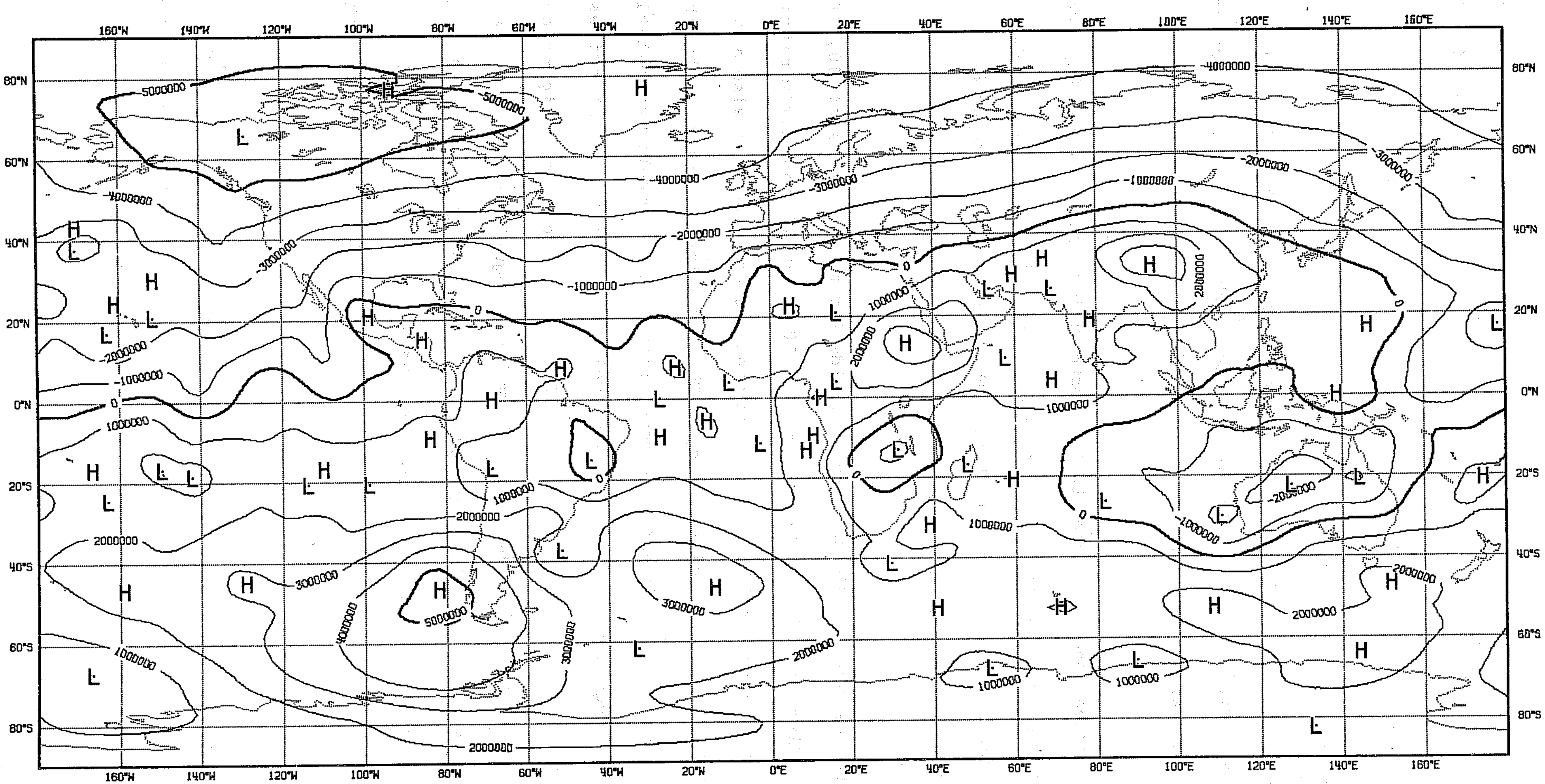

Fig. 12 Mean stream function initialisation increment at 50 mb for operational 12 GMT analyses for May 1981. The contour interval is $10^6 m^2 s^{-1}$ so that 1 interval in 10 degrees of latitude corresponds to 1 m/s.

that the problems identified here are indeed due to truncation error in the sigma system.

5.3 Stratospheric wind fields

This problem over the mountains is not the only problem we have found in the stratosphere. Figure 12 shows the mean change in the 50 mb stream function, due to initialization, for May 1981.

There is a marked zonal character to the northern hemisphere change (typical magnitude about 2 m/s in wind). The reason for this is unclear and we can only offer suggestions for investigation. The same feature is found in initializations of individual data sets.

In our system we generate the background fields for the analysis at 10, 20, 30 mb by adding the 10-50, 20-50 and 30-50 mb wind and temperature differences to the models 50 mb height and wind fields. There is also a weak (about 2%) blending with climatology. Thus the increments made by the analysis are not necessarily entirely consistent with the models hydrostatic equation. This equation for the top σ-level has a certain arbitrariness because we are at the boundary of the grid. If the analysed heights and winds are inconsistent with the models thermal wind equation then some part of the analysis increment will be rejected. A study of the sensitivity of the initialization changes at 50 mb to the arbitrary element in the hydrostatic equation at the top level would be interesting and perhaps illuminating.

For the moment we have no explanation for the difference in character of the results from the northern and southern hemispheres but perhaps it is related to the difference in the climatologies at that time of year.

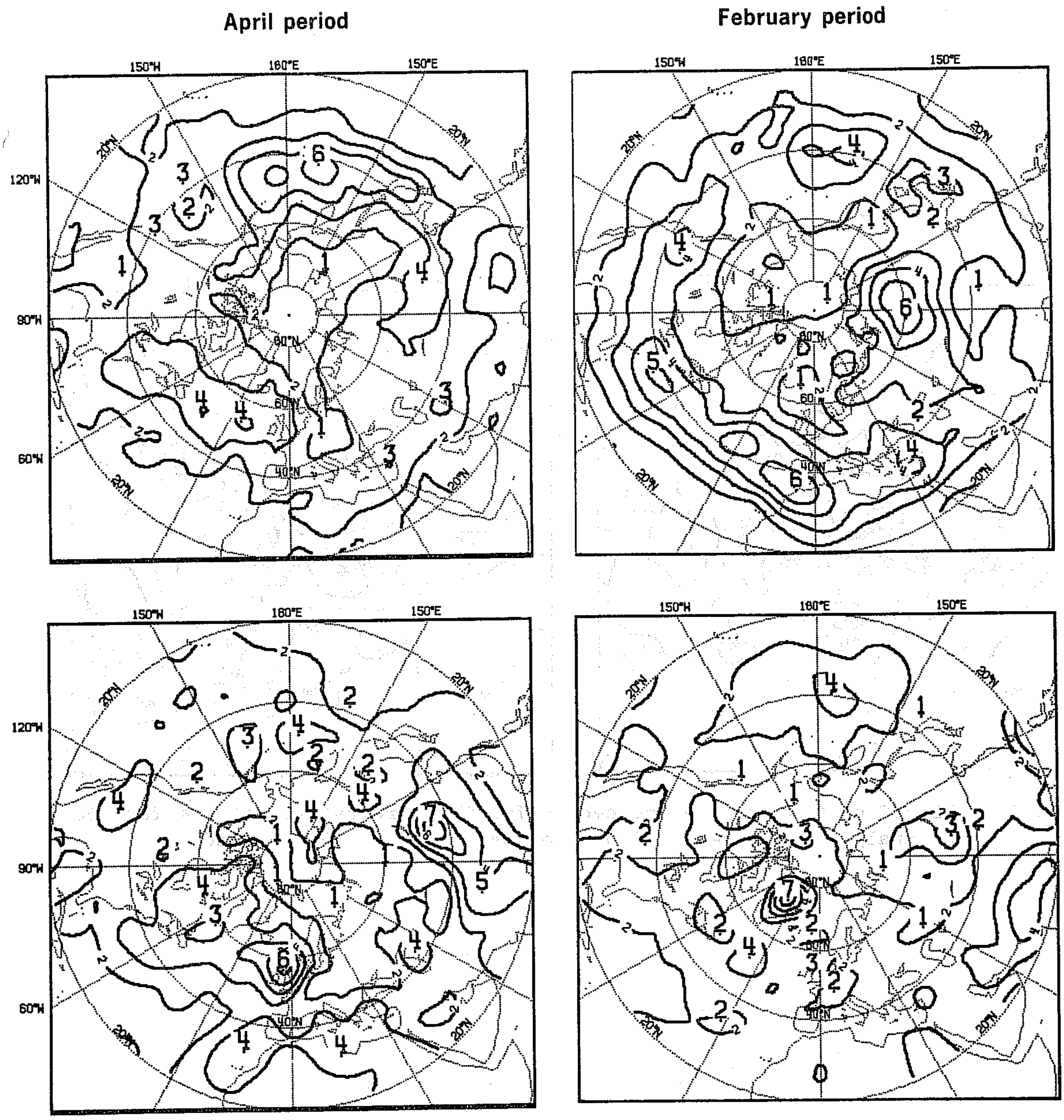

Fig. 13 Maps of the RMS 200 mb height increment for the forecast increments (top) and analysis increments (bottom) for the April 1979 period (left) and the February 1979 period (right). Units are dam.

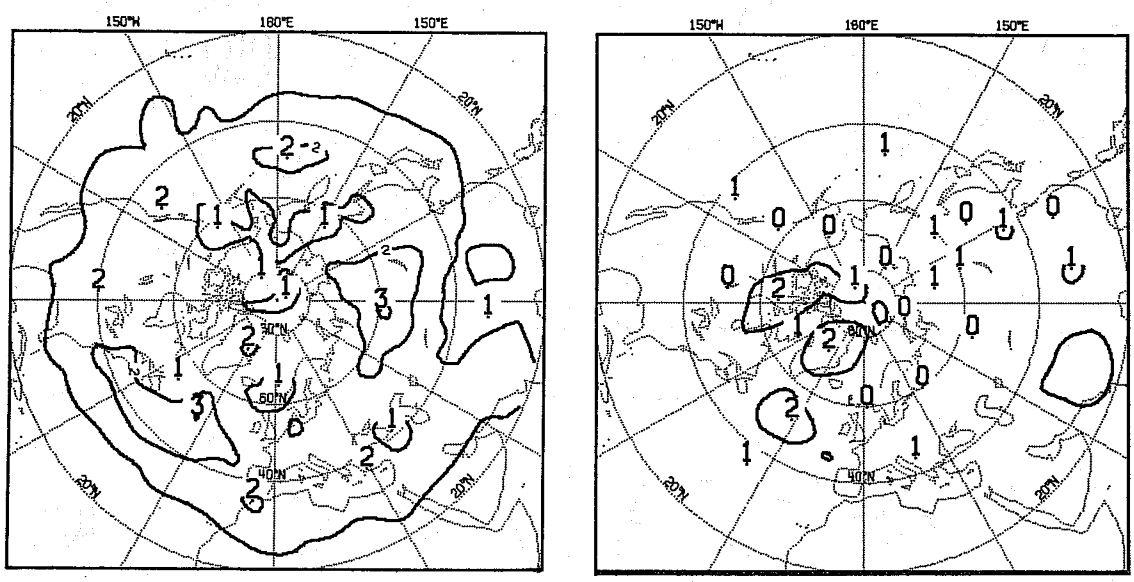

Fig. 14 RMS 500 mb temperature increments for the forecast (left) and for the net analysis increment (right) for the February 1979 period.

6. FIT TO OBSERVATIONS

So far in this study we have not considered the extent to which the analysis fits the observations. This is a large area and will be the subject of later reports. However, it is appropriate to consider the extent to which the patterns of variance in the forecast, analysis and initialization increments reflect the impact of the observations.

Figure 13 shows the rms 200 mb geopoential increments for the forecasts and analyses for both the April 79 and February 79 periods. The maxima in the forecast increments clearly reflect the storm tracks in the respective periods. The most striking features of the analysis increments, the bullseyes over the Northeast Atlantic, have already been identified as due to radiosonde problems. For our present discussion we wish to focus on the relative magnitudes of the analysis and forecast increments over the oceanic storm tracks. There is no clear correspondence of the patterns of the distributions. For example, in the Western Pacific along 40 N in April the forecast increment is over 4 dam while the analysis increment is close to or less than 2 dam. Again, over the Eastern Atlantic in February there is a clear storm track along 40 N with rms forecast increments in excess of 5 dam while the rms analysis increments are between two and three dam.

The fact that the patterns do not correspond is a little disturbing and might suggest that we are not using optimally the available aircraft data and more importantly the satellite temperature data.

One might be tempted to draw the same conclusion for the 500 mb temperature field (Fig. 14) where the areas of largest rms temperature increment from the analysis do not occur in the storm tracks. The only solution to this impasse is that the forecast temperature changes are in fact accurate.

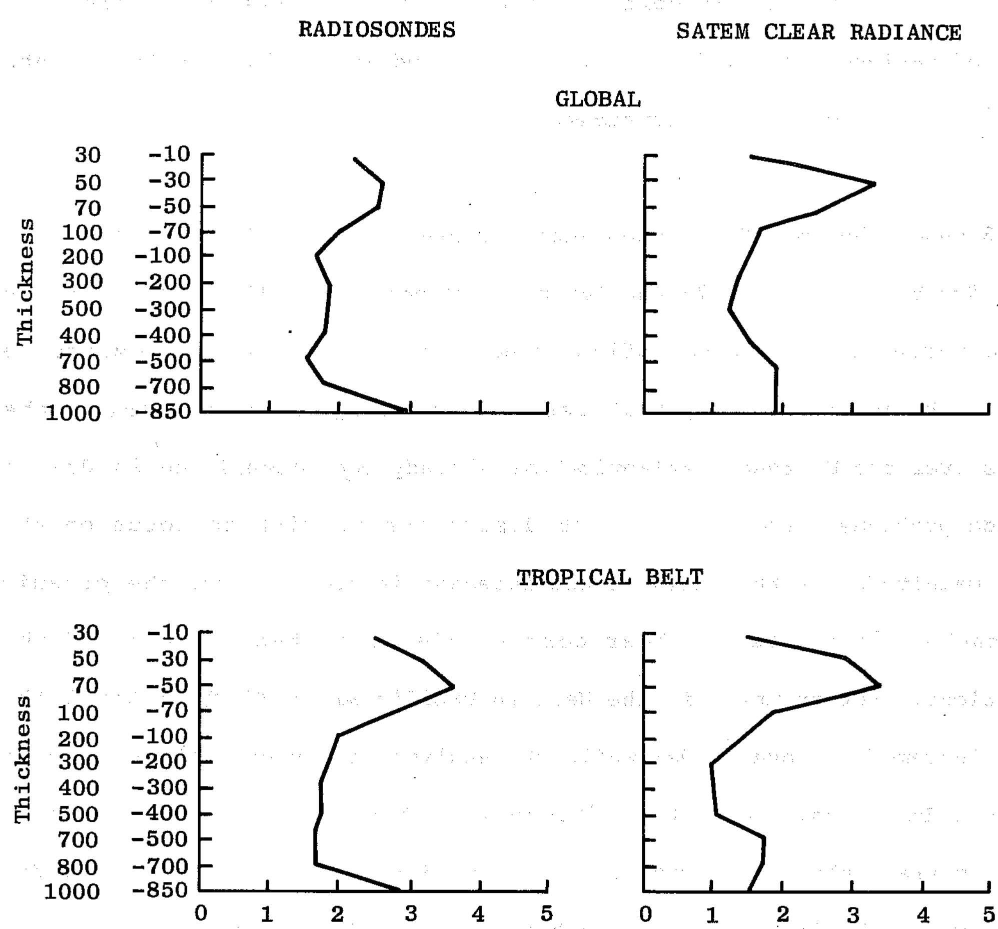

Fig. 15 Plots of the fit of the first guess to accepted virtual temperature reports from radiosondes (left) and SATEM clear radiances (right) for the globe (top) and for the tropical belt (bottom). The results are averaged over the FGGE production for June 1979. (Björheim pers. comm.)

This interpretation is supported by the results of Bjorheim (pers.comm.) for the FGGE analyses for June 1979 (Fig. 15). These results suggest that the first guess fits the SATEMS very closely(~1.2 degrees) in the upper troposphere.

This optimistic deduction must be treated with caution, pending further investigation, because of the known fact (Schlatter 1981, Phillips 1979) that the SATEM data tend to underestimate the atmospheric temperature variance.

7. SUGGESTIONS FOR FURTHER WORK AND SUMMARY

The results presented above suggest a number of areas where further work is necessary to remove, or investigate further, significant biases in the assimilation system.

1) Clearly we need to correct the radiosonde height data so as to reduce as much as possible the impact of radiosonde bias on the height and wind analysis. This will inevitably require separate treatment of 00Z and 12Z dates.

2) We need to study the effect of initialization on the thermal tides in order to estimate the extent to which they are being aliased onto Rossby modes in the present assimilation system (which has no diurnal cycle). This work may also explain, at least in part, the clear tendency for the initialization to bring the analysis closer to the first guess.

3) We must examine the effect of initialization on the stratospheric temperature field over mountains in the hybrid coordinate system. We expect that the results will show an improvement on those found in σ-coordinate models.

4) We need to study the sensitivity of the initialization changes in the 50 mb wind to the form of the models hydrostatic equation in the top σ-level. Here again we will expect an improvement with the hybrid model.

Overall the results clearly indicate that the model is an essential ingredient in the assimilation. Our conclusion on the relative magnitudes, F > A > I, will undoubtedly be strengthened if we can cure some of the biases listed above. In particular the hybrid coordinate and the diurnal cycle should help to reduce the biases. However we feel a little uncomfortable with the fact that the forecast increments are so much larger than the analysis increments over the oceanic areas. This will be a particular focus of the upcoming SATEM impact study.

Finally we mention that the existence of these biases must be taken into account in the calculation of first guess error statistics where they will have important consequences on the correlation functions and the error amplitudes.

Acknowledgements

S. Uppala, P. Lonnberg, D. Shaw and W. Wergen were generous with their time in helping with various aspects of this work.

References

McInturff et al 1979 Day-night differences in radiosonde observations of the stratosphere and troposphere. NOAA Technical Memorandum NWS NMC63.

Schlatter, T.W. 1981 An assessment of operational Tiros-N temperature retrievals over the United States. Mon.Wea.Rev., 109, 110-119.

Sundquist,H. 1975 On truncation errors in sigma system models. Atmosphere, 13, 81-95.