DIAGNOSTIC EVALUATION OF ANALYSIS AND FORECASTS:
CLIMATE OF THE MODEL

by

K. ARPE

1. INTRODUCTION

Medium range forecasts have been produced operationally at ECMWF for three years. Although considerable improvements have been made during this time many of the main problems discussed by Hollingsworth et al (1980), in a study of early forecasts, still remain. The main object of this lecture is to describe these problems in more detail. As the analyses are a very important component of the whole system, some diagnostic studies of the analyses are also included.

2. SYSTEMATIC ERRORS OF THE MODEL

The model has a large number of variables, at many levels, but the most commonly studied is still the 500 mb height field and therefore we will concentrate on this. It will, however, be shown that the systematic error in one variable is in many cases connected with errors in others and it is often difficult to judge cause and effect.

2.1 500 mb height field

Shown in Fig. 1 top panel are the analyzed 500 mb height fields averaged for all three available winters and the corresponding 7-day forecasts. The forecasts are, on the whole, more zonal, having less intense ridges over the eastern Atlantic and over the Rockies. The most important error in the European area is an underprediction of the diffluent flow.

The differences between forecast and analysis fields are shown in the middle left panel which displays a dipole structure with the forecast height field being too low over Europe and Alaska. This is a very persistent pattern in the monthly mean error fields. The remaining panels show examples for January 1980, 1981 and 1982. Although there are variations from month to month, they all have negative values around the British Isles and around Alaska. There is an increase in the extrema of the errors in the day 10 forecast but they still remain much smaller than those found by Hollingsworth et al (1980). The pattern, however, has hardly changed. They had two series of experiments with

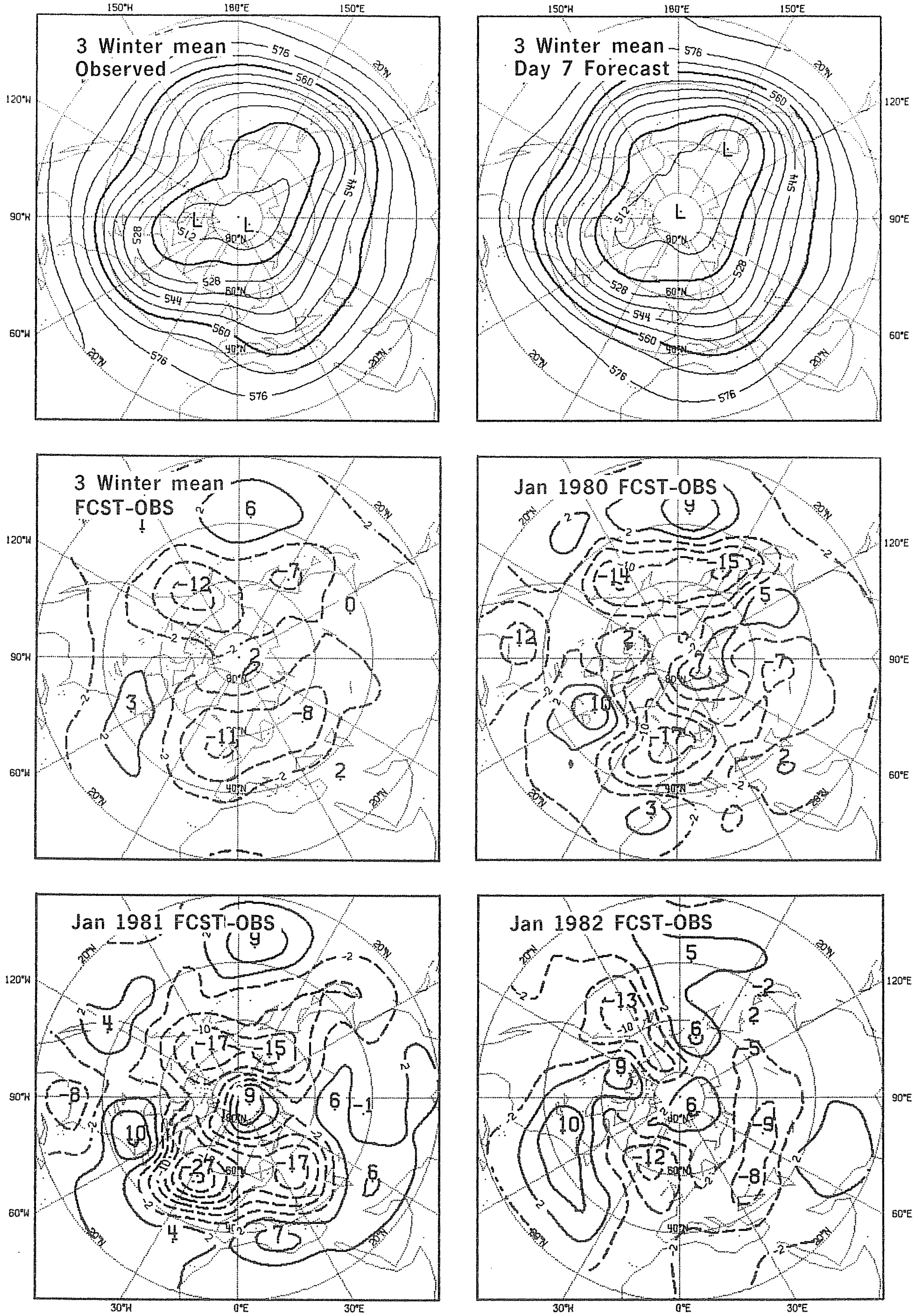


Fig. 1 Mean 500 mb height fields at the Northern Hemisphere. Units: 10 m.

quite different parameterizations of subgrid scale processes but, nevertheless the systematic error in both series was quite similar.

The systematic error depends on the initial conditions only in so far as they might influence its growth rate. All forecasts which have been carried out beyond 10 days (mostly up to 50 days) exhibit a systematic departure from climatology which is similar to the error field shown here (Tiedtke, 1982).

Comparative forecasts performed with the operational gridpoint model and with an experimental spectral model by Jarraud et al (1981) showed that the horizontal differencing scheme is not responsible for this error. The only dependency was found when Cubasch (1981) investigated different horizontal resolutions. His model with the lowest resolution, a spectral T21, had clearly smaller amplitudes in the error pattern than all the higher resolution models. Other (operational) models show similar error patterns (Bengtsson and Lange, 1981). The systematic error is clearest in winter and we will therefore concentrate on this season.

The southern hemispheric winter forecasts have monthly mean errors which are of the same order of magnitude as those occurring in the northern hemisphere, as shown in Fig.2. In July 1981 the mean error is similar in amplitude to the northern hemisphere mean error due to a more zonal circulation in the forecasts. In July 1982, however, the analysed circulation is already very zonal and the mean forecast circulation is less zonal. In particular, the error near South Africa is due to a more intense trough in the forecasts. This indicates that the systematic error may be due not so much to a lack of forcing but due to an incorrect form of forcing.

2.2 Prospects for error correction through statistical methods

As the monthly mean error pattern is so persistent, it seems worthwhile to investigate a statistical correction of the model output. The simplest way

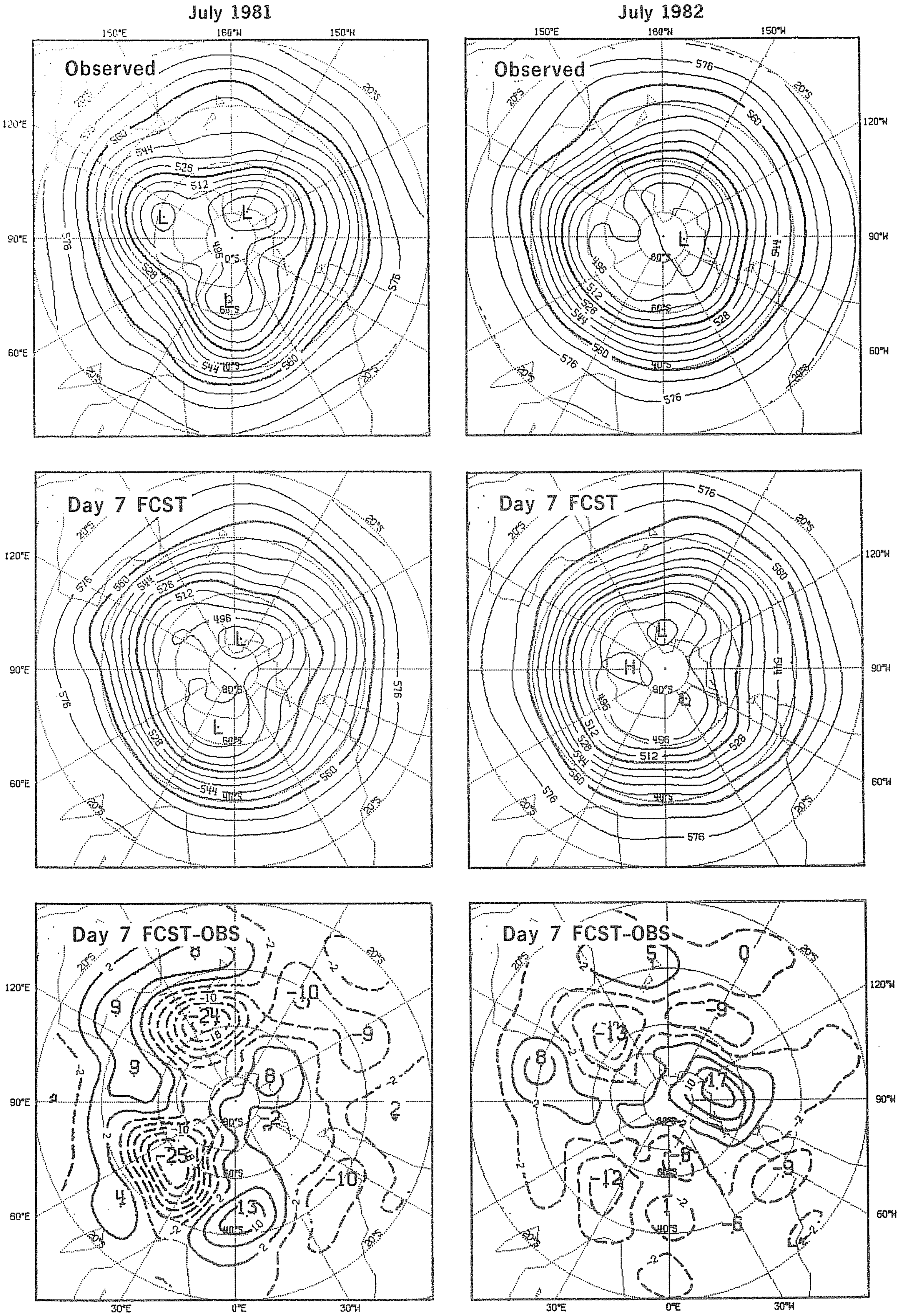


Fig. 2 Mean 500 mb height fields at the Southern Hemisphere. Units: 10 m.

would be to subtract a mean correction field. Fig.3 shows the impact of such a correction on the monthly mean skill scores for January 1981. Three anomaly correlation coefficients of the 500 mb height fields are shown. The short-dashed line gives the scores of the direct model output. It crosses the 60% level after 5 ¼ days, which is often regarded as the range of predictability. The long- dashed line shows the scores when the mean error of the same month is taken away from each forecast. This is of course of no practical use because it would not be available for operational work but it shows the maximum possible gain with such a simple correction method.

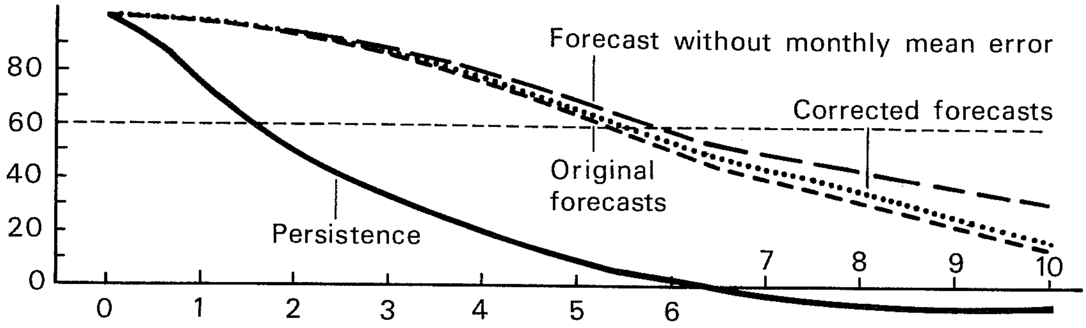


Fig. 3 Anomaly correlation coefficients of the 500 mb height field north of 20°N during January 1981.

———— persistence forecasts
– – – – – original forecasts
……………… forecasts after correction with the winter 1979/80 mean error
— — — forecasts after removing the monthly mean error of the same month.

The dotted line shows the improvement by subtracting the previous winters (1979/1980) mean error. Small increases of the scores throughout the forecast period are obvious, increasing the predictability by ¼ day. Only the very long waves are noticeably improved while shorter waves are hardly affected.

In Fig. 4 the mean 1000 mb height fields for January 1981 and the impact of the correction is shown. The main errors of the day 7 forecasts are that the Icelandic low exends too far to the east instead of to the north east, the anticyclone over the western North America is too weak and the Aleutian cyclone is also extended to the east instead of to the north east. On all these points

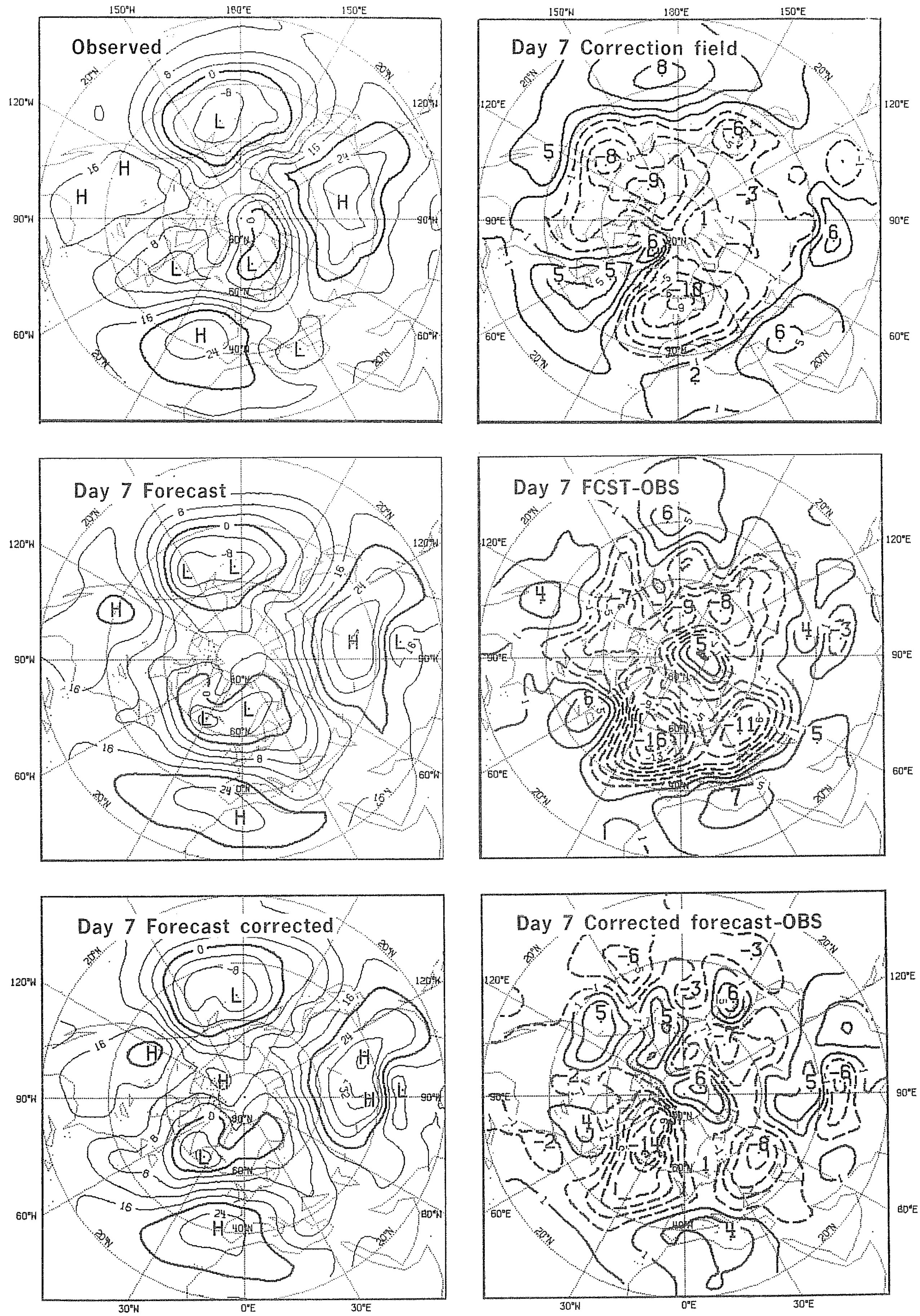


Fig. 4 January 1981 mean 1000 mb height fields. Units: 10 m. The correction field is the mean error field during Winter 1979/80.

the correction gives an improvement. The two lower right panels, indicating differences between the forecasts and the analyses, show further the overall improvement with only small areas where the forecasts have been deteriorated.

The effect of the correction on the scores of single forecasts is illustrated in Fig. 5 as a scatter diagram. Scores on the diagonal (dashed line) indicate that the skill was unaffected by the correction, scores below the diagonal indicate an improvement and those above a deterioration of the forecast skill after applying the correction. The better forecasts (anomaly correlation coefficients of more than 50%) were hardly affected while the bad forecasts

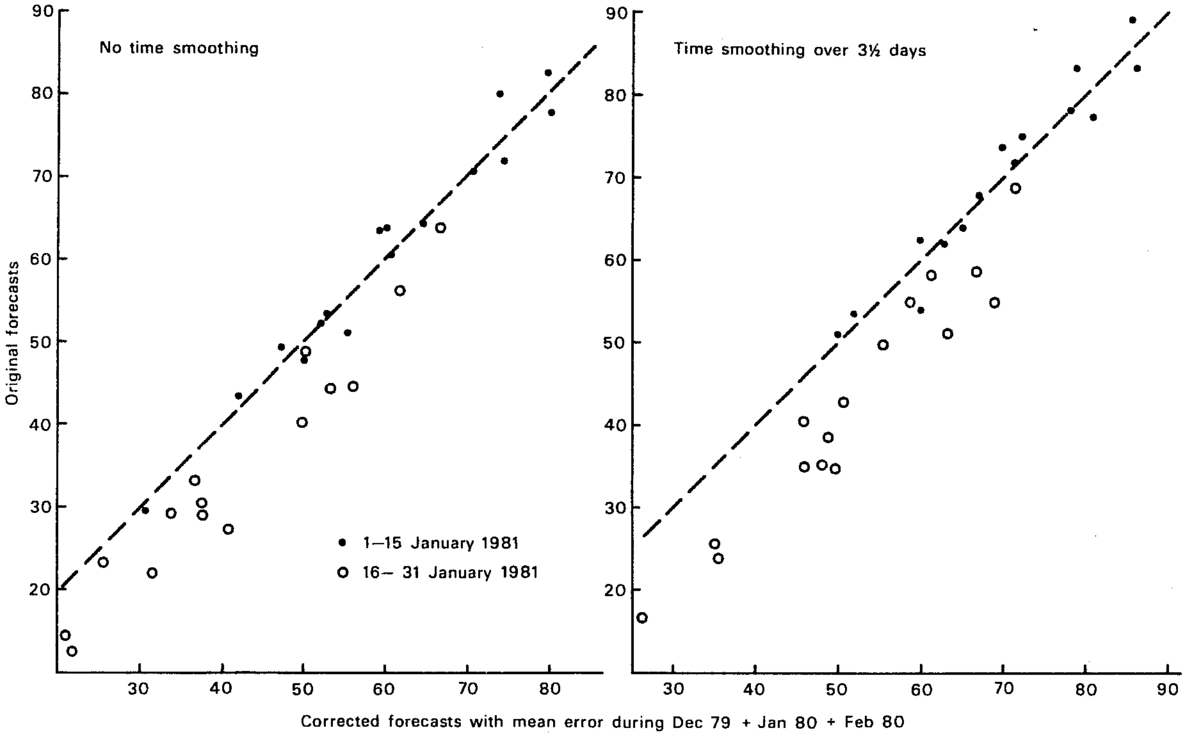


Fig. 5 Anomaly correlation coefficients of the 500 mb height field of the day 7 forecasts during January 1981. Original forecasts are compared with corrected ones.
Forecasts which attained a higher score after the correction are positioned below the dashed line and those with lower scores are placed above.
Left panel: coefficients of daily data.
Right panel: coefficients for $3\frac{1}{2}$ day mean fields.

generally show increases of the scores of the order of 10%. This tendency is clearest in the scores for time averaged fields shown in the right panel.

An interesting by-product of this figure are the differences in the scores between the first and second half of the month. The first half of the month, giving higher scores, was dominated by a stationary wave pattern while during the second half the transient waves were more pronounced (and more difficult to predict). Thus there is a sensitivity of the scores to the circulation type.

A similar improvement was gained when the same correction procedure was applied to January 1982 data.

It has been shown earlier that the mean error has a very large scale pattern represented by zonal wavenumbers 1, 2 and 3, while the scales which are affecting our weather most are the shorter baroclinic waves. It is therefore doubtful, as will be demonstrated below, that a correction to the large scale fields would really improve the usefulness of the forecasts.

Fig. 6 shows the number of cyclones (mean sea level pressure minima) in boxes (5° x 10°) during the winter 1980/81 and 1981/82. The top panel with the observed values shows the main cyclone track going from the eastern U.S. coast to the Norwegian Sea, and a second track over the Mediterranean Sea. These cyclone tracks are largely steered by the 500 mb windfield. In Fig.1 it was shown that the mean circulation in the forecasts has serious errors. The influence of this error on the cyclone tracks is demonstrated in Fig. 6 lower panel for the day 10 forecasts. The forecast cyclone track over the Atlantic is much more west-east oriented than in the analyses and penetrates much further into Europe. Cyclone occurrences in the far north and south of Europe are, however, largely reduced. It is not only the cyclone track itself which is controlled by the mean 500 mb flow but also cyclogenesis.

The day 5 forecasts in the middle panel already give an indication of the error but to a lesser extent.

In Fig. 7 the impact of the forecast change in the cyclone tracks on the mean precipitation rate over Europe is shown. The top left panel gives the average precipitation amount from October 1981 to February 1982 as received from the GTS. All observations occurring within a 1.875° x 1.875° box (which is representative of a grid point of the model) are averaged. The lower left panel shows the corresponding day 8 to 9 forecast values. The deficiency in the forecasts, which is connected with the incorrect positioning of the cyclone tracks, can be clearly seen. It leads to much too high precipitation over Britain and central Europe and to an underprediction in the Mediterranean area.

A bias is found over mountainous areas where the model predicts too little precipitation on the windward sides and too much on the lee-sides. The difference maps in the upper right panel shows this clearly over Scandinavia and the Iberian peninsula. The model obviously cannot create such sharp gradients in the precipitation amounts as occur in the real atmosphere. This bias can probably be easily corrected in the model output by a statistical method.

Both deficiencies can be found already in the shorter range forecasts as indicated in the day 3-4 error map in the lower right panel. The bias near mountains is as large as after 8 days (this bias is already of the same order of magnitude after the first 24 hours), while the bias due to poor positioning of the cyclone tracks grows steadily through the forecast.

A correction with a mean error field cannot correct for errors in the transient features and therefore cannot correct individual cyclones. Perhaps more sophisticated methods of statistical correction will be able to improve individual features.

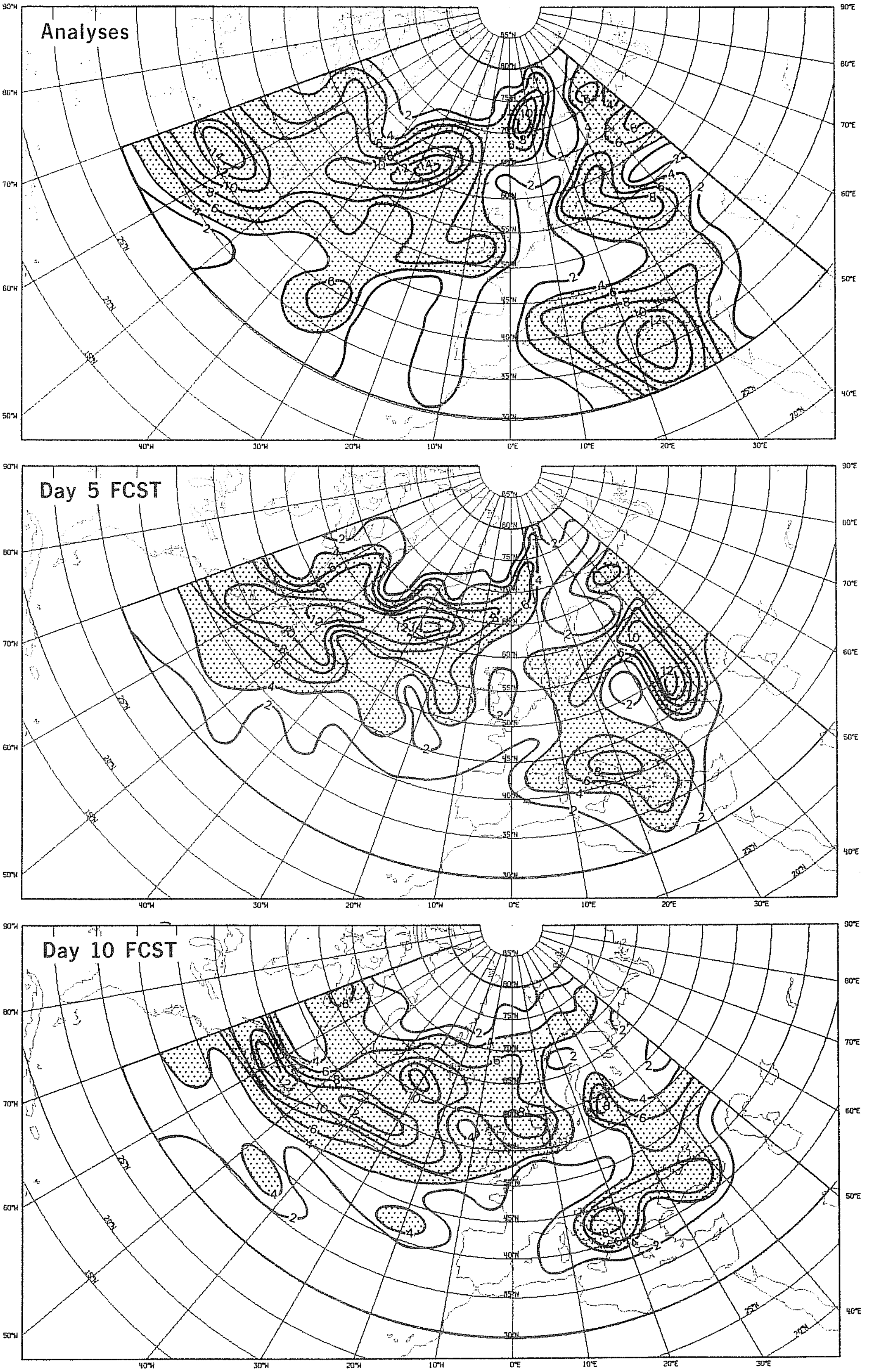


Fig. 6 Occurence of cyclones (positions of relative minima in sea-level pressure) in boxes of 10° longitude x 5° latitude during winter 1980/81 and 1981/82 (from Akyildiz, 1983).

The day 5 forecasts in the middle panel already give an indication of the error but to a lesser extent.

In Fig. 7 the impact of the forecast change in the cyclone tracks on the mean precipitation rate over Europe is shown. The top left panel gives the average precipitation amount from October 1981 to February 1982 as received from the GTS. All observations occurring within a 1.875° x 1.875° box (which is representative of a grid point of the model) are averaged. The lower left panel shows the corresponding day 8 to 9 forecast values. The deficiency in the forecasts, which is connected with the wrong positioning of the cyclone tracks, can be clearly seen. It leads to much too high precipitation over Britain and central Europe and to an underprediction in the Mediterranean area.

A bias is found over mountainous areas where the model predicts too little precipitation on the windward sides and too much on the lee-sides. The difference maps in the upper right panel shows this clearly over Scandinavia and the Iberian peninsula. The model obviously cannot create such sharp gradients in the precipitation amounts as occur in the real atmosphere. This bias can probably be easily corrected in the model output by a statistical method.

Both deficiencies can be found already in the shorter range forecasts as indicated in the day 3-4 error map in the lower right panel. The bias near mountains is as large as after 8 days (this bias is already of the same order of magnitude after the first 24 hours), while the bias due to poor positioning of the cyclone tracks grows steadily through the forecast.

A correction with a mean error field cannot correct for errors in the transient features and therefore cannot correct individual cyclones. Perhaps more sophisticated methods of statistical correction will be able to improve individual features.

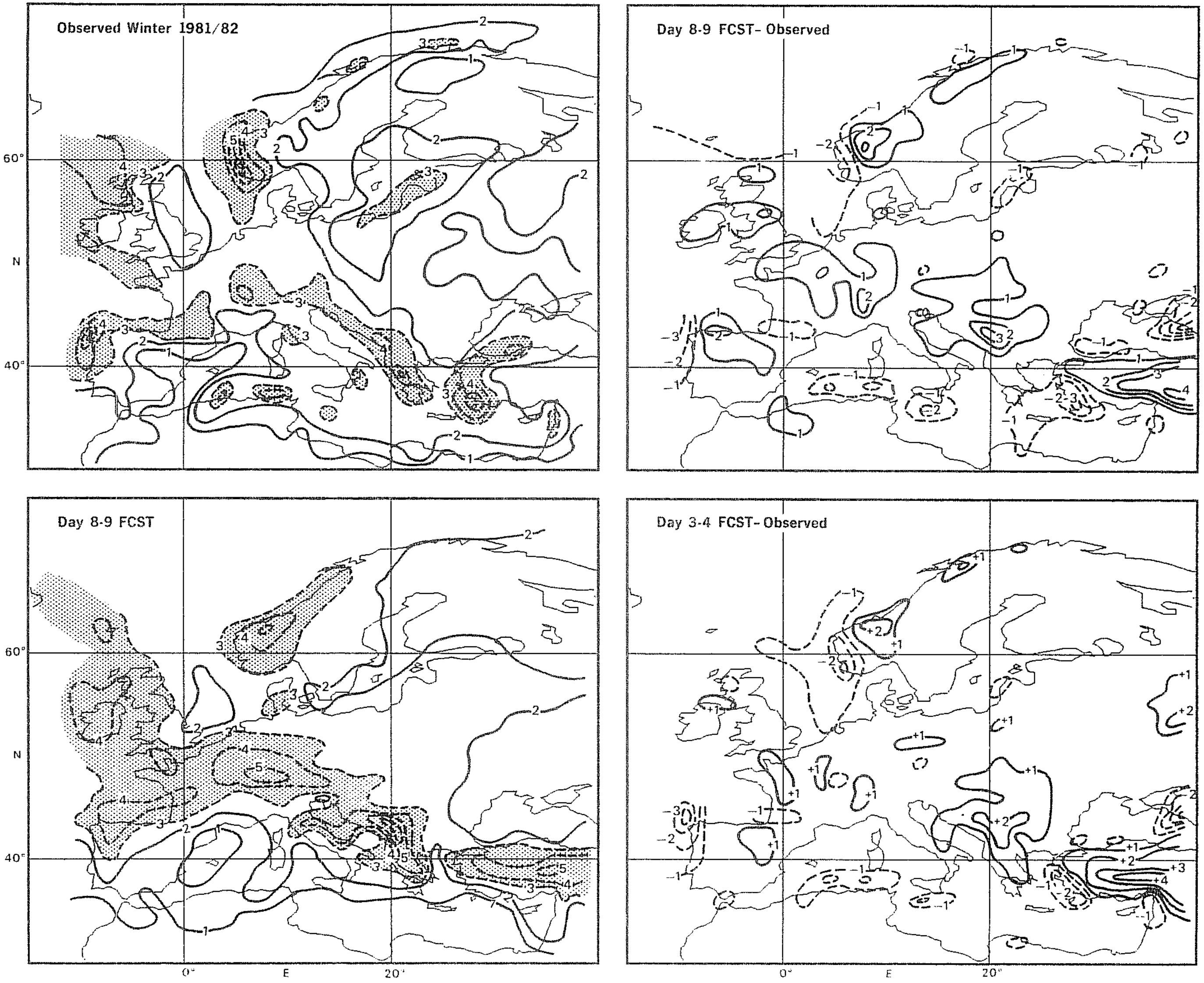


Fig. 7 Mean precipitation during winter 1981/82. Units: mm/day.

2.3 The evolution of the systematic error with time and connections with other parameters

It is important to eliminate or reduce the systematic error if major improvements in medium range forecasts are to be achieved. Displaying connections with errors in other parameters may be helpful. To avoid secondary effects such as the changes in cyclone tracks we will concentrate on the evolution of the error during the first few days of forecast.
The error growth of the 1000 mb height field is displayed in Fig.8. The fastest growth is found near the Himalayas and near the Rockies. Too low heights on the windward side with too large values on the lee side gives a pattern which, from vorticity arguments, suggests that the mountains are too low, as pointed out by Wallace et al (1982). The impact on the mountain torque will be shown later.

There are two areas with negative height errors which intensify steadily during the first two days, i.e over southern Europe and over Alaska. The error pattern of the 1000 mb height field resembles that of the 500 mb height field indicating that the systematic error has a strong barotropic component.

From day 2 to day 5 one finds in addition to a further growth of the error pattern an overall increase of negative areas north of 45° and increase of positive areas south of 40°. This increase in the mean meridional gradient of the 1000 mb height field is more clearly exhibited in Fig.9. The analysis has a plateau in the zonally averaged height field at around 40°N and strong gradients north and south of it. The forecasts exhibit this structure only at the beginning. By day 5 only one very strong gradient remains in the forecasts. The differences in the gradients between forecasts and analyses are largest between 35°N and 50°N. The resulting differences in geostrophic winds are 2 m/s at day 5.

More detailed differences in the mean zonal winds are displayed in Fig.10. During the course of the forecast a steady increase in the wind errors can be seen. The mean zonal wind between 35° and 50°N is forecast to be too strong. In Fig.9 it was shown that approximately half of the maximum error can be

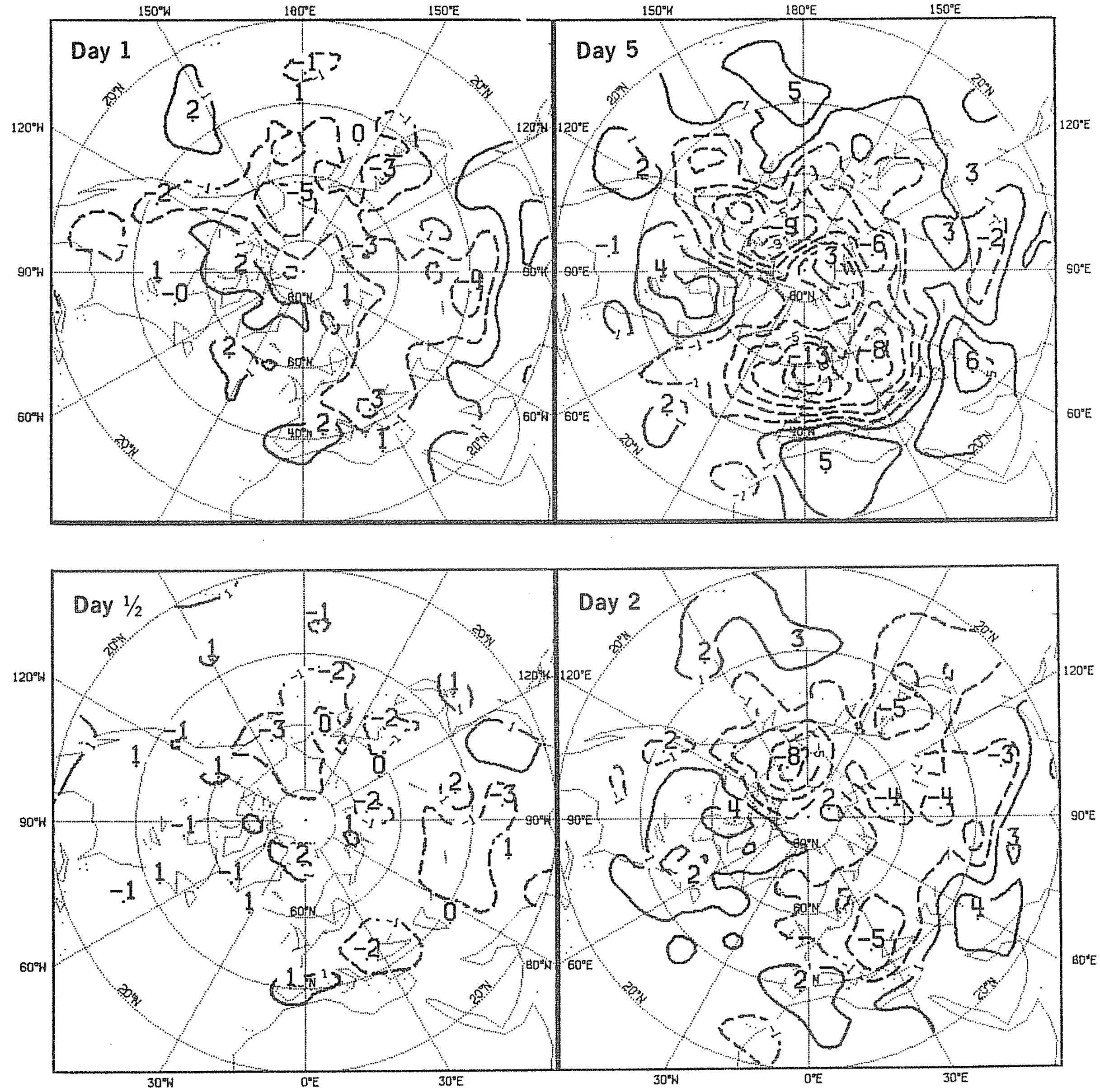


Fig. 8 Mean 1000 mb height difference fields between forecasts and analyses during January 1981. Units: 10 m.

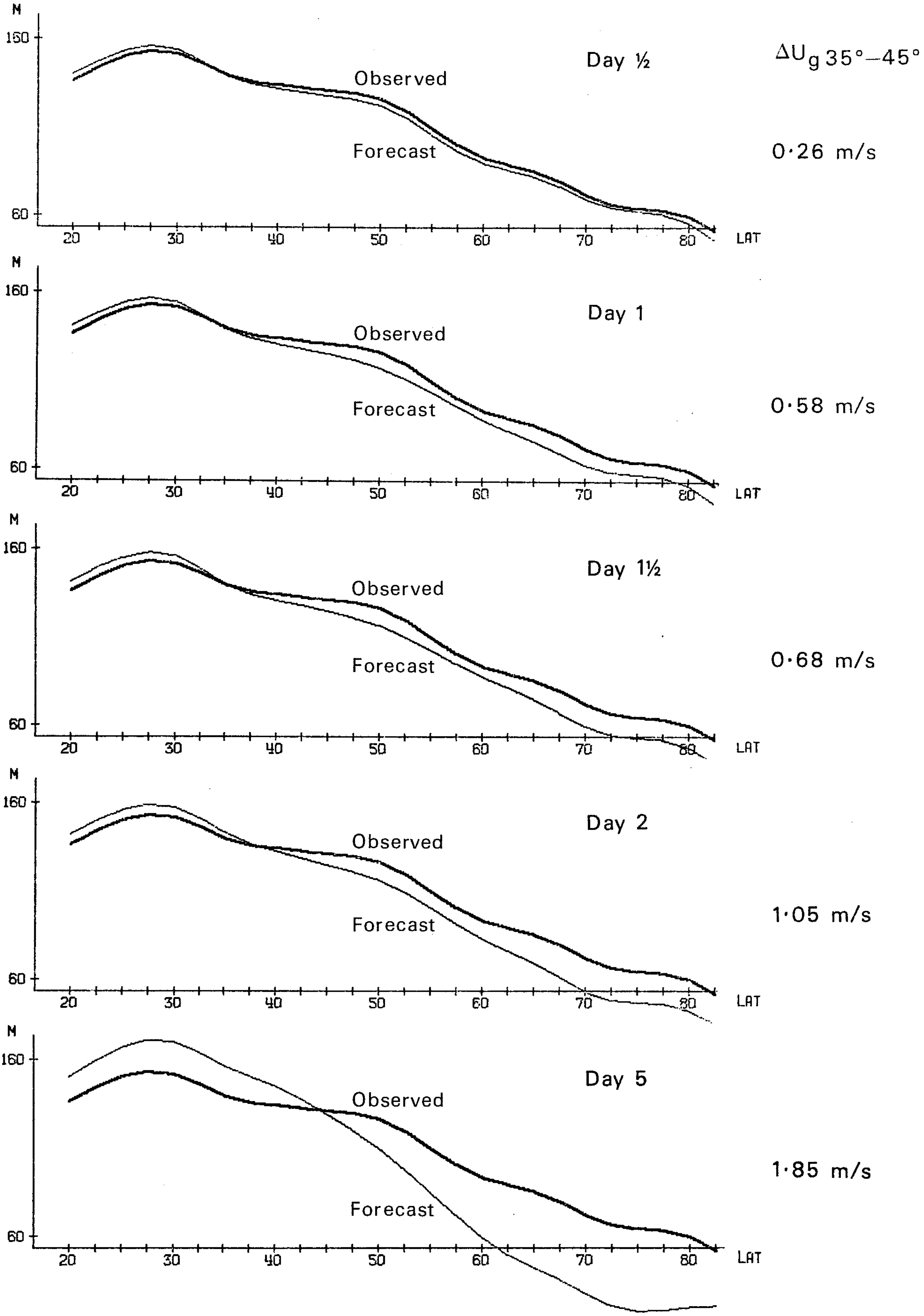


Fig. 9 Zonal and monthly mean 1000 mb height fields during January 1981.

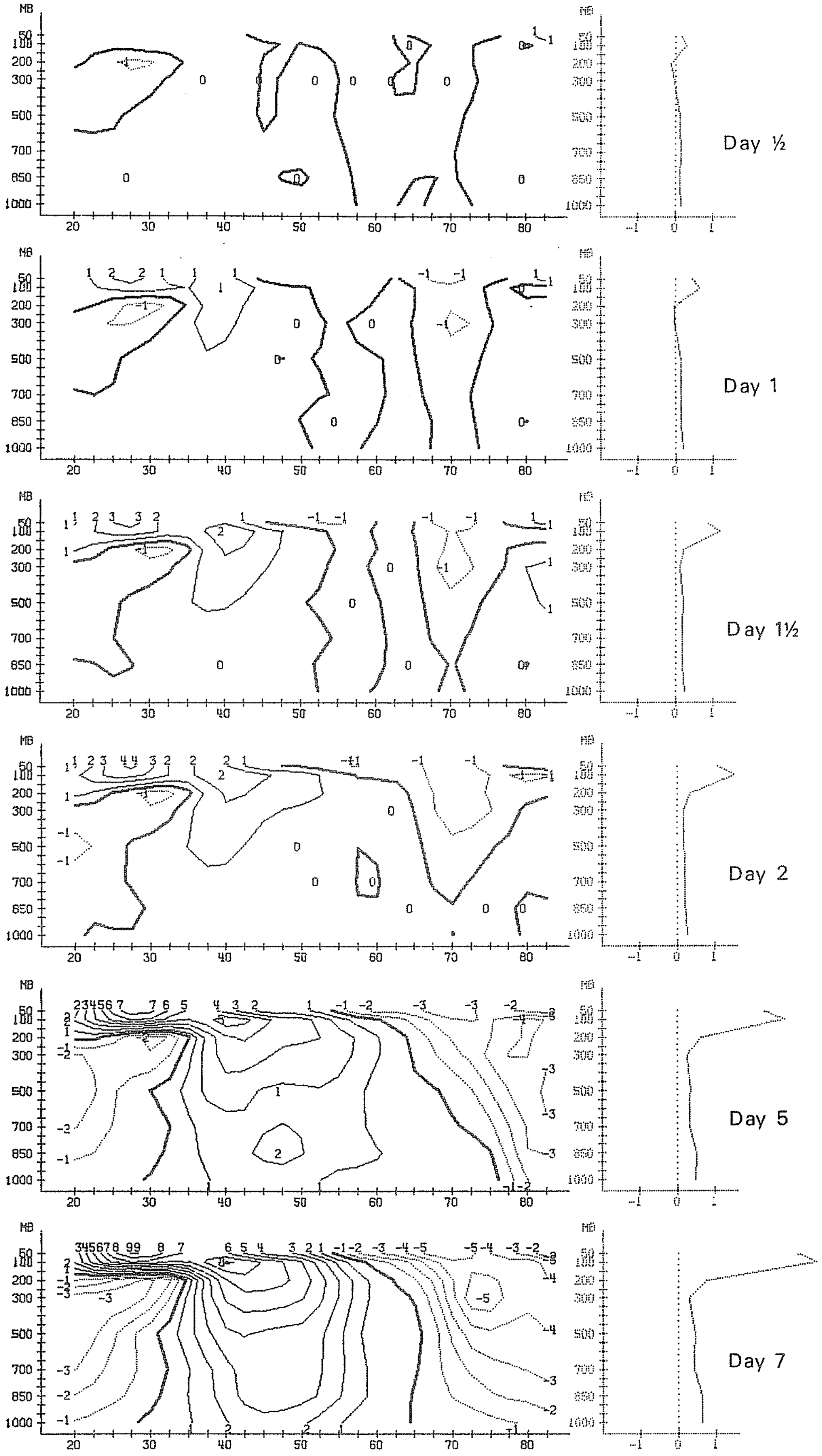


Fig. 10 Zonal and monthly mean zonal wind deviations between forecasts and analyses. Panels to the right give area averages north of 20°N.

attributed to the error of the zonal mean 1000 mb height field while the other half must be due to errors in the thermal structure.

A simultaneous decrease of the mean zonal wind south of 30°N and north of 70°N indicates that the subtropical and polar-front jet streams are getting closer together. In Fig.1 this was apparent as an absence of diffluent circulation, mainly over Europe and to a lesser extent over the eastern Pacific. On cross-sections of the mean zonal wind this often appears as a poleward shift of the subtropical jetstream.

A dipole pattern is apparent in the error fields at about 30°N between 200 and 50 mb which indicates an upward shift of the subtropical jet. Both movements of the jet, upward and poleward, can be found in both hemispheres, and are strongest in the winter hemisphere.

The diagrams on the right show the area average of the zonal wind between 20°N and 82.5°N. They show clearly an overall increase of the mean zonal wind which is independent of the shifts of the jets.

Budget calculations of the vertically and zonally averaged momentum equation reveals further connections with this error. An important term is the eddy meridional flux of momentum $\overline{u'v'}$ which is shown in Fig.11 for January 1981. A steady increase with forecast time is obvious with maxima occurring around 32°N. The corresponding southward flux on the southern hemisphere does not show such an increase, in agreement with a less obvious systematic error on the summer hemisphere. The increased divergence of the flux between 20° and 30°N and its convergence between 35° and 50°N is of a magnitude which can explain the poleward shift of the subtropical jet. It will be shown later by means of the eddy kinetic energy statistics that the amplitudes of the wind components do not increase during the forecast but rather decrease. The increase in the fluxes must therefore be due to increases in the horizontal tilts of the troughs and ridges and this has already been demonstrated by Klinker (1983) in a statistical investigation and by Haseler (1982) in case studies. Haseler showed that in most of the cases she studied the increased momentum flux is caused by errors in the tropical regions.

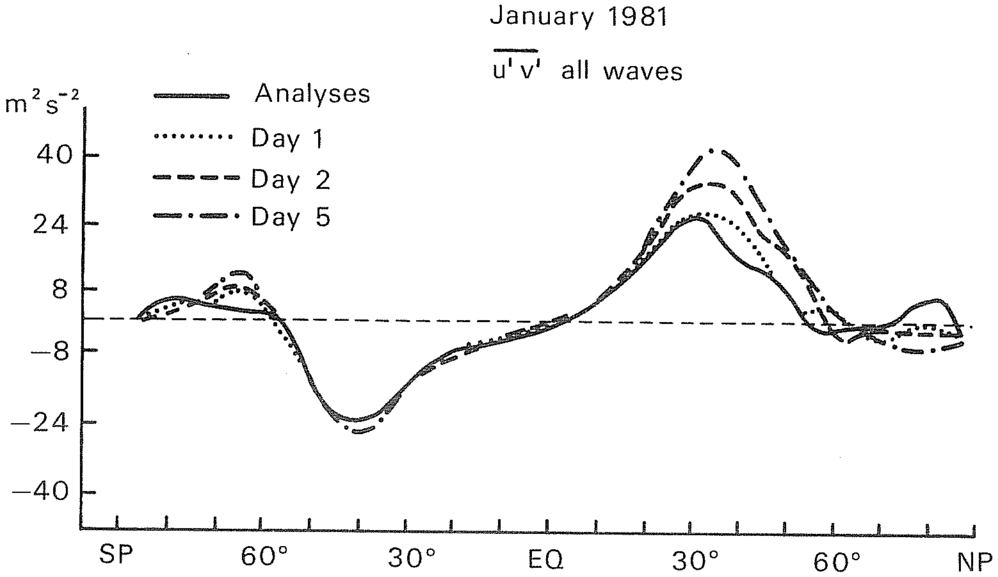
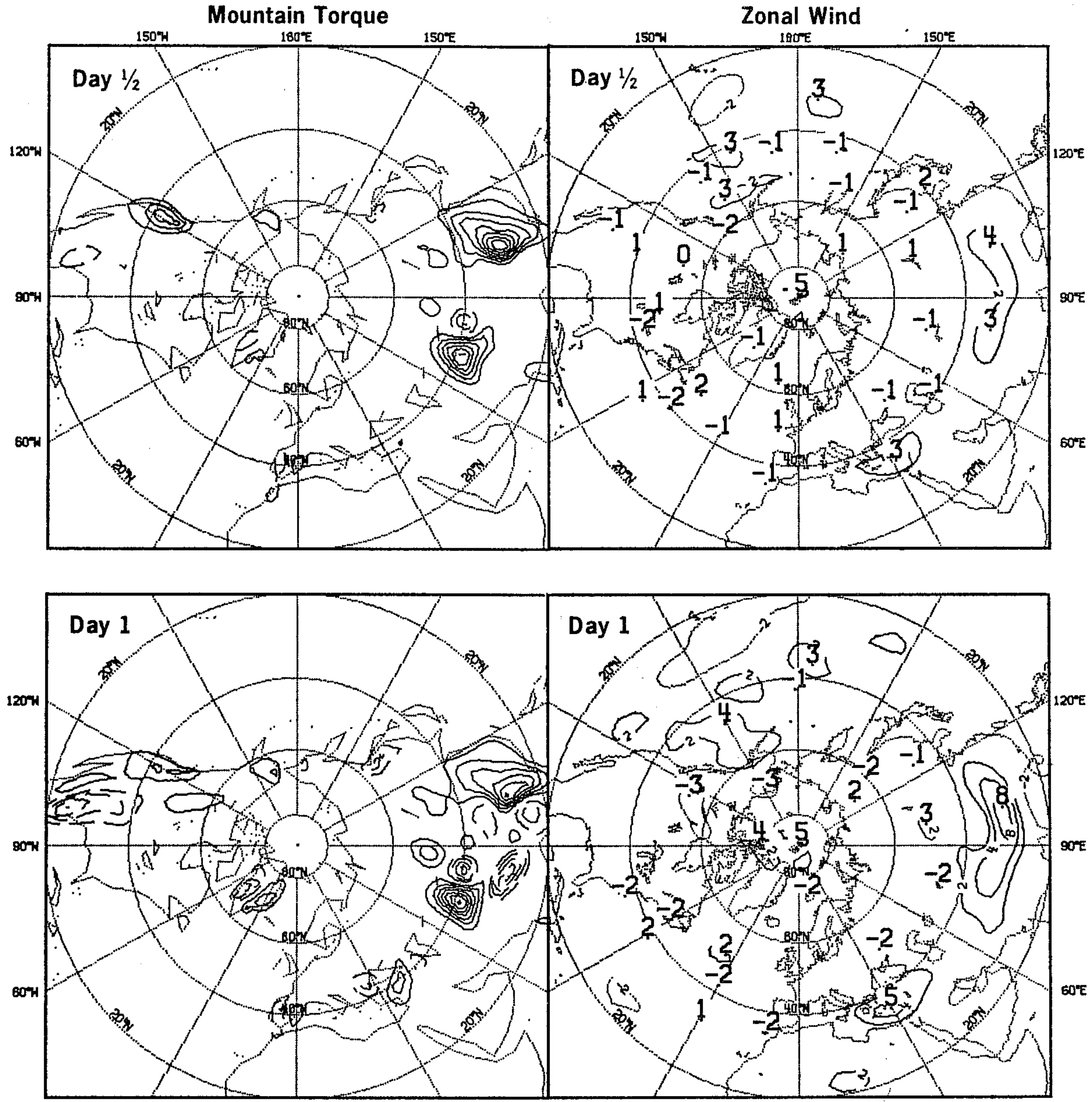


Fig. 11 Zonal mean of eddy meridional flux of zonal momentum $\overline{u'v'}$. Vertical averages between 1000 and 50 mb during January 1981 in the forecasts and analyses.

A separation of the eddy momentum flux into contributions by wavenumber groups (not shown) indicates that during the first 2 days of the forecasts the increase is only due to the very long waves. After two days the medium, mostly barolinic, waves start to contribute. The increase of the eddy momentum flux of the baroclinic waves may be a secondary effect because it happens later than other effects.

The increase of eddy momentum flux can shift the maximum of zonal wind but cannot increase its mean as found in Fig.10. For this the mountain torque, which is the negative correlation between the surface pressure and the mountain slope, gives an interesting contribution. In Fig.12 difference maps between forecasts and analyses of the mountain torque and of the zonal wind component are shown. The clearest signal can be seen over the Himalayas where both the up- and down-slope parts contribute to the mountain torque in the sense of an acceleration of the zonal wind; at the same place one also finds an increase of the zonal wind.



Mountain Torque

Zonal Wind

Fig. 12 Difference of mountain torque (left) and of zonal wind (right) between forecasts and analyses during January 1981.

As is often the case in diagnostic investigations, one does not know which of these effects is the primary one. The momentum equation requires the mean zonal wind to increase when a positive mountain torque increment is found. But if there is an increase of zonal wind over the Himalayas, as shown here, which is not balanced by meridional wind components it would cause a divergence field which would decrease the surface pressure to the west of the mountain and increase it to the east of the mountain. This would lead to a mountain torque map as shown here.

The benefit of these diagnostic tools is to show through sensitivity experiments if the desired effect did occur, e.g. Wallace et al (1982) could show that by increasing the mountain barrier this typical pattern with positive mountain torque on both sides of the Himalayas in difference maps can be destroyed together with a reduction in the systematic error of the height- and wind-field.

2.4 Baroclinic processes

An overview of the atmospheric energetic processes can be gained from Fig.13. This shows the annual mean energetics on the northern hemisphere for the different steps in the analysis-forecast system.

Most of the terms show little variation during the analysis cycles except the eddy available potential energy. Its apparent semi-diurnal cycle in the analyses may be due to a tidal wave or to lack of observational data at 06Z and 18Z. Its amplitude is only 2% of the mean amount and in terms of temperature one has to resolve differences in temperature amplitude in the order of 0.1K which is difficult to achieve. There is, however, a clear semi-diurnal tide to be seen in the 1000 mb height field which has an amplitude of about 10 m during a test period of 10 days in September 1982. It was evaluated by comparing mean fields at 06Z and 18Z with those at 00Z and 12Z. This wave has positive maxima

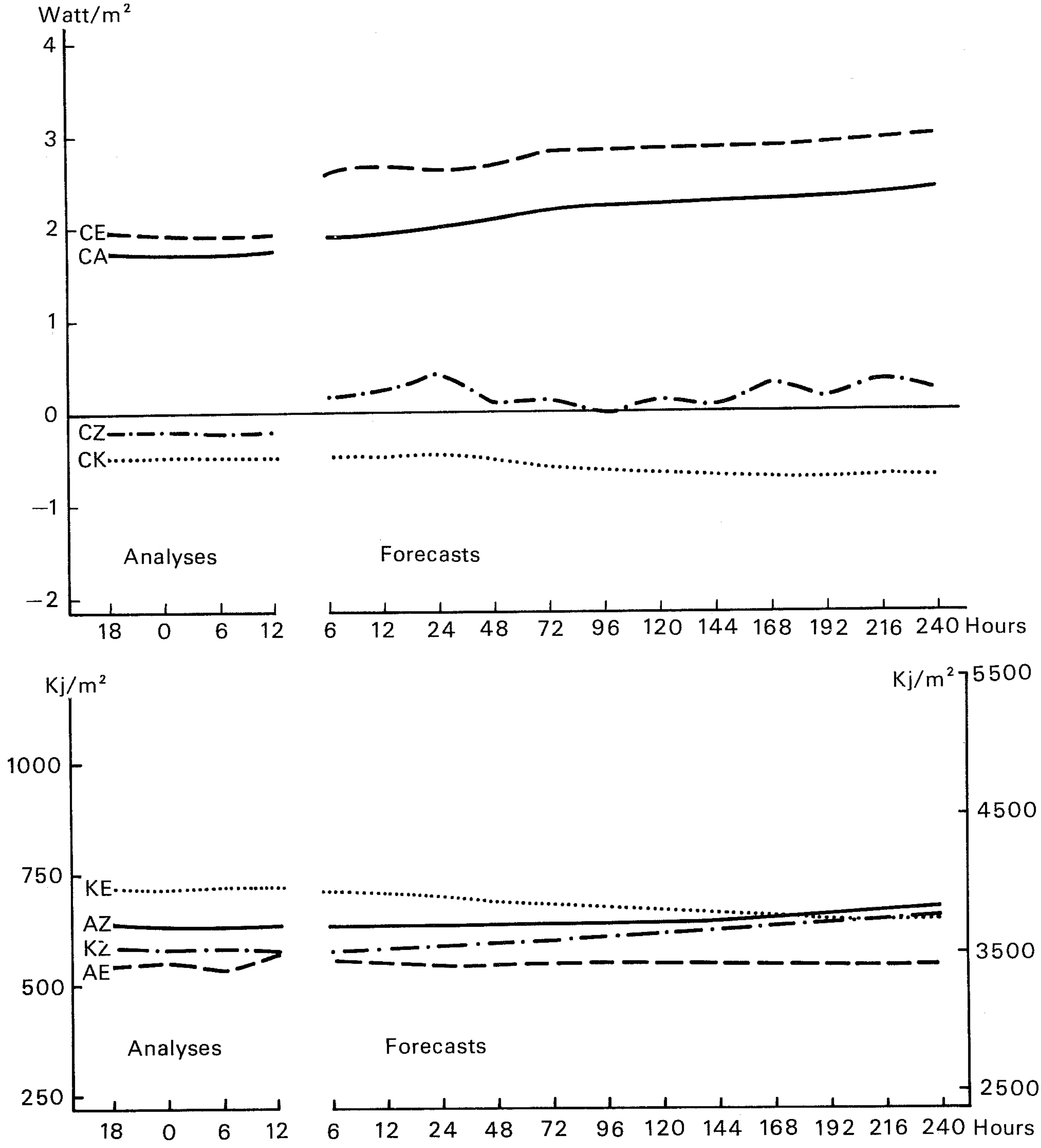


Fig. 13 Energy conversions and contents of analyses and forecasts. Average for 1981 and the Northern Hemisphere between 1000 and 50 mb (after Oriol, 1982a). Left scale of lower panel for KZ, KE and AE, the right one for AZ.
AE = eddy available potential energy, AZ = zonal available potential energy, KE = eddy kinetic energy, KZ = zonal kinetic energy, CE = conversion AE → KE, CA = conversion AZ → AE, CZ = conversion AZ → KZ, CK = conversion KZ → KE. For exact formulations see Oriol (1982b).

at 06Z and 18Z over central America (115-125°W) and over India (55°-65°E). The resulting adiabatic temperature variation would be around 0.1K at 1000 mb, large enough to explain the semidiurnal fluctuation in the eddy available potential energy, but other fluctuations in the temperature field during the 10 day period were much larger.

During the course of the forecast one finds a steady increase of zonal energies and decrease of eddy energies. This is consistent with the zonalisation of the mean circulation shown in Fig.1 but the main contribution to the drop in eddy energies comes from the transient flow. In the southern hemisphere the zonal energies increase also and the eddy available potential energy stays fairly constant.

The jump from the analyses to the forecasts in the energy conversions need not indicate an error in the model but rather a problem in the analysed data. For the energy calculations the initialized analyses have been used for consistency within the mass- and wind-field. The initialization has a strong smoothing effect on the vertical wind field, especially in the tropics and therefore one finds large jumps from the analyses to the forecasts in the conversions which involve the vertical component of the wind, i.e. CE and CZ.

In the conversion from zonal available potential energy to zonal kinetic energy, CZ, there is a strong fluctuation with a period of two days. This may be connected with gravity waves being excited in the same way at the start of each forecast. They can be found in any monthly mean, with similar phase and amplitude. In the southern hemisphere a similar oscillation has a longer period of up to 3 days.

The most interesting feature in Fig.13 is the increase of the conversions CE, CA and CK during the forecasts. The increase of CK is connected with the increase of the eddy momentum flux due to stronger horizontal tilts of troughs and ridges in the forecasts, as shown in Sect.2.3. The increase in CE and CA can be shown to be connected with an increase in baroclinicity of the waves.

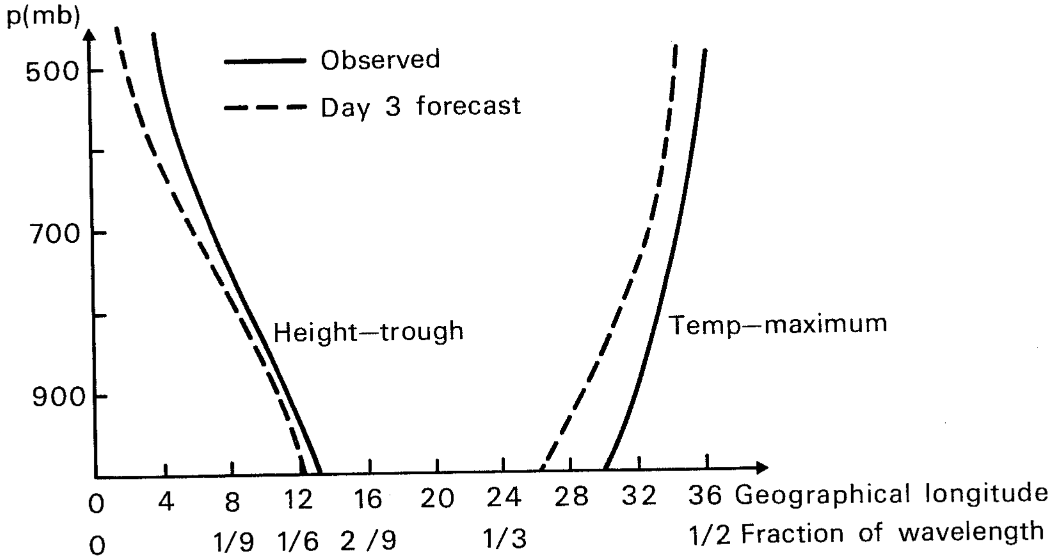


Fig. 14 Mean structure of the wave with zonal wavenumber 5 and period 5.3 days during winter 1980/81 (from Klinker, 1983).

Fig.14 displays for the, largely baroclinic, waves with zonal wavenumber 5 and periods of 5.3 days the phase of the temperature-and height waves at each level for the day 3 forecasts and the analyses. The waves in the forecasts are lagging behind the ones in the analyses, i.e. the waves in the model are propagating too slowly. The mean lag in the height field is 1/72 of a wavelength or 1° longitude which is less than the resolution of the model. Consistency throughout the troposphere suggests, however, that it is a real effect. The importance of this figure for the energy conversion is in the increased vertical tilts. It means that the waves in the forecasts are more baroclinic and more efficient in converting energy, as shown in Fig.13.

Defant (1974) demonstrated the importance of the static stability and of the meridional gradients of temperature for the strength of vertical velocity in baroclinic waves. This is a very important ingredient for the baroclinic waves; these parameters are shown in Fig.15. Both upper left panels show immediately that the model produces stratospheric temperatures that are too low and tropopauses that are too weak, especially in the tropics. The lower left panel shows the differences between forecast and analysis. Except at 850 mb, where there is very little bias, the model has temperatures throughout that are too low.

Of interest for the baroclinic development is the strong cooling north of 50°N, which leads to an enhanced meridional temperature gradient in the model. Vertical differences in the cooling rates lead to strongly decreased static stabilities for the lower troposphere and slightly increased static stabilities in the upper troposphere, as shown by $\partial\theta/\partial p$ in the right panels. The typical decrease in the lower troposphere is from -55 K bar^{-1} in the analyses to -45 K bar^{-1} in the forecast while the maxima increase for the layer 500 to 300 mb is only from -65 to -70 K bar^{-1}. The decrease of the static stability grows steadily from the beginning like the other errors shown above.

There is of course variability from month to month in the mean errors of the temperature field, e.g. in winter 1981/82 it was not the zonal mean 850 mb temperatures but the 700 mb temperature which were forecast correctly while the 850 mb temperatures were forecast too warm. But nevertheless the other errors stayed almost the same especially those for the static stability.

In Fig.16 the evolution of the 850 mb temperature error during the winter 1981/82 is shown. As the corresponding maps for the 700 mb temperature show

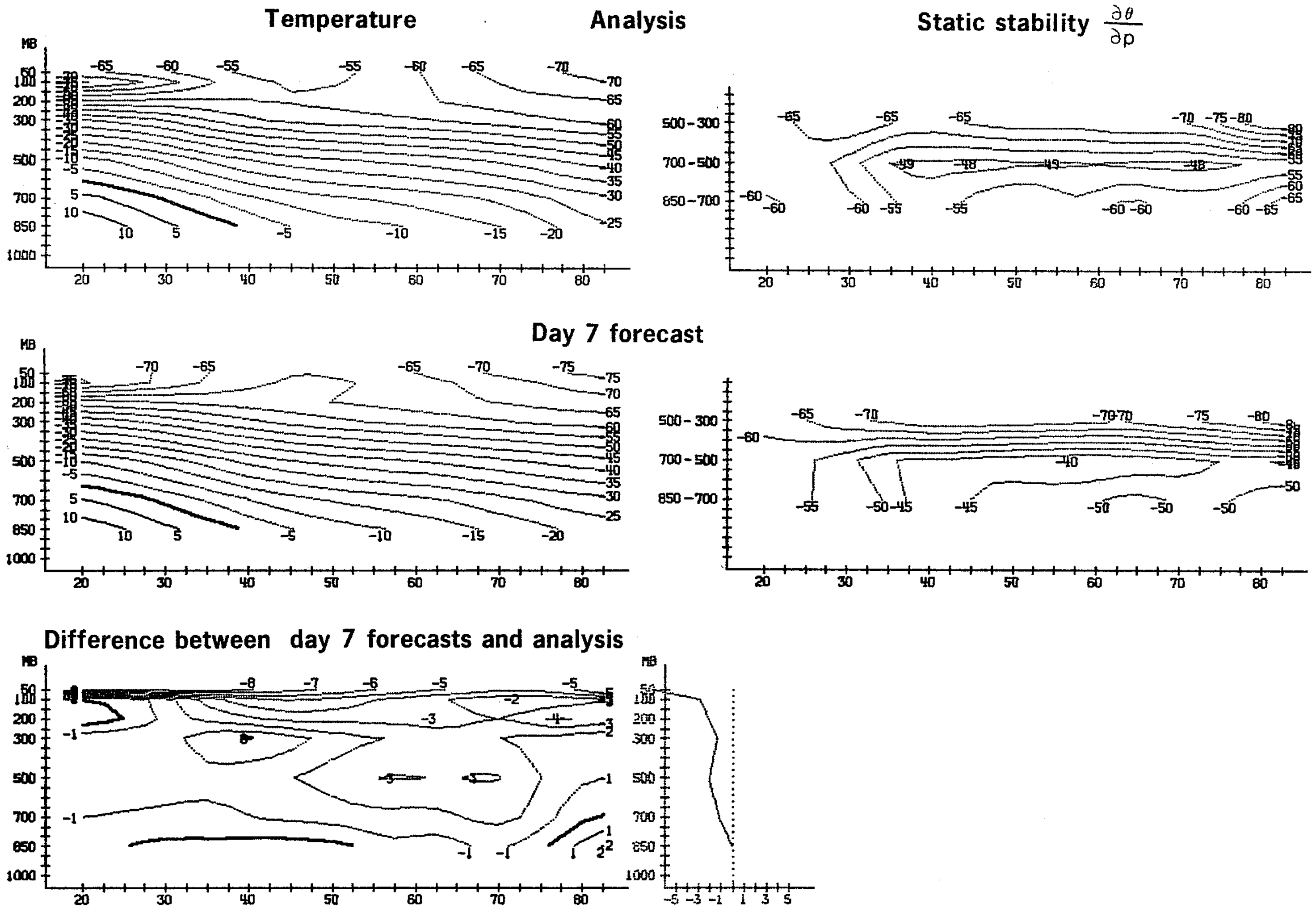
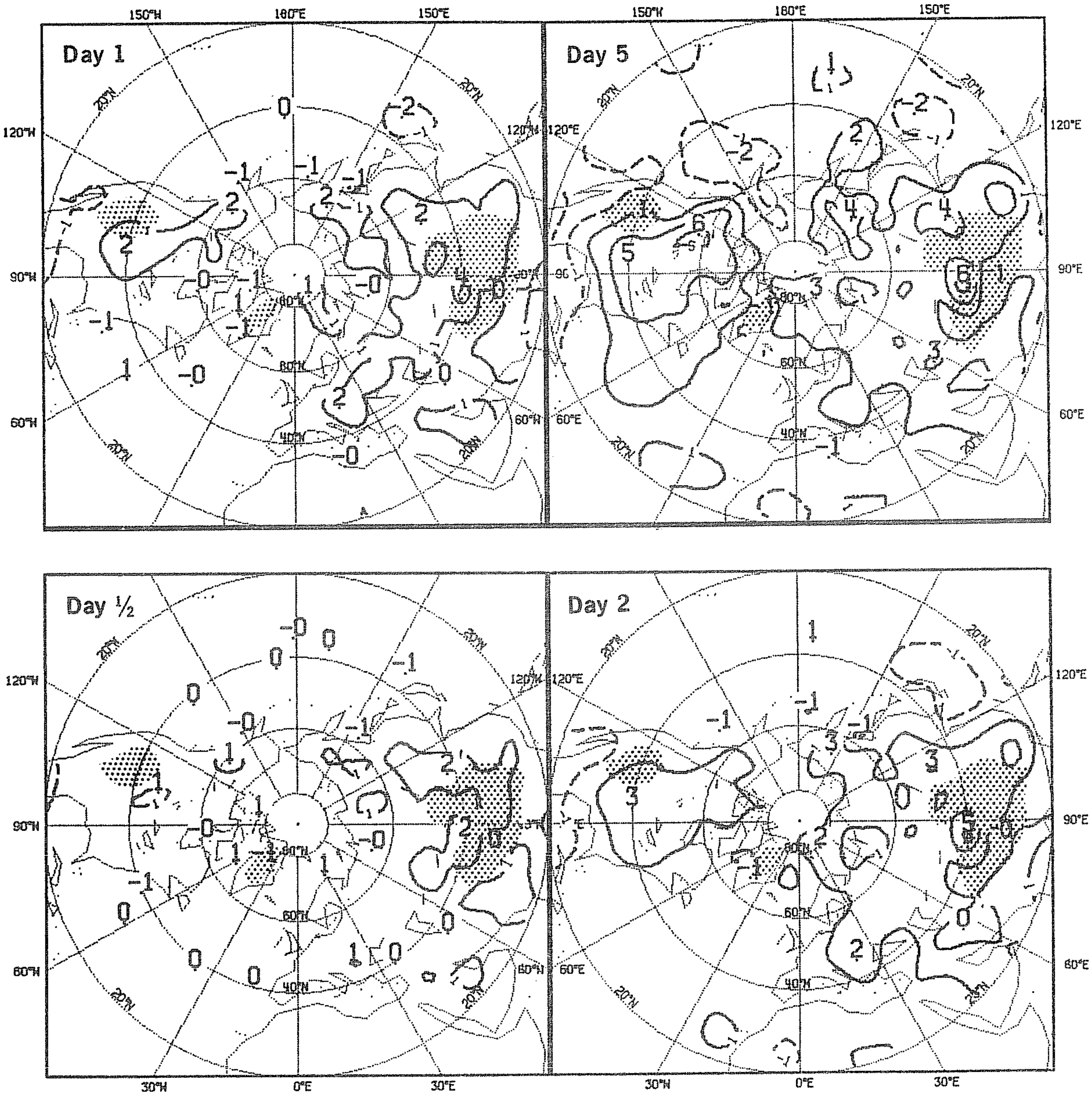


Fig. 15 Meridional cross-sections of zonal mean temperature and static stability (∂θ/∂p) during January 1981. Units: °C and K/bar.



Day 1

Day 5

Day ½

Day 2

Fig. 16 Mean 850 mb temperature difference fields between forecasts and analyses during the winter 1981/82.
Areas with the 850 mb level below ground are shaded.

much smaller errors, Fig.16 also indicates the error growth of the static stability for the 850-700 mb layer.

It is interesting to note that the errors in the 850 temperature occur first over the mountaineous areas, these over very high ground having to be disregarded because they are only extrapolated values. The largest errors, mostly on the lee side of the mountains, are growing very fast in the beginning and are giving the largest contributions to the mean error in the static stability. Over the American continent it is connected with a weakening of the strong horizontal gradient. This is also consistent with the cooling over the Caribbean Sea and Mexico.

Again an incorrect treatment of the flow around the mountains seems to be connected with this systematic error. The fact that the worst errors are occurring over the continents but not over the oceans where the cyclone tracks are found, weakens the suggestion that the increased baroclinicity of the waves and the decrease of zonal mean static stability are connected.

The problem of baroclinic waves being too active in the forecasts can also be shown by comparing the filling rates of cyclones. All cyclone developments with clear deepening and filling phases during the winter 1981/82 are used for a comparison between the forecasts and the analyses (Akyildiz, 1983). This comparison had to be restricted to the first third of the forecast period in order to be sure that the cyclones in the model and the forecasts were the same features. The clearest signal is presented in Fig.17 left panel which shows the filling rate of the cyclone central surface pressure between the day of the minimum pressure and the following day. In nine cases the filling was forecast correctly but more often the cyclones were forecast to fill too slowly, on average by about 4 mb/day. This reflects the same problem as the over large baroclinicity shown in Fig.13.

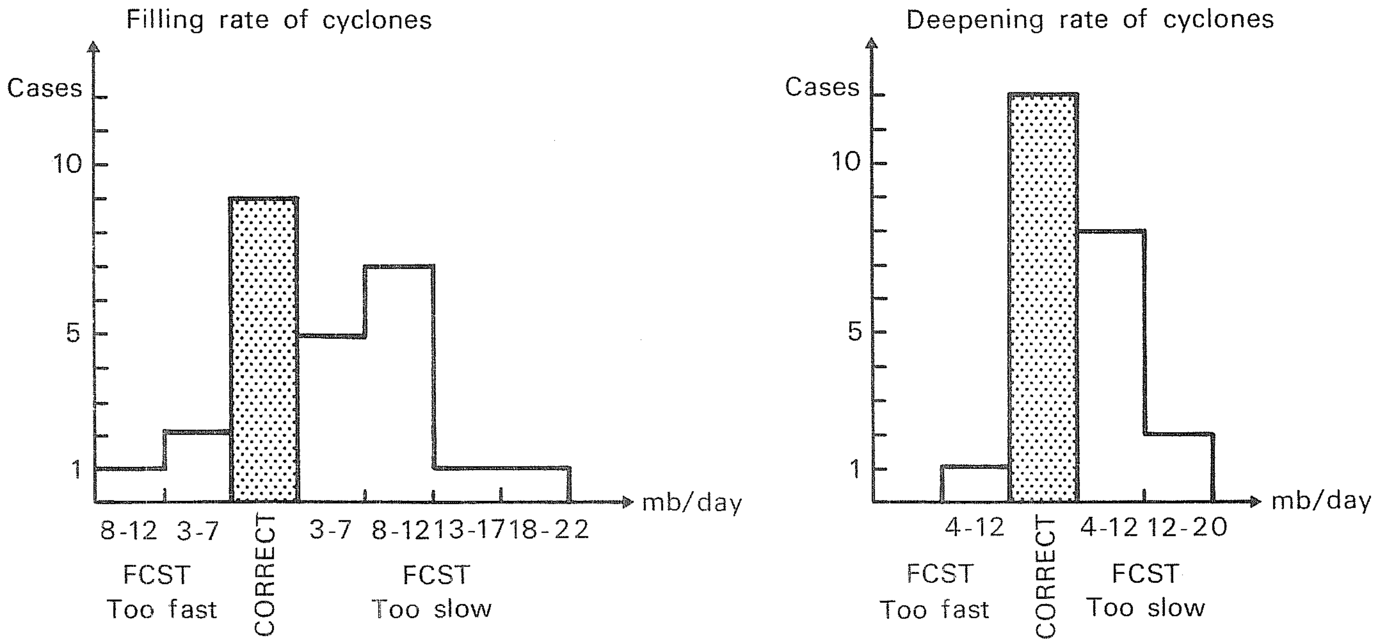


Fig. 17 Comparison of filling and deepening rates of cyclones in the short-range forecasts and analyses (from Akyildiz, 1983).

For the user of the forecasts this slow filling of cyclones might be viewed as an overdevelopment of cyclones because he will find that the foreast cyclone after 3 days is likely to be 5 mb too deep. Clear overdevelopments of cyclones have, however, been found in the model after 5 days with cyclones which were not present in the initial data.

Fig.17 also gives information on the deepening rate of the cyclone central surface pressure between the first day of appearance and the following day. 14 cases were correctly forecast and 10 cases showed too slow deepening on average by 1 mb/day while only 1 case was found where the cyclone deepened more rapidly in the forecast.

This problem of slow deepening is largest in the very early stage of cyclone development and results in delayed pressure changes. Simmons (1981) showed in a case study that this error can be reduced with an increased horizontal resolution of the model.

2.5 Errors in the tropics

The tropics are a significant source of mid-latitude errors as has been demonstrated by Haseler (1982) and further suggested by Wallace et al (1982). To exhibit the problems of tropical forecasts monthly mean windfields during July 1982 are shown.

In Fig.18 the mean day 10 forecast of the 850 mb wind is compared with the corresponding analyses. Largest differences occur over the equatorial central Pacific and over Africa. In both areas almost no wind was observed but strong winds forecast, leading to a dominant wavenumber 1 in the error field of the u-component. The map of differences show also that the spurious westerlies over Africa are extended to the Atlantic due to underprediction of easterlies there. This error field results in a large scale increase of convergence over Indonesia and a divergence around central America. This low-level divergence field is compensated by a divergence field at 200 mb with opposite signs not

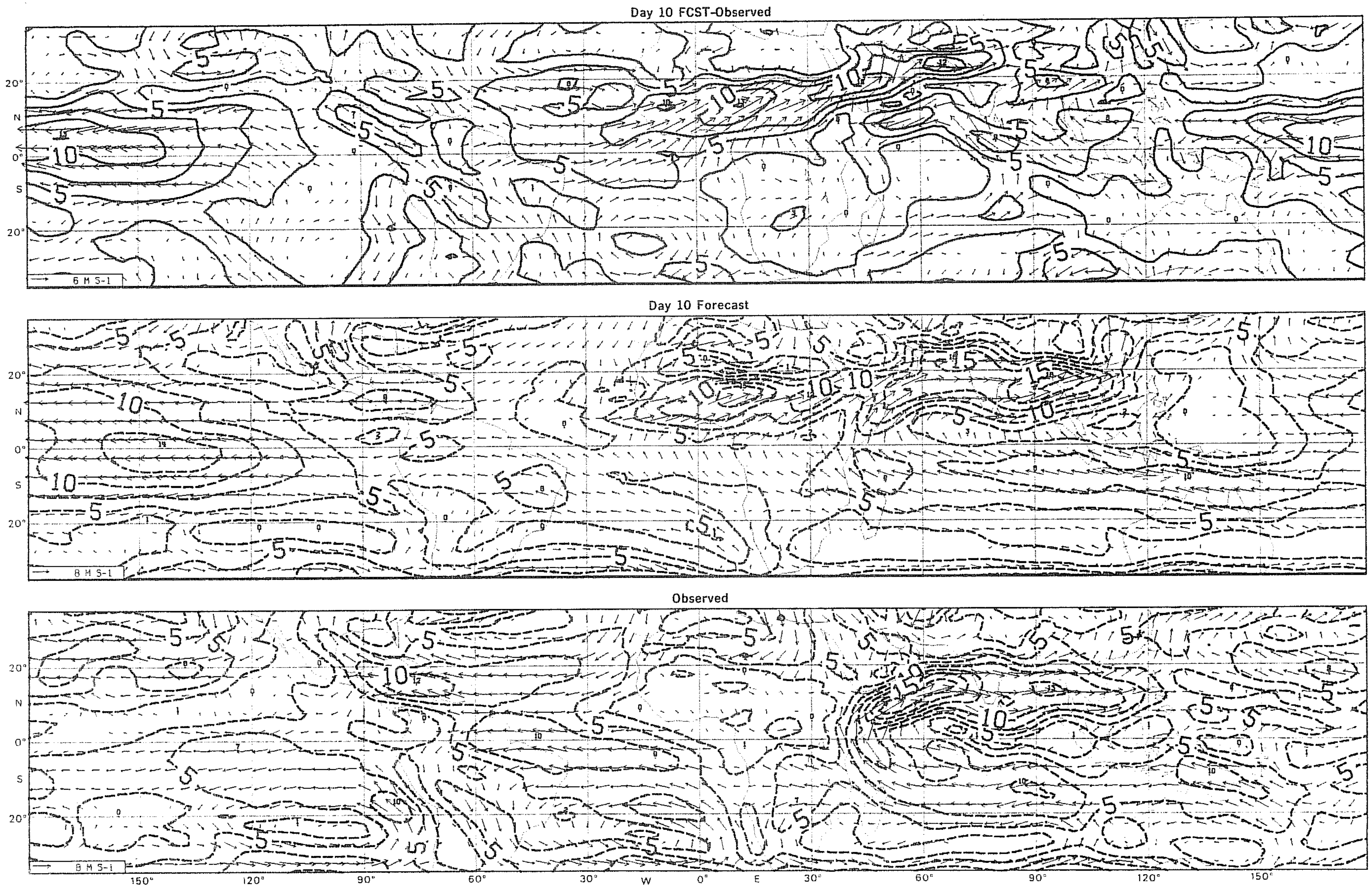
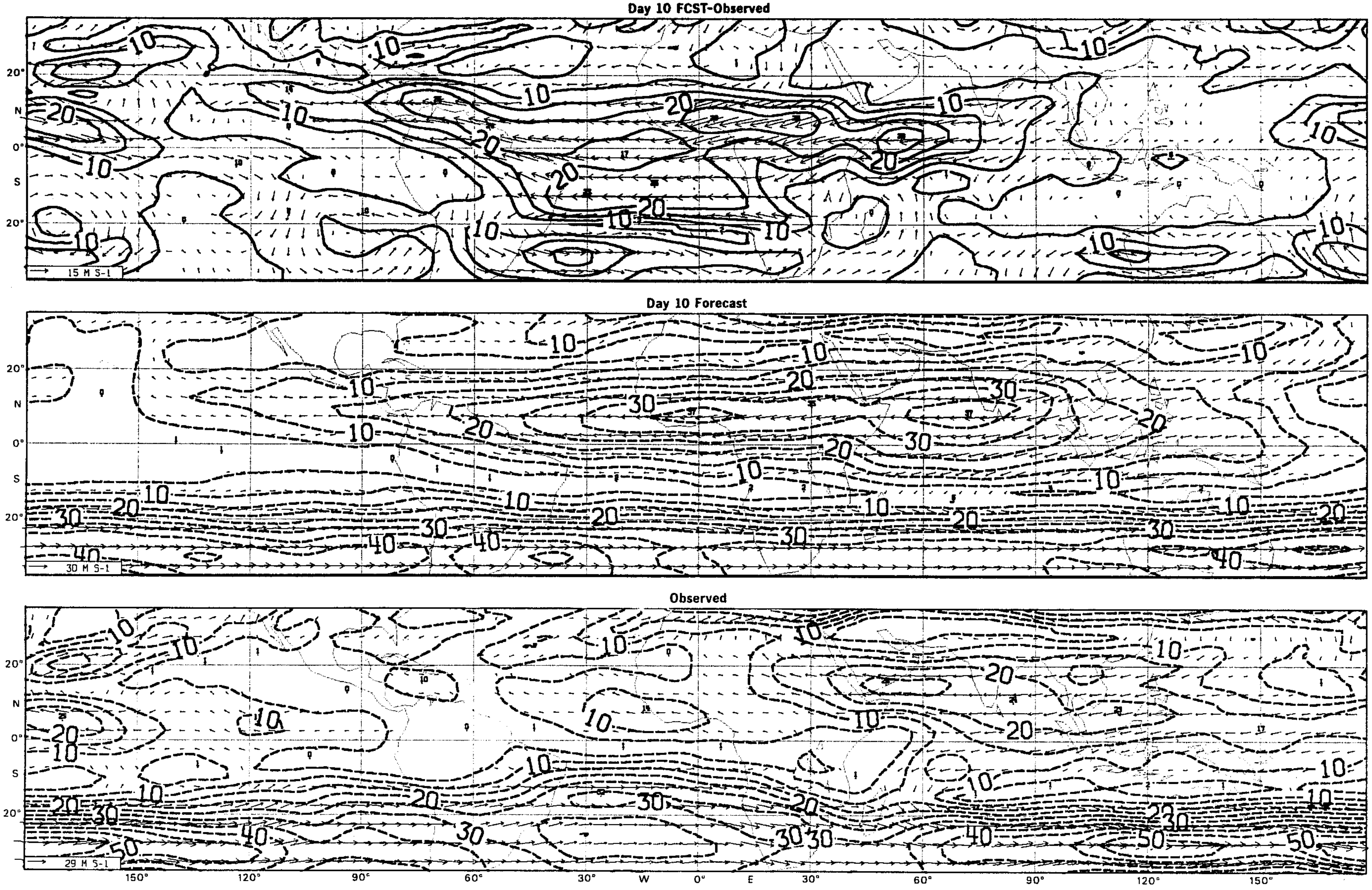


Fig. 18 July 1982 mean 850 mb wind field in the tropics. Contour lines give the wind speed.


Day 10 FCST-Observed

Day 10 Forecast

Observed

Fig. 19 Same as Fig. 18 for 200 mb wind field.

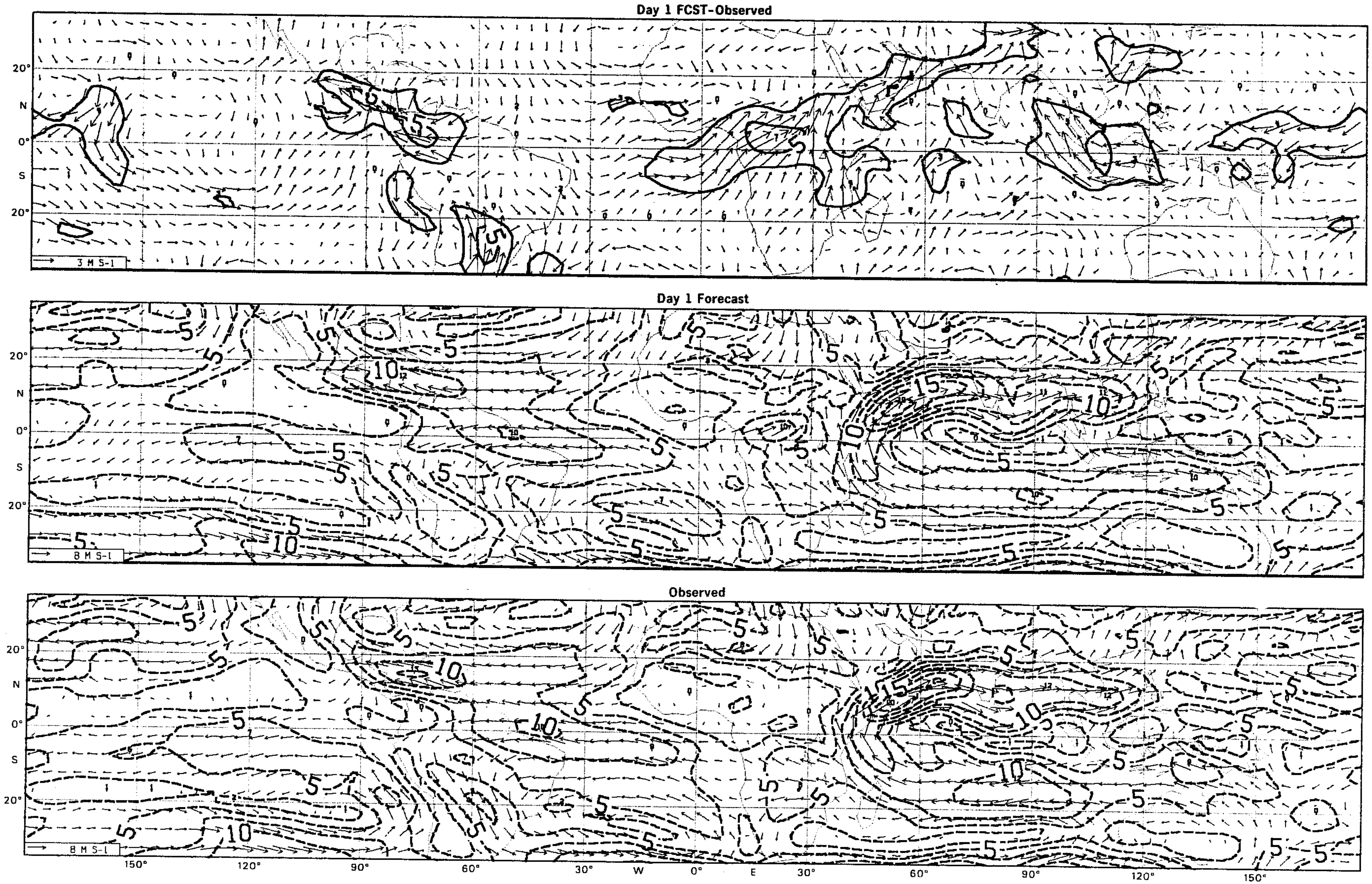
only in the divergence field but also in the u-field (Fig.19). In addition to the wavenumber 1 pattern, one finds an overall increase of easterlies which reflects the tropical extension of negative biases in the zonal mean zonal wind shown in Fig.10.

Another area with large errors at the 850 mb level (Fig.18) lies around the Indian Ocean. The monsoon circulation there is shifted in the forecasts to the north. Changes in the pattern are a shift of the wind maximum of 18 m/sec into the Arabian Sea (with a reduction to 16 m/s) and the build up of a further wind maximum of 18 m/s in the Bay of Bengal. This error has been shown to be sensitive to the parameterization scheme for the convection (Simmons, 1981).

Less dramatic but still important is a weakening of the subtropical anticyclones. In the difference map this is exhibited as cyclonic circulations at the positions of the anticyclones, perhaps clearest shown over the South Atlantic.

In the day 1 forecasts many of the errors described above can already be found (Fig.20). The spurious westerly over central Africa reaches 7 m/s at day 1 which is half of what was found on day 10. Also the weakening of the subtropical anticyclones is indicated.

Over South America large errors grow quickly but then remain steady up to day 10 but are not emphasized in Fig.18 because other features are dominant. Mean error fields at other months show errors of the same size but with different patterns. One can guess that incorrect forcing by the convection or by the mountains or by a combination of both are causing this problem. Variations in the error pattern during the month seem to be connected with variations in the precipitation patterns but it has not yet been fully explored. Heckley and Gill's (1983) results suggest that a large part of the errors in the tropics during the forecasts can be associated with adjustments due to an imbalance between the initial data and the model diabatic forcing.



Day 1 FCST-Observed

Day 1 Forecast

Observed

Fig. 20 Same as Fig. 18 for 850 mb wind field at day 1.

The fact that the errors in the tropics are growing so quickly may indicate that they could contribute to the systematic error in mid-latitudes. Further work is needed to confirm this hypothesis.

3. DIAGNOSIS OF THE ANALYSES

The analysis scheme is an integral part in the ECMWF forecasting system. It will be shown that the quality of the analyses is very important for the quality of the forecasts. To get indications of problems in the analyses, data which has been prepared with two different analysis schemes, will be compared. Forecasts from these different analysis sets can amplify the differences.

For the period 15.2.71 to 19.2.79 analyses from ECMWF and from NMC (National Meteorological Centre, NOAA, USA) were available; both had been offered the same observational data (FGGE-IIb). Differences in the final analyses are therefore only due to the analysis scheme, which includes data checking, a model produced first guess, and many other features, which are quite different in both schemes.

Fig.21 top left panel shows the mean 500 mb height field during the period of 4 days. This field has a clear correlation with the mean difference field between both data sets. The troughs are markedly stronger in the NMC-analyses. A cause for this may be a systematic underprediction of eddy kinetic energy in the ECMWF model as shown in Fig.13 but this is still under investigation.

The map of rms differences due to non-stationary waves between both analysis sets (lower panel) shows maxima where the data coverage is poor and variability is high as might have been expected. The typical rms difference for the total fields between the analysis sets is about 20 m for the 500 mb height field and 15 m for the 1000 mb height field calculated between 20°N and 82.5°N. A similar comparison between NMC analyses and those by the German Weather Service during a period in February 1976 gave values of 30 m and 20 m respectively (Arpe, 1980). Knowledge of the state of the atmosphere has

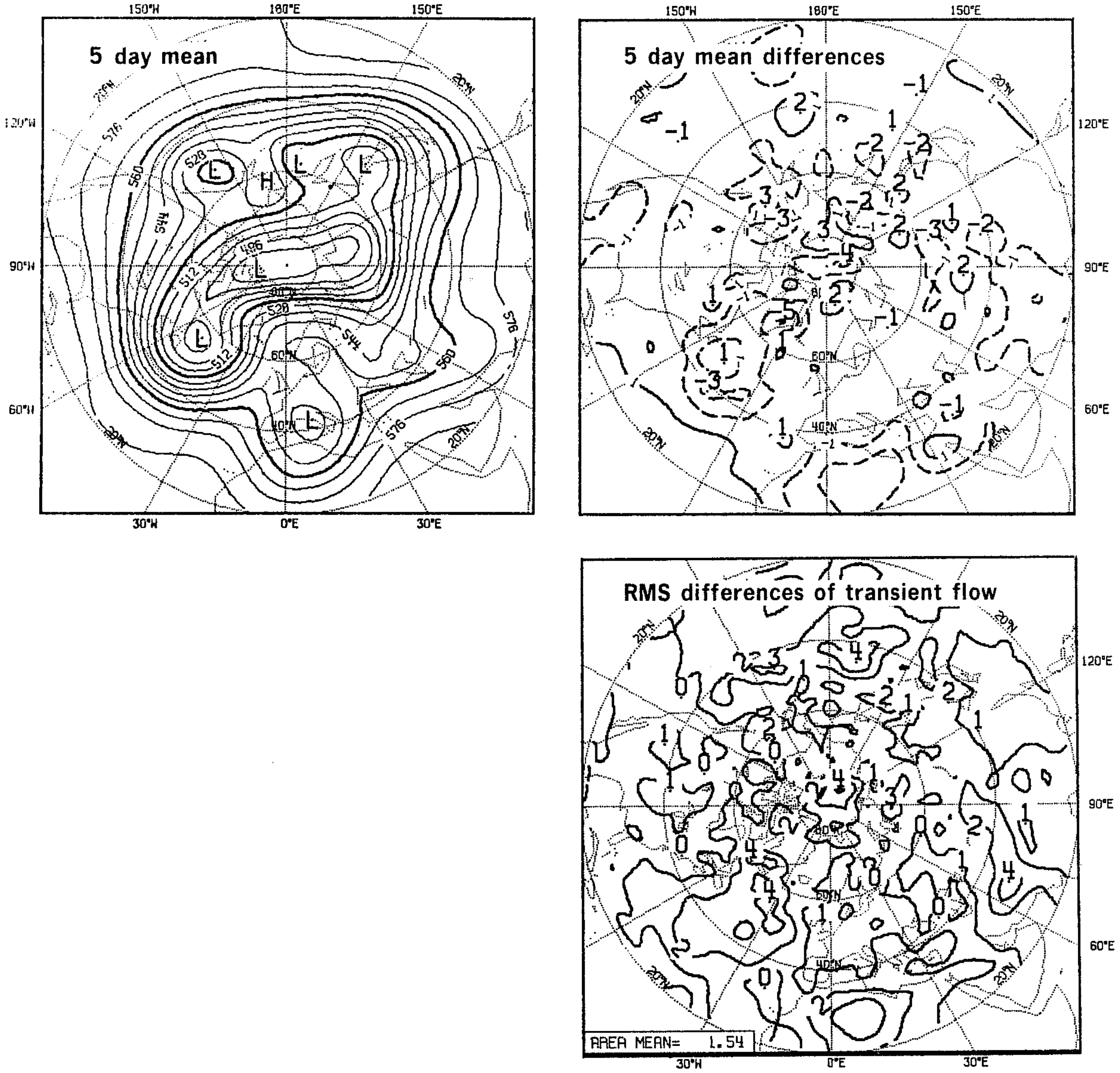


Fig. 21 Difference of 500 mb height analyses prepared by different analyses systems during 15 to 19 February 1979. Units: 10 m.

obviously improved due to better analysis methods and also due to a better data coverage during FGGE. In the southern hemisphere the rms-differences between 20°S and 60°S are about 27 m and 15 m for the 500 mb and 1000 mb height fields respectively, showing the impact of less radiosonde-data there compared to the northern hemisphere.

These differences between, or uncertainties in, the analyses will have an impact on the forecasts. To show this four forecast pairs from both analysis sets have been carried out. The results of these experiments are summarized in Fig.22 and 23. When comparing both forecasts-series with the analyses one finds only small differences in the forecast skill (dotted and dashed line in Fig.23) but this does not mean that the forecasts are the same using both sets of analyses as demonstrated by the solid line. It shows the anomaly correlation coefficients between the two sets of forecasts, differing only in the initial data. It shows that the differences between both forecasts on day 6 is as large as the differences between the analyses and the forecasts on day 5. One can interpret this curve as the ultimate skill one can reach with a perfect model but without any further improvement in the analyses. Taking the intersection of the anomaly correlation coefficient with the 60% level as a measure for predictability it would mean that improvements of the forecast model alone can only increase the predictability from 5 1/4 days to 6 1/2 days. These numbers should only be used to demonstrate the importance of the analyses for the forecasts, because only four forecast pairs have been used and because an improvement of the model would also improve the quality of the analyses.

The measure of skill has been chosen here arbitrarily. If one choses the 40% anomaly correlation level to define predictability one finds that 9 1/2 day forecast might be possible with the same skill as the present 7 day forecasts. This possible gain by improved models is only slightly below the estimate by Lorenz (1982) for the northern hemisphere winter in a more thorough investigation on a much larger data sample. The advantage of the present study

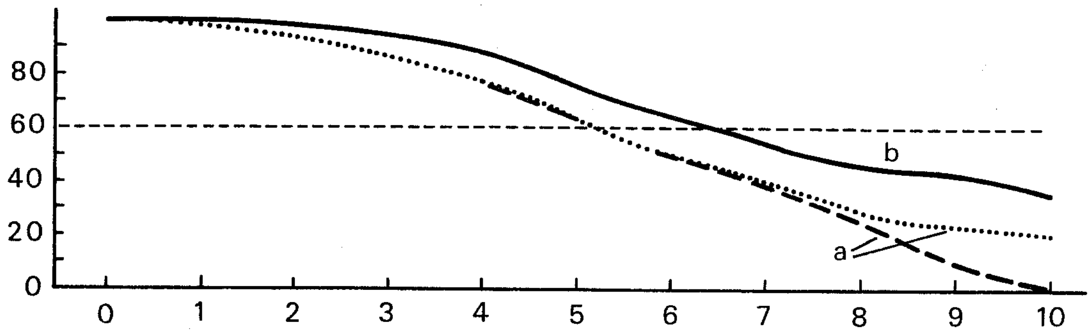


Fig. 22a Ensemble mean anomaly correlation coefficients between analyses and forecasts on two sets of initial data (dotted and dashed line) and b. ensemble mean anomaly correlation coefficients between both sets of forecasts (solid line).

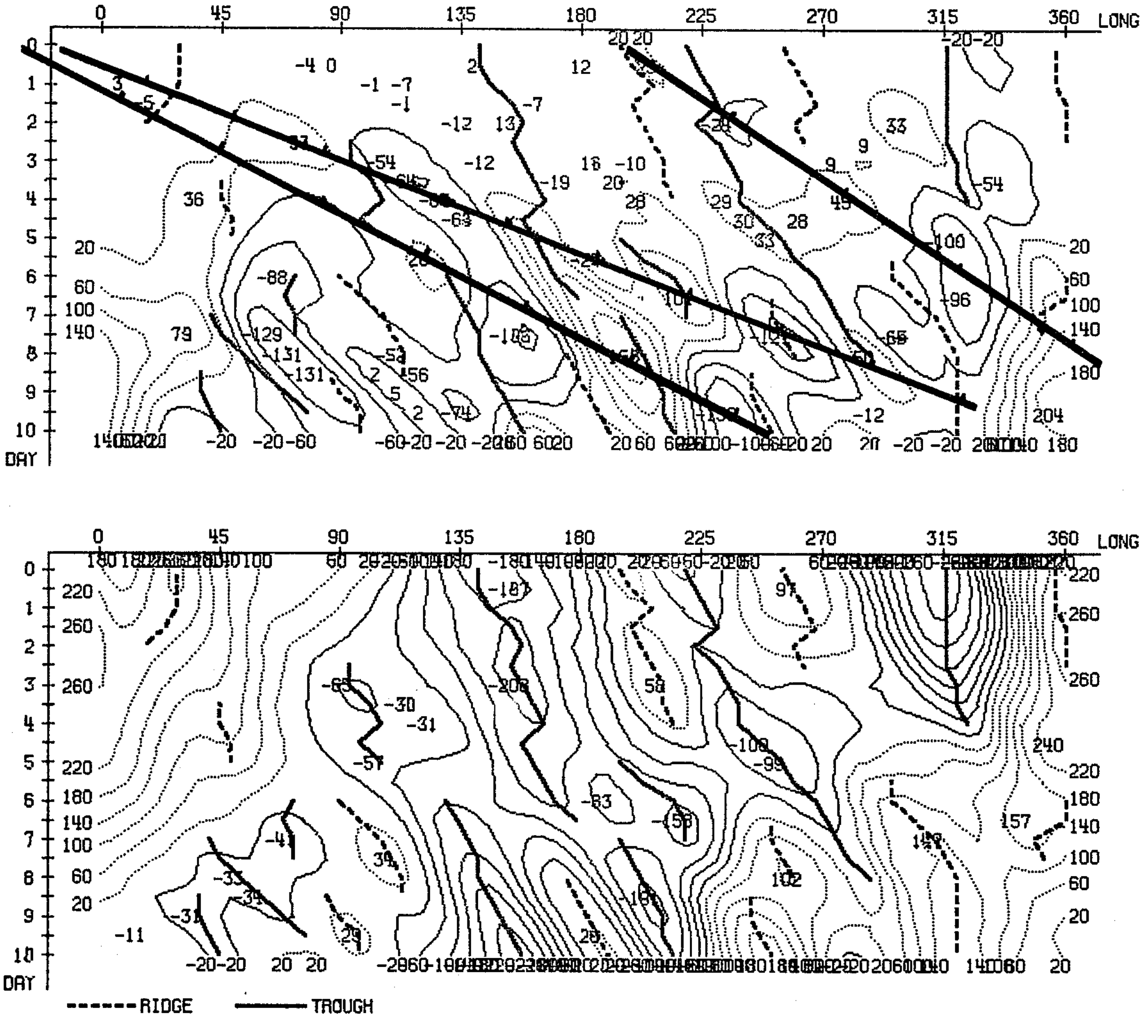


Fig. 23 Trough-ridge diagram of 500 mb height fields between 40°N and 60°N.

Top panel: Difference between both sets of forecasts.
Lower panel: Ensemble mean of one of the forecast sets.

Trough and ridge lines from lower panel are copied to top panel.
Heavy solid lines: Wavetrains.

is that the initial errors are analyses differences while his initial errors were somewhat like typical prediction errors in one, two, etc, day forecasts.

In Fig.23 one can see in a trough ridge diagram how the initial differences influence the 500 mb height field during the forecast. Both sets of analyses differ most over the oceans at 200°E and 330°E. From these one can find rays with positive and negative deviations in constant time intervals. They propagate with a group velocity which is larger than the velocity of the troughs and ridges or the mean zonal flow and can circulate round the earth within 10 days. After 6 days the differences reach amplitudes which are of the same magnitude as the amplitudes of the forecast troughs and ridges shown in the lower panel.

4. CONCLUSIONS

Although the ECMWF analyses-forecast system is one of the most advanced, it has been demonstrated that there are still deficiencies which influence the quality of the forecasts. Here several errors which are common to most forecasts have been listed and connections between them have been drawn. The cause of the systematic errors is not yet known but we have indications that the forcing of the atmosphere by mountains and also by convection is important. We further saw that the tropical atmosphere reacts very quickly to incorrect forcing which may well change the horizontal tilt of the mid-latitude troughs and ridges. The final effect of this was a shift of the subtropical jet. It was also shown that the baroclinic waves became too active during the course of the forecast which was exhibited by too slow filling of cyclones and by an incorrect vertical tilt of the temperature waves. This may be connected with a simultaneous decrease of the static stability.

This listing of systematic errors will guide the diagnostic work in future sensitivity experiments which will be needed to understand the specific defects of the model.

Further errors such as the change of cyclone tracks and the over-prediction of precipitation over Central Europe are probably secondary effects which hopefully will be alleviated when the primary errors are overcome.

Difficulties such as the absence of a contrast precipitation between both sides of the mountains can only be expected to be solved with a much higher resolution in the forecast model and a corresponding improvement of the representation of orography but can most likely be corrected in the output by statistical methods.

It has also been shown that the analysis scheme is crucial for the quality of the forecast.

Acknowledgements

I would like to thank my colleagues at ECMWF for providing me with all the information needed and for many discussions during the collection of this material. I am thankful to Dr. W.Heckley for repairing a good deal of my broken English.

REFERENCES

Akyildiz, V., 1983: A synoptic analysis of cyclones in ECMWF analyses and and forecasts for the winter seasons 1980-81 and 1981-82. ECMWF Tech.Memo., in preparation.

Arpe, K., 1980: Confidence limits for verification and energetics studies. ECMWF Tech.Rep.No.18, 21p.

Bengtsson,L. and A.Lange, 1981: Results of the WMO/CAS Numerical Weather prediction data study and intercomparison project for forecasts for the northern hemisphere in 1979-80. WMO program on weather prediction research (PWPR), 78p.

Cubasch, U., 1981: The performance of the ECMWF model in 50 day integrations, ECMWF Tech.Memo.No.22, 74p.

Defant, Fr., 1974: Das Anfangsstadium der Entwicklung einer baroklinen Wellenstörung in einem baroklinen Grundstrom. Berichte aus dem Institut für Meereskunde Kiel, Nr 4, 103p.

Haseler, J. 1982: An investigation of the impact at middle and high latitudes of tropical forecast errors. ECMWF Tech.Rep.No.31, 42p.

Heckley,W. and A.E.Gill, 1983: Some simple analytical solutions to the problem of forced equatorial long waves. Submitted to Quarterly Journal Roy.Met.Soc.

Hollingsworth, A., Arpe, K., Tiedtke, M., Capaldo, M. and H.Savijärvi, 1980: The performance of a medium range forecast model in winter - impact of physical parameterizations. Mon.Wea.Rev., p1736-1773.

Jarraud, M., Girard, C., and U.Cubasch, 1981: Comparison of medium range forecasts made with models using spectral, or finite difference techniques in the horizontal, ECMWF Tech.Rep.No.23, 96p.

Klinker, E., 1983: Errors of transient waves in the ECMWF medium range forecasts. Forthcoming ECMWF Tech.Rep.

Lorenz,E.N., 1982: Atmospheric predictability experiments with a large numerical model. Tellus, 34, p505-513.

Oriol, E., 1982a. El cido energetico de la atmosfera en las predicciones de ECMWF. Proceedings from the 19th Reunion Bienal RSEFQ, Santander, Spain, 119-120.

Oriol,E., 1982b. Energy budget calculations at ECMWF, Part 1. Analyses 1980-81. ECMWF Tech.Rep.No.35, 114p.

Simmons, A.J. 1981: Current problems in medium range forecasting at ECMWF Seminar on Problems and Prospects in long and medium range weather forecasting, p113-140.

Tiedtke,M. 1982: Winter and summer similations with the ECMWF model, ECMWF Workshop on Intercomparison of large-scale models used for extended range forecasts, 30 June-2 July 1982, p263-314.

Wallace, J.M., Tibaldi, S. and A.J. Simmons, 1982: Reduction of systematic forecast errors in the ECMWF model through the introduction of an envelope orography. ECMWF Workshop on Intercomparison of large-scale models used for extended range forecasts, 30 June - 2 July 1982, p.371-434.