EVALUATION AND USE OF ECMWF FORECASTS

O. Åkesson

European Centre for Medium Range Weather Forecasts
Reading, U.K.

1. INTRODUCTION

In this study forecasts from two winter periods (Nov.80-Mar.81 and Nov.81-Mar.82) are used; attempts are made to isolate typical features in the medium range of the forecasts. During the 3 years of operational forecasting at ECMWF, tuning and improving the model has meant that some characteristics of the forecasts have changed, especially for parameters such as surface fields and precipitation. The over-deepening of surface lows frequently observed in early ECMWF forecasts is now a rare occurrence.

For this study the 500 mb geopotential height fields are used. The reason for choosing mainly the 500 mb rather than the surface level is not just traditional, but it gives a clearer overview of the large scale pattern, which can be expected to be reasonably predictable in the medium range. The link with the surface level is well established and the error patterns are in fact highly barotropic.

The so called systematic forecast error, emerging when the individual forecast errors are averaged over a month or a season (see Arpe, 1982), will be further examined for individual forecasts. Although these errors fluctuate about the mean a great deal, the fluctuations show a more organised pattern than do completely random processes. For periods of typically 3-8 days the error at a single location often remains well above or below the mean, indicating that certain flow types might be associated with certain characteristic error patterns. The objective approach will be to seek teleconnections for the error at discrete points over Europe with height derived parameters in the Northern Hemisphere. This is, in a sense, an approach to predict the 120 hr forecast error. Rinne (1973) and Bennet and Leslie (1979 used principal component analysis for prediction of short range forecast errors in baratropic models.

Finally the use of the forecasts will be briefly discussed in general terms based partly on the results obtained from a study relating forecast consistency with the quality of the forecast in question. This problem is related to the consideration of whether there is information in previous forecasts that could enhance the usefulness of an individual forecast.

2. SYNOPTIC CHARACTERISTICS OF THE ECMWF FORECASTS

From subjective assessments made both at the ECMWF and in the Member States, some characteristic features of the forecasts emerge. These are listed below.

1. A phase lag exists especially for shorter waves over central and northern Europe. The temperature pattern sometimes lags even further behind than the geopotential. For southern Europe and the Mediterranean phase errors are more random. A frequency wavenumber analysis gives quantitative values for global data, as shown by Klinker (1982).

2. For quasi stationary or retrograding long wave systems, for example blocking, the forecast has a tendency to progress too quickly to the east. This seems to be associated with a longitudinal compression of these systems. (Salomonsson, 1981).

3. Cut off lows in the vicinity of the Iberian Peninsula are often less developed, fill and move eastwards or to the north-east too quickly. (Salomonsson, 1981).

4. Forecast cut off lows in the Bay of Genoa or around the Balkan Peninsula are also less or not at all developed. Instead of a cut-off, the forecast sometimes produces a deep through over eastern Europe, often somewhat displaced to the east of the true position and no split in the jet is found over western Europe. A cooling is observed in the upper troposphere. Dell'Osso and Tibaldi (1982) showed that Genoa cyclones are developed better when the mountain barrier of the Alps is increased in the model.

5. Observed central and north European jets are often displaced to the south over Europe and penetrate too far into western Asia. Surface lows in the forecasts are therefore steered on too southerly tracks across central and northern Europe. (Arpe, 1982)

6. Developed lows over the northern part of Europe have a tendency to occlude and fill too slowly. (Akyildiz, 1983)

7. Trends in consecutive forecasts are observed rather often, i.e. forecast events are gradually approaching the real event as demonstrated below.

8. The onset of blocking is normally more accurately predicted than the break-down phase, when the forecasts often retain the systems too long. (Salomonsson, 1981)

9. There is often a time lag for rapid developments in the early stages of the forecast. When the onset has been captured, occasional over-shootings take place. (Akyildiz, 1983). Simmons (1981) showed in an example that this problem can be reduced by a higher resolution of the model.

10. In static, low index situations, the forecast tends to have a higher predictability than for zonal weather types, as shown by Arpe (1982) for January 1981.

11. Strong inversions observed close to snow covered ground are normally too smooth in the forecast and surface temperatures tend to be too high. This is not only a problem of the forecast but already in the analyses as demonstrated by Klinker (1982).

One has however to be very cautious with generalizations ; there are always exceptions to be found. For example, Fig. 1 shows the 500 mb analysis for 4 February 1981 and forecasts verifying on that date. In point 5 above, it is noted that the baroclinic zone is often displaced to the south of the true position by the medium range forecasts. In Fig. 1, 9 consecutive forecasts are shown where the jet is gradually being displaced further and further to the north the longer the forecasting time considered. Notice the trend in these forecasts, which is mentioned in point 7.

3. THE DAILY VARIATION OF THE 500 MB HEIGHT ERROR IN 5 DAY FORECASTS OVER EUROPE

The daily variation of the 500 mb height error at specific locations for the winter season 1981-82 (Fig. 2) exhibits characteristics similar to those found for the variation of forecast scores such as anomaly correlation, with frequent periods of typically 3-8 days when heights remain above or below the observed mean. These variations indicate that they might be associated with certain flow types. Considering the position 54N, 00E just off the east coast of England with a high negative mean error for the 81/82 winter season as can be seen in Fig. 2, one finds mostly well defined but different flow types for each period. Fig. 3 shows two periods with higher forecast than observed geopotential at this location, the first example associated with an under-developed cut off low near the Pyrenees with the build up of too high pressure all over western Europe and the second associated with a slight eastward displacement of a blocking high west of the British Isles.

There is one particular flow type that systematically produces too low geopotential heights. This happens when there are observed cut-off lows over the south Adriatic and Greece with an upstream split of the jet and a long wave ridge positioned over western Europe and the eastern part of the Atlantic. (See Fig. 4). Arpe (1981 *) found large errors of this kind in connection with downstream energy dispersion.

4. ERROR RELATIONSHIPS

Simultaneous relationships between meteorological variables at widely separated points on earth are often referred to as teleconnections. For instance, Wallace and Gutzler (1981) have studied teleconnections in the geopotential height field in order to identify standing oscillations in the planetary waves. In a similar way, in this study relationships have been sought among height derived instantaneous parameters, (the predictors), both at the initial time step of the forecast (D+0) and at other

---

* Presented at the IAMAP conference in Reading, 1981.

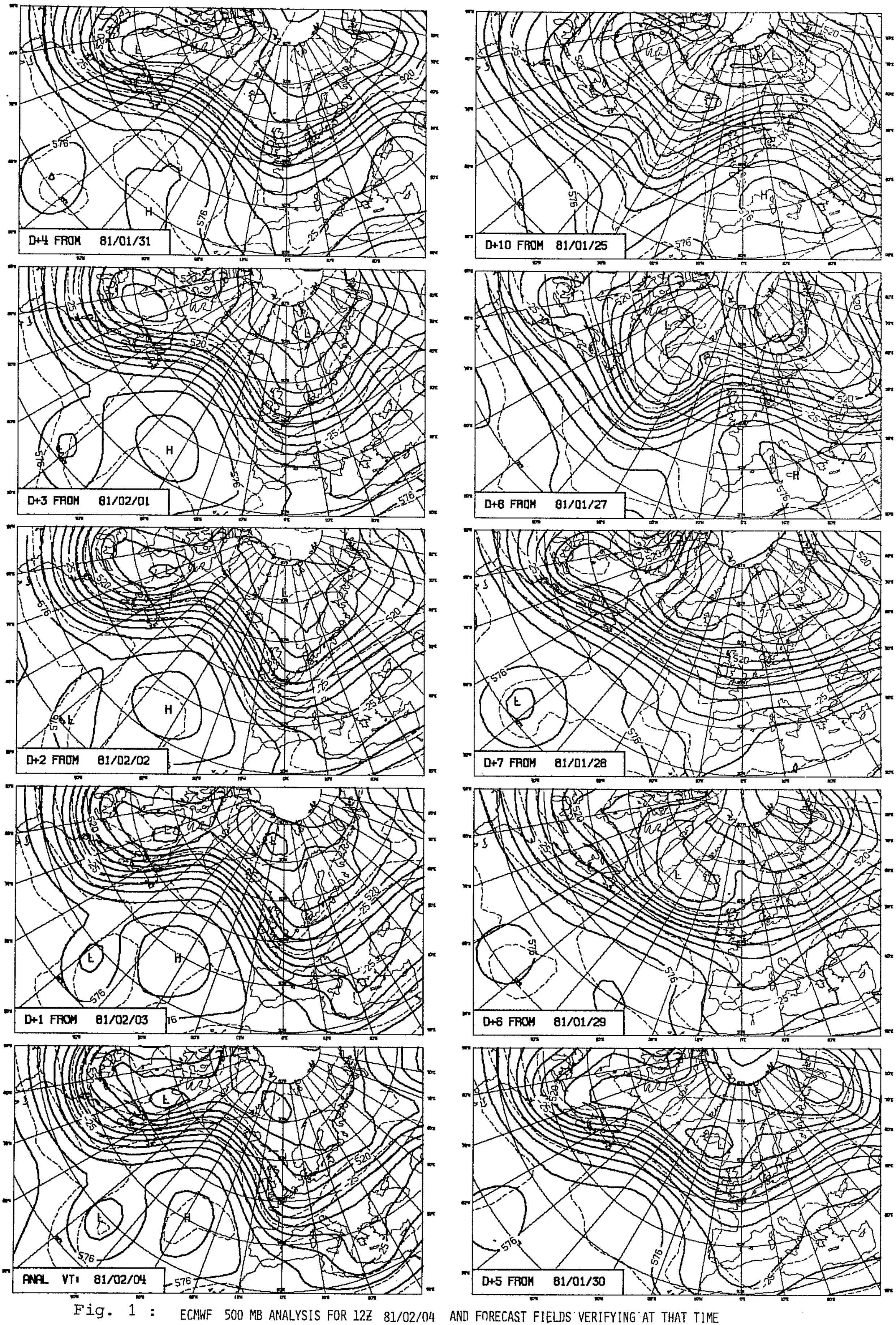


Fig. 1 : ECMWF 500 MB ANALYSIS FOR 12Z 81/02/04 AND FORECAST FIELDS VERIFYING AT THAT TIME
CONTOUR INTERVAL 8DAM(THICK LINES),5K(DASHED LINES)

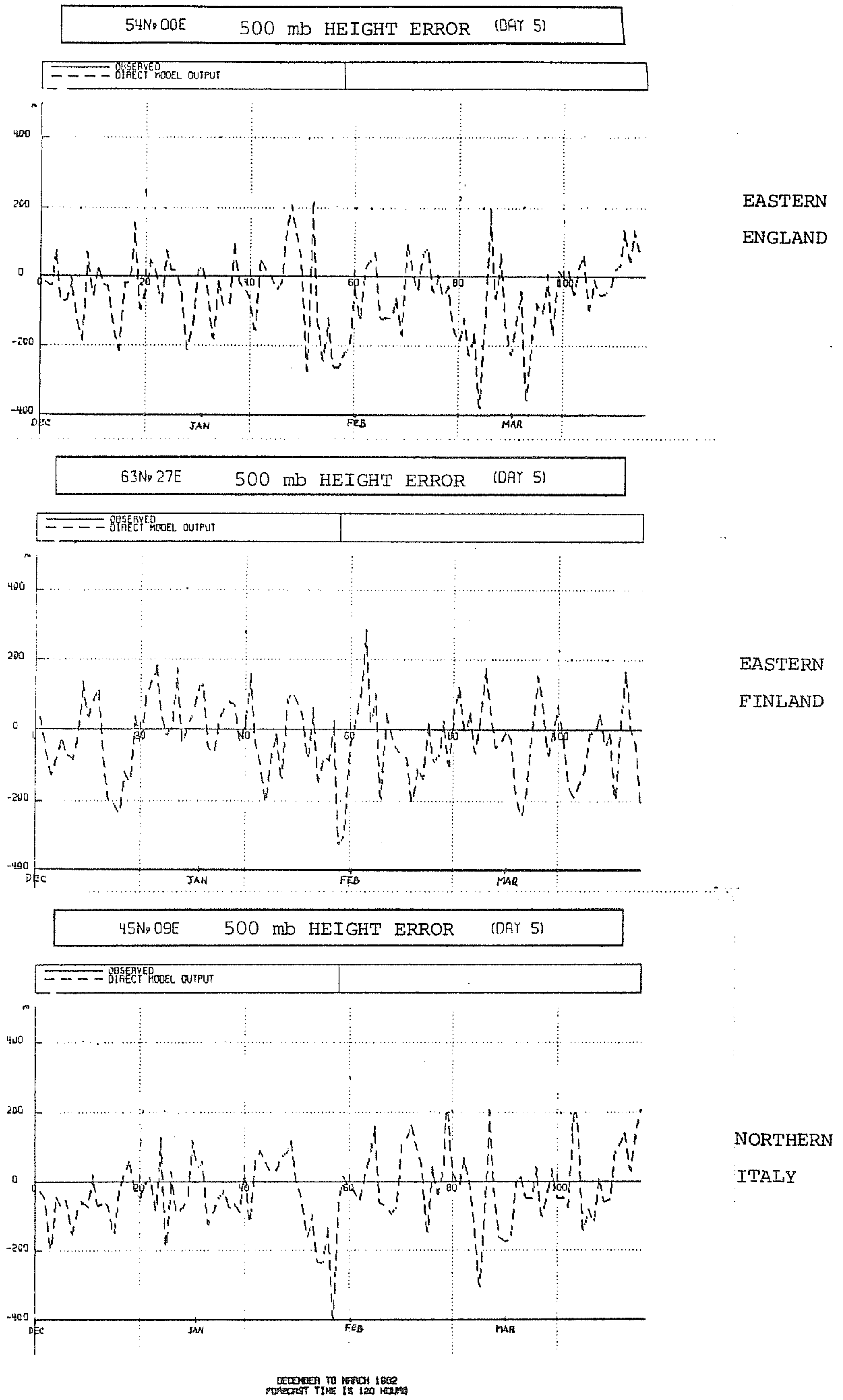


Fig. 2 : Meteograms of 120 hour 500 mb height errors at 3 European locations for the time period Dec.81 to Mar.82

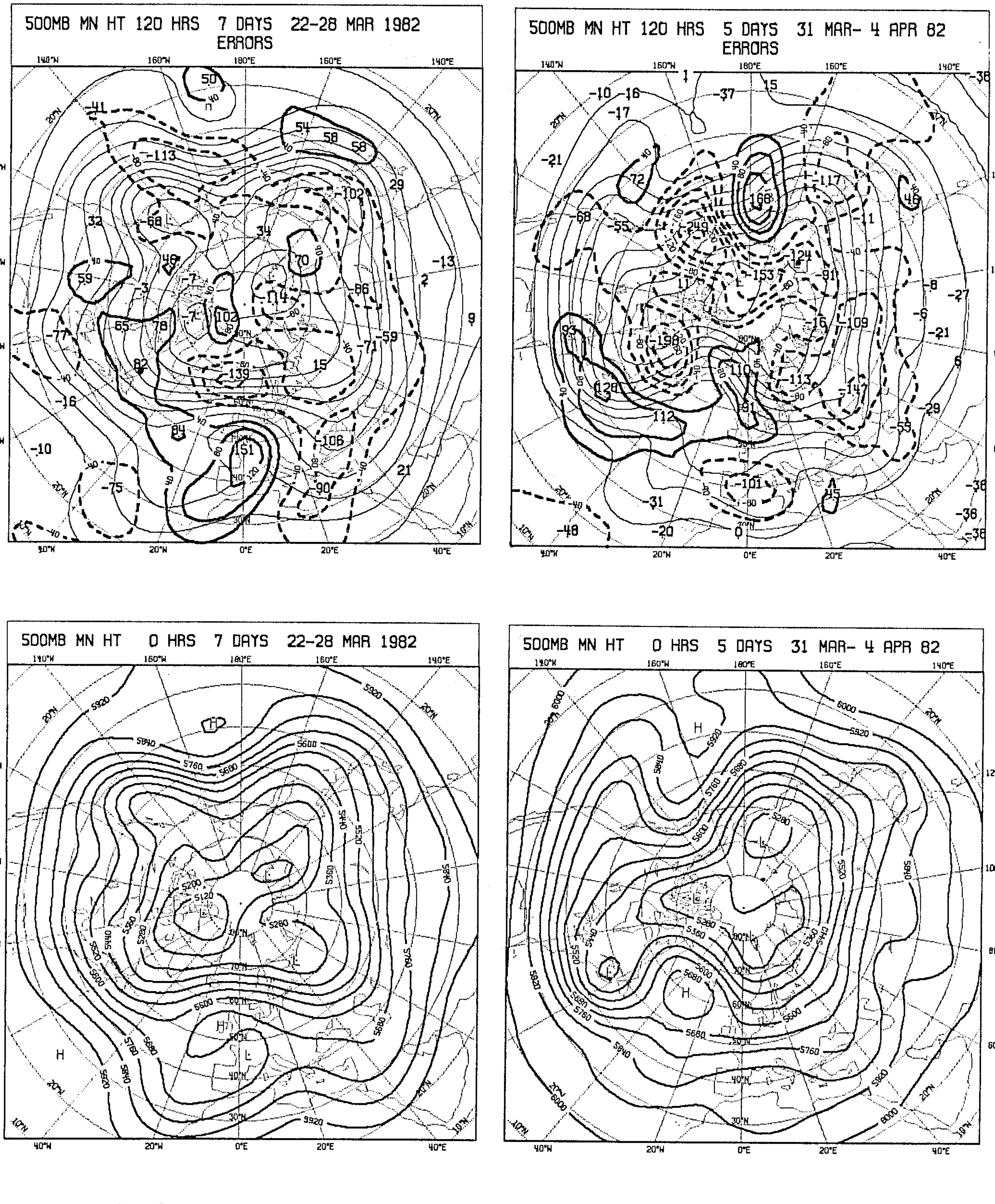


Fig. 3 : Mean 500 mb fields (thin solid lines) for D+5 forecast (top) and verifying analysis (bottom). The mean error field is displayed in the top figure as thick solid lines (positive) and dashed lines (negative). Unit : m.
The figures on the left are from the time period : 22-28 Mar. 1982.
The figures on the right are from the time period : 31 Mar.- 4 Apr.1982

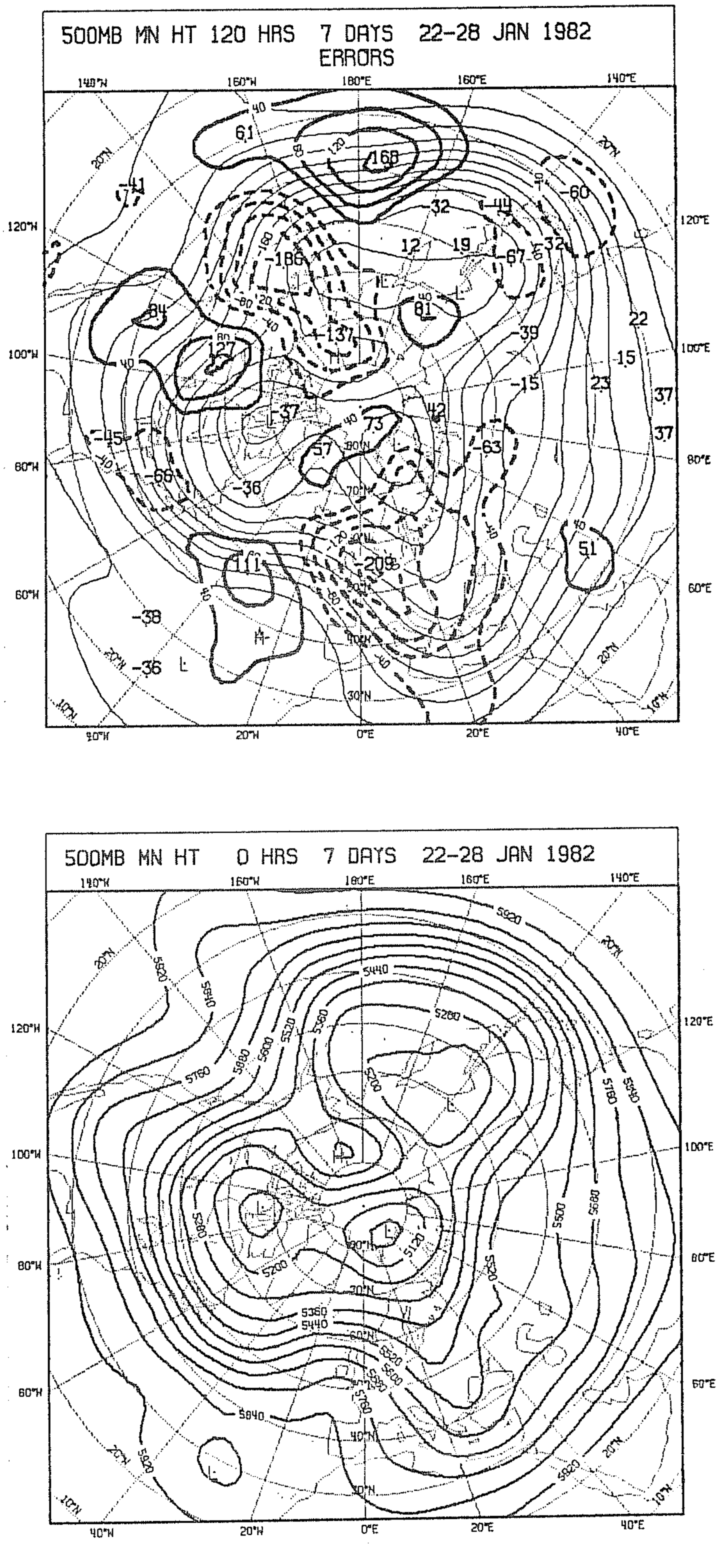


Fig. 4: Mean 500 mb height fields (thin solid lines) for the D+5 forecast (upper figure) and the verifying analysis (lower figure) for the time period 22-28 January 1982.

The mean error field is displayed in the top figure with positive errors in thick solid lines and negative errors in dashed lines.

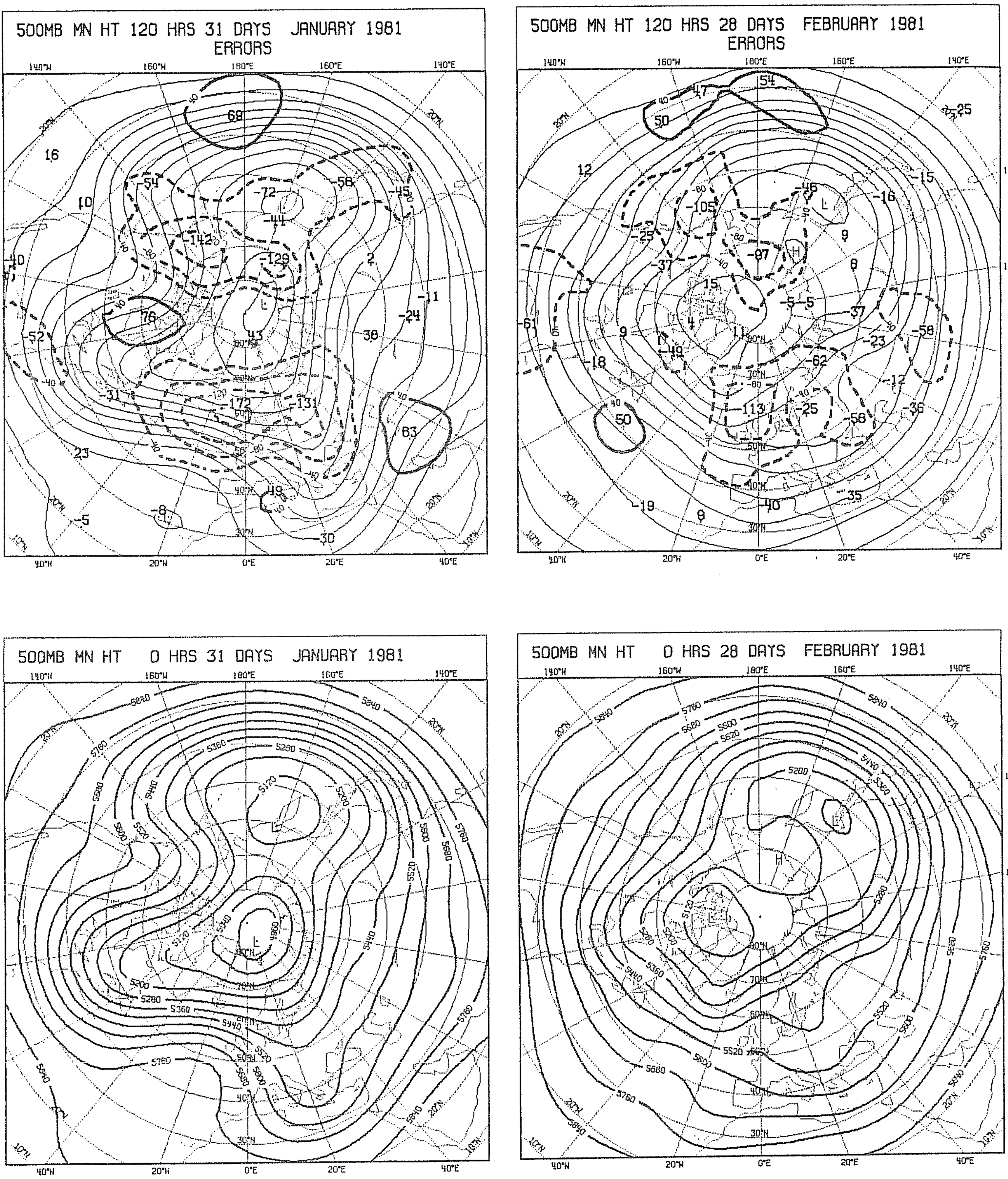
Unit: m

timesteps into the forecast (D+1 or D+5), to be related in some way with the D+5 500 mb geopotential error (the predictand) at single points across Europe. Wallace et al (1983) found intuitively a connection between the position of too low height fields in the forecasts near the Rockies and the position of the jet stream. To do this systematically the correlation has been calculated between the error at one point with the other parameter in points distributed over the entire Northern Hemisphere with a spacing of 9x9 degrees of latitude and longitude. Two 5-month winter periods have been considered: 1 Nov. 1980-31 Mar. 1981 and 1 Nov. 1981-31 Mar. 1982, one period being treated as an independent sample to test the reproducibility of the findings from one period to the next. Potential 500 mb predictors to be tested include geopotential height, height anomaly and D+5 and D+6 forecast consistency, where D+6 refers to the forecast issued 1 day earlier so that both forecasts refer to the same time of validity. Furthermore, meridional flow, D+0 to D+5 height tendency and the D+1 and D+5 error autocorrelation have been tested. A one to one correspondence was not expected between the error and any other variable, but the level at which noise is separated from signal is often difficult to judge.

4.1 Relationships of initial state parameters

For a D+5 forecast, the corresponding errors over Europe can be shown in one example to originate from upstream regions as far away as from the Pacific (see section 5). Hollingsworth et al (1983) show in forecast experiments on different analyses that such errors in the forecasts could be due to analyses errors over the Pacific. To illustrate the belief that certain flow regimes at D+0 might be associated with the error 5 days later, consider Fig. 5. This figure shows the mean D+5 forecasts and corresponding error pattern (upper half) and the mean verifying analysis (lower half) for two months - Jan. 81 and Feb. 81. Despite a 5 day time lag, the verifying mean monthly analysis at the lower part of the figure can also be expected to be fairly representative for the mean initial states for the two corresponding months. The error characteristics can be seem to be very different for the two months and so can the flow-pattern over most of the hemisphere, Jan. 81 displaying strong negative errors over much of Europe and with a predominance of low index circulation, whereas the opposite is observed for Feb. 81.

Rather disappointingly a test with height anomaly as a candidate predictor of the errors comes out weak and rather noisy. This is shown for two points in Europe at 54N,0E and 45N, 18E in Fig. 6a and b. At 54N,0E the response is generally somewhat lower than at 45N, 18E which shows negative correlations of -0.3 to -0.4 over the northern and eastern Atlantic and positive correlations over the point itself, indicating that anomalous flow of the initial state over Europe and the Atlantic has a tendency to lead to larger errors over eastern Europe. A ridge-trough situation from the eastern Atlantic to eastern Europe at the initial state would then tend to lead to negative errors and the trough to over-develop and vice-versa.



500MB MN HT 120 HRS 31 DAYS JANUARY 1981
ERRORS

500MB MN HT 120 HRS 28 DAYS FEBRUARY 1981
ERRORS

500MB MN HT 0 HRS 31 DAYS JANUARY 1981

500MB MN HT 0 HRS 28 DAYS FEBRUARY 1981

Fig. 5 : 500 mb height fields and error for a) Jan. 81 (left)
b) Feb. 81 (right)

See Fig. 3 for further explanations

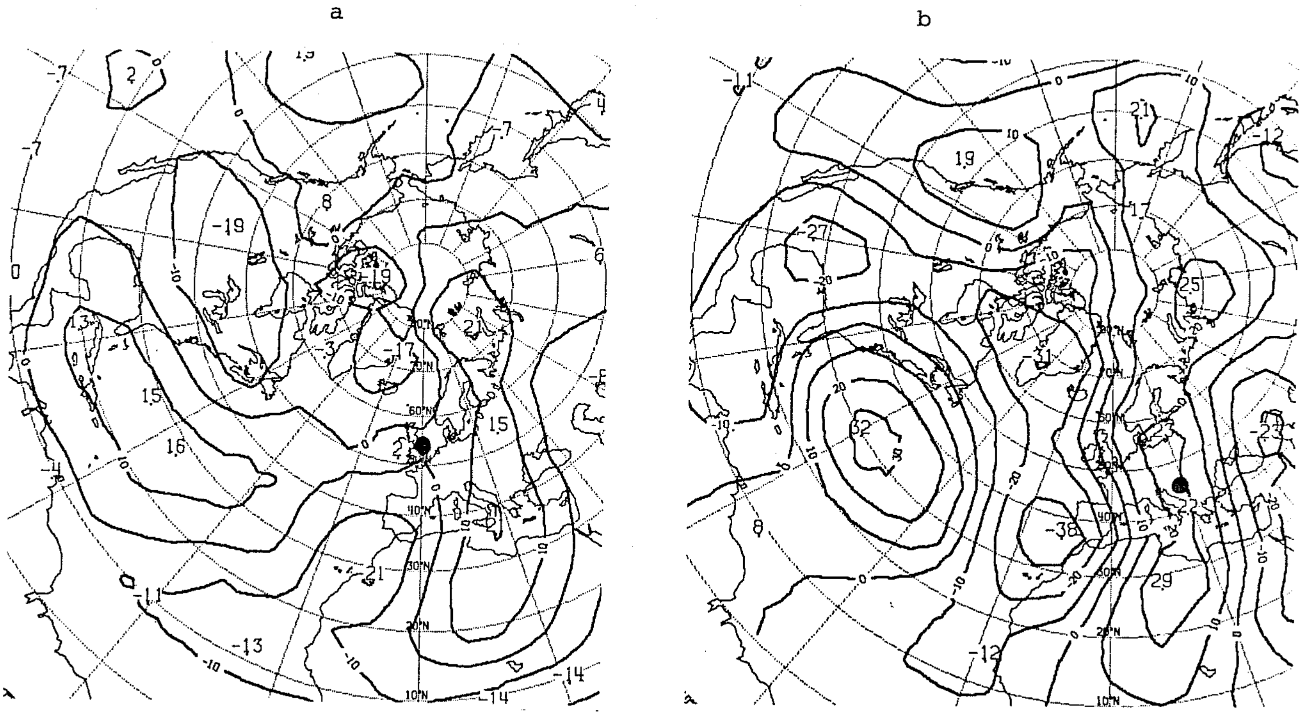


Fig. 6: Correlation isolines (%) between the D+5 500 mb geopotential height error at a): 54N, 0E (left) and (b): 45N 18E (right) and the D+0 height anomaly for the 81/82 winter season. The error position is shown by a dot.

Correlations further west are weaker and no conclusive evidence is given as to whether a certain type of anomalous flow there gives rise to certain error distribution over Europe 5 days later. The response from meridional flow is, however, somewhat stronger. The meridional flow has been divided into its southerly (positive) and northerly (negative) flow components. The correlation between southerly winds with initial data and the forecast errors for a single point in eastern Europe for the two winter seasons are shown in Figs. 7a and 7b. An area of negative correlation over western Canada using a southerly wind component as predictor for the error at point 45N, 27E for the winter period 80/81 (Fig. 7a) has been displaced southwards for the following winter period. This suggests that negative errors at 45N, 27E over Rumania have often been preceded by southerlies over and along the Rocky Mountains 5 days earlier. The northerly component (not shown) shows a weak negative correlation for both periods in the same area, indicating a possible link between positive errors in Rumania and northerly winds over the western parts of Canada and the US in the initial state. However, the mean analysis for Jan. 81 (Fig. 5a) shows only a ridge with small meridional components. Even northerly flow regimes can be found preceding other periods of strong negative errors over eastern Europe. Despite a weak predominance of southerlies, there tends to be quite a variability in the initial flow pattern and that can probably explain why the signals using height anomaly are weak.

Testing other points in Europe yields weaker but mainly similar relationships. In some cases, however, there are inconsistencies between the two periods.

4.2 Error correlation

In an attempt to relate the error at an early stage (D+1) of the forecast to the error at D+5 for a certain location, the error correlation was calculated. The correlation between the D+5 error at 54°, OE and the error elsewhere in much of the Northern Hemisphere at D+1,2,3 and 4 is shown in Fig. 8 a-d respectively. Only at D+3 and 4 are there areas of significant relationships to be found, one position near the point at 54°, OE and another area of negative correlation further west, just south of Greenland, which can be traced to the east coast of Canada at D+2 and with some uncertainty to the area of Hudson Bay at D+1.

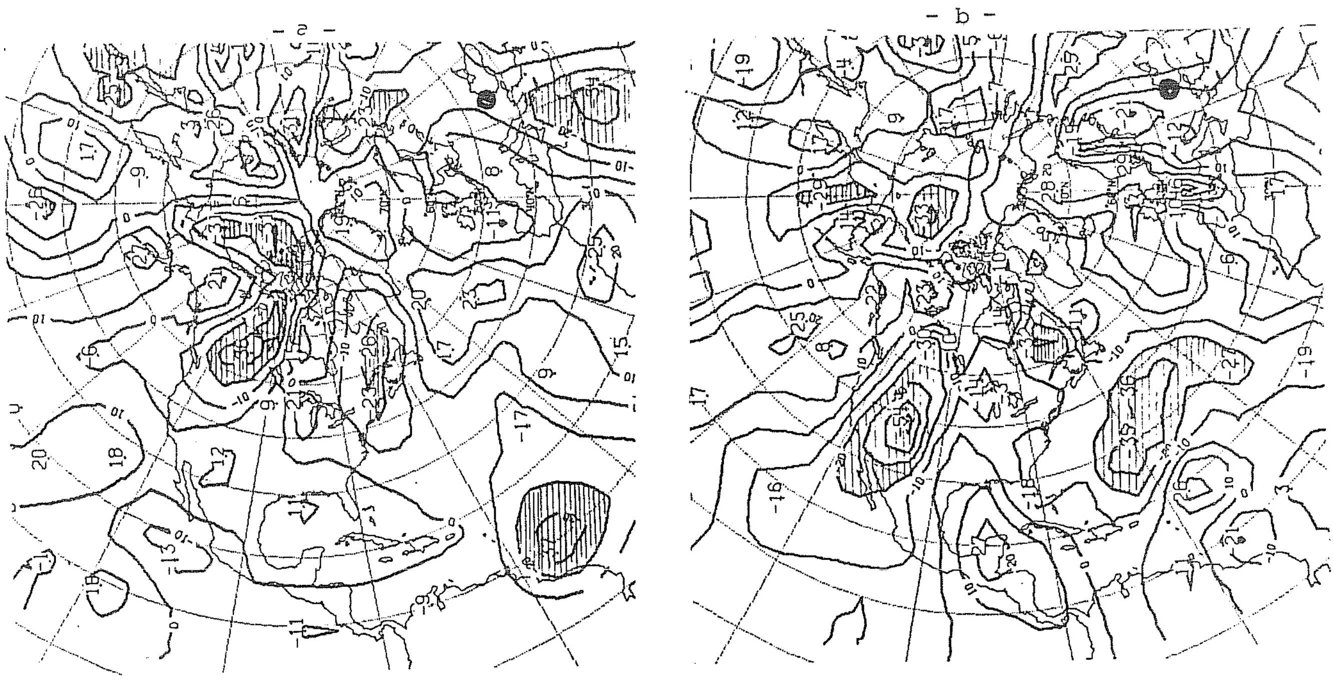


Fig. 7 Correlation isolines (%) between the D+5 500 mb geopotential height error at 45 N, 27 E and the D+0 500 mb southerly wind component. The error position is shown by a dot.

Figure 6a shows correlation for winter 80/81 and figure 6b for winter 81/82.

Areas with correlation coefficient ⩽o.2 are hatched.

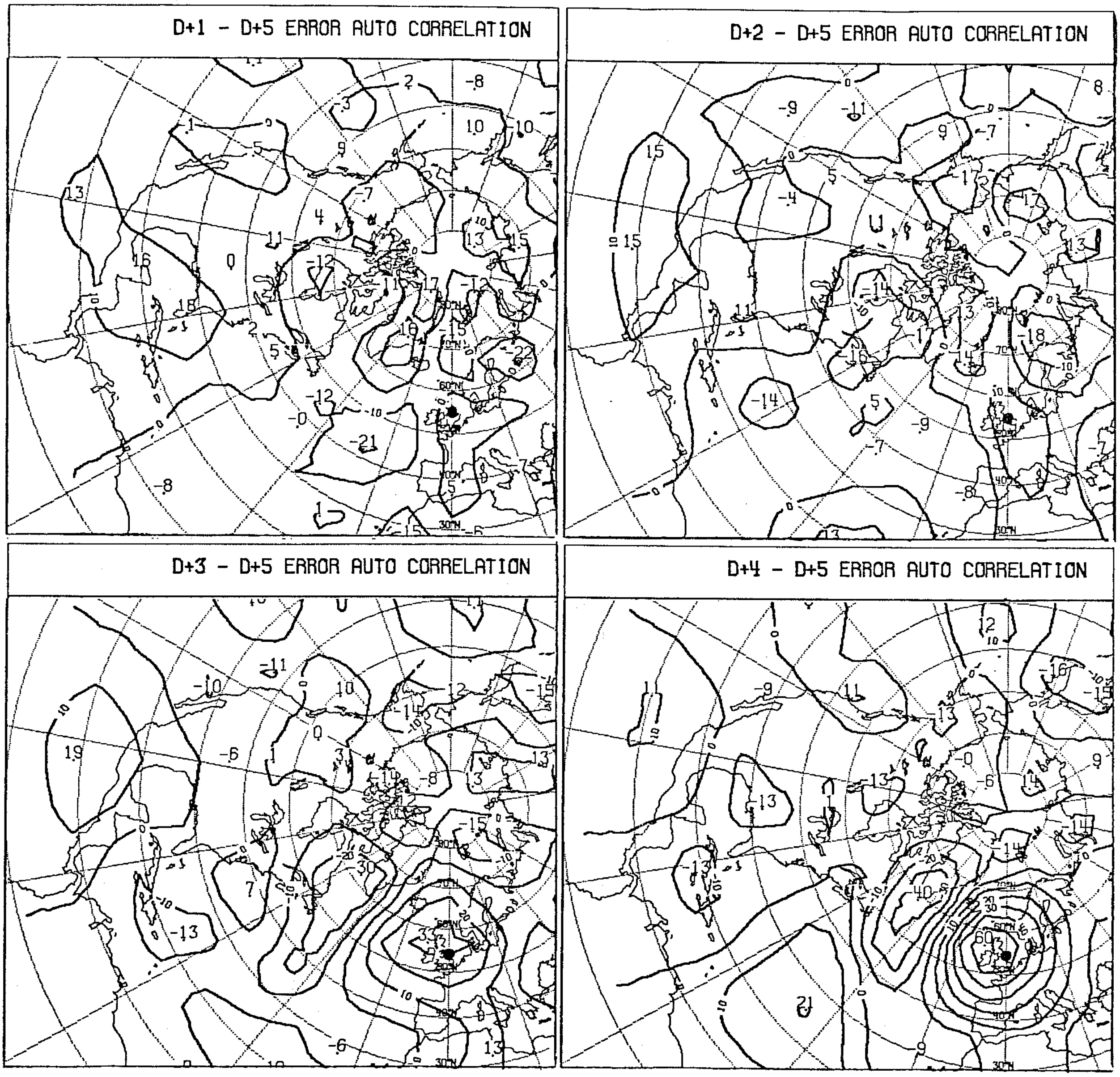


Fig. 8: Error correlation isolines (%) between the D+5 500 mb geopotential height error at 54N, 0E and the spatial height error at a: D+1, b: D+2, c: D+3 and d: D+4 for the 81/82 winter season. The error position is shown by a dot.

4.3 Relationships of D+5 parameters

Compared to the weak initial state teleconnections, the D+5 predictors give, in general, a stronger response. The spatial variability is however large with fairly strong relationships observed over eastern Europe, whereas western and northern Europe are mainly associated with feeble correlations. Nonsignificant relationships are obtained using consistency (for a further discussion, see section 6).

4.3.1. D+0 to D+5 geopotential height change (tendency).

If the ECMWF model was in general a cautious and persistent model and could not forecast major changes with enough amplitude, positive errors would be connected to negative height changes (and vice versa) and the resultant correlations would be negative. Positive correlations indicate consequently a tendency of the forecast to exaggerate and over-develop lows and highs. Positive correlation is found over eastern Europe associated with the error in the same area, whereas points in western and northern Europe show only weak signals.

Both the positive and the negative parts of the tendency have been tests as predictors. For points in eastern Europe the negative height change is found to be a better predictor than the positive change. Spatial differences are large. Fig. 9 shows the relationship between negative height change and the error at two different locations: 54 N, 0 E (Fig. 9a) and at 45 N, 18 E (Fig. 9b). Notice that the maximum response for point 54 N, 0 E is found over the Alpine region, indicating that decreasing height over the Alps is often associated with negative D+5 height errors over Britain. Arpe (1981, see above) tried to explain such different positioning of lows over Europe but also over other parts of the world by the failure of the model to produce a southward guidance of the downstream energy dispersion. This leads to the conclusion that overdevelopment of depressions is more typical than overprediction of anticyclones.

Fig. 9 Correlation isolines (%) between the D+5 500 mb geopotential height error at a.: 54 N, 0 E (left) and b.: 45 N, 18 E (right) and the negative part of the D+0 to D+5 500 mb height change for the 81/82 winter season. The error position is shown by a dot.

4.3.2 Height anomaly

Again the response is located mainly over eastern Europe, indicating that forecast extremes have a tendency to be more frequently overdeveloped in this region. This is illustrated by Fig. 10 which shows stronger correlation values for 45N, 18E located over that position and with isolines extending to the north. At 54N, 0E there is also positive but weaker correlation around the point and extending south-eastwards.

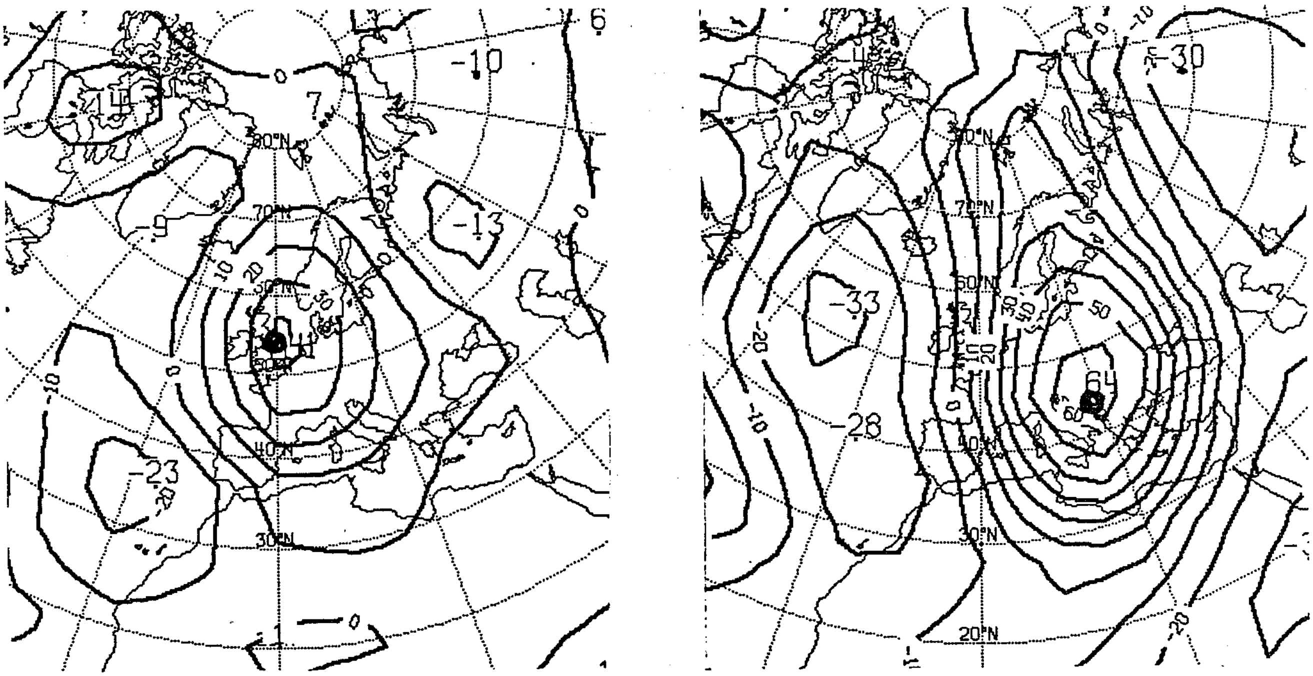


Fig. 10: Correlation isolines (%) between the D+5 500 mb geopotential height error at a: 54N, 0E (left) and b: 45N, 18E (right) and the D+5 height anomaly for the 81/82 winter season. The error position is shown by a dot.

4.3.3. Meridional flow components.

Meridional flow, subdivided into southerly (positive) and northerly (negative) components show stronger response for the error in eastern Europe than anywhere else in the European area (cf Figs. 11 and 12). The maximum relationships are not situated right over the point of the error, as is mainly the case for tendency and anomaly predictors. Instead the maximum or minimum area of correlation is located between 10 to 20 degrees latitude to the west or to the east of the error point. Thus, the forecast of southerlies over the westernmost parts of Europe has a tendency to be associated with the overdevelopment of highs for points at 45N, 18 E and 45 N, 28 E, whereas the forecast of northerlies over the western part of the Alps tends to be associated with overdevelopment of troughs and lows for the same two points. Note also the stronger response and consistency for the northerly component between the two seasons than for the southerly component, again indicating the fact that the overdevelopment of highs are not as consistent and frequent as the overdevelopment of lows.

The results from the 3 predictors tested above for the negative errors indicates consequently a flow regime similar to that seen in Figs. 4 and 5a, where strong negative errors over most of Europe are associated with an overdeveloped forecast trough positioned over eastern Europe and a suppressed ridge over the westernmost parts of Europe and the eastern Atlantic. The observed pattern often exhibits a cut-off low over the central and eastern Mediterranean and a split in the jet over Europe. The maximum response located over the Alps for the error at 54 N, 00E (Fig. 9a) using tendency as predictor, indicates a link between falling pressure over central Europe with too low forecast geopotential height values over Britain, where usually the suppressed forecast ridge is observed. It is therefore likely that the too weak ridge over Britain is not amplified sufficiently and, in association with that, a cut-off low over the Mediterranean is not produced. The response, albeit weak, from the Rocky Mountain area using D+0 parameters indicates that incorrect mountain forcing might be one possible reason for the systematic error, but incorrect radiation or forcing from other sources and areas cannot be ruled out either.

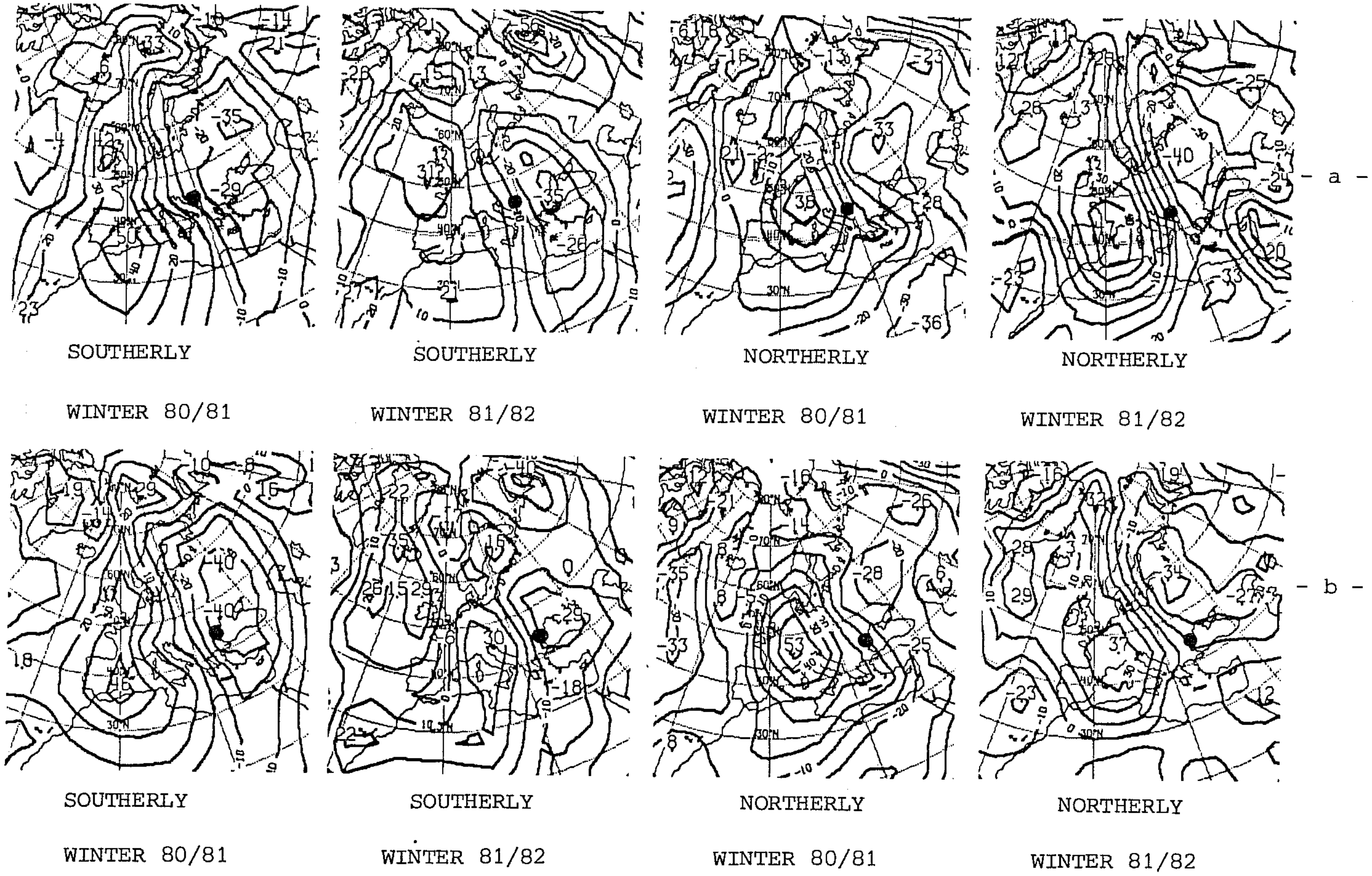
Figure 13 shows the surface analysis and forecasts verifying on 21 January 1981, with an unusually strong cyclone just to the south-east of Italy. Of the forecasts longer than 4 days, only the D+6 prediction by the forecast from the 15 January managed to capture the correct event well. All the other medium range forecasts show various signs of the typical forecast deficiencies listed above for this kind of situation.

This low originated in the lee of the Alps near Genoa one day earlier and moved south-eastwards. Genoa lows are also to an extent associated with a similar error structure to that described above, but with less amplitude.

Typical Genoa lows do not emerge as strongly in the relationships; this may be because:

1. as mentioned in point 4 in the list of known typical features, section 2, troughs and lows over the Mediterranean or over southern Europe tend to be less developed and have an eastward phase shift compared to the observed events and the results above are related to the forecast position which consequently is normally found to be to the east or south-east of the true position. This feature might be due to incorrect forcing by the Alps.

2. The Genoa lows seems to be of more transient character in the model and in reality than the persistent depressions further to the east or south-east, at least for the two winter seasons tested.



SOUTHERLY

WINTER 80/81

SOUTHERLY

WINTER 81/82

NORTHERLY

WINTER 80/81

NORTHERLY

WINTER 81/82

SOUTHERLY

WINTER 80/81

SOUTHERLY

WINTER 81/82

NORTHERLY

WINTER 80/81

NORTHERLY

WINTER 81/82

Fig. 11 : Correlation isolines (%) between the D+5 500 mb geopotential height error at a.: 45 N, 18 E (top row) and b.: 45 N, 27 E (bottom row) and the D+5 500 mb meridional wind components for two winter seasons. The error position is shown by a dot. The four figures to the left show correlation using the southerly wind component and the four figures to the right show correlation using the northerly wind component.

SOUTHERLY WINTER 80/81

SOUTHERLY WINTER 81/82

NORTHERLY WINTER 80/81

NORTHERLY WINTER 81/82

- a -

SOUTHERLY WINTER 80/81

SOUTHERLY WINTER 81/82

NORTHERLY WINTER 80/81

NORTHERLY WINTER 81/82

- b -

Fig. 12 : Same as Fig. 11 but for forecast error points a.: 63 N, 27 E and b.: 54 N, 0 E.

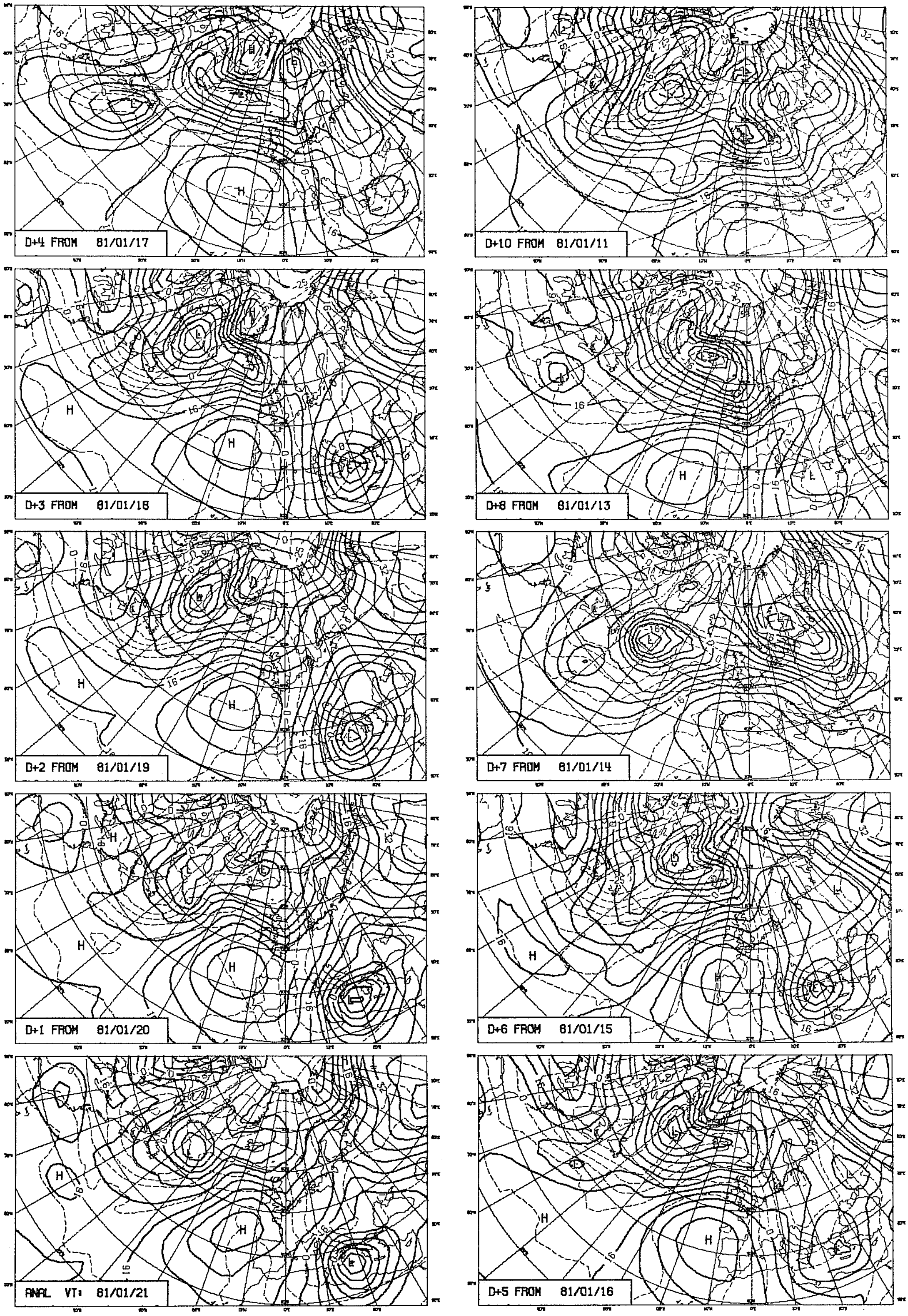


Fig. 13 ECMWF 1000 MB ANALYSIS FOR 12Z 81/01/21 AND FORECAST FIELDS VERIFYING AT THAT TIME
CONTOUR INTERVAL 4DAM(THICK LINES), 5K(DASHED LINES, 850 MB TEMPERATURE)

5. EXAMPLES OF DATA SENSITIVITY, ERROR GROWTH AND ERROR PROGRESSION

In some cases comparisons of consecutive ECMWF forecasts or of ECMWF forecasts with forecasts from other sources show large inconsistencies already at an early stage into the forecast. This could indicate that data problems have caused fatal, but not necessarily large errors which then grow dramatically, for example if the error is positioned in a developing baroclinic wave. Hollingsworth et al (1983) give a few examples of forecast differences due to analyses differences by differing schemes. Differences in the forecasts could be traced back in time to the initial data and the reasons for the differences were investigated.

This method can also be used in the case of large differences appearing comparatively early in two consecutive forecasts. One such case is found in Fig. 14 showing differences between the two ECMWF forecasts. Significant differences at D+5 over the Atlantic can be followed back in time with some confidence and the probable area of origin is the region both south of Alaska and east of the Canadian Rockies where significant errors are found already after 12 hours (lower left in Fig. 14). In an experiment performed at ECMWF by Cats and Åkesson (1983) a section of the 3-dimensional 24 hr forecast from 2 September 1982 from this area was transplanted into the analysis from 3 September. A forecast run from this modified analysis shows similarities with the operational forecast from the later date and the difference field so obtained exhibits about the same pattern and sign as, but with less amplitude than, the difference west of Ireland in Fig. 14.

Growth of prognostic differences has also been described by Atlas (1982) and Baede et al (1982) performed sensitivity experiments using FGGE analyses. Downstream propagation has been investigated by Simmons and Hoskins (1979).

6. FORECAST CONSISTENCY

The accuracy of forecasts normally improves as they get closer to the event, i.e. a D+5 forecast is on average more accurate than the corresponding D+6 forecast issued one day earlier. The variability between consecutive forecasts is a reflection of the inaccuracies that can be expected from a forecast at that projection in time. In the medium range (4-10 days) the variability can sometimes be very large so as to produce a completely new weather type compared with the previous forecast. At least one of the two consecutive forecasts must be misleading, often both. It could be argued that in such situations always the latest available forecast should be considered only, since that forecast normally is the more accurate of the two. On the other hand assuming that the errors are random, the subsequent forecast estimates for a particular region or point would normally converge towards the correct event,

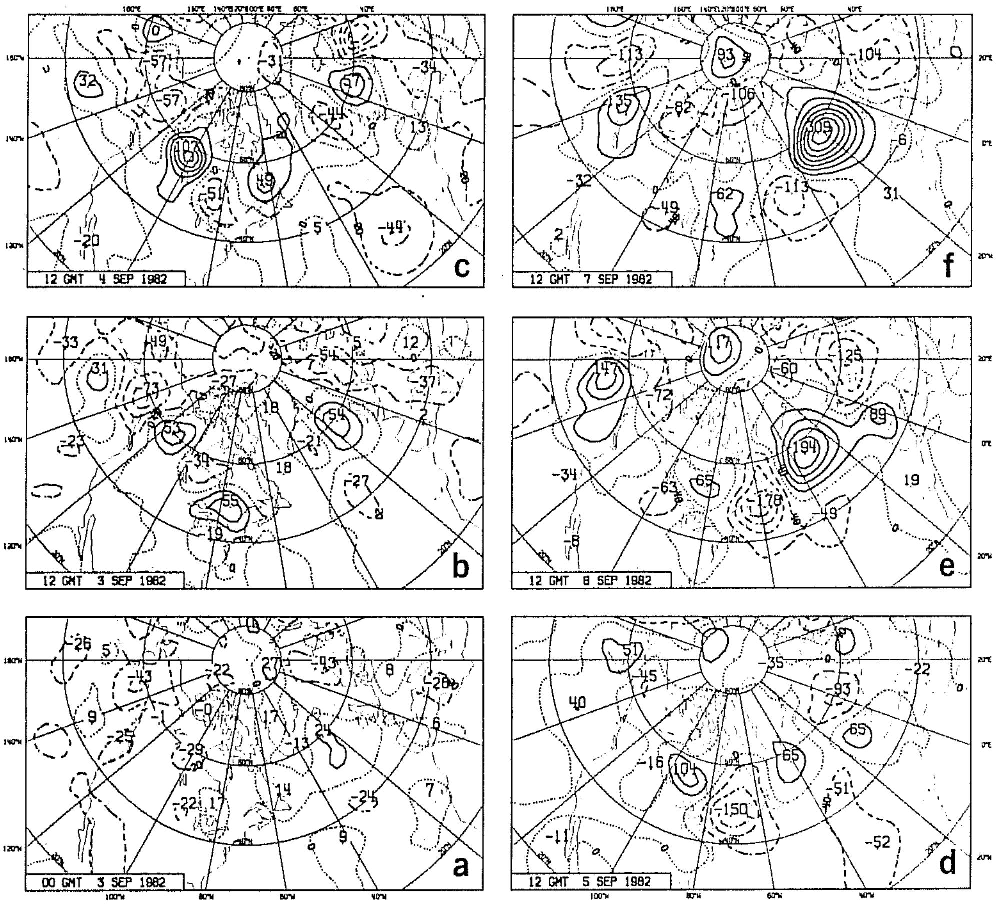


Fig. 14

500 mb geopotential height difference (m) between ECMWF forecast from 12 GMT, 2 September 1982 and:

(a) ECMWF analysis 0 GMT, 3 September 1982,

(b) ECMWF analysis 12 GMT, 3 September 1982;

(c) through (f) 1 through 4 day ECMWF forecast from 12 GMT, 3 September 1982;

at corresponding verification times. Solid contours: positive differences; dashed contours: negative differences; dotted contours: zero differences. (Contour interval: a-c: 20m; d-f: 40m).

so that errors in the different estimates would cancel out to some extent in a combined estimate, i.e. treating them as independent forecasts. Thompson (1976) has discussed the improvement of accuracy by combining independent forecasts. However, the presence of systematic errors means that individual estimates are not independent on average. In fact, the errors in consecutive 500 mb forecasts are fairly highly correlated in the medium range: 0.62 for D+5 and D+6 forecasts and 0.48 for D+5 and D+7 forecasts for a 4 month period from December 81 to March 82. This implies that the forecast often produces similar errors day after day. The information in a combination of estimates will therefore not be expected to be significantly more accurate than in a single medium range forecast alone. Results of ensemble forecasts will be discussed at the end of this section.

If, however, the different forecast solutions agree to a large extent, will the credibility of the predicted event be enhanced by the support given by previous estimates, although they are not entirely independent? Fritsch and Chelius (1982) have discussed this subject for probability of precipitation forecasts. Here, the consistency between all 500 mb D+5 forecasts from the 4 month period and the corresponding D+6 forecasts is related to the skill of the D+5 forecasts in scatter diagrams (see Fig. 15). The standard of measuring skill and consistency was STD of the error and difference respectively. Three different sized areas were used: 1. Europe + surrounding areas, 2. northern Europe and 3. the Scandinavian Peninsula, to find out if area size has any effect on the relationship. As can be seen in Fig. 13 there is a slight, positive relationship for these areas.

Relationship between consistency and skill expressed as STD of difference

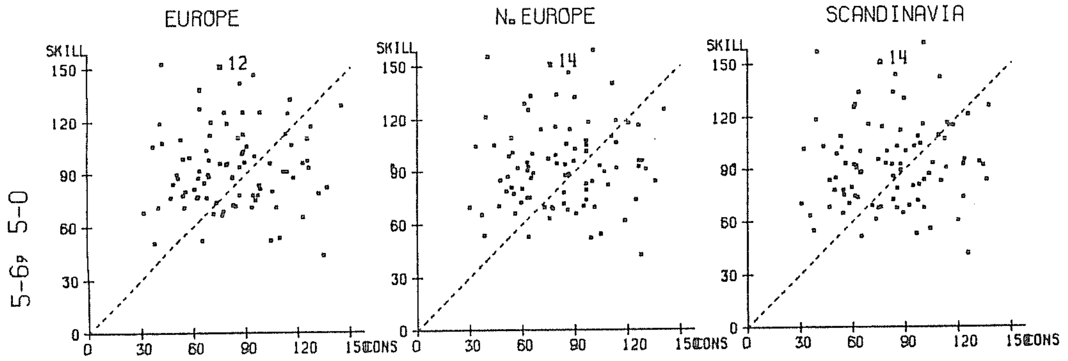


Fig. 15 Skill of D+5 500 mb height forecasts as a function of consistency (CONS) between consecutive D+5 and D+6 forecasts for the time period Jan.- Apr. 82. Skill is expressed as STD of the error and consistency as STD of the differences (m). The left-hand frame is for Europe and surrounding areas, the middle frame is for northern Europe and the right-hand frame is for the Scandinavian Peninsula. The number in each frame gives the correlation of the relationship.

Choosing anomaly correlation instead of STD as a standard of measurement (Fig. 16), a somewhat higher relationship can be seen. Shown at the lower part of Fig. 16 is a combination of D+5, D+6 and D+7 forecasts with weights calculated according to $a_\ell = \sigma_\ell^{-1} / \sum_{n=1}^{3} \sigma_n^{-2}$ (ℓ=1,3) where σ_ℓ is the average standard deviation of estimate ℓ with the (incorrect) assumption of independent estimates. Using different combinations of weights on an empirical basis only improve the result marginally. Shown in Fig. 17 is the equivalent to Fig. 16 except that D+4 forecasts are considered instead of D+5. Consistency is thus measured from D+4, D+5 and, in the lower part of Fig. 17, also with D+6 forecasts.

Since anomaly correlation is mainly phase sensitive, and since there is only a slight relationship using standard deviation (see Fig. 15). consistency could perhaps be used to say something about the confidence of the longitudinal position of troughs and ridges but the amplitude and latitudinal position can not be judged.

Again treating the estimates as independent with weights calculated as above, the average scores in terms of absolute correlation and STD of error for the normal D+5 forecasts and for the weighted ensemble of D+5, D+6 and D+7 forecasts are the following :

	D+5	Ensemble	Climatology
Absolute correlation	.84	.85	.74
STD of error (mtrs)	92	87	108

TABLE 1 Average scores of absolute correlation and standard deviation of the error for D+5 forecasts (left), ensemble forecasts (D+5, D+6 and D+7) (middle) and climatology (right) for the time period 1 Jan. - 30 Apr. 1982.

As a reference, Table 1 includes scores using the mean (climatological) field for the same 4 month period Jan.-Apr. 1982. The ensemble STD scores corresponds to an improvement of about 6-9 hours in terms of forecast time, compared to individual D+5 forecasts.

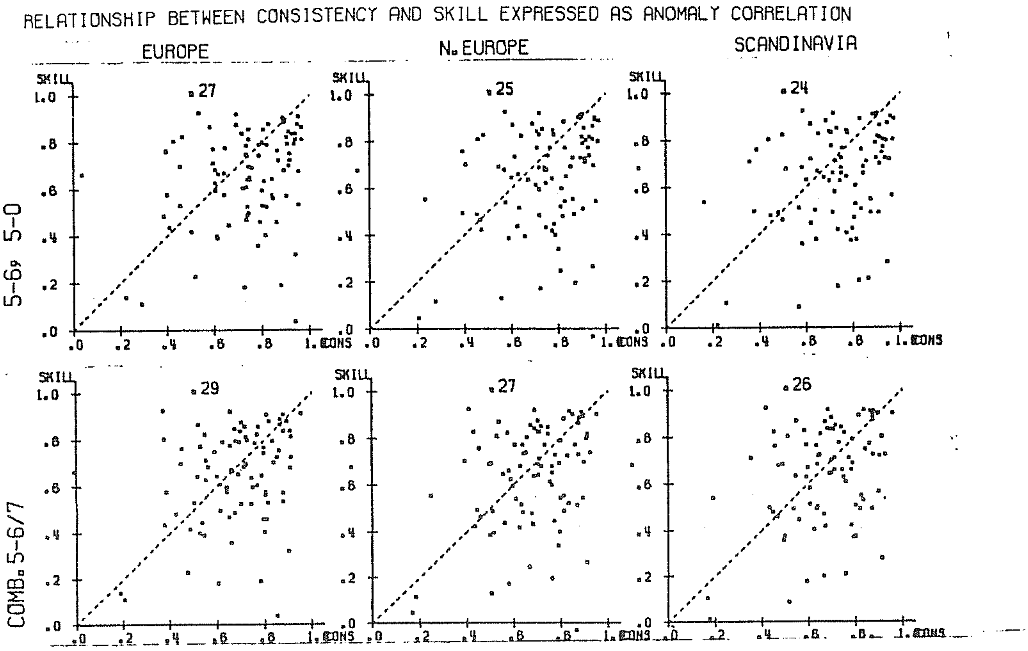


Fig. 16 : Upper part : same as Fig. 13 but with skill and consistency, expressed as anomaly correlation. Lower part : consistency is measured as a weighted combination of D+5, D+6 and D+7 forecasts.

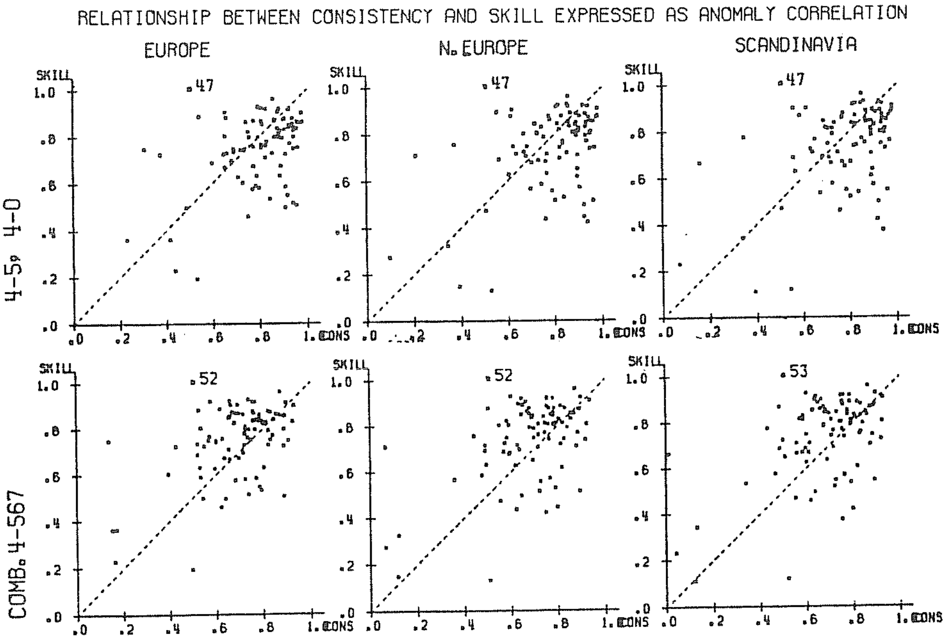


Fig. 17 Upper part: same as Fig. 14 but skill is for D+4 forecasts and consistency is measured between D+4 and D+5 forecasts.
Lower part: skill is for D+4 forecasts and consistency is measured as a weighted combination of D+4, D+5 and D+6 forecasts.

7. SUMMARY

An evaluation of ECMWF forecasts has been conducted concentrating on two winter seasons: Nov. 80 - Mar. 81 and Nov. 81 - Mar. 82. The main goal of this study has been to get a more comprehensive knowledge of deficiencies associated with certain flow types in the ECMWF medium range forecasts in order to use them appropriately. Some relationships have been found between the 500 mb D+5 forecast height error and D+5 variables such as height anomaly, meridional flow and D+0 to D+5 height tendency, which identify one characteristic flow type associated with unusually large negative errors (the systematic error) over most of Europe on D+5 and another less characteristic weather flow type associated with positive errors over eastern Europe on D+5. Knowledge of such deficiencies does not explain the cause of the errors, but used appropriately, could give forecasters a warning and possibly allow correction of known typical errors.

Relationships with the variables from the initial state did not give conclusive evidence of a flow pattern on D+0 associated with certain errors over Europe 5 days later. It gave however some indication that very strong negative D+5 errors over Europe often were associated with low index flow over the Rocky Mountains 5 days earlier.

Data sensitivity, error growth and progression were discussed with one example of how initial errors over North America and the Western Pacific affect the forecast detrimentally over Europe 5 to 6 days later.

The question whether consistency with previous forecasts means higher credibility and vice versa of the latest forecast has also been investigated.

Many of the points raised in this study are known to forecasters who have used these forecasts regularly. In addition to subjective evaluation, this study has attempted to quantify certain relationships statistically in order to draw objective conclusions on typical deficiencies of ECMWF forecasts and also to show alternative ways of using them.

REFERENCES

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

Arpe,K., 1982: Diagnostic evaluation of analysis and forecasts: Climate of the model. ECMWF seminar proceedings on interpretation of numerical weather prediction products, 13-17 September 1982.

Atlas, R., 1982: The Growth of Prognostic Differences between GLAS Model Forecasts from SAT and NOSAT Initial Conditions. Mon. Wea. Rev., 110, 877-882.

Baede, A.P.M., P. Kållberg and S. Uppala, 1982: Impact of Aircraft Data on ECMWF Analyses and Forecasts. ECMWF Technical Report (in preparation).

Bennet, A.F. and L.M. Leslie, 1979: Statistical correction of Dynamical Prognoses in the Australian Region. Mon. Wea. Rev., 107, 1254-1262

Cats, G. and O. Åkesson, 1983: An investigation into a marked difference between two successive ECMWF forecasts of September 1982. Submitted to Beiträge zur Physik der Atmosphäre.

Dell'Osso, L. and S. Tibaldi, 1982: Some preliminary results on an ALPEX case of Lec cyclogenesis, ALPEX preliminary scientific results, Geneva, December 1982, GARP-ALPEX No.7.

Fritsch, J.M. and C.R. Chelius, 1982: Forecast consistency: an important factor to forecast credibility. Proceedings of Ninth Conference on Weather Forecasting and Analysis, AMS, June 28-July 1 1982.

Hollingsworth, A., A.C. Lorenc, M.S. Tracton, K. Arpe, G. Cats, S. Uppala, P. Kållberg, 1983: The Response of Numerical Weather Prediction Systems to FGGE IIB Data : Analyses. (In preparation).

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

- , 1982: Local weather element guidance from the ECMWF forecasting system in the medium range II : Interpretation experiments. ECMWF Seminar Proceedings on Interpretation of numerical weather prediction products, 13-17 September 1982.

Rinne, J., 1971a: Investigation of the forecasting error of a simple barotropic model with the aid of empirical orthogonal functions. Part I. Geophysica, 11, 185-213.

-, 1971b: Investigation of the forecasting error of a simple barotropic model with the aid of empirical orthogonal functions. Part II. Geophysica, 11, 215-239.

-, 1972: Investigation of the forecasting error of a simple barotropic model with the aid of empirical orthogonal functions. Part III. Geophysica, 12, 57-59.

Salomonsson, G., 1981: Report on the Swedish Meteorological Hydrological Institute's assessment of the ECMWF forecast charts presented at Forecasters Meeting, ECMWF 18-19 May 1981.

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.

Simmons, A.J., and B.J. Hoskins, 1979: The Downstream and Upstream Development of Unstable Baroclinic Waves. J. Atmos. Sci., 36, 1239-1254.

Thompson, P.D., 1976: How to improve Accuracy by Combining Independent Forecasts. Mon. Wea. Rev., 105, 228-229.

Wallace, J.M. and D.S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere Winter. Mon. Wea. Rev., 109, 784-812.

– , S, Tibaldi 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.