TECHNICAL REPORT No. 36

OPERATIONAL VERIFICATION OF ECMWF FORECAST FIELDS AND RESULTS FOR 1980-1981

by

Rauno Nieminen

February 1983

CONTENTS

Abstract

The operational verification system using grid-point fields for objective assessment of the numerical forecasts at the European Centre for Medium Range Weather Forecasts (ECMWF) is described, and the performance of the Centre's model during 1980-1981 is discussed. Conventional objective scores are computed for geopotential height, temperature, wind and relative humidity at standard pressure levels over hemispheric and several limited areas of the globe. This system gives both a general overview of the performance of the ECMWF model and an estimation of the quality of the Centre's products in different parts of the world.

The objective scores are higher for the Northern Hemisphere than for the Southern Hemisphere reflecting in part differences in the data coverage. In the tropics the predictability is substantially lower than in the middle and high latitudes, indicating the difficulties in predicting the quasi-stationary circulation and the intense small scale features of the tropical atmosphere. During the two year period investigated, a steady improvement in the scores is found, associated with several revisions made to the data assimilation and forecasting system.

1. Introduction

Two objective verification procedures are used at ECMWF to assess the operational numerical forecasts. One, not reported in this paper, uses results which are stored in the form of spherical harmonic coefficients in the pressure level archive. Hemispheric scores, including anomaly correlation, standard deviation of error and mean error, are produced for height and temperature at standard pressure levels for different wave components (ECMWF Meteorological Bulletin M2.2/1, 1981).

The second operational verification system, called the field verification, makes use of grid-point pressure level data, and is the subject of this report. The field verification procedure in its present form has been carried out since the beginning of January 1980. The period from August to December 1979, when the operational forecasts were already running, was more or less a test phase in the field verification where only the European area was considered. In the current version of the verification system conventional objective scores, including correlation coefficient, rms-error, standard deviation of error and mean error, are computed for geopotential height, temperature, wind and relative humidity; the S1 skill score is also provided for height fields. The results are available for standard pressure levels from 1000 to 50 mb (only to 200 mb in 1980), and for 12 limited areas and 3 latitudinal zones. Surface parameters, such as precipitation, surface wind or temperature, are not included in the field verification. They are verified against observations, and several studies have been carried out at the Centre in assessing those direct model output products, including Åkesson (1981), Johannessen (1982), Åkesson et al.(1982), Böttger and Grønaas (1982) and Pümpel (1982).

Minor changes were made to the verification areas and more objective scores were added during 1981, and so the results for 1980 were rerun to establish a consistent data set for comparison of the two years. Furthermore, a proper plotting package was designed and used for the compact presentation of the large amount of results.

The purpose of this paper is first to describe the field verification procedure operationally used at ECMWF, and second, to present and evaluate the results for 1980-1981. A selection of the results for the European area have been published routinely (ECMWF Forecast Reports, 1980-1981).

2. Field Verification System

2.1 Verified data

The data assimilation procedure described by Lorenc (1981) is a part of the operational weather prediction system at ECMWF. This intermittent assimilation cycle is run four times a day, providing analyses for 00, 06, 12 and 18 GMT. For each analysis step a 6-hour forecast is introduced as a first guess. A non-linear normal mode initialisation is used to eliminate the influence of gravity waves on the analysed fields. A 10-day forecast is run once a day from the 12 GMT analysis, using the Centre's operational grid-point model with 15 σ-levels in the vertical and a staggered horizontal grid of 1.875 x 1.875 degrees of latitude and longitude (ECMWF Forecast Model Documentation Manual, 1982). A brief description of the Centre's operational forecasting system is given in Appendix 1. The analysis is performed on pressure levels with the model grid of 1.875 degrees, but in the post-processing the analysed fields are interpolated to a non-staggered grid of 1.5 x 1.5 degrees. Forecast fields are interpolated in the post-processing from the model output both vertically from σ-levels to pressure levels and horizontally from the model grid of 1.875 degrees to the 1.5 degree resolution. For the field verification a special data base is created by selecting every second grid-point along latitudes and longitudes from the post-processed fields. The verification data base thus consists of global grid-point fields at pressure levels with a 3 x 3 degree horizontal resolution. The fields for this data base have been stored on magnetic tapes since 1 January 1980.

It should be stressed that the forecasts are verified against the initialised analyses. It has been found that the normal mode initialisation used by the Centre introduces a minor smoothing on the analysed height fields. The effect of the initialisation is even more pronounced in the tropics where it reduces the intensity of the Hadley circulation. This is partly related to the fact that diabatic forcing was not included in the Centre's initialisation scheme during the period investigated (diabatic forcing was introduced in the operational initialisation in September 1982). The analyses are also indirectly influenced by the forecasts through the first-guess, especially in data sparse regions. Uninitialised analysis fields of geopotential height and wind components are included in the verification data base from 1 July 1981, and temperature and relative humidity fields for the uninitialised analyses are provided from 1 November 1982.

Daily operational forecasts have been produced at ECMWF since 1 August 1980, and before that from 1 August 1979 five days a week (i.e. excluding weekends). Therefore the sample number in the monthly means varies from 20 to 23 in the first part of the year 1980, until the daily routine started in August 1980. One exception to this occurs in May 1980 when the number of cases was reduced to 14, due to archive problems.

Table 1 summarises the content of the field verification data base. In January 1981 two more levels, 100 and 50 mb, and the complete set of 10-day forecasts for wind were added. As regards relative humidity, the first-guess (i.e. forecast) was used as a verifying analysis until the end of March 1981; this was changed to verification against analysed humidity on 1 April 1981.

Table 1: Levels, parameters, frequency and time-steps of the fields in the verification data base at ECMWF

global fields; $3^{\circ} \times 3^{\circ}$ horizontal resolution
1000,850,700,500,300,200 mb (from 1 January to 31 December 1980)
1000,850,700,500,300,200,100,50 mb (from 1 January 1981)
geopotential height
temperature
wind components (to +120h from 1 January to 31 December 1980,
to +240h from 1 January 1981)
relative humidity (analysed humidity from 1 April 1981)
5 days/week from 1 January to 31 July 1980 (but 7 days/week for analyses)
7 days/week from 1 August 1980
12 hour interval to +48h
24 hour interval from +48h to +240h

2.2 Verification areas

The forecasts at ECMWF are produced globally, and for the evaluation of the accuracy of products both hemispheric and limited area scores are provided. The verification areas, including 12 limited regions and 3 latitudinal zones, are defined in Table 2. Figure 1 shows a typical distribution of global observations which are used operationally at the Centre. Differences in coverage of the surface and upper air data can clearly be seen between the two hemispheres and also between the oceans and the continents. Satellite data, however, give a valuable complement to the surface-based observations. In order to provide as reliable analyses as possible for the verification, all the limited regions have been defined over relatively data-rich parts of the continents as the distribution in Figure 2 shows. The latitudinal zones, of course, cover the oceans too.

As is shown in Table 2, the horizontal grid resolution used in the verification procedure is 3 x 6 latitude-longitude degrees in the high latitudes and 3 x 3 latitude-longitude degrees in the sub-tropical and tropical areas, (see section 2.5 below).

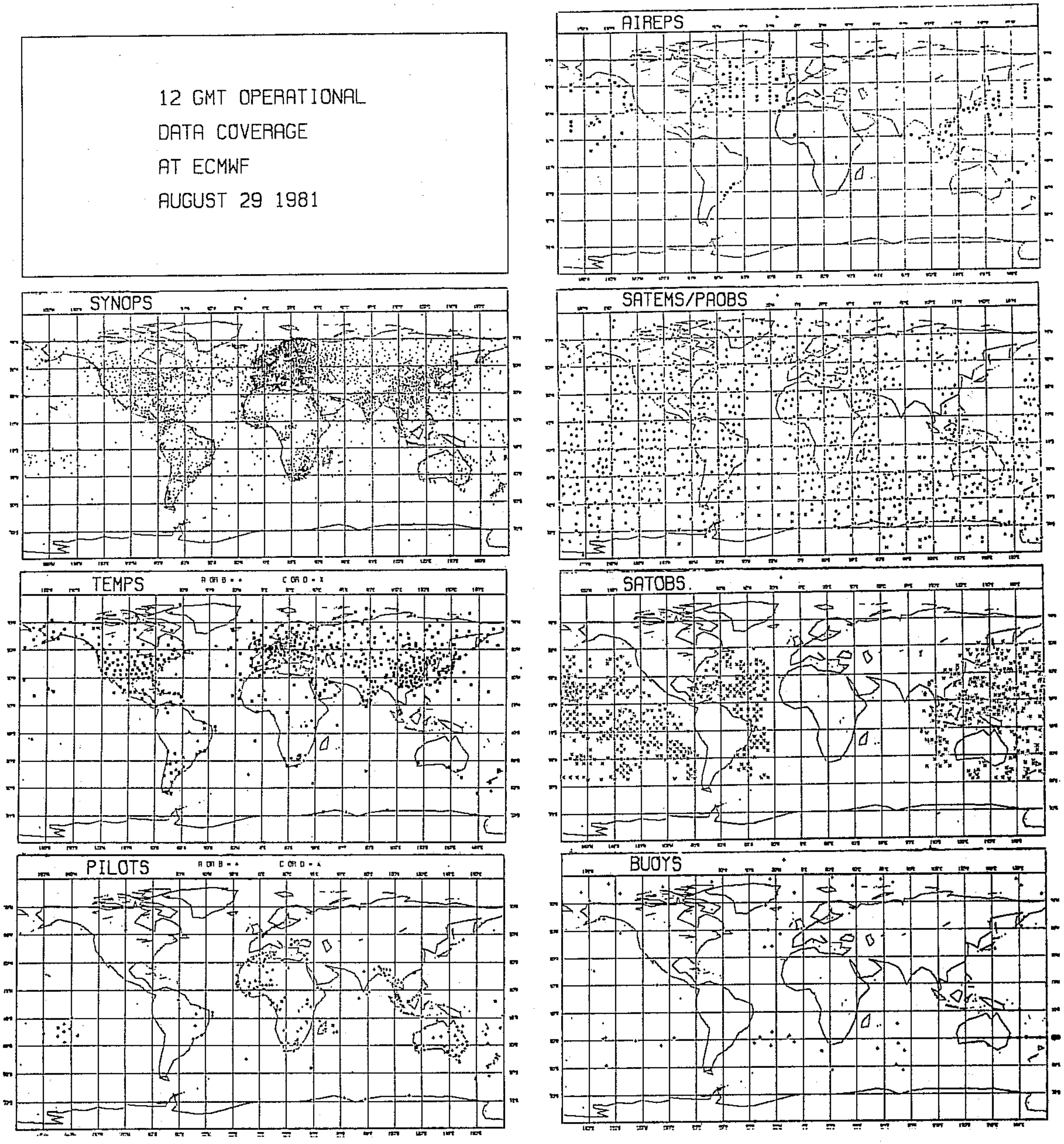


Figure 1: A typical data coverage of the observations used operationally at ECMWF

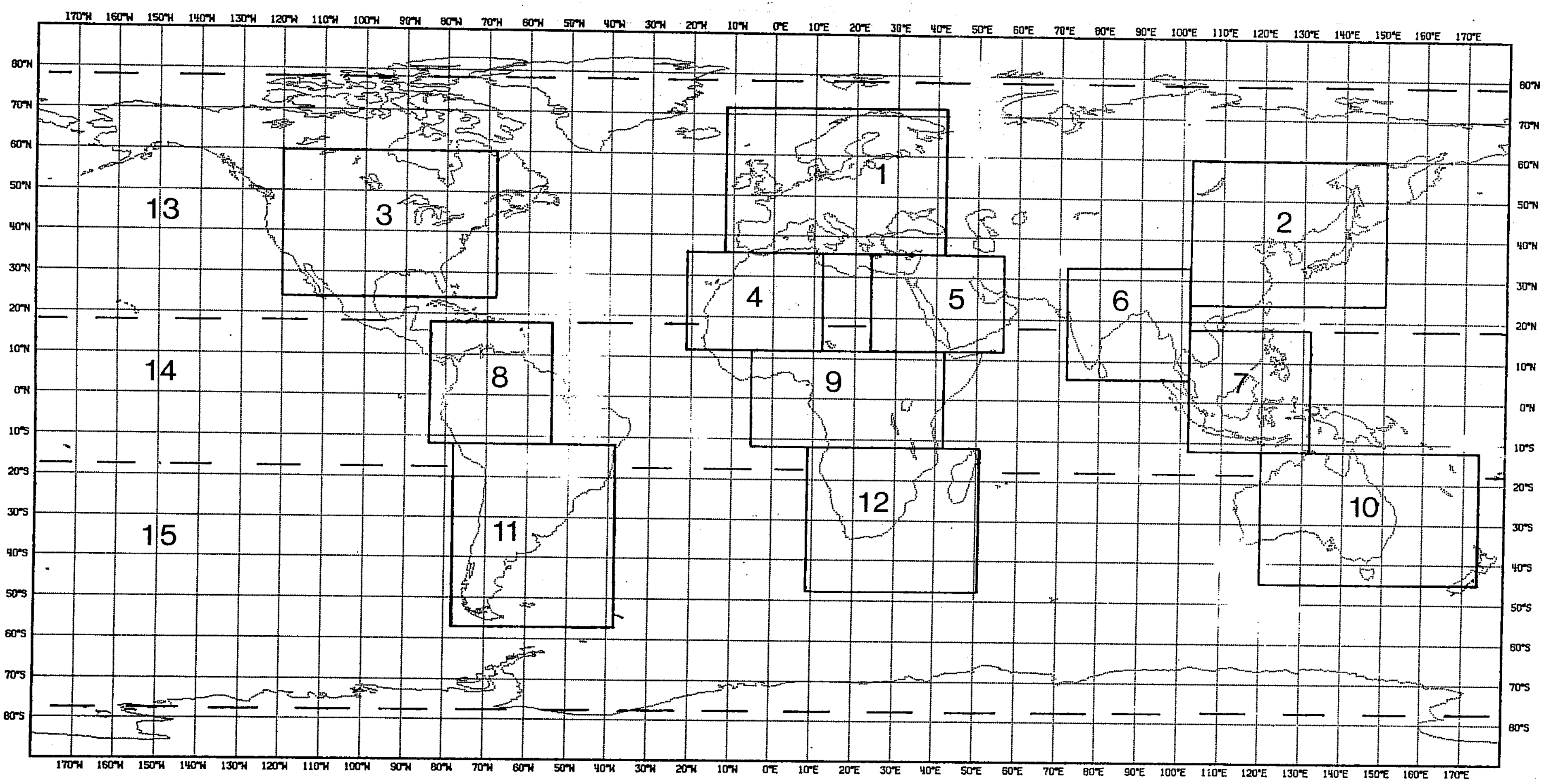


Figure 2: Areas used for the objective field verification of the operational forecasts at ECMWF, (boundaries of latitude belts are indicated by dashed lines).

For the latitudinal zones, the resolution is 6 x 9 latitude-longitude degrees in the extra-tropical hemispheres and 6 x 6 latitude-longitude degrees in the equatorial belt. For the area means, a latitudinal cosine weight is used to compensate for the decreasing grid interval along latitudes in the meridional direction from the equator to the poles.

Table 2: Verification areas, grid resolution and number of grid-points used in the field verification system of ECMWF.

1.	Europe	72N-36N, 12W-42E	3x6 deg.	130 points
2.	China-Japan	60N-24N, 102E-150E	3x6 "	117 "
3.	North America	60N-24N, 120W-72W	3x6 "	117 "
4.	N.W. Africa	36N-12N, 21W-12E	3x3 "	108 "
5.	Arabia	36N-12N, 24E-57E	3x3 "	108 "
6.	India	33N-06N, 72E-102E	3x3 "	110 "
7.	Indonesia	18N-12S, 102E-132E	3x3 "	121 "
8.	Latin America	18N-12S, 84W-54W	3x3 "	121 "
9.	Central Africa	12N-12S, 06W-42E	3x3 "	117 "
10.	Australia	12S-45S, 120E-174E	3x6 "	120 "
11.	South America	12S-57S, 78W-42W	3x6 "	112 "
12.	South Africa	12S-48S, 09E-51E	3x6 "	104 "
13.	Northern Hem.	78N-18N, 00E-351E	6x9 "	440 "
14.	Tropics	18N-18S, 00E-354E	6x6 "	420 "
15.	Southern Hem.	18S-78S, 00E-351E	6x9 "	440 "

2.3 Verification scores

The recommendation of the WMO/CAS Working Group on Weather Prediction Research (1978) for the verification procedures of NWP products has been taken as a guide-line for the selection of verification parameters, and the following objective scores are used in the operational field verification at ECMWF.

tendency correlation = correlation between predicted and analysed changes

anomaly correlation = correlation between predicted and analysed anomalies (i.e. deviations from climatology)

absolute correlation = correlation between predicted and analysed fields (used for wind)

root-mean square error for forecast and persistence

standard deviation of error for forecast and persistence

mean error (bias) for forecast and persistence

S1 skill-score of geopotential height for forecast and persistence

Using the following symbols:

A_o = initial analysis
A_v = verifying analysis
F = forecast
C = climatology
n = number of gridpoints in the verification area

$F-A_o$ = predicted tendency
A_v-A_o = verifying tendency
$F-C$ = predicted anomaly
A_v-C = verifying anomaly
A_o-C = persistence anomaly
$F-A_v$ = forecast error
A_o-A_v = persistence error

(overbar) $\overline{\qquad\qquad}$ = area mean

we get the expressions for the objective scores:

$$1/n \ \Sigma(F-A_v) = (\overline{F-A_v}) = \text{mean of forecast error}$$

$$1/n \ \Sigma(A_o-A_v) = (\overline{A_o-A_v}) = \text{mean of persistence error}$$

$$\sqrt{1/n \ \Sigma(F-A_v)^2} = \text{rms-error of forecast}$$

$$\sqrt{1/n \ \Sigma(A_o-A_v)^2} = \text{rms-error of persistence}$$

$$\sqrt{1/n \ \Sigma \left[(F-A_v)-(\overline{F-A_v})\right]^2} = \text{standard deviation of forecast error}$$

$$\sqrt{1/n \ \Sigma \left[(A_o-A_v)-(\overline{A_o-A_v})\right]^2} = \text{standard deviation of persistence error}$$

$$\sqrt{1/n \ \Sigma \left[(A_v-C)-(\overline{A_v-C})\right]^2} = \text{standard deviation of verifying anomaly}$$

$$\frac{\Sigma\{[(F-A_o)-(\overline{F-A_o})]\ [(A_v-A_o)-(\overline{A_v-A_o})]\}}{\sqrt{\Sigma[(F-A_o)-(\overline{F-A_o})]^2\Sigma[(A_v-A_o)-(\overline{A_v-A_o})]^2}} = \text{tendency correlation for forecast}$$

$$\frac{\Sigma\{[(F-C)-(\overline{F-C})]\ [(A_v-C)-(\overline{A_v-C})]\}}{\sqrt{\Sigma[(F-C)-(\overline{F-C})]^2\Sigma[(A_v-C)-(\overline{A_v-C})]^2}} = \text{anomaly correlation for forecast}$$

$$\frac{\Sigma\{[(A_o-C)-(\overline{A_o-C})]\ [(A_v-C)-(\overline{A_v-C})]\}}{\sqrt{\Sigma[(A_o-C)-(\overline{A_o-C})]^2\Sigma[(A_v-C)-(\overline{A_v-C})]^2}} = \text{anomaly correlation for persistence}$$

$$\frac{\Sigma\{[F-\overline{F}]\ [A_v-\overline{A}_v]\}}{\sqrt{\Sigma[F-\overline{F}]^2\Sigma[A_v-\overline{A}_v]^2}} = \text{absolute correlation for forecast}$$

$$\frac{\Sigma\{[A_o-\overline{A}_o]\ [A_v-\overline{A}_v]\}}{\sqrt{\Sigma[A_o-\overline{A}_o]^2\Sigma[A_v-\overline{A}_v]^2}} = \text{absolute correlation for persistence}$$

For vector wind we combine the results of the wind components as follows:

$$\begin{aligned}
\overline{V}_{mean} &= \sqrt{[u_{mean}]^2 + [v_{mean}]^2} \\
rms(\overline{V}) &= \sqrt{[rms(u)]^2 + [rms(v)]^2} \\
stdv(\overline{V}) &= \sqrt{[stdv(u)]^2 + [stdv(v)]^2} \\
correlation\ r(\overline{V}) &\approx \sqrt{1/2 * [r(u)^2 + r(v)^2]}
\end{aligned}$$

The S1 skill score is introduced for the gradient fields of geopotential height. Using the upper index $(\ )^i$ for the west-east direction and correspondingly $(\ )^j$ for the north-south direction, we obtain the following formula for the gradients:

$$\left.\begin{aligned}
F(i,j)^i &= F(i,j)-F(i+1,j) \\
F(i,j)^j &= F(i,j)-F(i,j+1)
\end{aligned}\right) = \text{predicted gradients}$$

$$\left.\begin{aligned}
A_v(i,j)^i &= A_v(i,j)-A_v(i+1,j) \\
A_v(i,j)^j &= A_v(i,j)-A_v(i,j+1)
\end{aligned}\right) = \text{verifying gradients}$$

Introducing these expressions we define the S1 skill score:

$$S1 = 100 * \frac{\sum_{i,j} \{F(i,j)^i - A_v(i,j)^i\} + \sum_{i,j} \{F(i,j)^j - A_v(i,j)^j\}}{\sum_{i,j} \{G_L^i\} + \sum_{i,j} \{G_L^j\}}$$

where

$$\begin{aligned}
G_L^i &= \max\,[\,(F(i,j)^i\ ,\ A_v(i,j)^i\,] \\
G_L^j &= \max\,[\,(F(i,j)^j\ ,\ A_v(i,j)^j\,]
\end{aligned}$$

2.4 Some aspects of objective scores

The evaluation procedure for the global model products by objective scores must be designed for a wide range of applications, including both short and medium range forecasts of different meteorological parameters which are used for various purposes. As was mentioned earlier, the recommendation of the WMO/CAS Working Group on Weather Prediction Research for the verification procedures offers a standardised and comprehensive set of objective scores for the field verification, and this has been followed at ECMWF.

The correlation coefficient of height fields is one of the most common scores in the evaluation of the model performance. The tendency correlation is suitable for short range forecasts, but for medium range forecasts the anomaly correlation is widely adopted. It was one of the scores which Miyakoda et al. (1972) preferred to use in assessing the predictability of medium range forecasts. The anomaly correlation seems to give a reasonable measure of the predictive skill, at least over large hemispheric areas. In smaller regions, however, some deficiencies are possible in the cases where the anomalies are very small. This subject will be discussed further in section 3.3.

The S1 skill score (Teweles et al., 1954) has been used for a long time, particularly in the United States, to verify pressure or height fields. A similar kind of measure is the absolute correlation of the vector wind (the combined correlation of the wind components). This is also used at ECMWF, where it has been found suitable for the equatorial regions.

One important aspect of objective scores is the definition of a limit beyond which no practical predictive skill exists. Comparison with persistence or with climatological normal values is one common way to define such a criterion. For the anomaly correlation, empirical values, for example 0.50 or 0.60 (sometimes given in percentages, as 50% or 60%), are used at ECMWF (Hollingsworth et al., 1980) as a limit of a useful forecast. Those values seem to agree fairly well with the subjective assessments.

There are cases when the anomaly correlation values for individual forecasts are below this limit, but where the forecasts still can provide some useful information, for example, about the general weather type (Grosswetterlage). On the other hand, high anomaly correlations related to high anomaly values can sometimes overestimate the usefulness of forecasts. In addition to this, scores like anomaly correlation, when computed over broad hemispheric areas, are basically measuring large scale features, whereas in limited regions small scale processes dominate score values. Partly for this reason the hemispheric scores are generally higher than the corresponding scores for limited areas. The hemispheric scores can also be influenced by the fact that the analyses over the oceans may lose a part of the real variation due to sparse data coverage.

2.5 Influence of the grid resolution on the verification results

For practical reasons the horizontal resolution used for the verification system is different from the model resolution. To investigate possible consequences, a test run over a three month period was carried out, in which the operational verification grid interval was modified. The European area was selected to represent limited regions, and the operational grid of 3 x 6 latitude-longitude degrees was reduced to a 3 x 3 degree resolution. A similar rearrangement was made for the Northern Hemisphere where the 6 x 9 degree verification grid was changed to a 6 x 6 degree horizontal resolution. This modification means that the number of grid points along latitudes was doubled in the European area and increased by a factor of 1.5 in the hemispheric area.

Tables 3 and 4 show the anomaly correlation and standard deviation of error for geopotential height at the 1000 mb level in Europe and at the 500 mb level in the hemispheric area for the operational and test scores. The results confirm that the verification parameters are not sensitive to these kind of changes of the horizontal resolution.

In this connection another comparison was made with the spectral verification system, also used operationally at ECMWF. In the spectral system, the Fourier components are calculated up to zonal wavenumber 20, roughly corresponding to the 9 degree interval along latitudes used in the grid point field verification for the hemispheric areas. The spectral verification was produced over the latitudinal zone from 20°N to 82.5°N, and the corresponding area in the field verification was from 18°N to 78°N. Table 5 shows the average hemispheric anomaly correlation of height at the 1000 and 500 mb levels in 1981 for both operational verification systems. The results agree very well, demonstrating that these independent procedures measure the quality of forecasts in a consistent way.

Table 3: Anomaly correlation (upper table) and standard deviation of error (lower table) of 1000 mb height for different grid resolutions averaged over the period of January - March 1981 for the European area.

1000 mb	D+1	D+2	D+3	D+4	D+5	D+6	D+7
3 x 6 deg.	.96	.89	.79	.65	.51	.41	.32
3 x 3 deg.	.96	.89	.79	.64	.51	.41	.32

1000 mb	D+1	D+2	D+3	D+4	D+5	D+6	D+7
3 x 6 deg.	22	36	51	64	76	86	93
3 x 3 deg.	22	36	50	64	76	86	92

Table 4: Anomaly correlation (upper table) and standard deviation of error (lower table) of 500 mb height for different grid resolutions averaged over the period of January - March 1981 for the Northern Hemisphere (18N-78N).

500 mb	D+1	D+2	D+3	D+4	D+5	D+6	D+7
6 x 9 deg.	.98	.94	.88	.80	.70	.60	.50
6 x 6 deg.	.98	.94	.88	.80	.70	.60	.50

500 mb	D+1	D+2	D+3	D+4	D+5	D+6	D+7
6 x 9 deg.	22	38	54	70	85	98	109
6 x 6 deg.	23	39	55	71	86	99	111

Table 5: Anomaly correlation of 1000 mb and 500 mb height averaged over the year 1981 for the Northern Hemisphere for the spectral and grid-point verification of ECMWF.

1000 mb	D+1	D+2	D+3	D+4	D+5	D+6	D+7
spectral (20 compon.)	.95	.88	.81	.71	.61	.51	.42
grid-point (6 x 9 deg.)	.95	.88	.80	.70	.60	.50	.41

500 mb	D+1	D+2	D+3	D+4	D+5	D+6	D+7
spectral (20 compon.)	.98	.93	.87	.78	.67	.56	.47
grid-point (6 x 9 deg.)	.98	.93	.87	.78	.68	.57	.47

3. Performance of the ECMWF operational model, 1980-1981

The verification scores of the global model products demonstrate a wide temporal and spatial variability in predictive skill. A reasonable time series is needed for the evaluation of such results, and the period of two years investigated in this paper provides us with adequate material to estimate the model performance.

3.1 Systematic error

Averaging the daily forecast error over a time period, for example over one month, reveals an important systematic error of the ECMWF model. The predicted mean values of geopotential height and temperature are too low at the upper levels of the model atmosphere. Figures 3 and 4 show that this global error grows with the time integration. According to Bengtsson (1981), the ECMWF model obtains its full heat balance after 50 days of integration, and during that time the model atmosphere has cooled by 4°C, probably due to unsatisfactory heating by convection. This systematic error seems to reflect a gradual shift towards the model "climate", and it is shared by most other general circulation models.

The largest negative values of the mean geopotential height errors occur during the winter months in the Northern Hemisphere where the mean height error has two regional maxima, one in the Eastern Pacific and another in the Eastern Atlantic and Europe. This kind of concentration of the mean error is probably connected to defects in the large scale orographical or thermal forcing, as discussed by Bengtsson and Simmons (1983) and Wallace et al. (1983).

The mean error of the u-component of wind is presented in Figure 5, where a slight intensification of the zonal flow in the extratropical areas can be seen at the lower levels of the stratosphere. In the tropics, on the other hand, the zonal wind component is systematically predicted to be too weak in the upper troposphere, when the average over the tropical belt is considered. This effect seems to be stronger during the period of December-February than in June-August.

As regards the mean error of relative humidity, it can be seen from Figure 6 that the predicted humidity values are consistently too high in the upper troposphere and lower stratosphere. The lower troposphere is predicted to be too dry, particularly in the tropics. Note, however, that the humidity analysis is very inaccurate due to lack of observations, and consequently the humidity field is to a great extent, generated by the model In the analysis of humidity at the Centre, observations are taken into account up to the 300 mb level, and in the stratosphere humidity is extrapolated towards climatology. However, Figure 6 clearly reveals that the mean

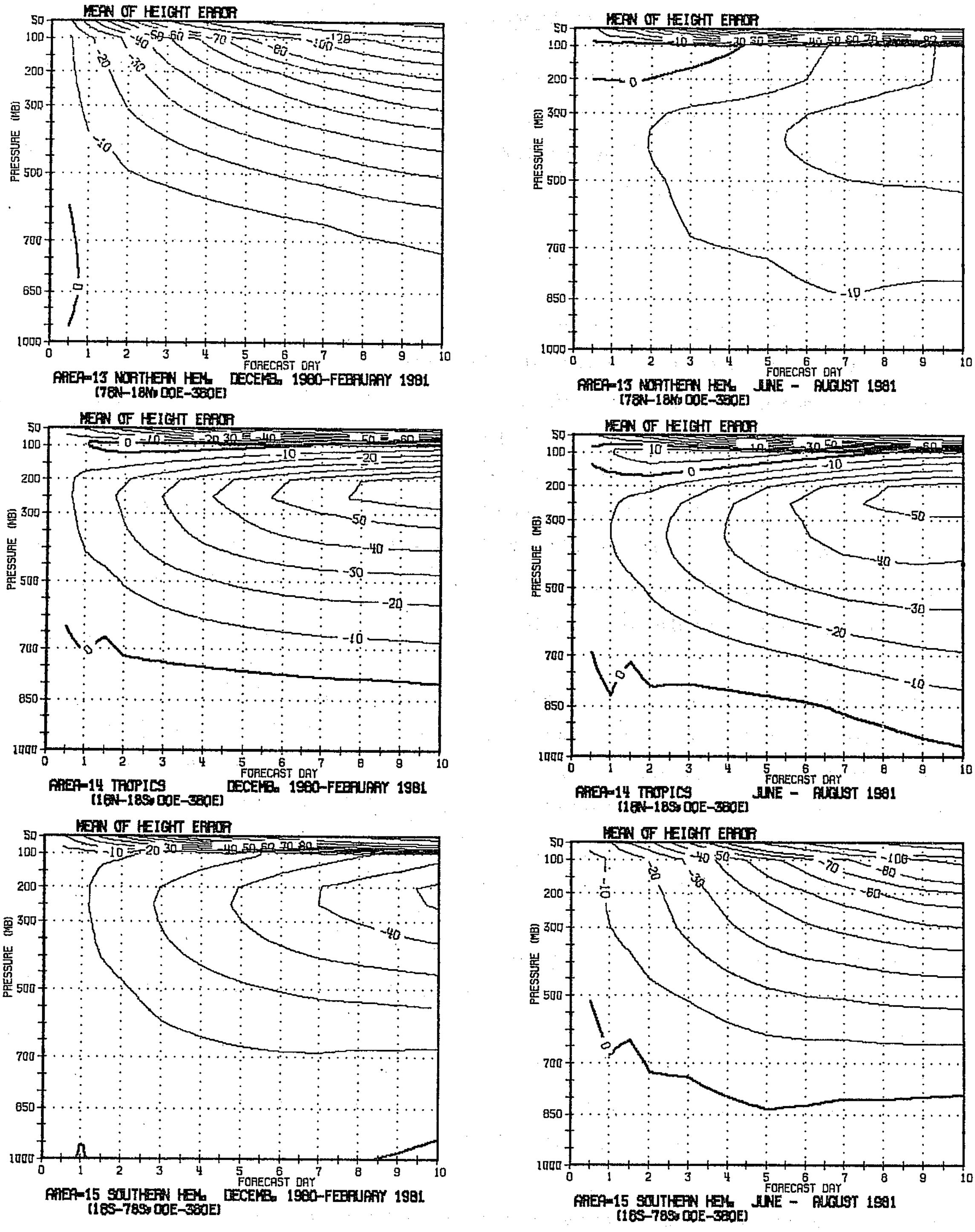
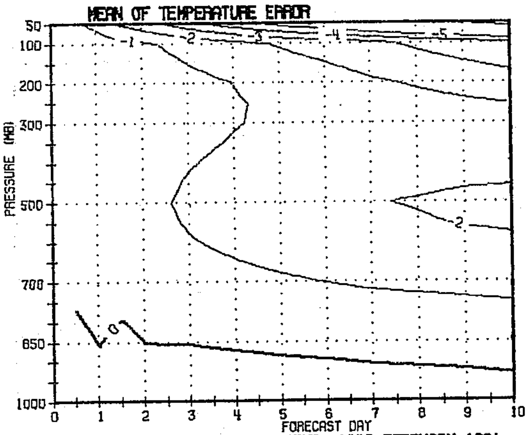
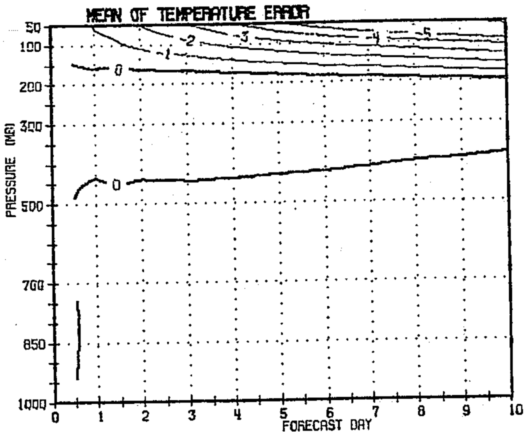


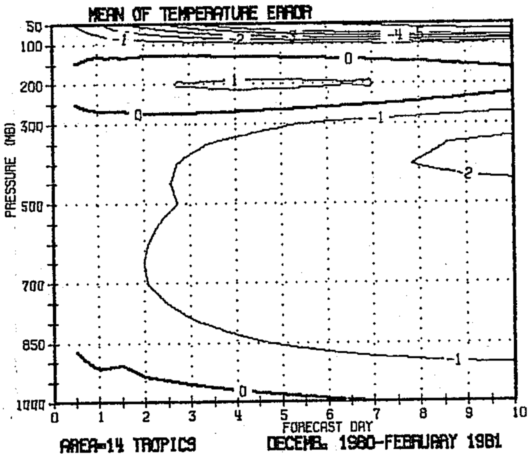
Figure 3: Vertical time-sections of mean errors of geopotential height forecasts for the Northern Hemisphere (top), the tropics (centre) and the Southern Hemisphere (bottom) averaged over two different three month periods. (Note that the winter scores for the Southern Hemisphere are at bottom right).



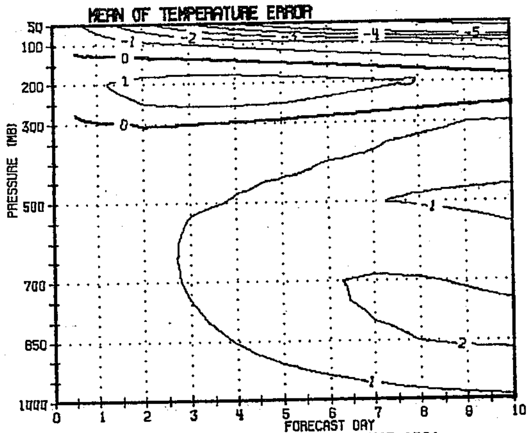
AREA=13 NORTHERN HEM. DECEMB. 1980-FEBRUARY 1981
(78N-18N; 00E-360E)



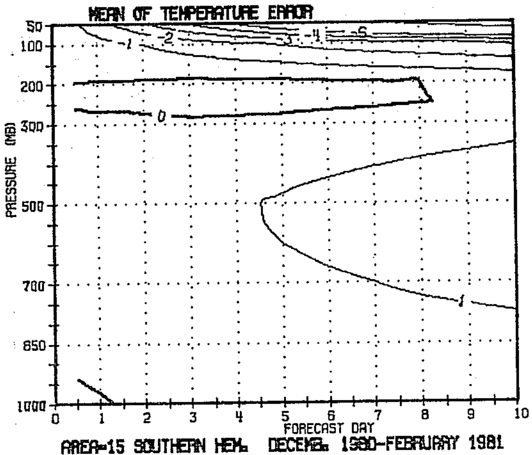
AREA=13 NORTHERN HEM. JUNE - AUGUST 1981
(78N-18N; 00E-360E)



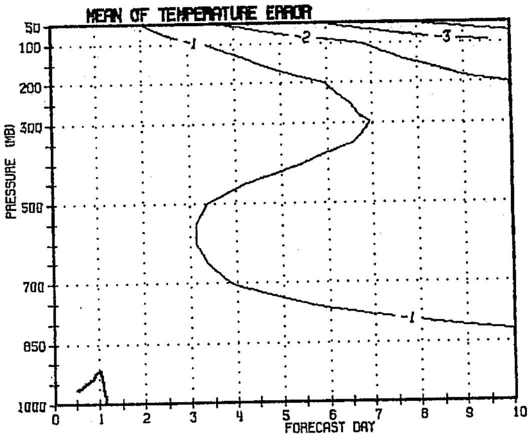
AREA=14 TROPICS DECEMB. 1980-FEBRUARY 1981
(18N-18S; 00E-360E)



AREA=14 TROPICS JUNE - AUGUST 1981
(18N-18S; 00E-360E)



AREA=15 SOUTHERN HEM. DECEMB. 1980-FEBRUARY 1981
(18S-78S; 00E-360E)



AREA=15 SOUTHERN HEM. JUNE - AUGUST 1981
(18S-78S; 00E-360E)

Figure 4: As Figure 3 but showing vertical time-sections of mean errors of temperature forecasts.

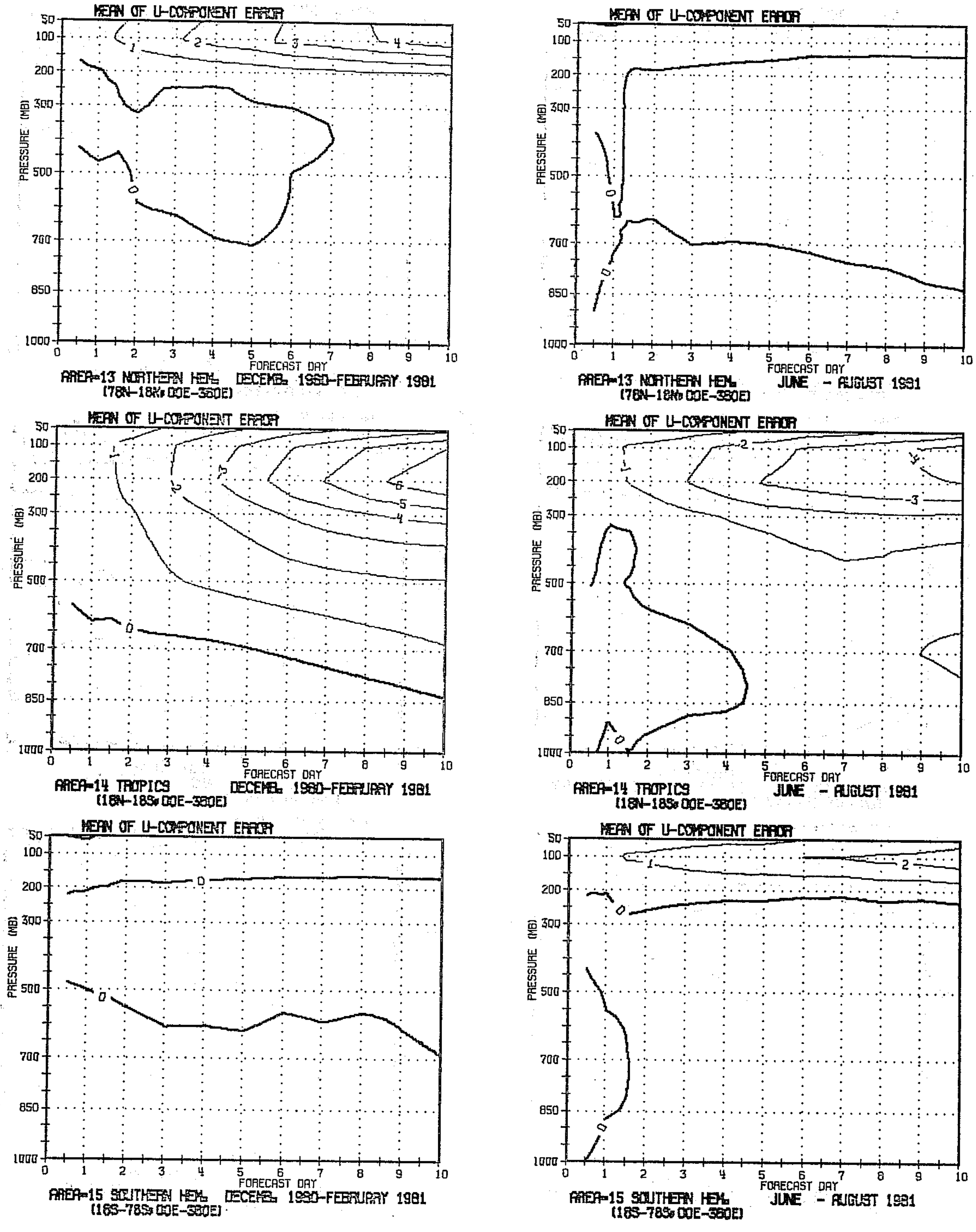
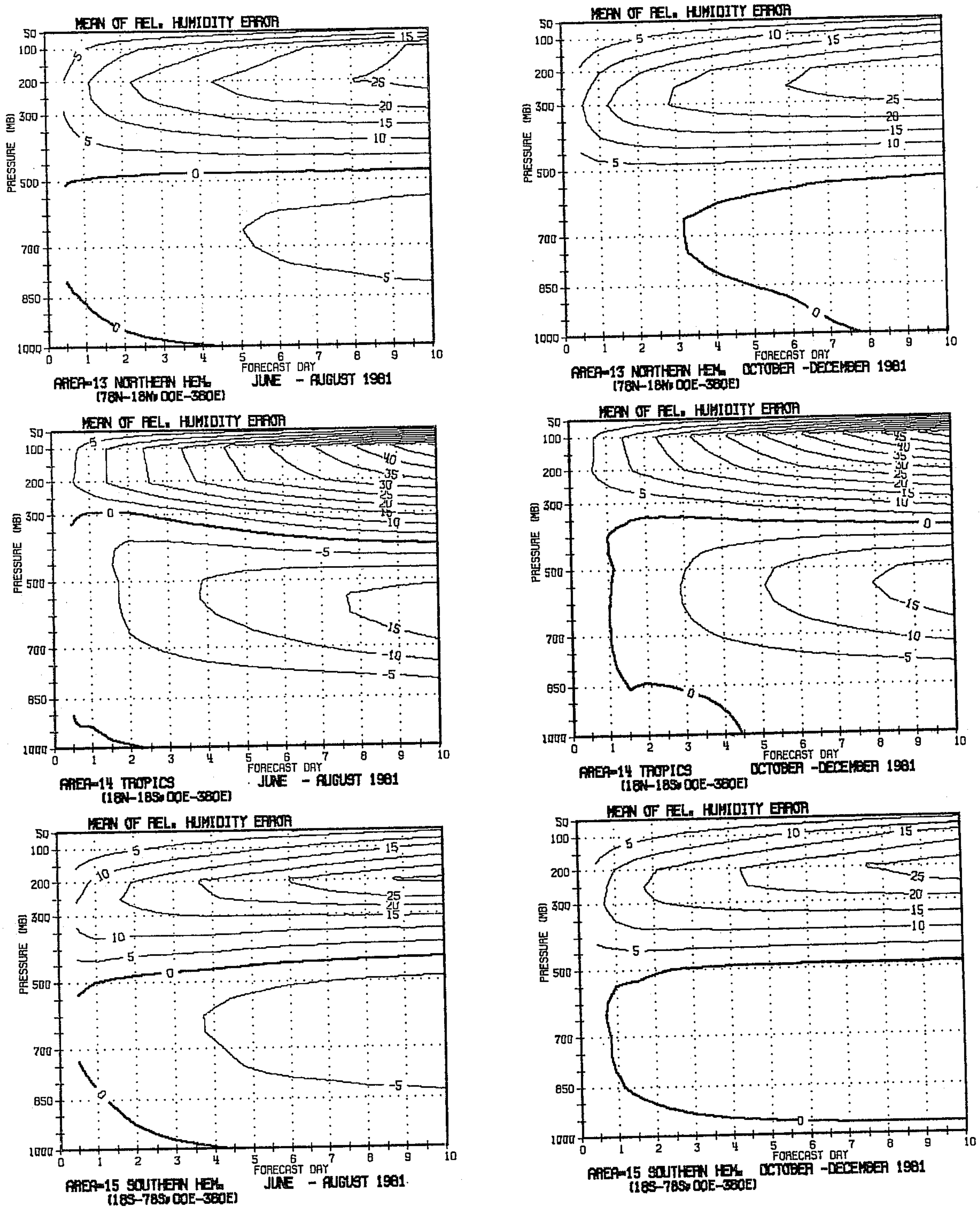


Figure 5: As Figure 3 but showing vertical time-sections of mean errors of zonal wind forecasts



MEAN OF REL. HUMIDITY ERROR

AREA-13 NORTHERN HEM. (78N-18N, 0OE-360E) JUNE - AUGUST 1981

MEAN OF REL. HUMIDITY ERROR

AREA-13 NORTHERN HEM. (78N-18N, 0OE-360E) OCTOBER -DECEMBER 1981

MEAN OF REL. HUMIDITY ERROR

AREA-14 TROPICS (18N-18S, 0OE-360E) JUNE - AUGUST 1981

MEAN OF REL. HUMIDITY ERROR

AREA-14 TROPICS (18N-18S, 0OE-360E) OCTOBER -DECEMBER 1981

MEAN OF REL. HUMIDITY ERROR

AREA-15 SOUTHERN HEM. (18S-78S, 0OE-360E) JUNE - AUGUST 1981

MEAN OF REL. HUMIDITY ERROR

AREA-15 SOUTHERN HEM. (18S-78S, 0OE-360E) OCTOBER -DECEMBER 1981

Figure 6: As Figure 3 but showing vertical time-sections of mean errors of relative humidity forecasts; these are averaged over the periods of June-August (left) and October-December (right) in 1981.

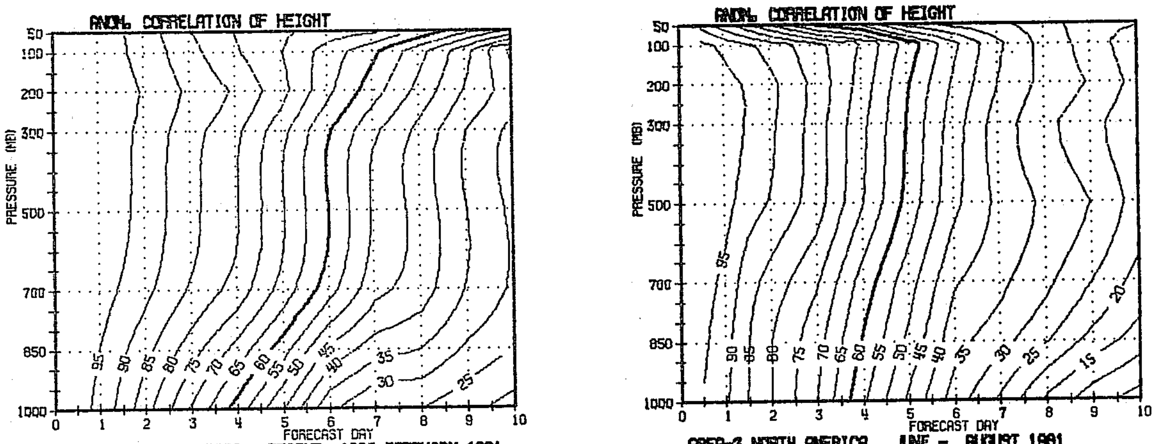
error for humidity increases with the forecast length, and this therefore indicates some deficiencies in the moisture balance of the model.

The cooling of the model atmosphere and the concentration of the mean error over certain areas is a characteristic also of other operational models (Bengtsson and Lange, 1982). Much effort has been made to reduce these errors in the ECMWF model. At the end of 1981 a modified physical package was implemented for the Centre's model, and the first results show some positive impact. Experiments with model orography also indicate that modifications to this may lead to reduction in the systematic errors (Tibaldi, personal communication).

3.2 Spatial and monthly variability

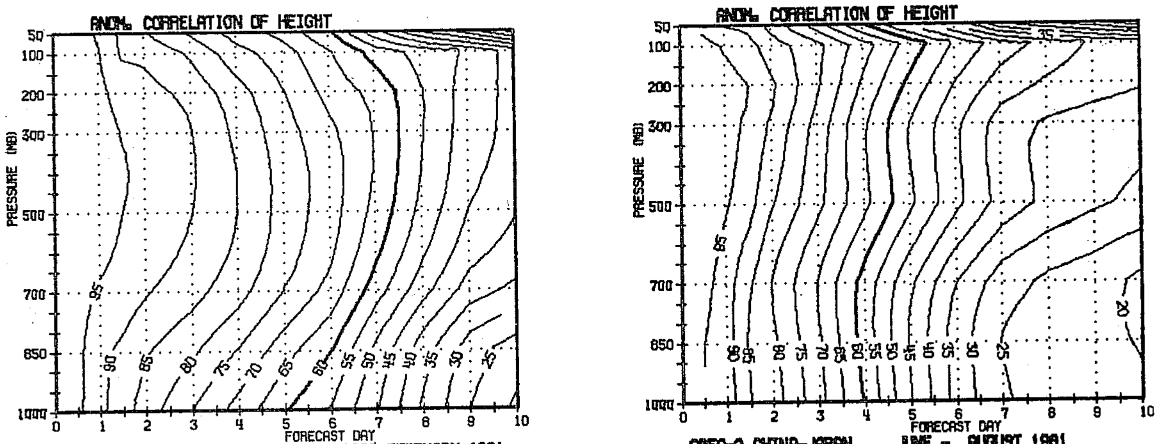
The verification results in the extratropical areas show considerable variation from region to region and from month to month. The spatial variability is very substantial during the winter season in the Northern Hemisphere, as can be seen in Figures 7 and 8 which show the averaged anomaly correlations of height for a winter and a summer period for three different areas of both hemispheres. By comparing those averaged anomaly correlations at the 500 mb level in winter, it can be seen from Figure 7 that the 60% score was reached at day 6 for the North American area, at day 7.3 for the East Asian area and at day 5.3 for the European area. The large systematic error of height affecting the European regions (see section 3.1 above) is probably the main reason for the lower scores for Europe. For the Southern Hemipsheric areas (see Figure 8) the regional differences are smaller than in the Northern Hemisphere. Generally, the objective scores for the anomalies are considerably higher for the Northern Hemisphere compared to the Southern Hemisphere. The seasonal variation of the anomaly correlation is also rather significant in the Northern Hemisphere (Figure 7), whereas the differences in the Southern Hemispheric scores are much smaller between the winter and summer periods (Figure 8).

The monthly variability of the limited area scores can be seen in Figure 9, which shows the monthly averages of the height anomaly correlations at 500 mb for the years 1980 and 1981. Three continental areas from both hemispheres are included. The late spring or the early summer seems to be favourable for prediction in Europe. A blocking pattern occurs often in some areas of Europe during that time, and it has been found that the Centre's model is generally good in predicting blocking situations. Therefore the high score values in the spring can be related to blocking or, more generally, to a strong meridional flow. Although the monthly variation makes the evaluation of Figure 9 difficult, small improvements of the forecasts in 1981 compared to 1980 can be seen, at least in the Northern Hemispheric areas.



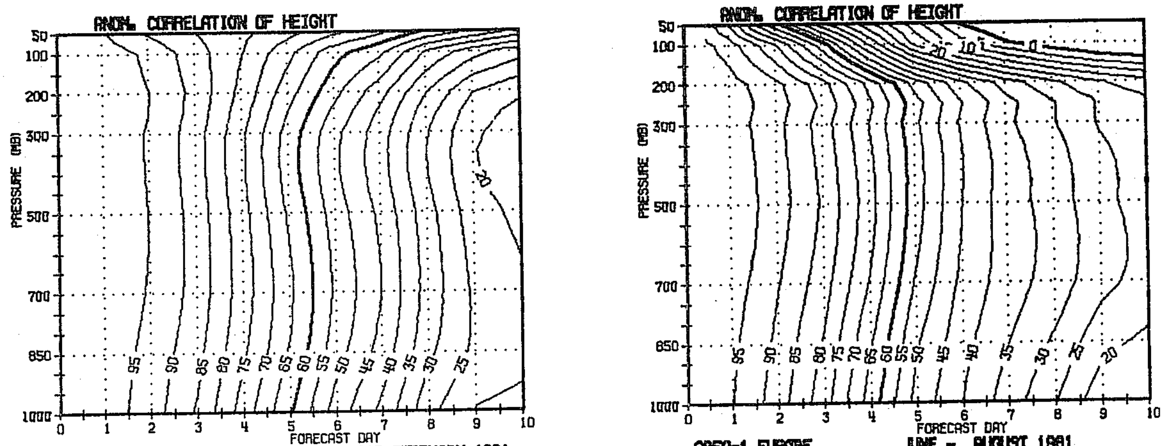
AREA-3 NORTH AMERICA DECEMB. 1980-FEBRUARY 1981
(60N-24N, 120W-72W)

AREA-3 NORTH AMERICA JUNE - AUGUST 1981
(60N-24N, 120W-72W)



AREA-2 CHINA-JAPAN DECEMB. 1980-FEBRUARY 1981
(60N-24N, 102E-150E)

AREA-2 CHINA-JAPAN JUNE - AUGUST 1981
(60N-24N, 102E-150E)



AREA-1 EUROPE DECEMB. 1980-FEBRUARY 1981
(72N-36N, 12W-42E)

AREA-1 EUROPE JUNE - AUGUST 1981
(72N-36N, 12W-42E)

Figure 7: Vertical time-sections of height anomaly correlations for three Northern Hemispheric areas. The left sections are the averages for the period of December 1980 - February 1981 and the right sections for the period of June - August 1981.

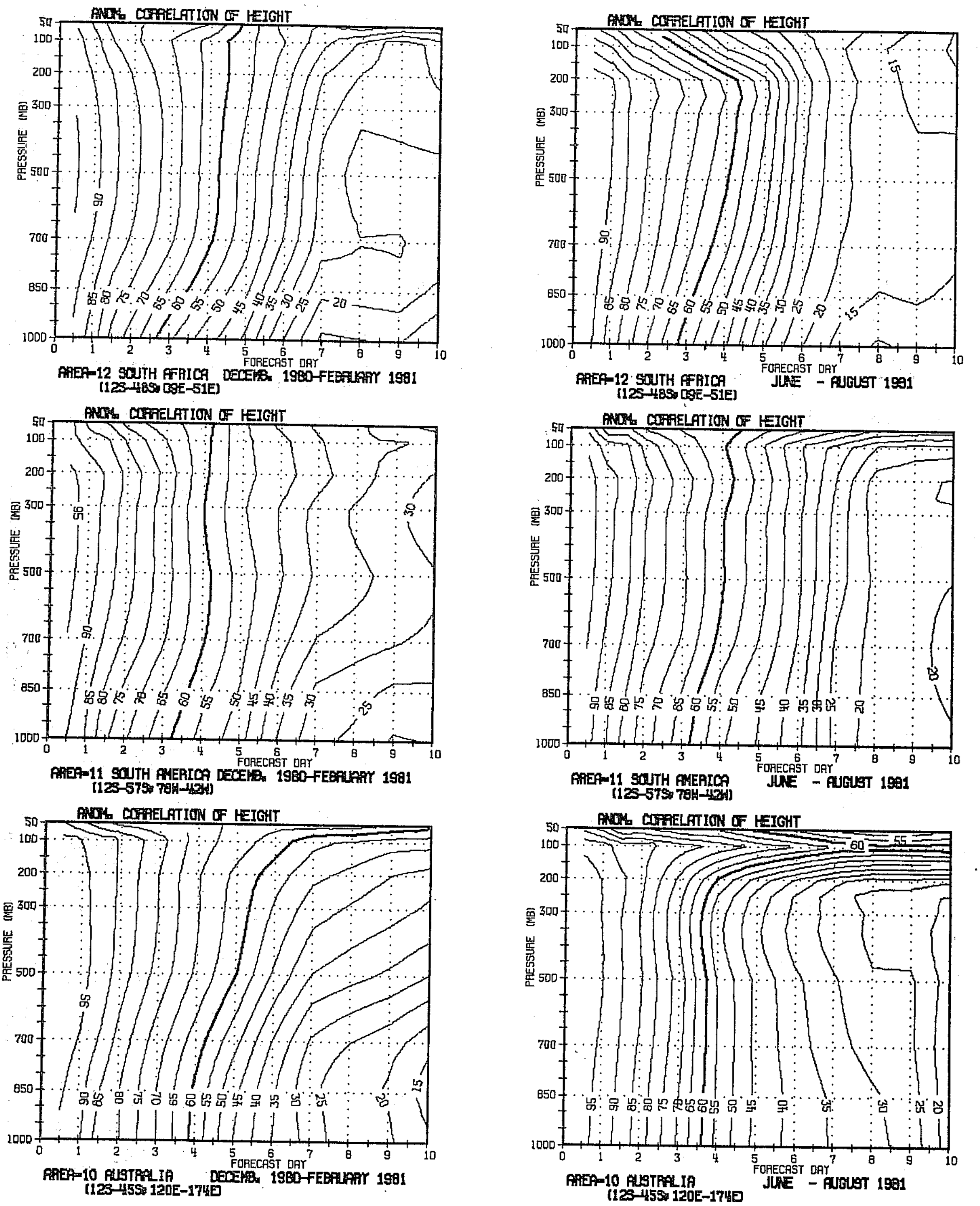


Figure 8: As Figure 7 but showing vertical time-sections of height anomaly correlations for three Southern Hemispheric areas.

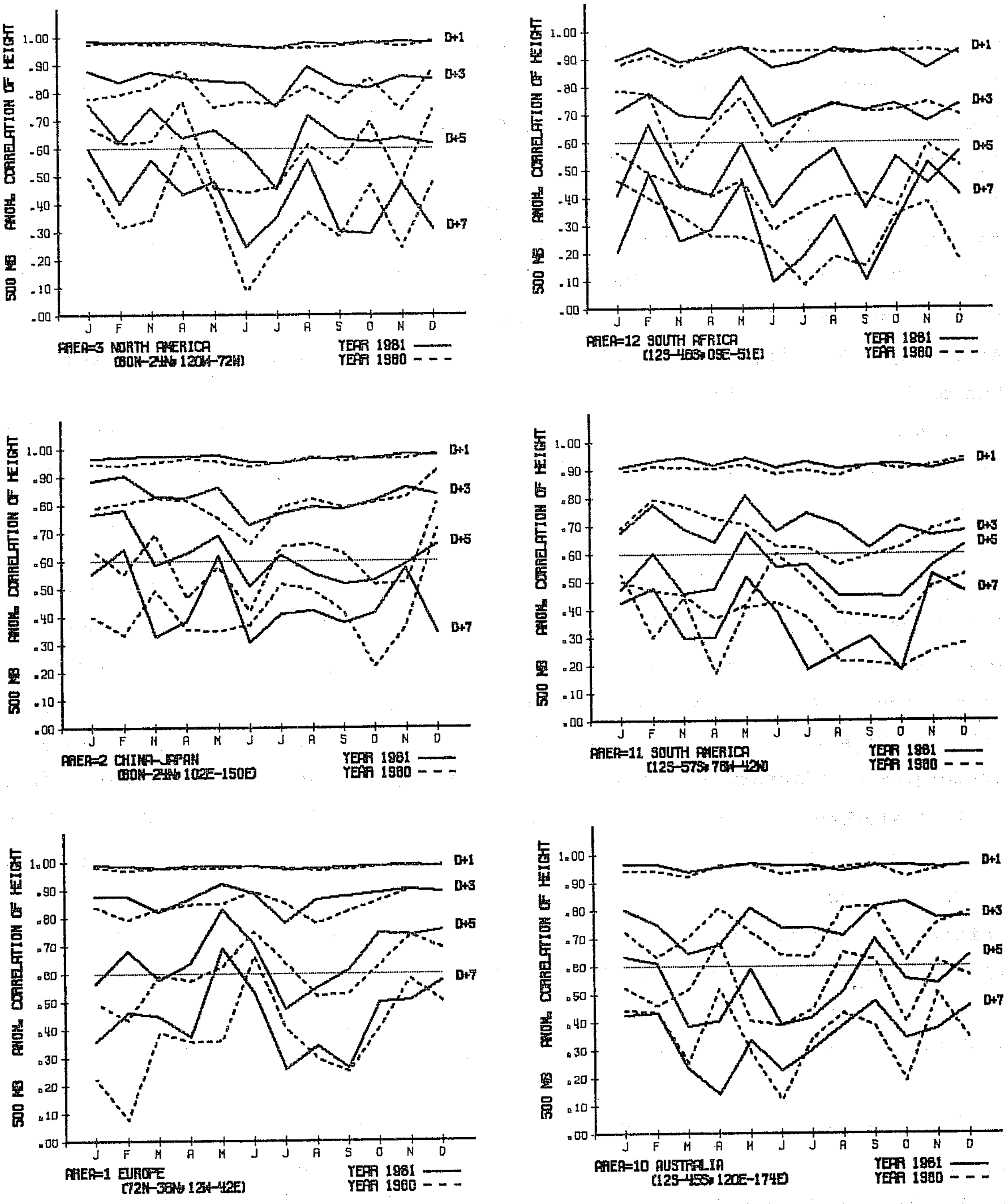


Figure 9: Monthly averages of 500 mb height anomaly correlations for three Northern Hemispheric areas (left) and for three Southern Hemispheric areas (right) for the years 1980 and 1981.

3.3 Daily variation

The accuracy of the forecasts varies considerably from day to day, especially in limited areas. Figure 10 presents, as an example, the daily scores for the European region from 1 October to 31 December 1981. The date of the plots corresponds to the initial time of the forecast and not to the verifying time. The anomaly correlation and the standard deviation of error for the 500 mb geopotential height shows that the one day forecast is uniformly of a very high quality. We see from Figure 10 that the 3-day forecasts are above the level of 0.80 of the anomaly correlation for most of the time; this level can be taken as a limit for good forecasts. After day 3, however, the variability of the predictive skill increases remarkably. The daily scores in Figure 10 show some poor 5-day forecasts, below the limit of 0.60, and a number of extremely bad 7-day forecasts, but on the other hand also several longer episodes where the 5 or 7-day forecasts reach reasonably high score values.

The standard deviation of the height anomalies for the analyses which verify the 7-day forecasts is also plotted in Figure 10. When comparing the verifying anomalies to the anomaly correlations, we observe that the good forecasts in many cases occur when large anomalies exist and the poor cases often with low anomaly values. This kind of relationship is natural, because the large scale motion, which can be quasi-stationary at times, is easier to predict than the fast moving small scale disturbances which are associated with low anomalies. The same relation is illustrated by Figure 11 which represents a scatter diagram between the anomaly correlation scores and the observed anomalies of the 5-day forecasts for the 500 mb height from the same example period presented in Figure 10. The major part of the higher scores are gathered around the anomaly values of 100-140 meters. The range of the observed anomalies corresponding to the poor forecasts is much wider, although a certain trend towards smaller anomalies can be noticed.

Figure 10 shows very clearly that there is a high correspondence between the anomaly correlation values and the forecast error, so that the error tends to increase while the correlation decreases. This relationship between high correlation scores and small errors and low scores and large errors, respectively, is demonstrated also in Figure 12 where we have a scatter diagram for the anomaly correlation and the standard deviation of the 5-day forecast error of the 500 mb height, again covering the 3 month period of October-December 1981. The same plot also illustrates the difficulties in defining an exact limit for a useful forecast. The range of error for the 5-day forecast, in Figure 12, is largest just in the area where the anomaly correlation reaches values around 0.50-0.60, which are regarded as lower threshold levels of predictive skill. On the other hand, it could perhaps be argued that the usefulness of forecasts with such a wide range of

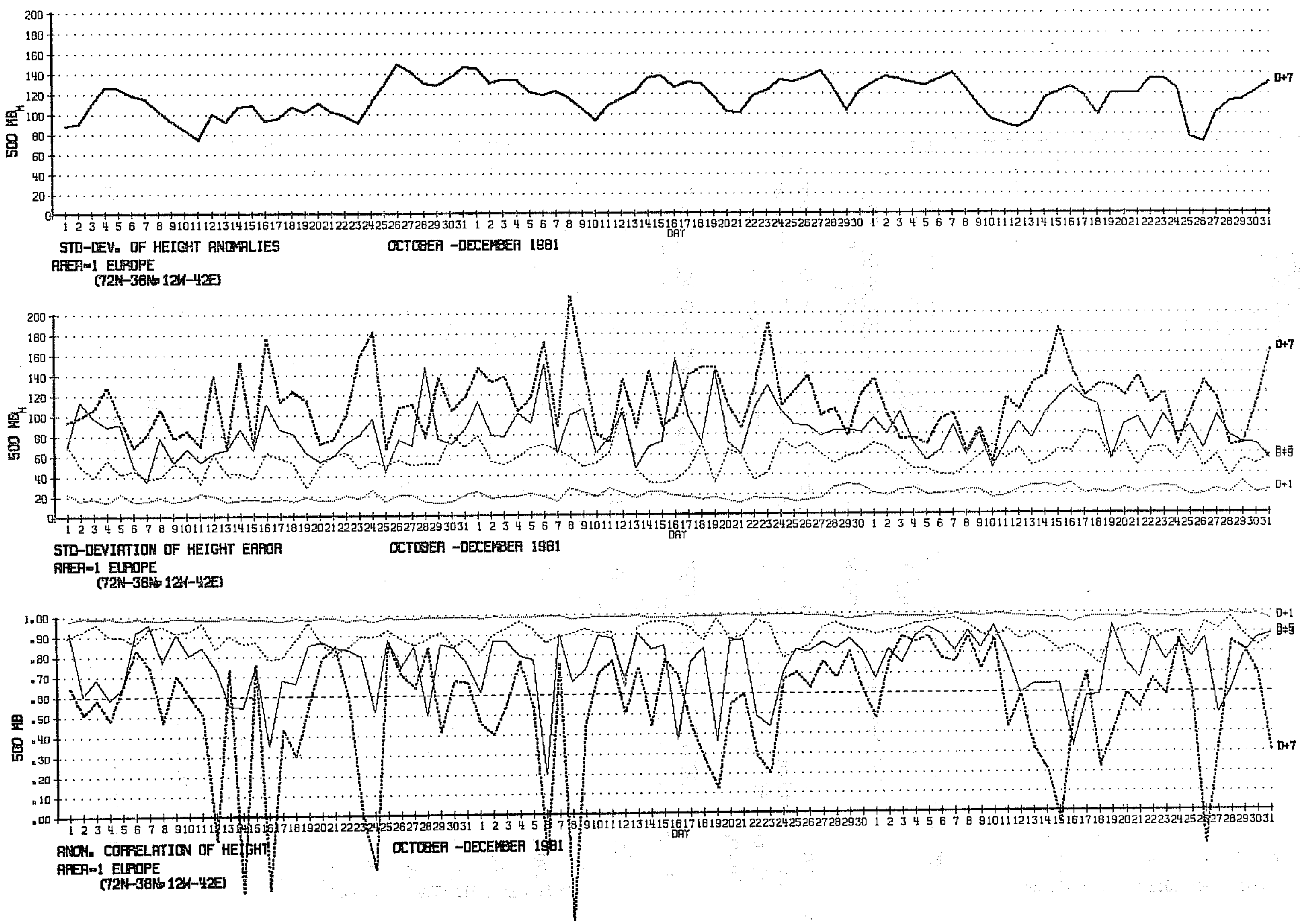


Figure 10: Daily variation of the anomaly correlation and the standard deviation of error of the 500 mb height for forecast days 1,3,5 and 7 from 1 October to 31 December 1981 in Europe. The standard deviation of the observed height anomalies which verify the 7-day forecasts is shown at the top.

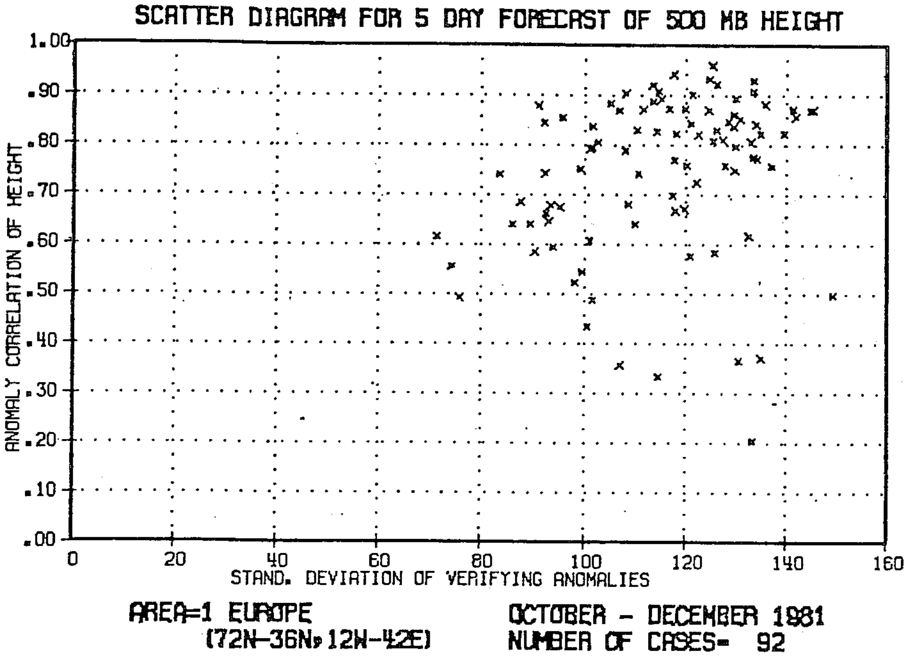


Figure 11: Scatter diagram showing the relationship between the anomaly correlation and the observed anomaly for the 5-day forecasts of the 500 mb height for the period October-December 1981 in Europe.

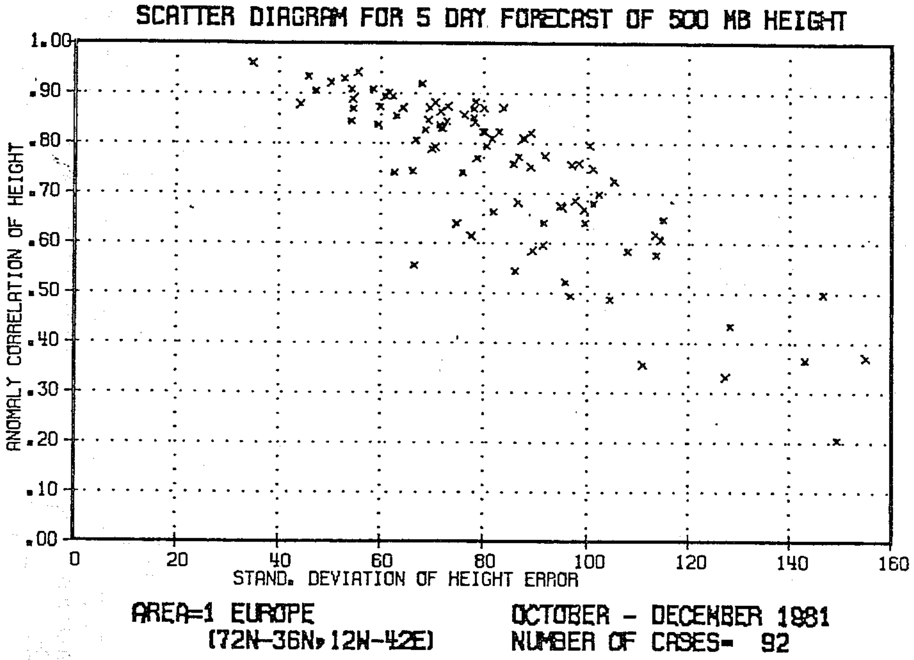


Figure 12: As Figure 11 but showing the relationship between the anomaly correlation and the standard deviation of error.

errors is severely reduced, thus justifying the definition of the anomaly correlation values of 0.50-0.60 as a limit for useful predictability. In this connection we must remember that the daily scores presented here measure the accuracy of individual forecasts over a limited area. In some cases, however, a series of medium range forecast maps can still give useful information, for example, about the general flow type, although the phase errors of migratory waves reduce the regional score values.

3.4 Predictive skill in extratropical hemispheric areas

Figures 13 and 14 show the averaged performance of the ECMWF model for the two hemispheres in 1980-81. The monthly means of the anomaly correlation of geopotential height (Figure 13) and temperature (Figure 14) at the 1000, 500 and 200 mb levels for the years 1980 and 1981 suggest that the modifications of the analysis and forecasting system made during that time have had a positive impact on the products. Most of the time the scores of 1981 are slightly better than the corresponding results of 1980.

The complete list of the modifications made during 1980-1981 to the Centre's global forecasting system can be seen in Appendix 2. One major change concerning the data assimilation took place in December 1980, when the increment interpolation technique was implemented (i.e. the analysed deviations between the first-guess and observations, instead of the full fields, are interpolated vertically to σ-levels). The other change affecting the Southern Hemisphere happened in March 1981, when Paobs, (manually derived surface observations from satellite data), were introduced in the analysis system. From the several modifications for the forecast model the following events could be highlighted: changes to the parameterisation and vertical diffusion in the beginning of 1980, a new orography and some other surface fields in April 1981, a revised interpolation scheme from σ- to p-levels for the analysed and forecast fields in June 1981, and finally, a modified physical package in December 1981.

The differences between the years 1980 and 1981 are more obvious at the 1000 mb level than at the upper levels. This probably reflects the improvements in the surface analysis and in the boundary layer of the model. The most significant differences can be seen in the temperature scores at the 1000 mb level, as shown by Figure 14. The sudden increment of the temperature correlation in the Northern Hemisphere indicates that this remarkable change is partly related to the influence of the revised interpolation scheme on both the analysed and predicted fields. On the other hand, the pronounced improvement in the temperature forecasts, especially in the Southern Hemisphere, starts at the end of 1980 and is therefore associated with the implementation of the increment interpolation technique.

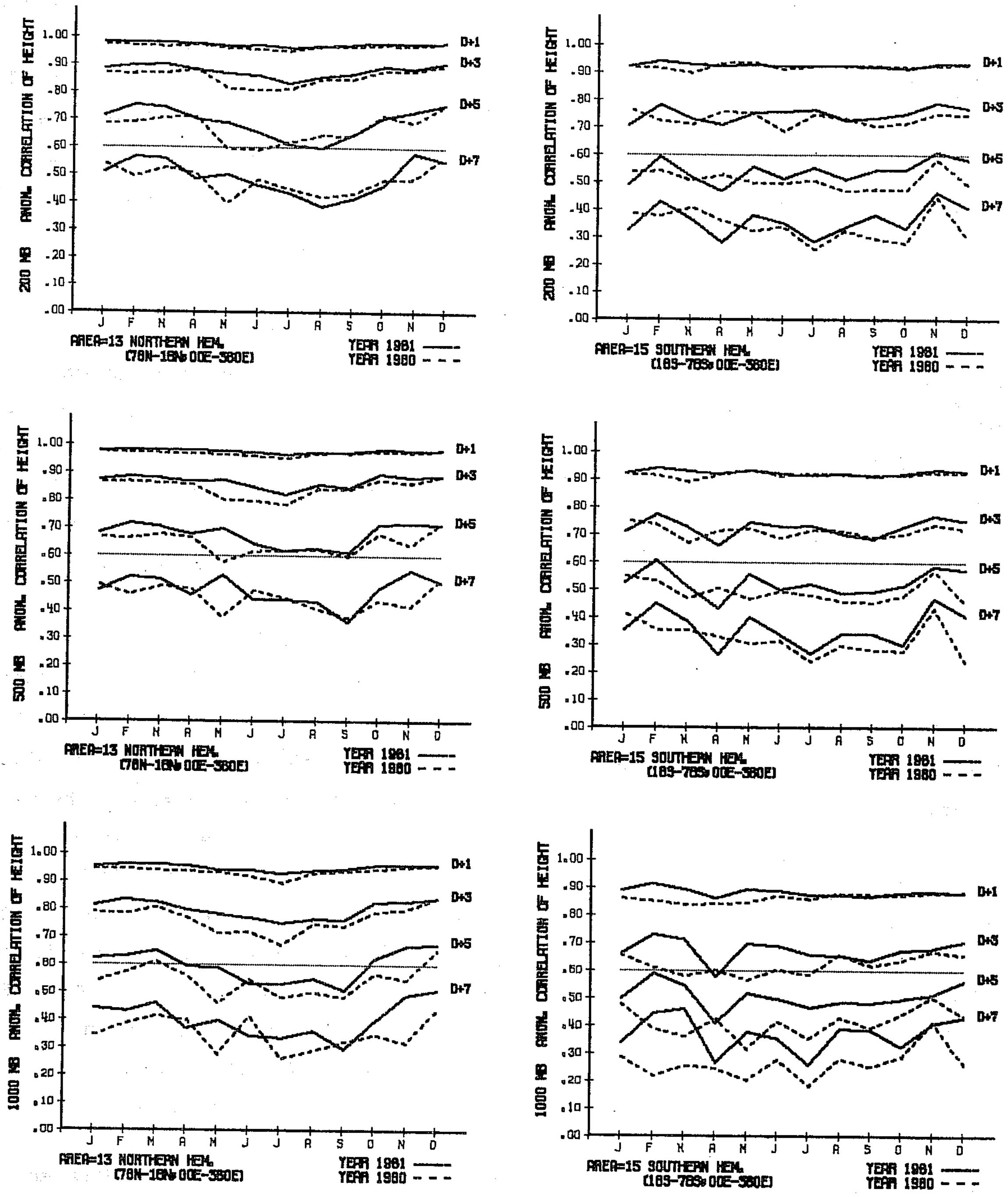


Figure 13: Monthly averages of anomaly correlations for height at 1000, 500 and 200 mb in the Northern Hemisphere (left) and in the Southern Hemisphere (right) for the years 1980 and 1981.

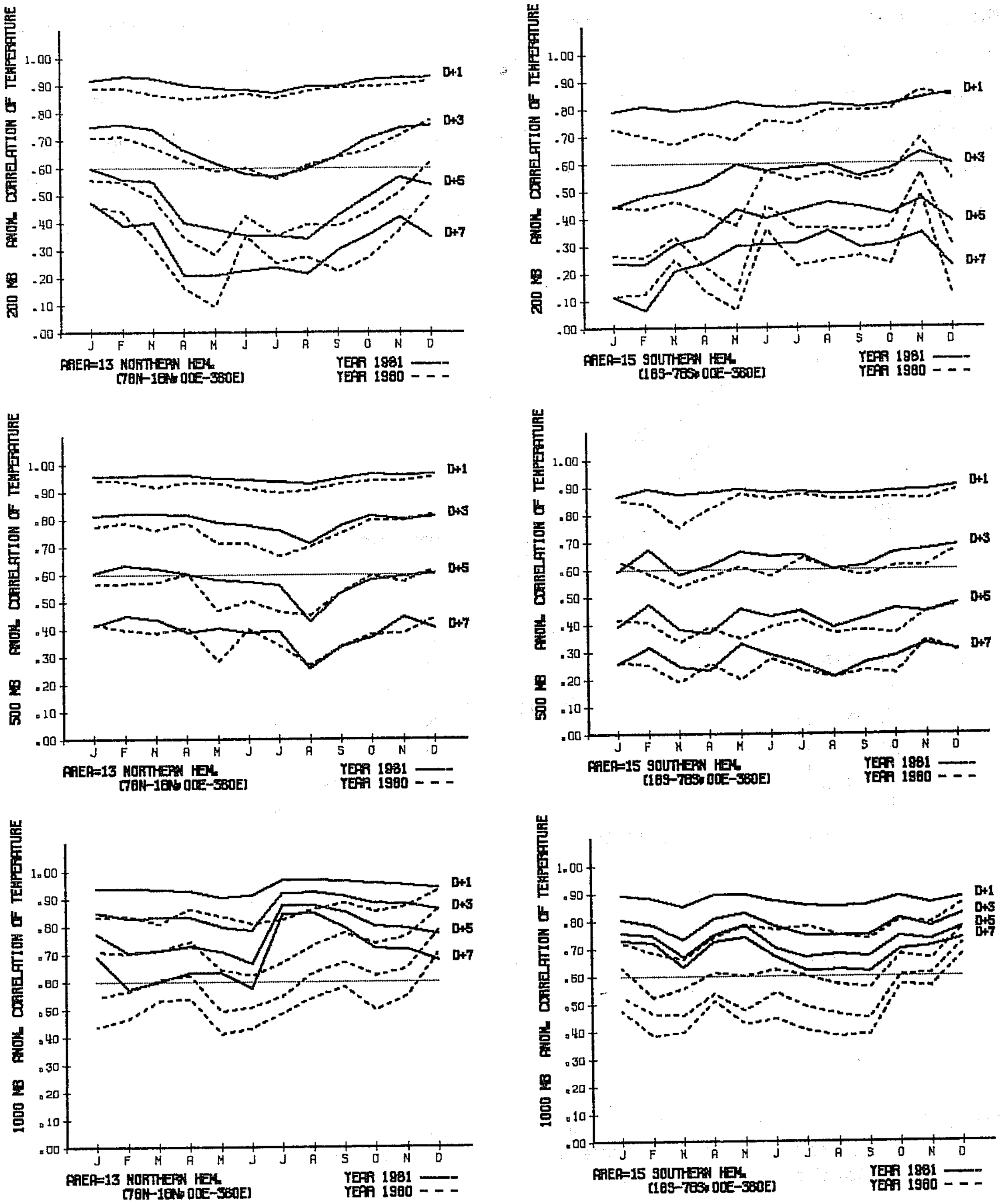


Figure 14: As Figure 13 but for temperature

The overall positive development is more clearly seen in Figure 15, which shows the averaged anomaly correlation of 500 mb height for forecast days 1 to 10 for the years 1980 and 1981. The figure also illustrates the difference of the forecast quality in both hemispheres. The predictive skill is 1.5 to 2 days longer in the Northern Hemisphere. The considerably denser data coverage in the Northern Hemisphere is clearly an important contributer to this difference.

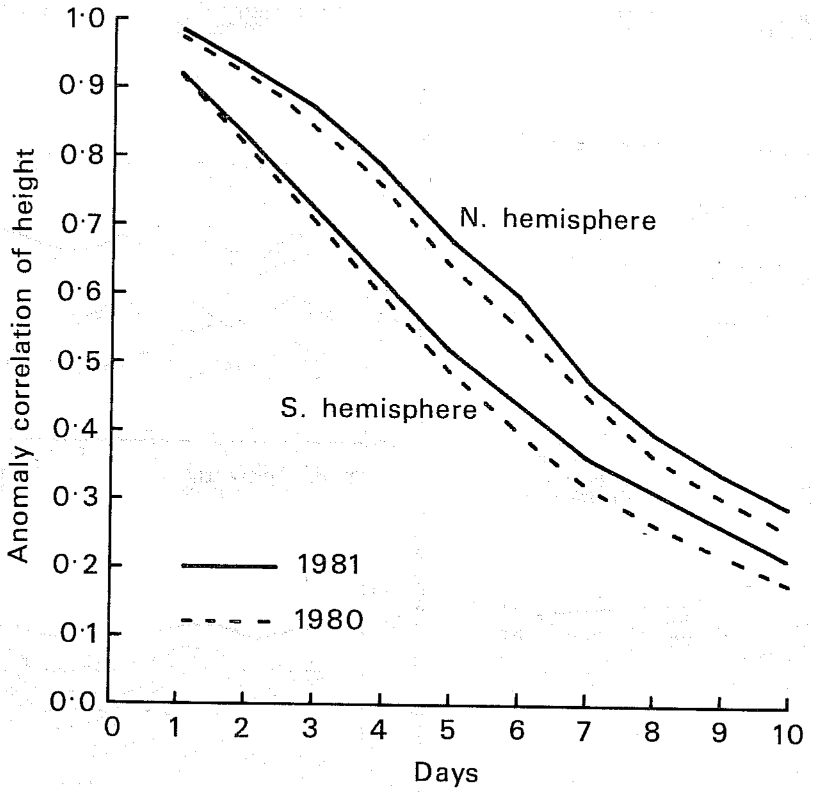


Figure 15: Anomaly correlation of 500 mb height for forecast days 1 to 10 averaged over the years 1980 and 1981 for the Northern and Southern Hemisphere.

Figure 16 shows the cumulative frequency distribution for the anomaly correlations of the 500 mb height for both hemispheres during 1980 and 1981, demonstrating the range of the predictive skill of the ECMWF forecasts. From Figure 16 we can see, for instance, that in the Northern Hemisphere all one day forecasts were on or above the level of 0.90 of the anomaly correlation. Taking the anomaly correlation value of 0.80 as a measure for "good" forecasts and the value of 0.60 to define the limit for "useful" forecasts, Figure 16 shows that in 1981 about 90% of the 3-day forecasts of the 500 mb height for the Northern Hemisphere were good, and almost 50% of the 6-day forecasts were useful. In 1980 the corresponding frequency numbers for the Northern Hemisphere were slightly lower, when 80% of the 3-day forecasts could be classified good and 35% of the 6-day forecasts useful.

The lower predictive skill in the Southern Hemisphere compared to the Northern Hemisphere can be seen from Figure 16 which shows the frequency distributions for the Southern Hemispheric scores too. Using the same criteria as above, we see from Figure 16 that in 1981 the 2-day forecasts of the 500 mb height were good in 75% of the cases (increased from 70% in 1980), and 60% of the 4-day forecasts were useful (compared to 50% in 1980).

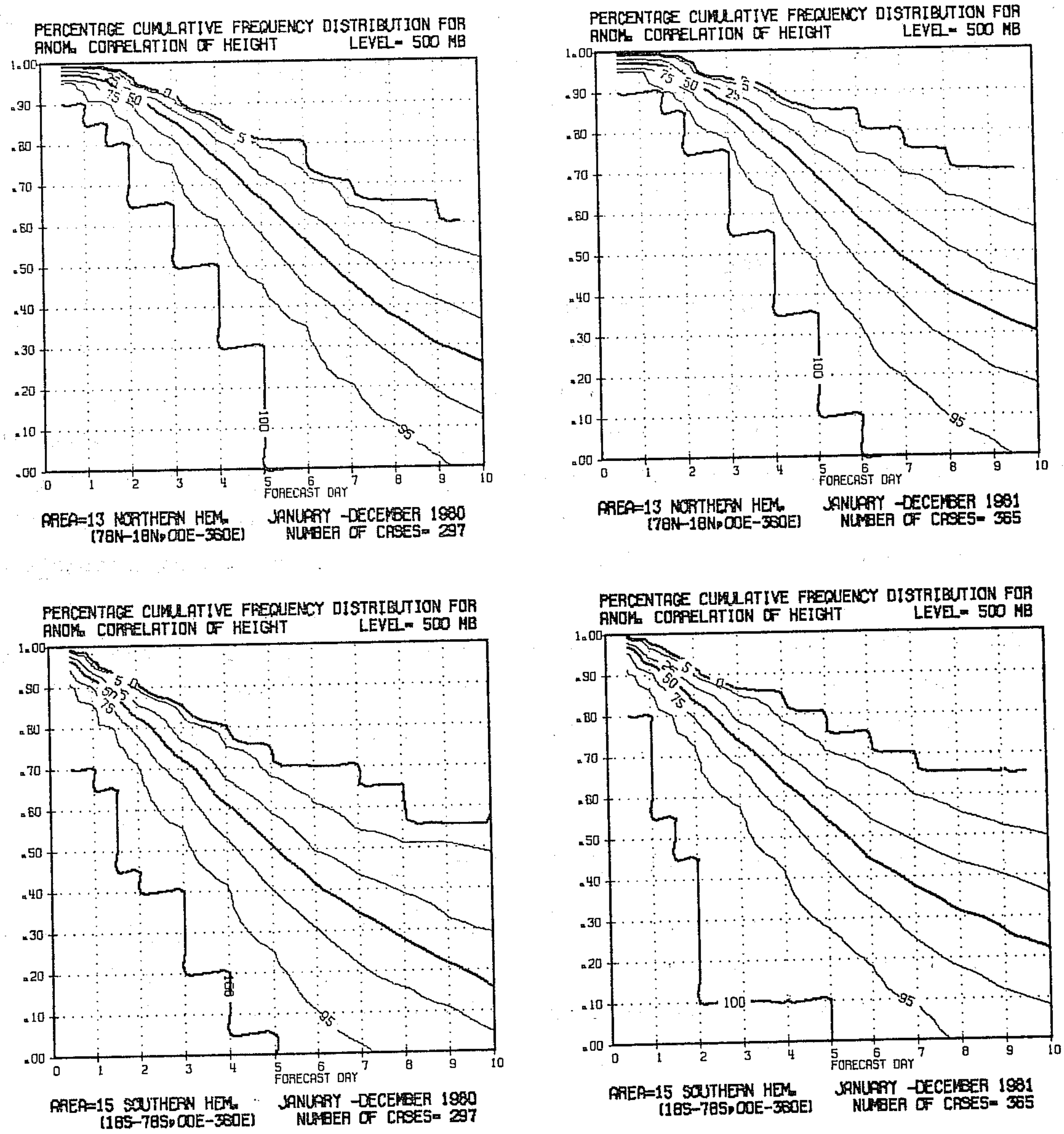


Figure 16: Cumulative frequency distribution of height anomaly correlation at 500 mb for the years 1980 (left) and 1981 (right) for the Northern Hemisphere (top) and for the Southern Hemisphere (bottom)

3.5 Predictive skill in Europe

Some of the scores for the European area have been introduced in sections 3.2 and 3.3 where the spatial and temporal variation of predictability was considered. Figure 17 represents the percentage cumulative frequency distribution of height anomaly correlation at the 1000 and 500 mb levels for the years 1980 and 1981 in Europe. The positive development which was seen in the hemispheric scores is also evident in those for the European area. As in the previous section, we take the anomaly correlation values of 0.80 and 0.60 to define the limits for "good" and "useful" forecasts, respectively. Then, Figure 17 shows that in 1981 the 3-day forecasts for the 500 mb height in the European area were good in 80% of the cases, (this percentage increased from 75% in 1980). From the 6-day forecasts, 50% were useful in 1981, (this increased from about 40% in 1980). At the 1000 mb level, the corresponding frequency numbers are, as we can expect, lower. From Figure 17 we can see that for the 1000 mb height, 55% of the 3-day forecasts were good in 1981, (increasing from 50% for 1980), and from the 6-day forecasts, about 40% were useful in 1981 (compared to 35% in 1980).

Figure 18 shows the cumulative frequency distribution for the standard deviation of vector wind error and the corresponding persistence error at the 1000, 500 and 200 mb levels for the year 1981. Figure 19 shows similar distributions for the temperature error and the comparison to persistence. These plots give us an estimate of the range of predictive skill of wind and temperature forecasts in Europe. Considering vector wind, the 50% frequency line shows that for half of the 3-day forecasts the standard deviation of error is smaller than 11 m/s at 200 mb, below 9 m/s at 500 mb and about 7 m/s at 850 mb. The corresponding persistence errors are 20.7 m/s for 200 mb, 16 m/s for 500 mb and 10.7 m/s for 850 mb. In a similar way we can see that the 50% frequency line for the 3-day temperature forecasts crosses the error value of 2.8°C at 200 mb and 2.5°C at both 500 and 850 mb, while the corresponding persistence error is 4.5°C for 200 mb and 4.0°C for both 500 and 850 mb.

3.6 Predictive skill in the tropics and sub-tropics

The predictive skill of numerical forecasts is substantially lower for the short and medium range forecasts in the tropics compared to the middle and high latitudes. The reasons for this have been extensively discussed by several authors e.g. Shukla (1981), and it is essentially due to the low variability of the tropical atmosphere and the rapid growth of circulation systems driven by the release of latent heat.

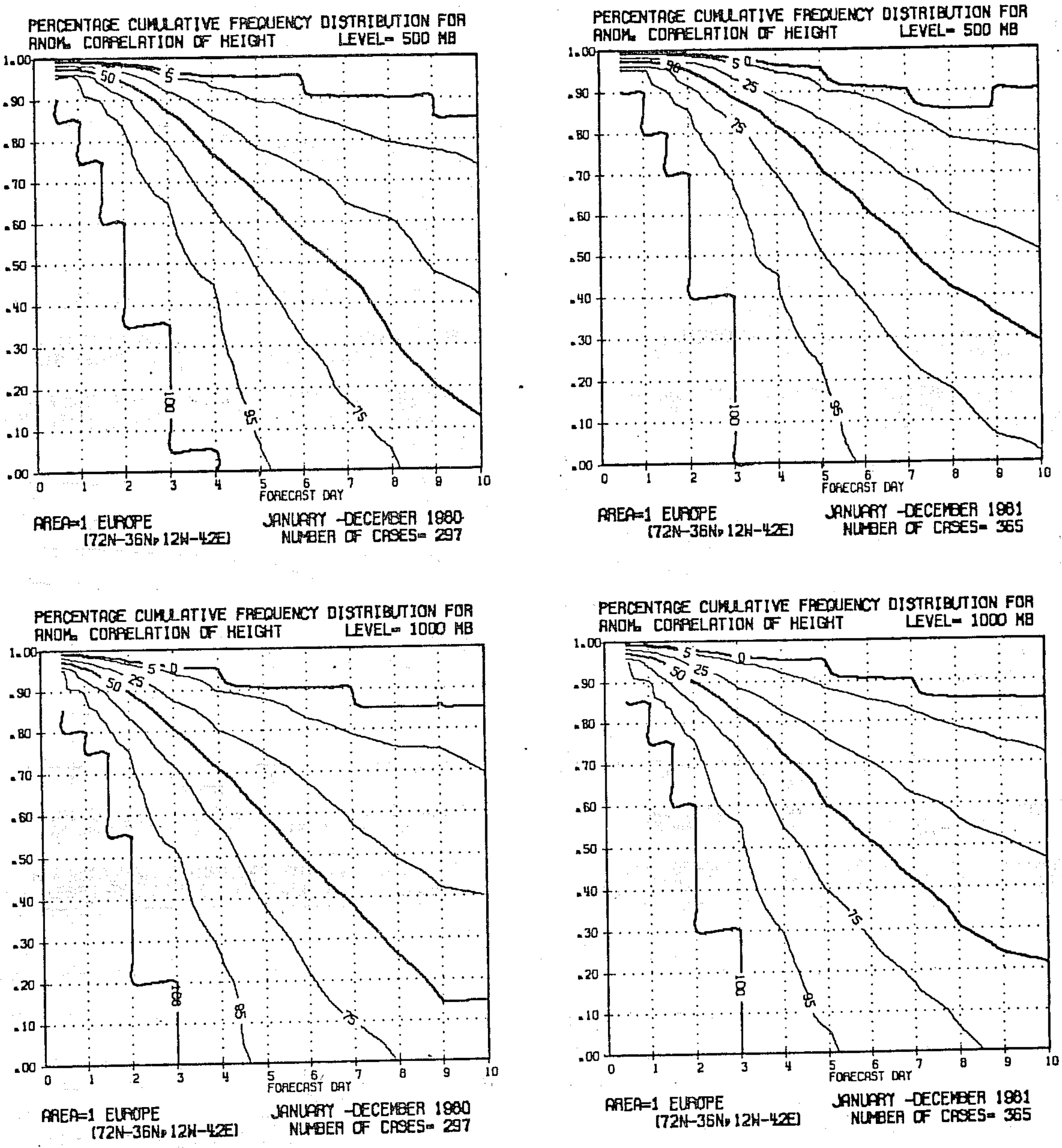


Figure 17: Cumulative frequency distribution of height anomaly correlations at 1000 mb (bottom) and at 500 mb (top) in the European area for the years 1980 (left) and 1981 (right).

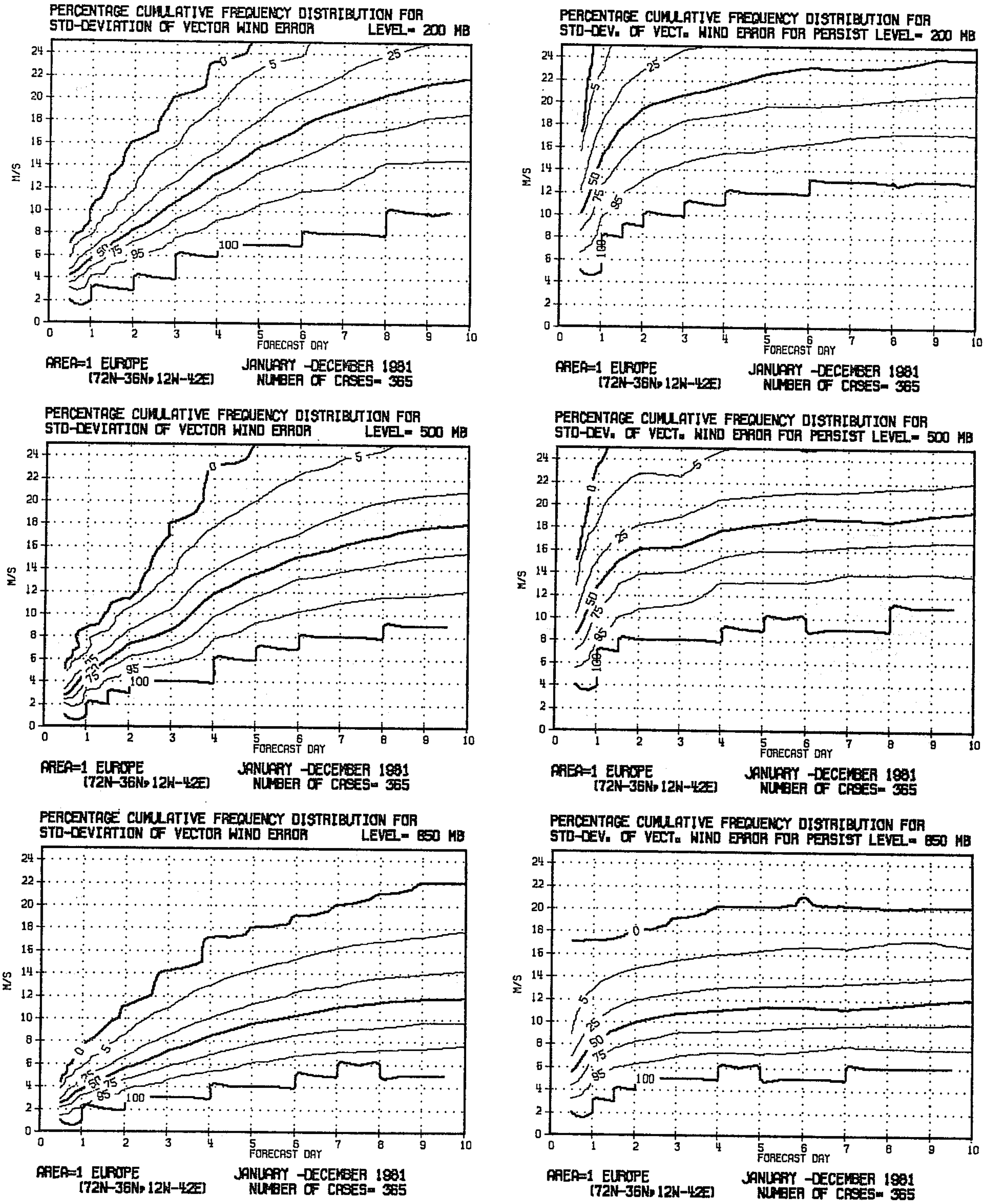
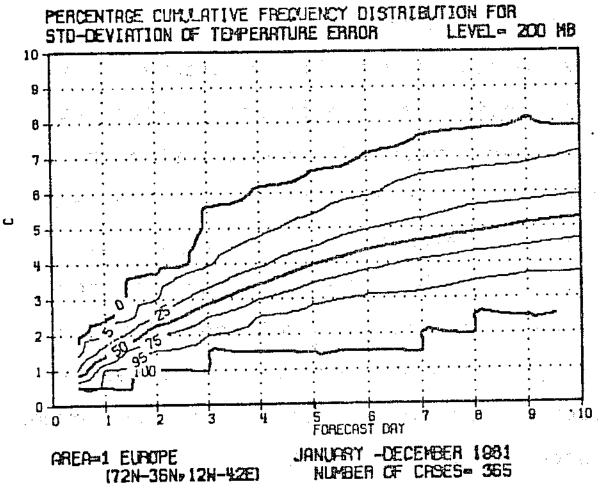
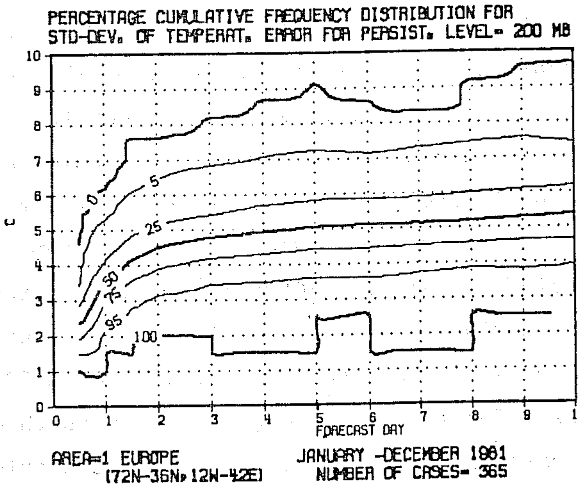
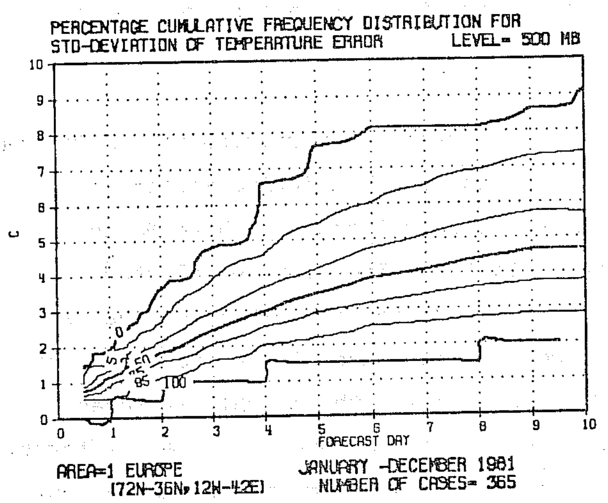
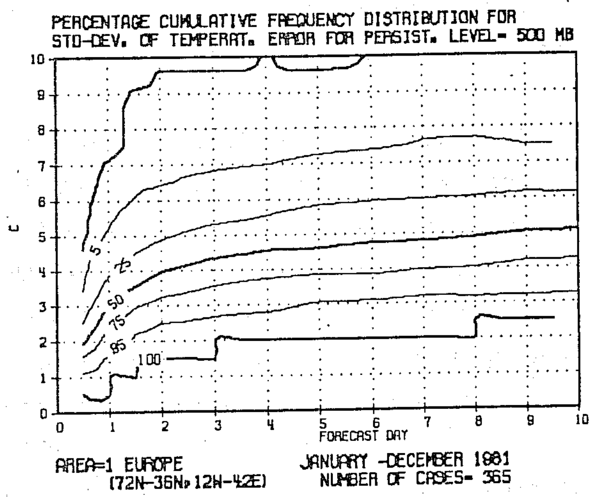


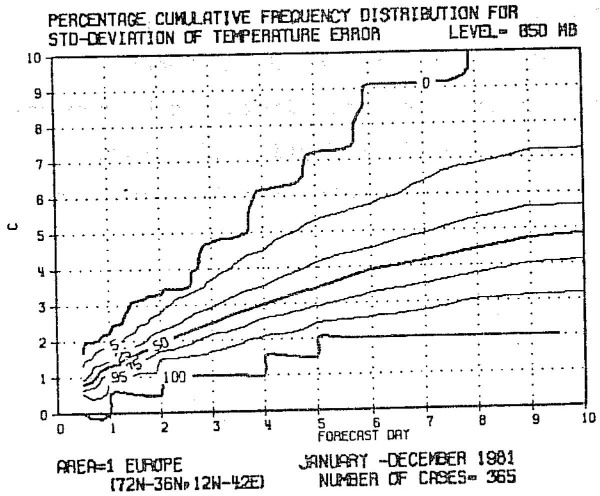
Figure 18: Cumulative frequency distribution for the standard deviation of vector wind error (left) with corresponding persistence error (right) at 850, 500 and 200 mb for the year 1981 in Europe.











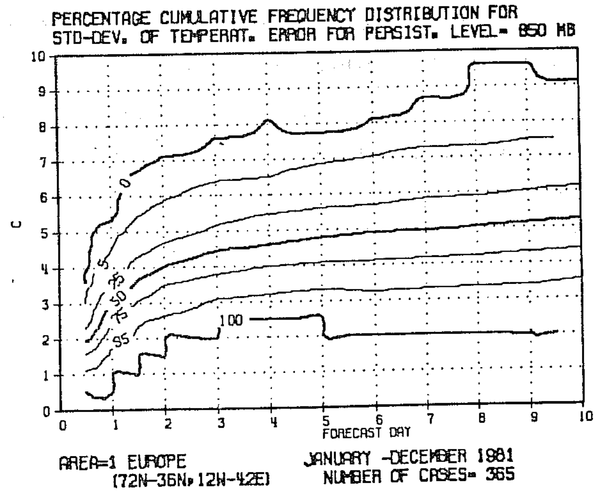


Figure 19: As Figure 18 but for temperature

Because of the different characteristics of the tropical atmosphere, the evaluation of tropical forecasts is different than that of the extratropical regions. The systematic error of the Centre's forecasts was already discussed in section 3.1, and it was found that the predicted height and temperature values are, on the average, too low in the tropics at the upper levels of the model atmosphere. Wind is one appropriate parameter for the objective assessment of the tropical and sub-tropical forecasts. The other important weather element for the tropics is rainfall, and the ECMWF forecasts of precipitation in the tropical areas have been investigated by Heckley (1981). Persistence has been used as a reference for the scores of the wind forecasts. Figure 20 shows the absolute correlation of vector wind for forecast and persistence in three tropical regions averaged over the summer period of June-August 1981. The comparison shows that in those tropical areas prediction has only little advantage over persistence, the forecast scores being higher mainly in the first two days.

In the regions located further from the Equator, the seasonal variation is substantial. The summer period resembles the conditions in the tropics, but during winter the westerly flow with transient waves affects the sub-tropical zones. Three Northern Hemispheric areas, North-West Africa, the Middle East and India, were selected to represent the subtropical areas. The plots in Figure 21 for the winter months of January-March 1981 show a considerable skill above persistence, in particular at the upper levels of the troposphere. Figures 22 and 23 present the monthly averages of the wind correlations at 850 and 200 mb for the years 1980 and 1981 over different tropical and sub-tropical areas. The one-day forecasts show some minor improvement for the 1981 scores in many of the regions, but the most positive development is found in the Arabian area at the 850 mb level where the skill of 3 and 5-day forecasts has also increased. Why this improvement occurred only in this sub-tropical area is difficult to explain, but may be related to the introduction of a more realistic soil moisture climatology, which possibly has affected the forecasts in 1981.

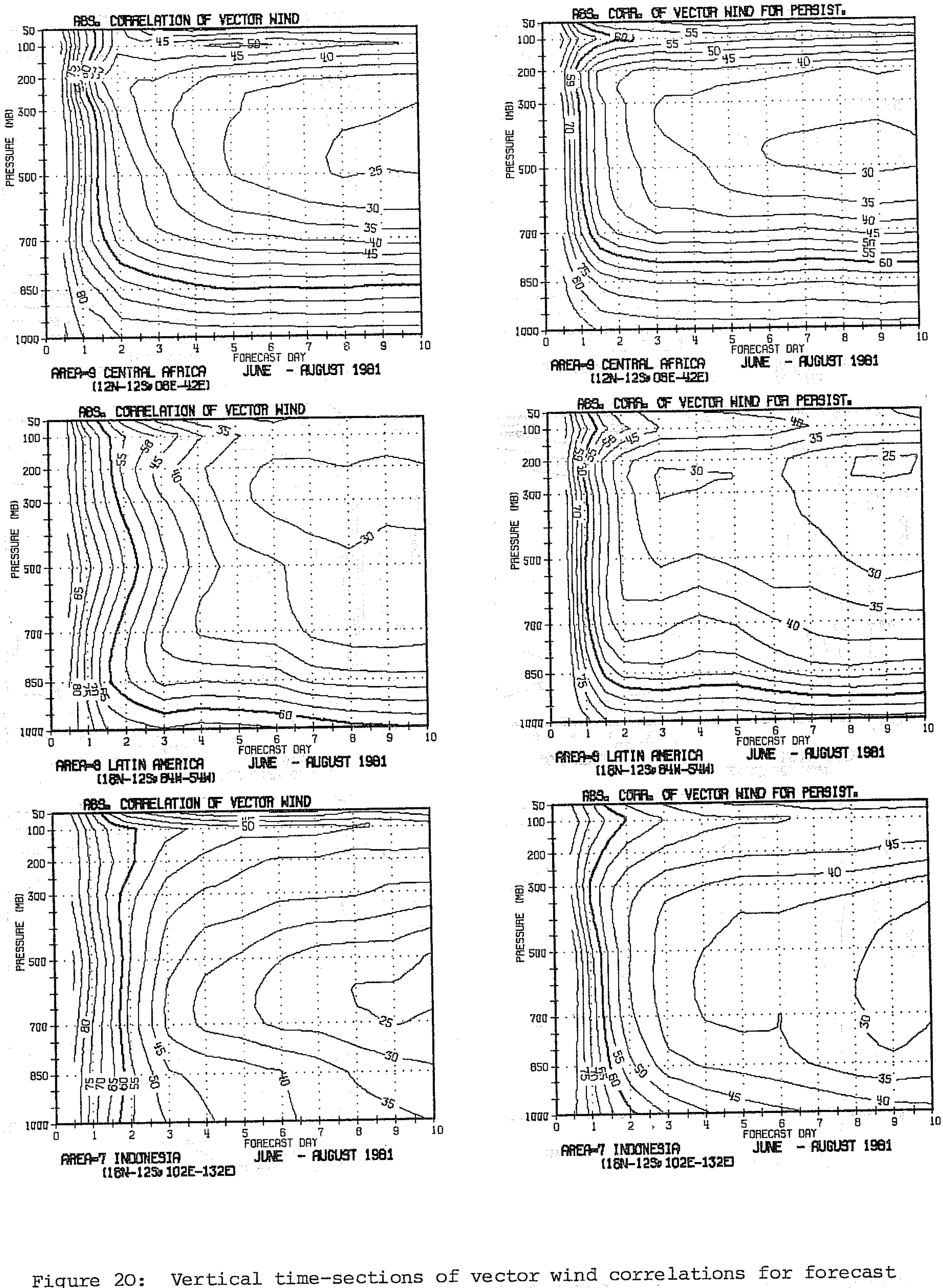


Figure 20: Vertical time-sections of vector wind correlations for forecast (left) and persistence (right) for three tropical areas averaged for the period of June-August 1981.

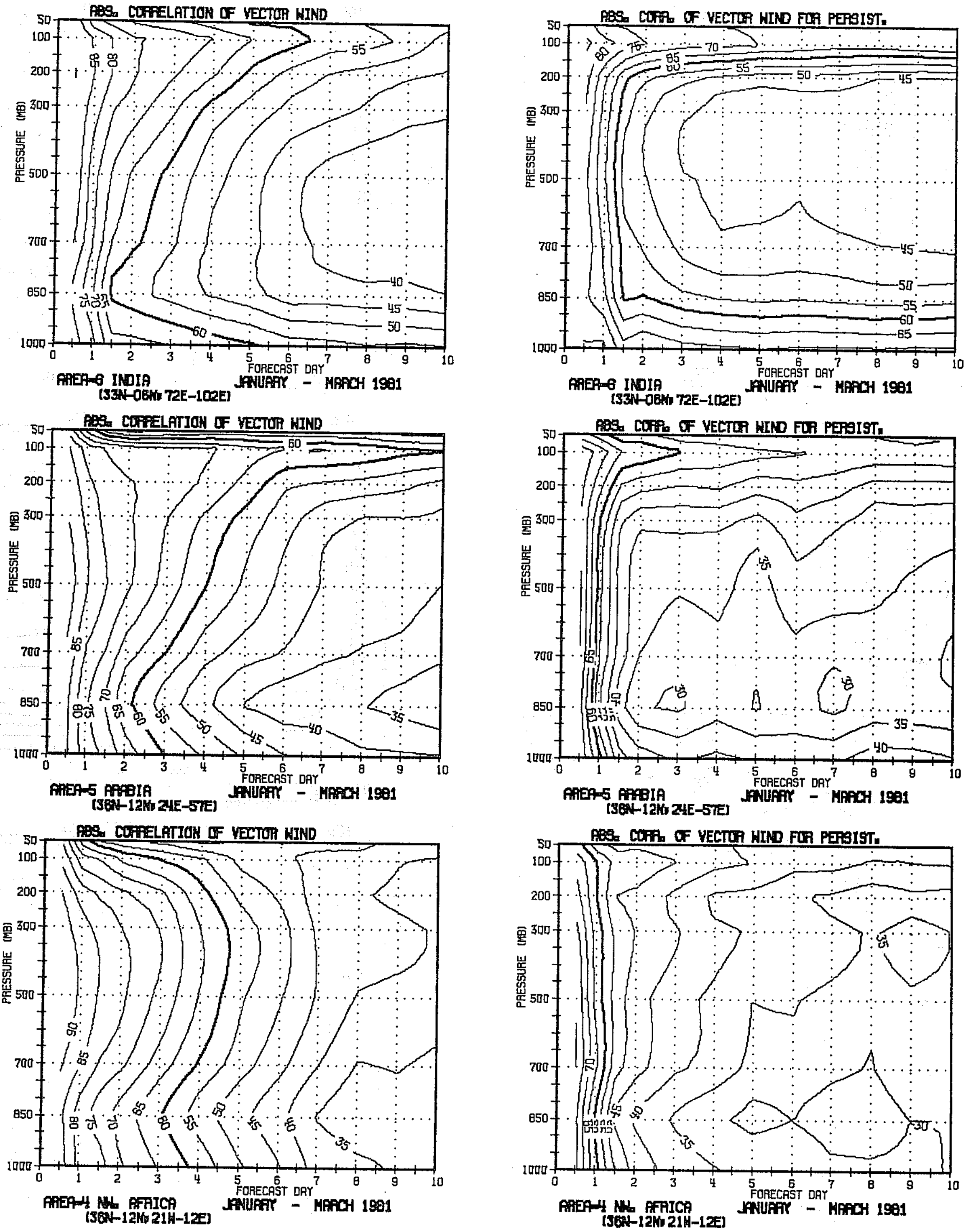


Figure 21: Vertical time-sections of vector wind correlations for forecast (left) and persistence (right) for three sub-tropical areas of the Northern Hemisphere averaged for the period of January-March 1981.

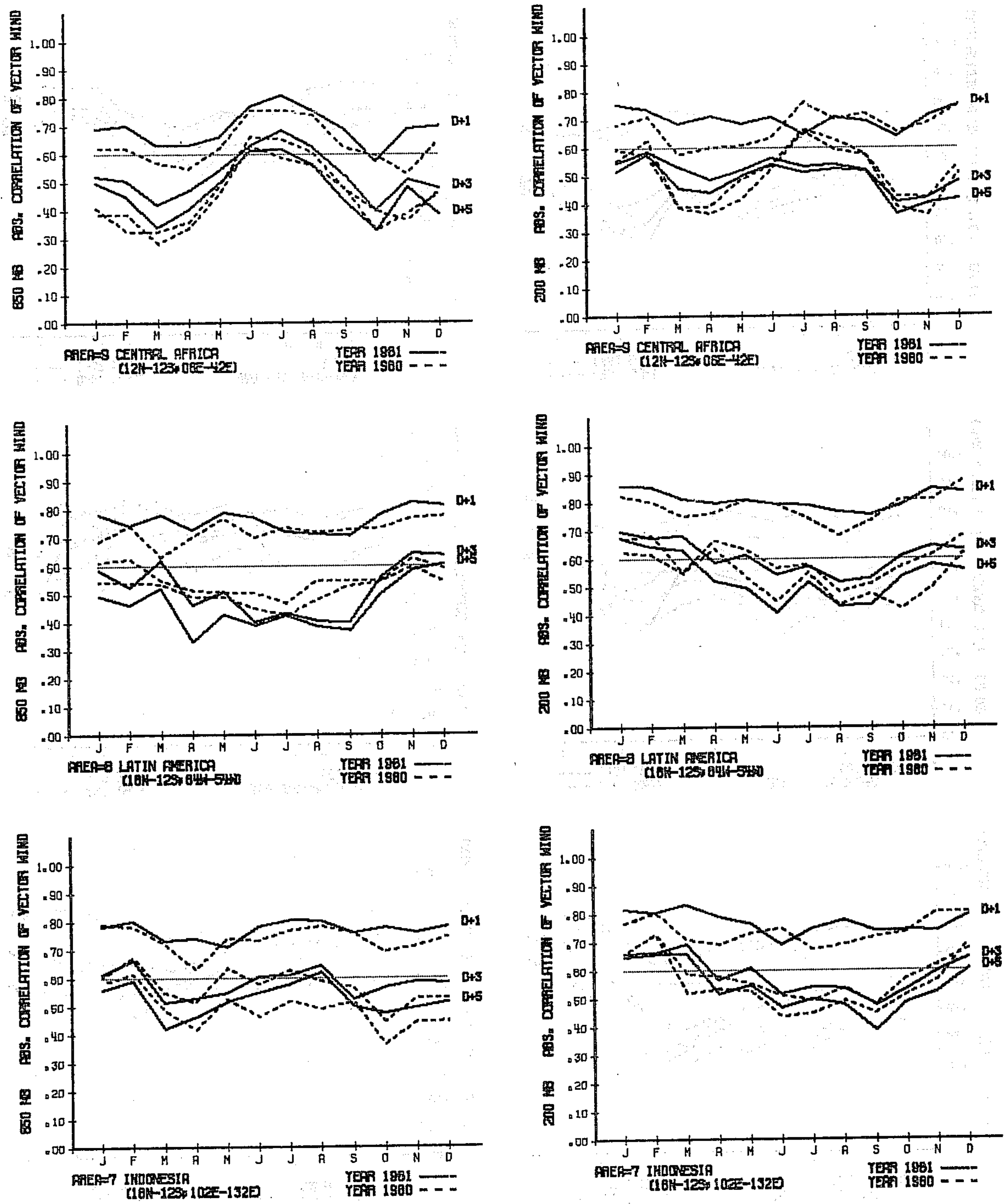


Figure 22: Monthly averages of vector wind correlations at 850 mb (left) and 200 mb (right) for three tropical areas for the years 1980 and 1981.

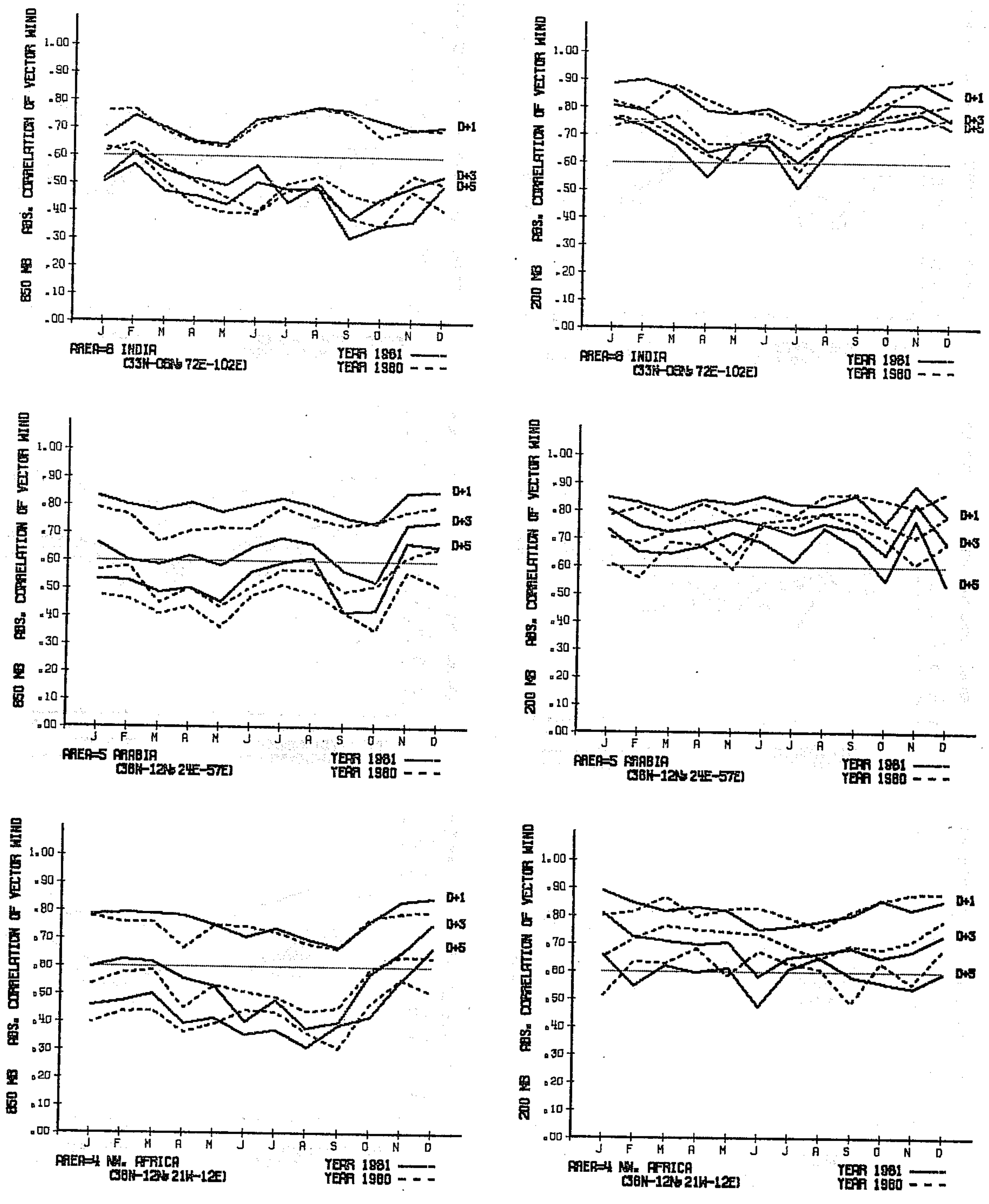


Figure 23: Monthly averages of vector wind correlations at 850 mb (left) and 200 mb (right) for three subtropical areas of the Northern Hemisphere for the years 1980 and 1981.

4. Summary

In the ECMWF operational field verification system described in this paper, grid-point values at standard pressure levels are used to compute objective scores, including rms-error, standard deviation of error, mean error and correlation coefficients for height, temperature, wind and relative humidity, and S1 skill score for height, for the two hemispheres, the tropical belt and 12 limited areas. The system thus provides a comprehensive set of objective scores for the evaluation of forecasts both on hemispheric scale and for limited areas.

The influence of horizontal resolution on verification results was tested by reducing the grid interval used in the operational verification from 6 to 3 degrees along latitudes in the European area and from 9 to 6 degrees in the Northern Hemispheric area. The higher resolution scores were compared with the operational scores, and no significant differences are seen in the results.

The consistency of the scores of the Centre's two objective verification systems was also tested. The anomaly correlation values from the spectral and grid-point verification over the Northern Hemisphere were compared, and the results agree very well.

The relationship between the anomaly correlation and the standard deviation of error was investigated; as would be expected, high correlations are associated with small errors and vice versa. However, the range of error scores increases around the anomaly correlation value of 0.60 which is often regarded as a critical limit of predictive skill.

The scores for 1980 and 1981 were examined in detail. The results show a considerable variability in time and space of the forecast accuracy in the extratropical areas. The differences in the predictive skill can be of the order of 1-2 days between different limited areas in the Northern Hemisphere, when the monthly averages of the winter scores are considered. The score values are generally lower in summer than in winter. Both the regional and the seasonal variation is larger in the Northern Hemisphere than in the Southern Hemisphere.

The accuracy of the Centre's forecasts was demonstrated by the cumulative frequency numbers of the objective scores. If we take the anomaly correlation values of 0.80 and 0.60 to define limits for "good" and "useful" forecasts, the results for 1981 show that for the Northern Hemisphere 90% of the 3-day forecasts for the 500 mb height were good and almost 50% of the 6-day forecasts were useful. In 1980 the corresponding frequency numbers were lower, about 80% for the 3-day forecasts and

35% for the 6-day forecasts. In the limited area of Europe, compared to the Northern Hemisphere, the predictive skill is slightly lower for the 3-day forecasts at the 500 mb level; about 80% of them were good, whereas over 50% of the 6-day forecasts were useful in 1981.

In the Southern Hemisphere the scores for the Centre's forecasts are lower, on the average, than those for the Northern Hemisphere. Using the same limits as above, 0.80 correlation for "good" forecasts and 0.60 for "useful" forecasts, the frequency numbers for 1981 show that 75% of the 2-day forecasts for the 500 mb height were good (70% in 1980), and 60% of the 4-day forecasts were useful in 1981 (50% in 1980). Generally, the predictive skill is lower by about 1-2 days on the Southern Hemisphere compared to the Northern Hemisphere.

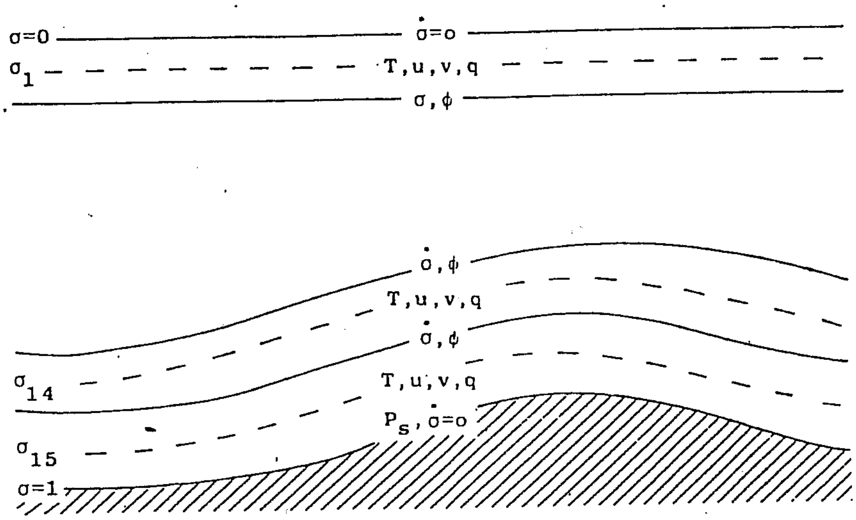
In the equatorial regions, wind forecasts were compared to persistence. The results show that in the tropical areas, prediction has only little advantage over persistence, the forecast scores being higher mostly in the first two days. In the sub-tropical areas, forecasts are clearly better than persistence in winter, but during summer the sub-tropical circulation is closer to the tropical conditions, and therefore the predictability is correspondingly reduced.

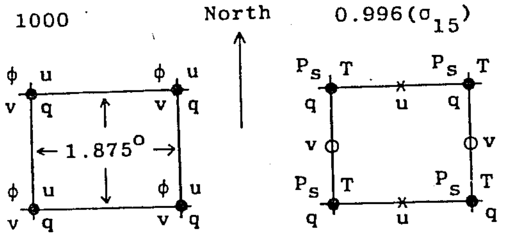
Important systematic errors of the Centre's model can be seen in the mean error values. The upper troposphere and the lower stratosphere of the model atmosphere cools systematically over all areas, indicating a lack of proper heat balance. The concentration of the mean error over certain areas is also very systematic, and this is probably connected to an inadequate orographical or thermal forcing. These types of systematic error are common for other high resolution models, too. For further discussion see Bengtsson and Simmons (1983).

Many revisions have been implemented in the data assimilation and forecasting system of the Centre during 1980-1981, and the verification results demonstrate that the forecasts have improved during that period. The introduction of a more realistic orography presently being investigated and the improvement of the parameterisation of sub-grid scale processes, particularly in the tropics, is likely to give a further reduction of the forecast error.

ECMWF - GLOBAL FORECASTING SYSTEM, 15-level grid point model
(Horizontal Resolution 1.875° Lat/Long)

ANALYSIS	PREDICTION
ϕ,u,v,q(for p>300)	T,u,v,q
p(mb)	σ
10	0.025(σ_1)
20	0.077
30	0.132
50	0.193
70	0.260
100	0.334
150	0.415
200	0.500
250	0.589
300	0.678
400	0.765
500	0.845
700	0.914
850	0.967
1000	0.996(σ_{15})





Vertical and horizontal (latitude-longitude) grids and dispositions of variables in the analysis (left) and prediction (right) coordinate systems.

ANALYSIS

Method	3 dimensional multi-variate (15-analysis levels, see above)
Independent variables	λ, φ, p, t
Dependent variables	ϕ, u, v, q
Grid	Non-staggered, standard pressure levels
First guess	6 hour forecast (complete prediction model)
Data assimilation frequency	6 hour (± 3 hour window)

INITIALISATION

Method	Non-linear normal mode, 5 vertical modes, adiabatic

PREDICTION

Independent variables	λ, φ, σ, t
Dependent variables	T, u, v, q, P_s
Grid	Staggered in the horizontal (Arakawa C-grid). Uniform horizontal (regular lat/lon). Non-uniform vertical spacing of levels (see above).
Finite difference scheme	Second order accuracy
Time-integration	Leapfrog, semi-implicit (Δt = 15 min) (time filter ν = 0.05)
Horizontal diffusion	Linear, fourth order (diffusion coefficient = $4.5 \cdot 10^{15}\ m^4 s^{-1}$)
Earth surface	Albedo, roughness, soil moisture, snow and ice specified geographically. Albedo, soil moisture and snow time dependent.
Orography	Averaged from high resolution (10') data set
Physical parameterisation	(i) Boundary eddy fluxes dependent on roughness length and local stability (Monin-Obukov) (ii) Free-atmosphere turbulent fluxes dependent on mixing length and Richardson number (iii) Kuo convection scheme (iv) Full interaction between radiation and clouds (v) Full hydrological cycle (vi) Computed land temperature, no diurnal cycle (vii) Climatological sea-surface temperature

APPENDIX 2

Modifications to the data assimilation and forecasting system at ECMWF during 1980-1981

15/01/80	Changes to forecast parameterisation scheme
19/02/80	Correction to vertical diffusion. Half precision routines in model physics.
27/02/80	New eigenvectors for initialisation.
29/02/80	Afilt after initialisation removed. Diffusion scheme in forecast changed.
11/03/80	Changes to Pregap. New climate tapes, with new correlations. Smoothed persistence used for 1st guess in stratosphere.
16/03/80	P* staggered for analysis vertical interpolation.
08/04/80	New observation errors used in analysis.
29/04/80	Fix for ill-conditioned matrices in gap. Corrections to pregap - 1) Direction of vertical variation of a B parameter 2) Treatment of cloud winds (non standard anal. levels).
16/06/80	Bug in WMO Master File routine fixed
29/07/80	Reduction of data used in analysis. Hough function approx. in stratosphere 1st guess field.
18/08/80	New soil moisture introduced to climate.
26/08/80	Q/C bug fixes - more data passed to analysis.
26/09/80	Omit tiros-N satellite data from analysis.
11/11/80	Virtual temperature changes to analysis and post-processing. Correction of error in lowest level wind staggering. Correction of error in satem extraction.
25/11/80	Overlapping boxes (analysis):- remove inconsistencies over box boundaries.
10/12/80	Introduce: A) analysis of increments B) changes to the interpolation library.
13/12/80	Symmetric multiple modifications made to CRAY-1.
08/01/81	Modification to stratospheric first guess.
17/02/81	Extra dry adjustment during initialisation to ensure static stability.
24/02/81	Addition of box numbers, source, and time of receipt to coded reports.
03/03/81	PAOB reports included in analysis.
04/03/81	SATEMS with base pressure 699 - 101 mean excluded from analysis.

10/03/81	Humidity analysis reintroduced.
17/03/81	Total tendency chopping space filter removed from 6 hr and 10 day forecast. Afilt removed from data assimilation.
18/03/81	Error removed from stratospheric humidity.
21/03/81	Changes made on 17/03/81 removed. Total tendency chopping reintroduced to forecasts, and afilt restored to data assimilation.
25/03/81	Errors removed from forecast. Total tendency chopping removed. Afilt removed from data assimilation. Deep soil temperature set to climatology.
31/03/81	PAOB observations restored.
01/04/81	New surface fields, especially topography, implemented.
07/04/81	Observation errors revised (FGGE values introduced).
16/06/81	Stratospheric interpolation scheme revised.
07/07/81	Optimisations, including 'SYMSOL', introduced into the analysis.
04/08/81	SATOB with latitude greater than 50 degrees rejected.
25/08/81	NOAA-7 SATEM data excluded from analysis.
10/09/81	SATOB eliminated outside the boundaries:- Latitude - -50 to +50 degrees Longitude - 90E to 170W (Japanese) 170E to 20W (GOES)
16/09/81	Modifications to surface fields to correct a bug in the new topography.
22/09/81	Change forecast time step from 15 min to 12 min.
29/09/81	Correct DEL4 diffusion scheme in forecast to overcome rainfall problem.
10/11/81	Change forecast time step from 12 min to 15 min.
08/12/81	Modifications to the model's parameterisation scheme - the first stage towards the introduction of the diurnal cycle.

Acknowledgements

I am very grateful to Dr. L. Bengtsson, who has encouraged me to write this paper, for his advice and many useful comments. I also thank Dr. S. Grønaas and Dr. A. Woods for helpful discussions and suggestions which led to improvements in the manuscript.

REFERENCES

Bengtsson, L., 1981: Medium range forecasting at ECMWF and remaining problems. Proceedings of the symposium on current problems of weather forecasting, Wien, June 23-26, 1981. WMO Publication No. 253, 115-124.

Bengtsson, L. and Lange, A., 1982: Results of the WMO/CAS NWP study and intercomparison project for forecasts for the Northern Hemisphere in 1979-1980. PWPR Report, WMO, Geneva, 26 pp.

Bengtsson, L. and Simmons, A., 1983: Medium range weather prediction - operational proceedings of the IAMAP Conference, Reading 3-7 August. To be published.

Böttger, H. and Grønaas, S., 1982: Optimal verification of ECMWF surface temperature forecasts at two locations in Europe. ECMWF Operations Department Technical Memorandum No. 64, 26 pp.

Heckley, W.A., 1981: Preliminary results of an investigation into the quality of ECMWF forecasts in the tropics. ECMWF Technical Memorandum No. 4, 46 pp.

Hollingsworth, A., Arpe, K., Tiedtke, M., Capaldo, M. and Savijarvi, H., 1980: The performance of a medium-range forecast model in winter. Impact of physical parameterisation. Mon. Wea. Rev.. 108, 1736-1773.

Johannessen, K.R., 1982: Verification of ECMWF quantitative precipitation forecasts over Europe from January 1980 to April 1981. ECMWF Research Department Technical Memorandum No. 51, 71 pp.

Lorenc,A., 1981: Global 3-dimensional multivariate statistical interpolation scheme. Mon. Wea. Rev., 109, 701-721.

Miyakoda, K., Hembree, G.D., Strikler, R.F., Shulman, I., 1972: Cumulative results of extended forecast experiments. I. Model performance for winter cases. Mon. Wea. Rev., 100, 836-855.

Pümpel, H., 1982: Regional and local verification of near surface weather parameters. ECMWF Seminar proceedings, 1982.

Shukla, I., 1981: Predictability of monthly means: Part II - Influence of the boundary forcing. Proceedings of ECMWF Seminar on problems and prospects in long and medium range weather forecasting, 261-312.

Teweles, S., Wobus, H.B., 1954: Verification of prognostic charts. Bulletin of Am. Met. Soc., Volume 35, 455-463.

Wallace, J.M., Tibaldi, S., Simmons, A., 1983: Reduction of systematic error in the ECMWF model through the introduction of an envelope orography. To be published in Quart. J. Roy. Met. Soc.

Åkesson, O., 1981: Evaluation of the ECMWF operational model precipitation forecast in October-November 1980. ECMWF Operations Department Technical Memorandum No. 24, 17 pp.

Åkesson, O., Böttger, H., Pümpel, H., 1982: First results of direct model output verification of near-surface weather parameters at 17 locations in Europe. ECMWF Operations Department Technical Memorandum No. 55, 46 pp.

WMO/CAS Working Group on Weather Prediction Research (Sub-group on Short and Medium range Weather Prediction). Final Report, Peking, 1978.

ECMWF Meteorological Bulletin M2.2/1: Objective assessment, 1981.

ECMWF Forecast Reports, 1980-1981: Published monthly (1980) and quarterly (1981) by ECMWF.

ECMWF Forecast Model Documentation Manual, Vols. I and II, 1981. Published by ECMWF.