TECHNICAL MEMORANDUM

97

Numerical prediction: some results from operational forecasting at ECMWF

A.J. Simmons

Research Department

February 1985

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

CONTENTS

1. INTRODUCTION

Prediction of atmospheric behaviour using comprehensive numerical models may be considered as falling into three categories. One is the 'deterministic' prediction of the instantaneous state of the atmosphere for as many days or weeks ahead as may be possible. Such predictions are generally referred to as 'short-range' or 'medium-range' weather forecasts. The other two categories form what can be regarded as climate predictions of the first and second kinds (Lorenz, 1975). The first kind is the prediction of statistical properties of the atmosphere, for example temporal or spatial means, for time ranges beyond the limit of deterministic predictability. Initial conditions in the atmosphere (and in the ocean and at the earth's land surface) may be in general as important in this case as they are in deterministic prediction, although different aspects of the initial state may be crucial. Climate prediction of the second kind concerns the prediction of the long-term impact of prescribed changes in the boundary conditions, composition or external forcing of the atmosphere. Statistics which are essentially independent of any experimental initial conditions are sought.

Results relating to the subject of large-scale anomalies of the general circulation have been obtained from all three categories of prediction, although in this contribution emphasis will be placed on practical experience with deterministic weather forecasting, for which a considerable body of data exists. Long-range prediction by numerical methods is at a much less developed stage than short- and medium-range prediction, although encouraging first results have been obtained from specific attempts at prediction (e.g. Miyakoda et al., 1983), and from study of the potential for skilful prediction (Shukla, 1981). In addition, climate studies too numerous to be referenced individually have examined the long-term response to anomalous boundary forcing, for example anomalous distributions of sea-surface temperature, and have thereby also contributed towards our understanding of the predictability of anomalies on a monthly time-scale and beyond.

Routine operational forecasting for a period up to 10 days ahead, using a global prediction system, began at the European Centre for Medium Range Weather Forecasts (ECMWF) on 1 August 1979. Archiving of both initial analyses and forecasts began one month later. Thus at the time of writing there exists an almost 5-year record of forecast results which can be examined to shed light on a number of questions concerning atmospheric predictability on a time-range of 10 days and beyond. The record may be used, for example, to define present levels of forecast accuracy, including its geographical and temporal variability, and dependence on the nature of the prevailing synoptic situation. Trends in predictive skill can be identified, and estimates made of the scope for further improvements. Information relevant to developing a capability for the accurate prediction of longer-term anomalies may be gathered from the treatment of such anomalies over the 10-day forecast period, and the general evolution of forecast error can be suggestive of some of the causes of this error, which may aid the improvement of numerical models used for all types of atmospheric prediction.

Some results from studies of the above type will be presented in this paper. Operational experience from the first two years of medium-range weather prediction at ECMWF has been summarized by Bengtsson and Simmons (1983), and much of what is reported here is in agreement with the earlier results. Although the discussion of forecast skill will not be restricted to cases of large-scale anomalies and blocking, particular attention will be devoted to these features.

The following section contains a short description of the ECMWF forecasting system, including the developments that have taken place over the past five years. Methods of assessment of the forecasts are then discussed. There follows a section in which the current overall level of forecast accuracy is summarized, and some specific discussion of the prediction of blocking and cut-off lows is given in Section 5. Section 6 then deals with the development of forecast skill, both that actually achieved in recent years and that expected in the future. The seventh section deals with the ability of the forecast model to maintain anomalous features of the general circulation present in the initial conditions over a monthly time-scale, and the overall question of systematic errors is briefly discussed in Section 8. This is followed by some concluding remarks.

2. THE ECMWF FORECASTING SYSTEM

2.1 The data assimilation

The ECMWF data assimilation system uses a three-dimensional multi-variate analysis scheme, described by Lorenc (1981), to produce initial analyses of geopotential and wind fields for the forecast model, and a correction method for the analysis of humidity. Analyses are produced every 6 hours, with the forecast model providing the necessary first guess for the analysis, which in turn provides the initial state for the 6-hour forecasting step required to produce the first guess for the following analysis. Observations are analyzed at 15 standard pressure levels from 1000 to 10 mb, and first guess data are interpolated (or extrapolated) to these levels from the terrain-following coordinate surfaces of the forecast model. In the version of the data assimilation system first introduced into operational use, complete pressure-level fields from the analysis were similarly interpolated back to the model levels, prior to non-linear normal-mode initialization (Temperton and Williamson, 1981; Williamson and Temperton, 1981).

Although the basic nature of the data assimilation has not changed over five years of operational implementation there have been numerous revisions. Perhaps the most significant of these have been the introduction late in 1980 of the interpolation from pressure to model levels of the difference between the analyzed and first-guess fields, rather than the full analyzed fields, and a recent substantial change in data control and selection, and in optimum interpolation statistics. The former change was such as to preserve the model boundary-layer structure through the analysis, since the lowest analyzed pressure levels of 1000 and 850 mb are insufficient to define such structure. The more recent change is the result of several years experience with the operation of the original scheme, drawing on the general accumulation of statistics of the performance of the scheme. Many of the other revisions have been individually minor ones concerned with the use of data, but changes worthy of note (particularly from the viewpoint of tropical prediction) are the introduction of a distinction between temperature and virtual temperature in November 1980, a revised vertical interpolation of humidity the following November, use of sea-surface temperature analyses produced by the National Meteorological Center, USA, in July 1982 and the introduction two months later of diabatic forcing of lower frequency gravity-wave modes in the initialization. An analysis of soil moisture and snow depth began in November 1983. In addition to such changes, the data analysis has also benefitted from the general development of the forecast model.

2.2 The forecast model

A major change in the numerical formulation of the forecast model has taken place over the five-year period considered here. The version chosen for the first phase of operational forecasting used a potential-enstrophy conserving finite-difference scheme for a staggered (C) grid, a sigma coordinate for the vertical representation, and a semi-implicit time scheme for the treatment of gravity-wave terms. A resolution of 1.875° in latitude and longitude, with 15 levels in the vertical, was adopted. Details have been given by Burridge and Haseler (1977) and Burridge (1979). Changes in the horizontal diffusion and the introduction of a more realistic representation of the orography and coastlines took place within the first two years of operational use.

The major change took place in April, 1983. The new version uses a spectral formulation in the horizontal, with triangular truncation at total wavenumber 63, a vertical coordinate (with 16-level resolution) that is terrain-following at low-levels but which reduces to pressure in the stratosphere, and a modified, more efficient, time-stepping scheme. The operational change to this version was accompanied by a change to a higher 'envelope' orography in the model. Papers relating to these changes include those of Girard and Jarraud (1982), Simmons and Burridge (1981), Simmons and Jarraud (1984), Simmons and Strüfing (1983) and Wallace et al. (1983). In the light of operational experience, minor adjustments of the orography, horizontal diffusion and time scheme have subsequently taken place.

The parameterization schemes described by Tiedtke et al. (1979) have been used with both versions of the operational model. They include a convection scheme following Kuo (1974), a stability-dependent representation of boundary and free-atmospheric turbulent fluxes (Louis, 1979), and a radiation scheme (Geleyn and Hollingsworth, 1979) which includes interaction with model-generated clouds. A number of minor adjustments of the parameterizations have taken place during operational use, but the most noteworthy change was the introduction of a diurnal radiative cycle in May, 1984. Other more substantial changes are at an advanced stage of testing and will be mentioned briefly in Section 9.

3. METHODS OF ASSESSMENT

For much of this paper attention will be concentrated on the results of objective verifications of forecasts carried out at ECMWF, using methods discussed by Arpe (1980) and Nieminen (1983). Results will be presented mostly for the anomaly correlation of the height field, the correlation between observed and predicted deviations from climatology of the height field at one or more levels in the atmosphere, evaluated over a certain region of the globe. This has generally been found to give a reliable indication of overall forecast skill, although as with any method of assessment care must be taken in interpreting results from very limited numbers of forecasts. Nieminen (loc.cit.) illustrates how within a particular season a good agreement is found between use of anomaly correlation and standard deviation of forecast error as measures of forecast accuracy.

Forecasts are also routinely subjected to synoptic assessment, particularly over the European area by the Member States of ECMWF. Results of one such assessment, of day 5½ forecasts of 500 mb height from January 1981 to July 1982, are compared with an assessment using anomaly correlations in Fig. 1. The subjective marking was carried out for forecasts for the neighbourhood of Scandinavia by the Swedish Meteorological and Hydrological Institute, while anomaly correlations were computed for a larger European area. The results indicate that an average subjective classification of the 5½ day forecasts is 'useful' and that this corresponds approximately to a 0.6 value for the anomaly correlation coefficient, in agreement with the previous use of 0.6 as a limit of useful predictability (e.g. Hollingsworth et al., 1980). Keeping in mind the different verification areas, it also appears from Figure 1 that the subjective and objective assessments are in broad agreement in respect of differences in forecast quality from month to month.

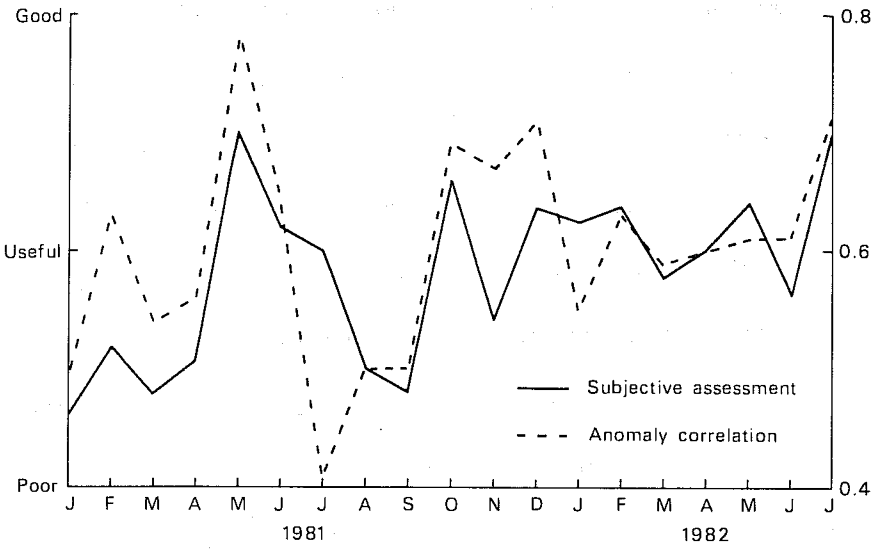


Fig. 1 Monthly-mean assessments of day 5½ 500 mb height forecasts. The solid line denotes results from a subjective assessment for the Scandinavian region, with averages computed from individual scores given on a scale of 1 to 5, with 4 representing a "good" forecast, 3 a "useful" forecast and 2 a "poor" forecast. The dashed line denotes means of anomaly correlations computed over a European area (12°W-42°E, 36°N-72°N).

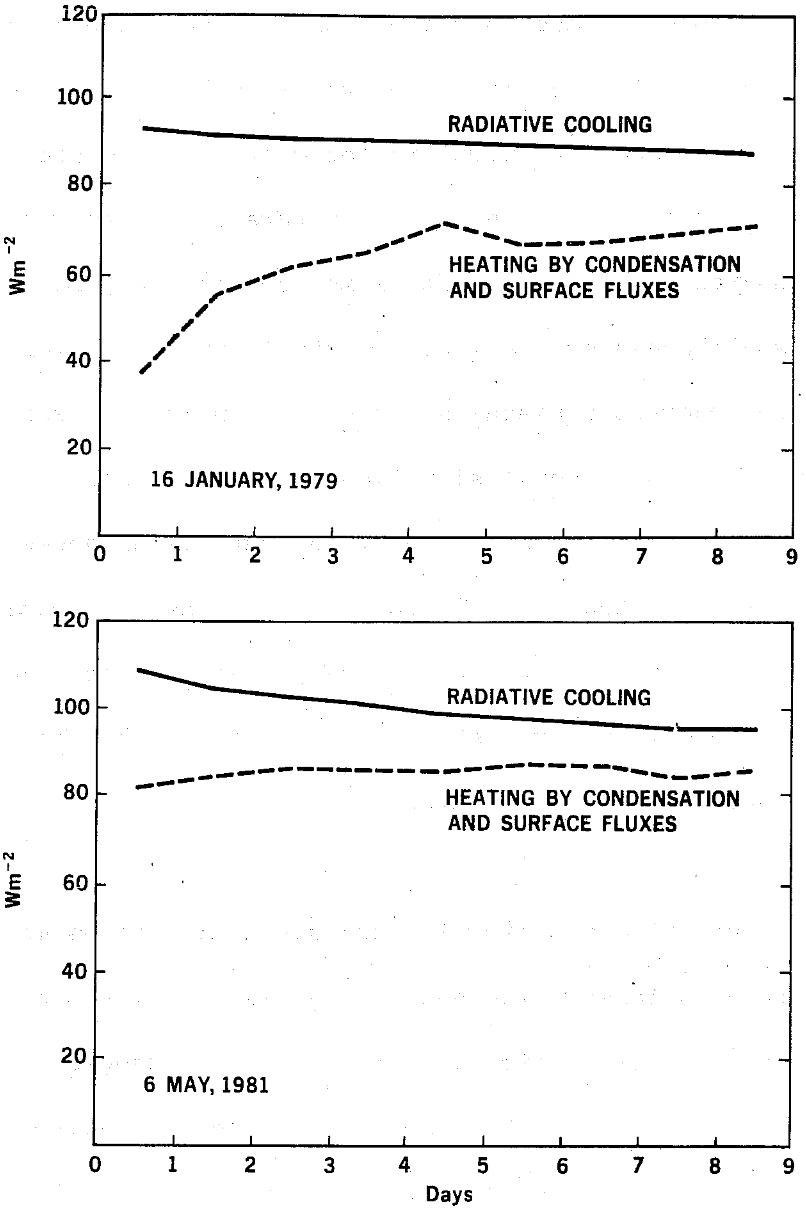


Fig. 2 Aspects of the global mean energy budget of the model atmosphere as functions of forecast range for two individual forecasts.

Despite such agreement it must nevertheless be recognized that judgement of the quality of particular forecasts, or spells of forecasts, can depend sensitively on the place for which the forecast is to be made, on the prevailing synoptic situation and on the nature of the forecast error. For example, medium-range predictions for some locations may be particularly sensitive to small phase errors in quasi-stationary situations involving meridional flow, whereas a useful medium-range outlook of weather type could be given in predominantly zonal situations for which there might be phase errors in the prediction of individual transient disturbances. Conversely, the forecasts may be rated badly in zonal situations at locations where systematic error in the latitude of the jet stream causes persistent biases in the track of cyclones. General statements concerning forecast skill must therefore be interpreted with caution.

In addition to objective verification and subjective assessment of forecasts, the performance of a forecasting system can also be monitored from the viewpoint of its overall physical consistency, for example through study of how the budgets of heat, momentum, moisture or energy evolve through the forecast period. Such assessment has also been carried out, partly on a routine basis and partly through case studies, for the forecasts discussed in this paper. Results will not be considered in any depth here, but by way of example Fig. 2 shows the evolution in time of the global-mean radiative cooling of the atmosphere of the forecast model, and the corresponding global-mean heating resulting from condensation and surface fluxes of sensible heat, for two particular forecasts. The upper plot is not for an operational forecast, but one carried out from FGGE data using a version of the forecasting system similar to that used at the beginning of operational prediction. Obvious deficiencies are an overall imbalance, which gave rise to a mean cooling of the model as the forecast proceeded, and a particularly deficient heating in the early stages of the forecast. Changes in the forecasting system noted in the preceding section were such as to reduce considerably this 'spin-up', as seen in the operational forecast for 6 May, 1981, shown in the lower plot of Fig. 2. The net cooling, though reduced, is still evident in this lower plot. The extent of this imbalance has been reduced further in more recent operational forecasts.

4. THE ACCURACY OF FORECASTS IN THE MEDIUM RANGE

Anomaly correlations calculated daily using height fields at standard pressure levels from 1000 to 200 mb for the extratropical Northern Hemisphere, and averaged over the year of 1983, are shown in the upper plot of Fig. 3 for the 10-day forecast range. Representing the mean over a large number of forecasts the plot shows a monotonic decline in forecast accuracy with increasing range, with the correlation coefficient of 0.6 reached after about 5½ days. A figure of 5.8 days is found for the 500 mb level alone, while accuracy is somewhat less at 1000 mb, the 0.6 level being reached at day 5.3.

The lower plot of Fig. 3 shows corresponding correlations computed separately for three groups of zonal wavenumber. Accuracy is evidently highest for the longest zonal scales, and there is a rapid growth of error in the shorter synoptic scales, here represented by zonal wavenumbers 10 to 20. Data for the winter of 1980/81 analysed by Kalnay (personal communication) shows that viewed in terms of the error variance of total wavenumber, n, rather than zonal wavenumber, error in scales with $n > 15$ on average reaches saturation within the first five days of the forecast. Synoptic analysis reveals both cases in which the large-scale pattern is well predicted despite error in the forecast of shorter transient disturbances, and cases in which the erroneous forecast of a small-scale feature is followed by deterioration of the forecast over a much larger area.

Anomaly correlations as in the upper plot of Fig. 3, but averaged only over the winter months of January, February and December or the summer months of June, July and August are shown in the upper panel of Fig. 4. Although seasonal differences are probably exaggerated by an unusually good performance in January 1983 which will be discussed later, these and other results (e.g. Bengtsson and Simmons, 1983) reveal a generally more accurate performance, as measured by anomaly correlation, in summer. In absolute terms, the growth of error variance is smaller in summer than winter, but when normalized by a measure of the actual atmospheric variability, for example the error variance of a persistence forecast (as shown for individual months in a later figure), the relative inaccuracy of the summer forecasts again becomes evident. A smaller absolute rate of error growth in summer is consistent with smaller growth rates for baroclinic instability in the weaker summer flow; relative inaccuracy may arise from a tendency for shorter scales to be more prevalent in summer and a greater sensitivity of summer circulation systems to processes requiring parameterization in the forecast model.

The accuracy of forecasts for the extratropical Northern and Southern Hemispheres, again for the year 1983, is also compared in Fig. 4. A clear difference in predictability is found, amounting to about 1½ days at a correlation of 0.6. Although some differences between the hemispheres could arise from interhemispheric differences in the nature of the general circulation, the rapid initial growth of error shown in Fig. 4 for the Southern Hemisphere is indicative of a much larger error in the determination of the initial state for this hemisphere. This is not surprising in view of the sparsity of some types of observational data in the Southern Hemisphere, and generally more accurate forecasts for this region have been reported using the enhanced data coverage of the FGGE year (Bengtsson, et al. 1982).

Verification statistics have also been accumulated routinely for smaller areas of the globe. In Table 1, predictability as measured by the time at which the monthly-mean anomaly correlation of 500 mb height reaches the value of 0.6 is shown for January and July of each operational year for three regions of the extratropical Northern Hemisphere. Such regional results show much more variability from month to month and year to year than those for the hemisphere

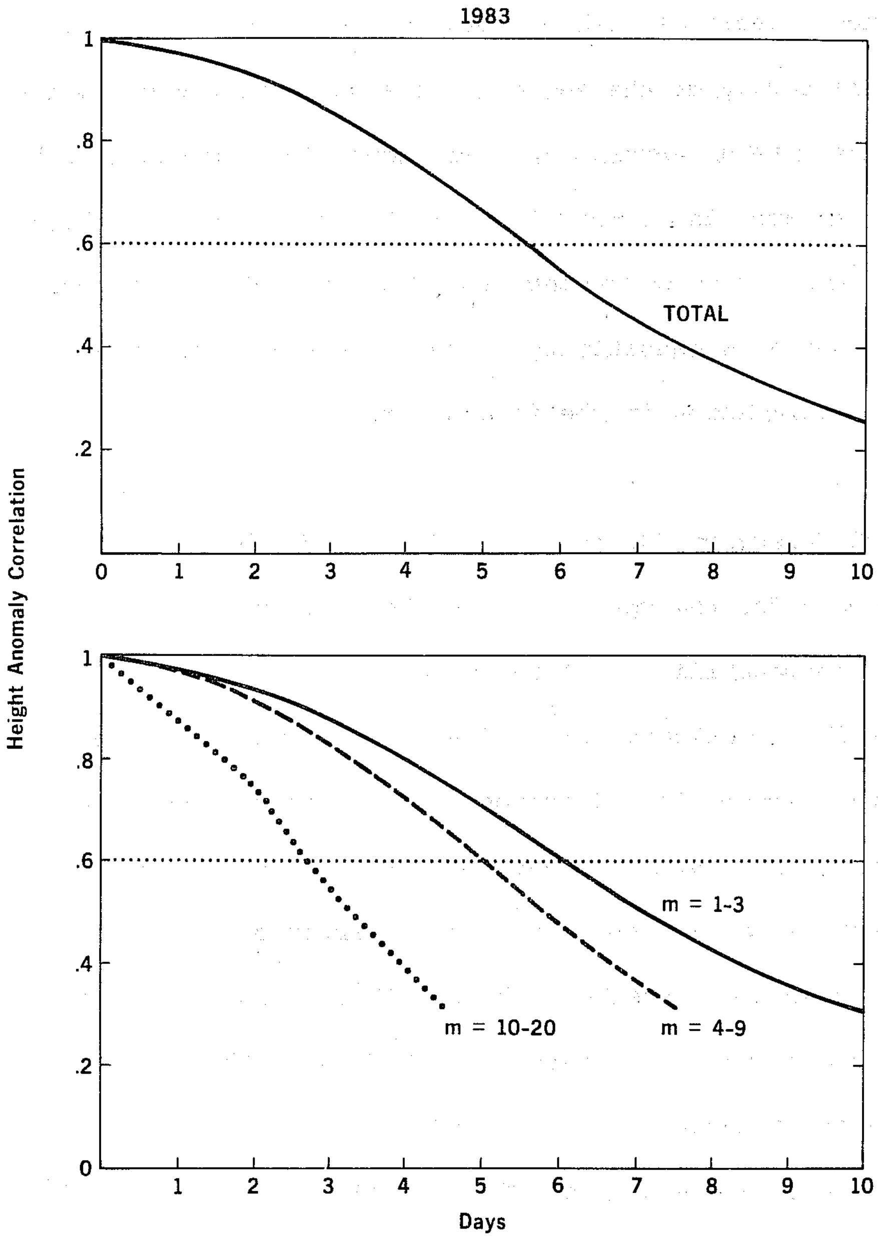


Fig. 3 Anomaly correlations of 1000-200 mb height for the area from 20°N to 82.5°N averaged over all operational forecasts carried out in 1983. Values are plotted as a function of forecast range. The upper panel is for the total fields, and the lower panel shows results for different groups of zonal wavenumber m.

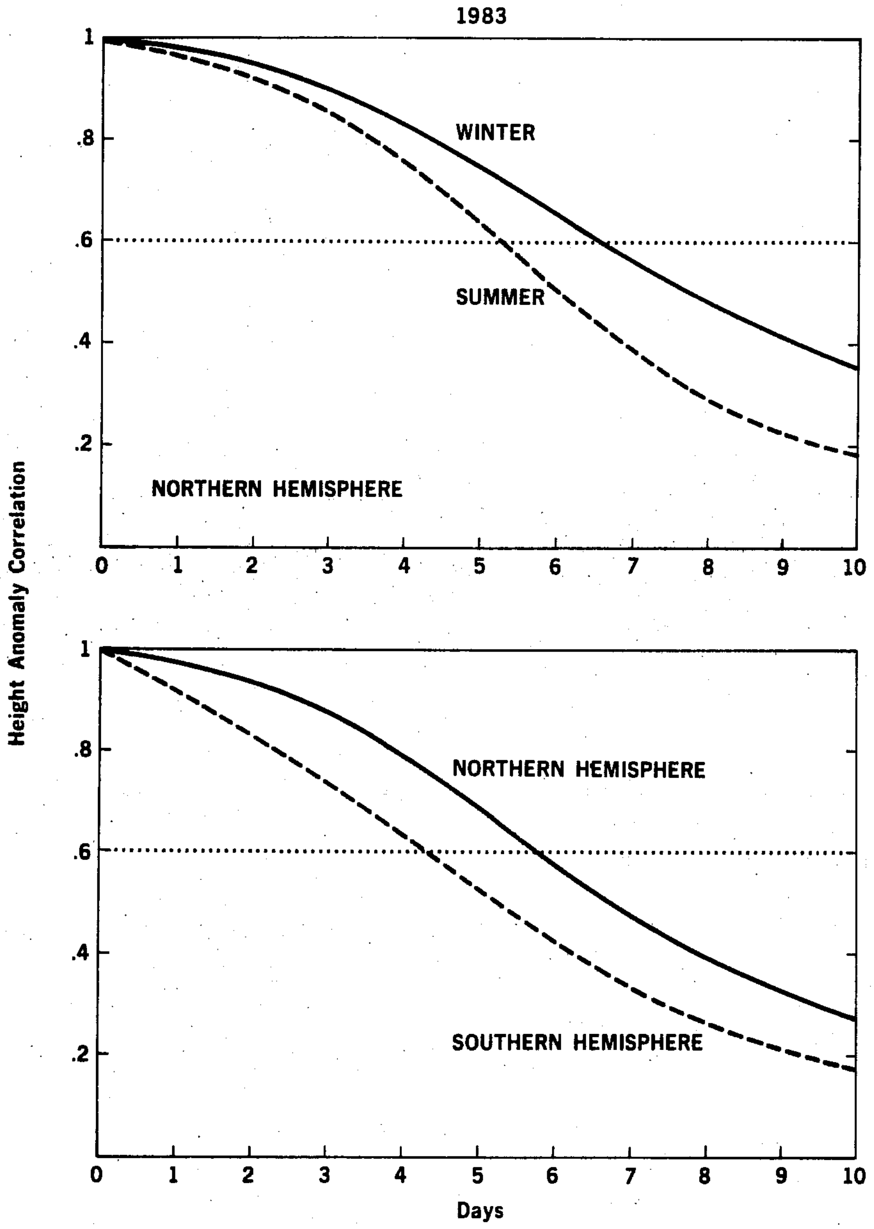


Fig. 4 Height anomaly correlations for 1983 as functions of forecast range for:

Upper - Means for January, February and December (marked winter) and June, July and August (marked summer) for the 1000-200 mb levels of the extratropical Northern Hemisphere.

Lower - Annual means for the extratropical Northern and Southern Hemispheres for the 500 mb level.

MONTH	EUROPE (36°-72°N; 12°W-42°E)	E. ASIA (24°-60°N; 102°-150°E)	N. AMERICA (24°-60°N; 120°-72°W)
Jan 1980	4.4	5.4	5.8
July 1980	5.5	5.7	4.0
Jan 1981	4.8	6.6	7.0
July 1981	4.1	5.2	3.9
Jan 1982	5.2	5.9	4.6
July 1982	6.4	6.1	4.7
Jan 1983	8.6	5.7	5.6
July 1983	5.3	3.8	5.2
Jan 1984	5.8	6.4	4.8
July 1984	5.7	5.5	5.4

Table 1 The forecast ranges (days) at which monthly-mean anomaly correlations of 500 mb height reach a value of 0.6 over different continental regions of the Northern Hemisphere.

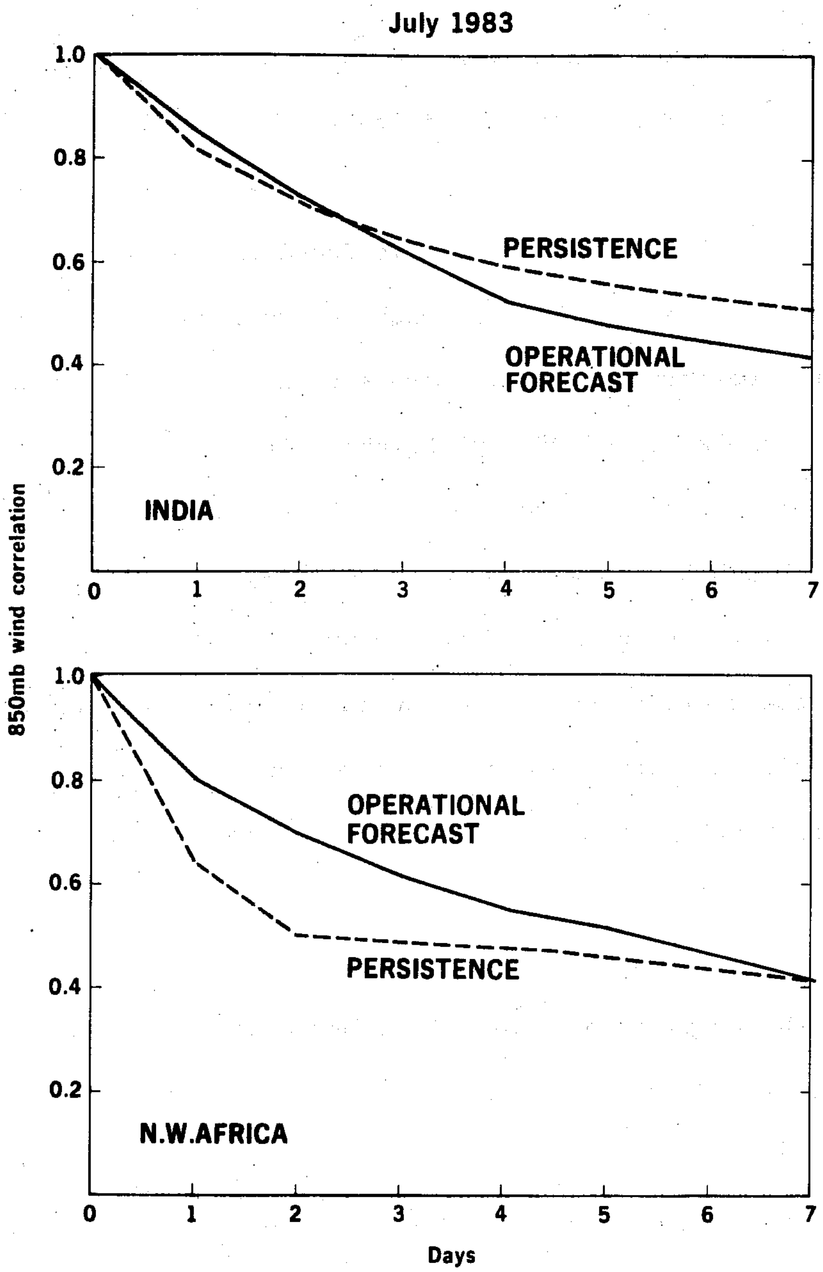


Fig. 5 Monthly means of the absolute correlation of 850 mb vector wind for July, 1983 as functions of forecast range. Results are shown for both operational numerical forecasts and persistence forecasts for two regions (6°-33°N, 72°-102°E, upper; 12°-36°N, 21°W-12°E, lower) spanning the tropics and subtropics.

as a whole, and general conclusions are difficult to draw. On the basis of the five-year record the winter forecasts are not found to be clearly worse over Europe than over the other continental regions, in contrast to the indication discussed by Bengtsson and Simmons (1983) on the basis of the first two years of operational prediction. An overall tendency for more accurate summer predictions at 500 mb over Europe is found from examination of Table 1 and from results for other months, although the more northerly latitude of this region should be borne in mind. More substantial differences are found for the anomaly correlation of 1000 mb height, with the 0.6 correlation coefficient reached on average about one day later in the forecast period over Europe than over North America. However, the extent of these differences is not confirmed by examination of another measure of skill, the standard deviation of forecast error, normalized by persistence, for the 850 mb wind. A definitive specification of regional differences in forecast accuracy over the extratropical Northern Hemisphere is thus difficult to give on the basis of such objective verification.

Forecasting for the tropics and immediate subtropics undoubtedly poses particular problems. There is not only in general a severe deficiency in data coverage, but also strong sensitivity to aspects of the data assimilation and parameterization schemes of the forecasting system. In addition, the strong persistence of some regional circulations emphasises difficulties. As an example, the upper panel of Fig. 5 shows the absolute correlation of the 850 mb vector wind for a region covering India and part of southeast Asia for July, 1983. Overall, the numerical forecasts barely improve on a persistence forecast for the summer monsoon flow, although individual cases of quite accurate forecasts of transient behaviour over several days can be found. The verification for Northwest Africa also shown in Fig. 5 indicates a much better performance relative to persistence, although the absolute correlation of the forecast evolves similarly for the two regions. Some further discussion of tropical forecast accuracy is given in Section 6.

The accuracy of forecasts for a given month or season can vary substantially from year to year even on a hemispheric domain. Extreme cases for the extatropical Northern Hemisphere, comparing anomaly correlations for January 1982 and 1983 and for July 1983 and 1984, are presented in Fig. 6. Although changes in the forecasting system over the two intervening one-year periods may have contributed to the differences, particularly between 1983 and 1984, the tests of the various changes that were made indicate that most of the differences seen in Fig. 6 are not due to development of the model or data assimilation. Rather, differences in predictability of up to two days at the 0.6 level of anomaly correlation appear to be mostly a consequence of differences in the circulation patterns from month to month. The extent to which this is due to a fundamentally higher predictability, due to a greater sensitivity to data coverage or analysis techniques, or due to a greater sensitivity to systematic model deficiencies, in certain synoptic situations is unclear.

The results shown in Fig. 6 are confirmed by corresponding plots of the standard deviation of forecast error. However, it is also found that standard deviations for persistence forecasts show some of the same variability. In Fig. 7 we plot a measure of forecast accuracy based on the standard deviation of the model forecast normalized by the standard deviation of persistence, for the same months as in Fig. 6. For the two Januaries there appears little difference in forecast quality out to day 4, according to this measure, although the superiority of 1983 still becomes evident in the second half of the forecast range. July 1984 appears clearly better than July 1983 beyond day 2. Thus although differences are not quite so pronounced for a measure normalized by the skill of a persistence forecast, they are nevertheless substantial in the medium range.

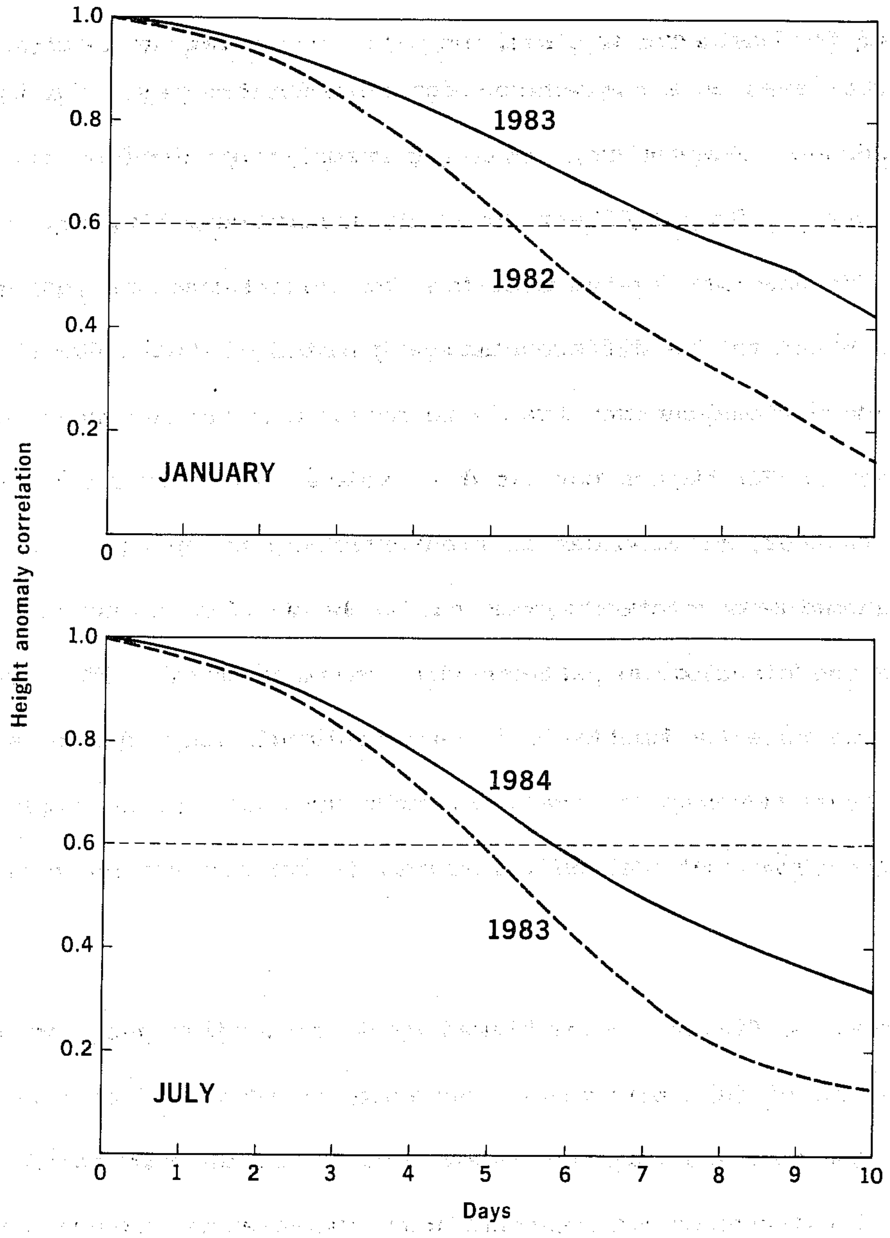


Fig. 6 Monthly-mean anomaly correlations of height for 1000-200 mb and the extratropical Northern Hemisphere as functions of forecast range for January 1982 and 1983, and July 1983 and 1984.

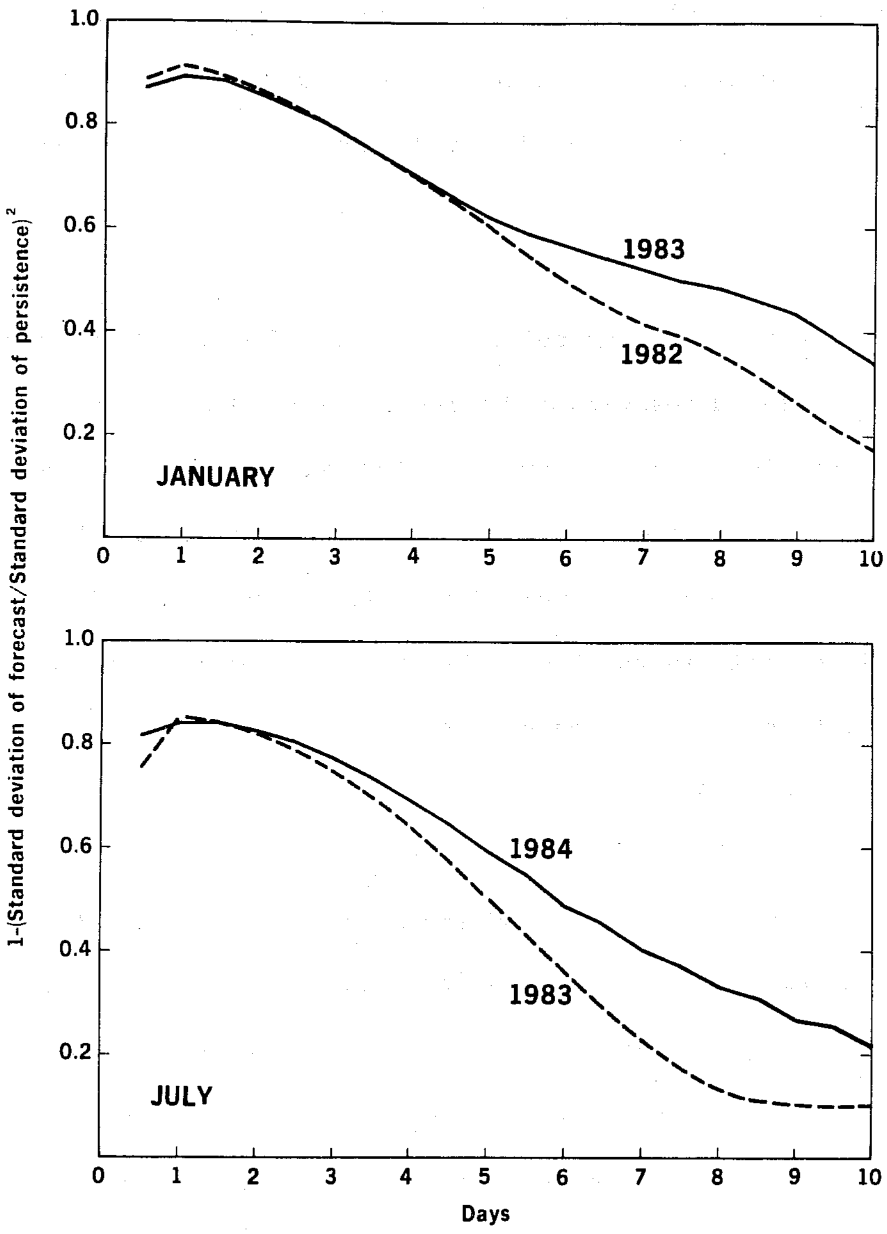


Fig. 7 As Fig. 6, but for a measure of forecast accuracy based on standard deviation normalized by persistence.

5. THE PREDICTION OF BLOCKING AND CUT-OFF LOWS

Spells of above and below average predictability are also found within a monthly time scale. Examples have been discussed by Bengtsson and Simmons (1983), and here we present a striking recent case, November 1983, as a basis for choice of specific examples of forecasts involving blocking and the formation of cut-off lows.

Figure 8 shows height anomaly correlations for 3-, 7- and 10-day forecasts performed from initial dates within November, 1983. Particularly evident in the 7-day forecasts is a slow variation in forecast skill over the course of the month, with below average performance at the beginning of the month and in the final week, and an accurate spell during the second and third weeks. This variation is discernible in the 3-day forecasts, and more so at day 10. Small daily variations in the accuracy of individual forecasts at day 7 are seen to be amplified at day 10.

To illustrate the range of accuracy of large-scale forecasts within this month, 5-day mean maps of 500 mb height from analyses and two forecasts are shown in Fig. 9. The forecast from the 11th shown in the left-hand column is classified as the most accurate of the month according to the anomaly correlations at days 7 and 10 shown in Fig. 8. The right-hand forecast plot shows the result from the forecast from the 26th, which ranks as the poorest on the hemispheric scale.

The forecast from November 11 is successful in its representation of most large-scale features present over the final 5 days of the forecast period. It has captured the retrogression of the block initially located over western Europe, the enhancement of the closed low to the north east, and the decay of the low over northern Canada. Conversely, major error is seen in the forecast from November 26. In particular its evolution has produced a predominant zonal wavenumber two pattern in the high-latitude height field rather than the strong wavenumber three pattern which occurred in reality. The forecast is successful in its formation of a cut-off low in the European sector, but for practical application suffers from an important error in the position of this low, and in the eastward extension of the zonal jet over the Atlantic.

It is of interest from the viewpoint of systematic model errors (to be discussed further in Section 8) to examine in further detail the development of the latter features of the forecast from 26 November. Figure 10 shows the initial 500 mb height analysis for the European/Atlantic region, together with the corresponding maps for the 5-day forecast and its verifying analysis. At the 5-day range a pronounced ridge extends over western Europe, and the cut-off low has just formed to the southeast in both the forecast and reality, although it is already seen to be located further east in the forecast. To the west, the forecast exhibits a characteristic error in that the trough over the eastern Atlantic exhibits a phase lag at southern latitudes, and fails to exhibit a weak cut-off west of the Iberian peninsula.

Figure 11 shows the 6-day forecast and verifying analysis. As in idealized studies of mature baroclinic waves (e.g. Simmons and Hoskins, 1978) the tilted eastern-Atlantic trough in the forecast has decayed and enhanced westerly flow in this sector, whereas in reality a weak cut-off remains, with high pressure over southern England. In addition, the major Mediterranean cut-off has drifted slightly eastward in the forecast, rather than becoming established over the south of Italy.

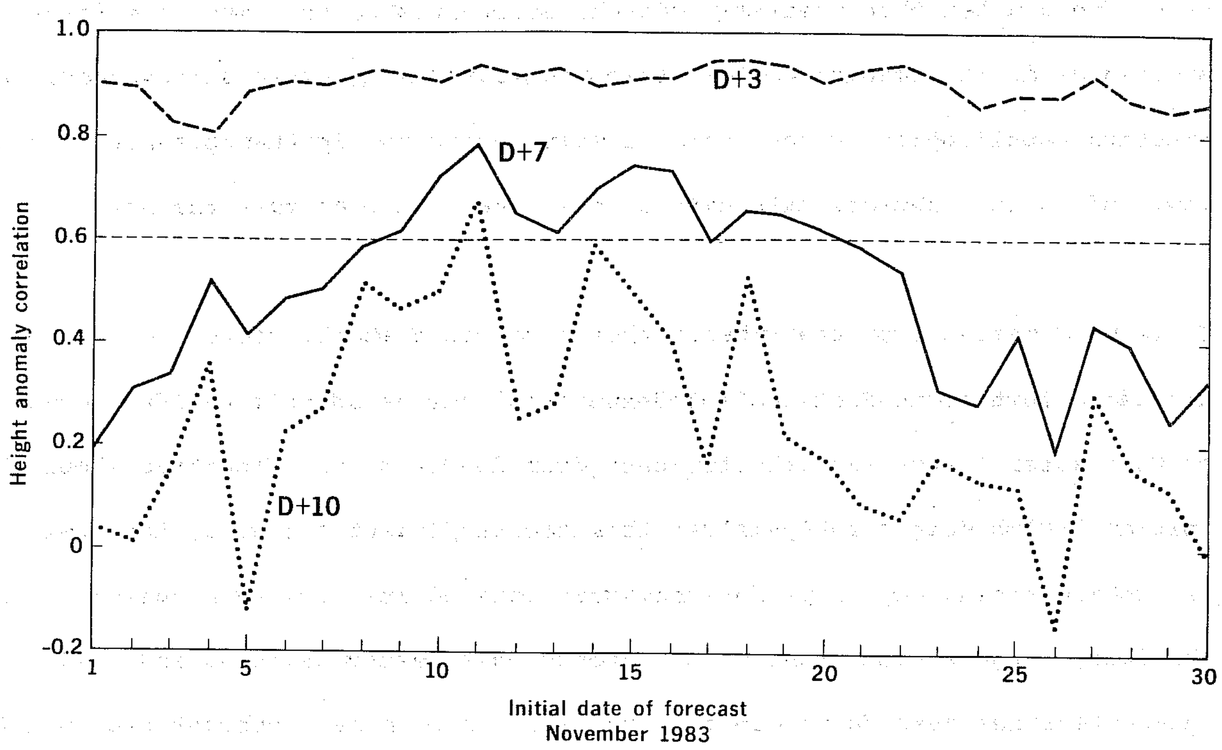
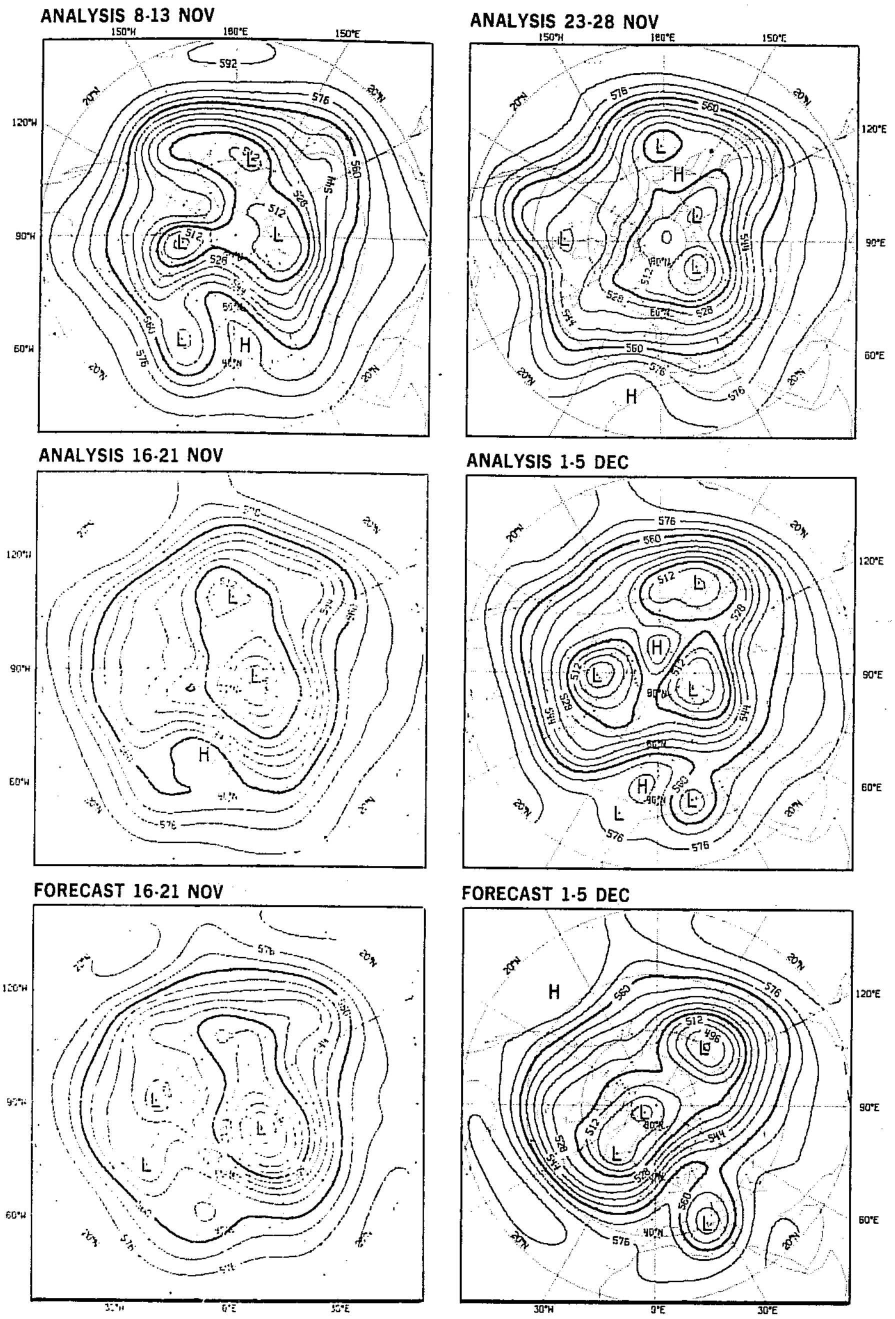


Fig. 8 Anomaly correlations of height for 1000-200 mb and the extratropical Northern Hemisphere for 3-, 7- and 10-day forecasts performed from initial dates within the month of November 1983.



ANALYSIS 8-13 NOV

ANALYSIS 23-28 NOV

ANALYSIS 16-21 NOV

ANALYSIS 1-5 DEC

FORECAST 16-21 NOV

FORECAST 1-5 DEC

Fig. 9 Mean 500 mb height analyses for the periods 8-13 November (upper left), 16-21 November (middle left), 23-28 November (upper right) and 1-5 December (middle right). Corresponding means for 16-21 November and 1-5 December are also shown for the forecasts from 11 and 26 November, respectively. The contour interval is 8 dam.

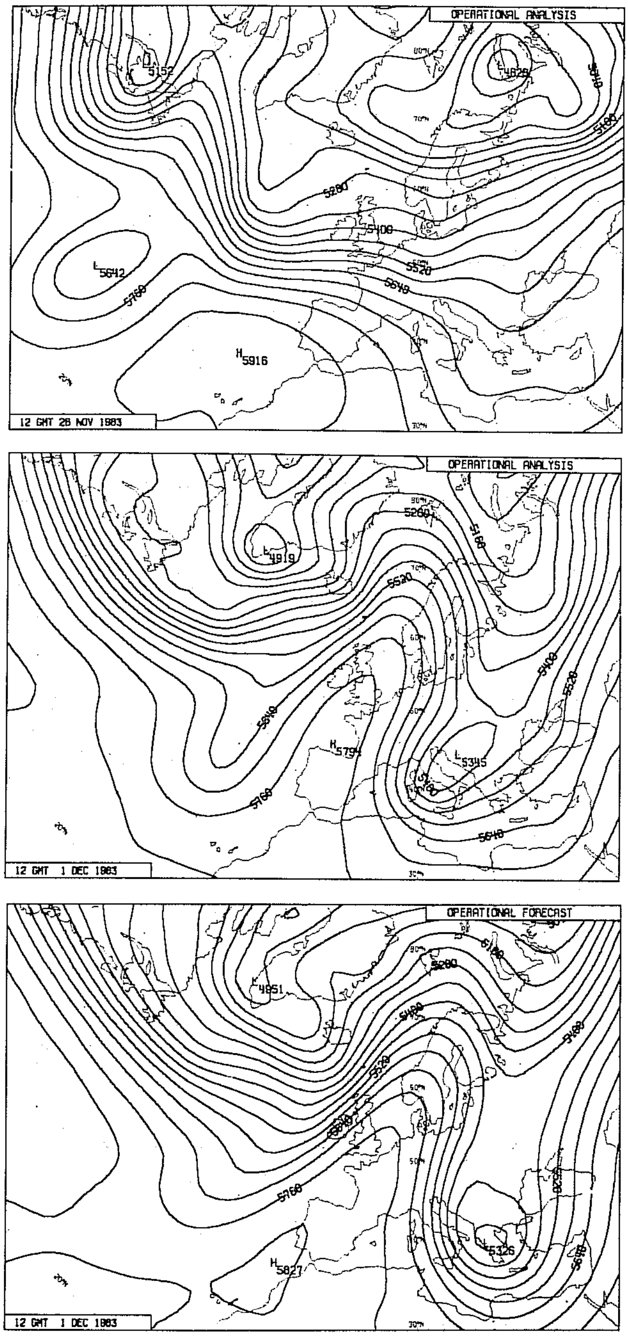


Fig. 10 500 mb height analyses for 12 GMT 26 November (upper) and 1 December (middle) and the 500 mb height field forecast for 1 December (lower) from initial conditions for 26 November. The contour interval is 6 dam.

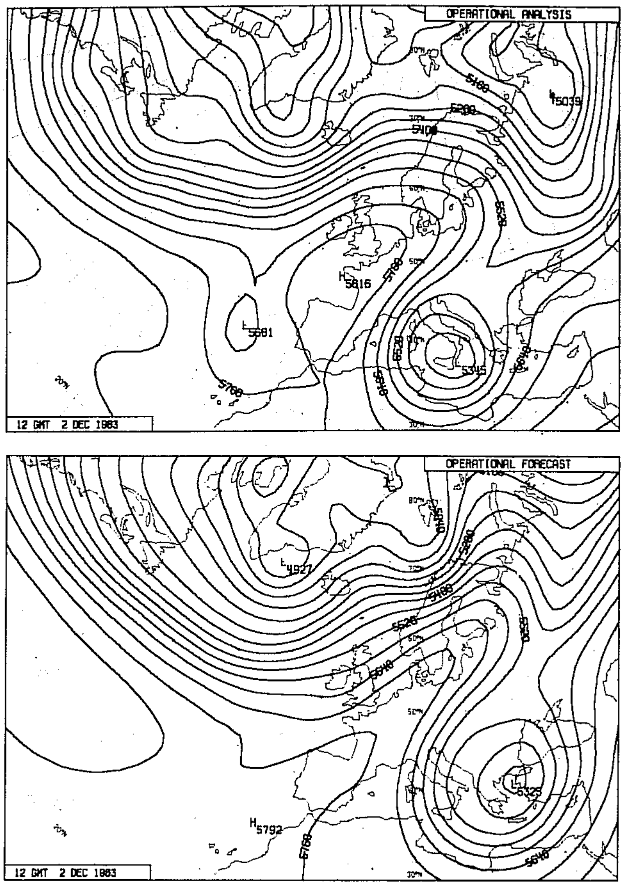


Fig. 11 The 500 mb height field for 2 December (upper) and the corresponding 6-day forecast (lower) from 26 November.

Grønaas (1983) has carried out an extensive study of forecasts for the European/Atlantic area for the years of 1980 and 1981. He identified 20 spells lasting 7 days or longer in which the anomaly correlation coefficient of 500 mb height for the European area indicated above or below average forecast quality. These spells were typed synoptically, and it was discovered that the best scores were found for situations characterized by some form of blocking, with poorest results in prevailing zonal situations. Mean scores for day 7 in blocking spells were comparable with those for day 5 in cases of zonal flow.

More specifically, Grønaas (loc.cit.) reports above average forecast scores when persistent large-scale synoptic features such as blocking and cut-off lows are present in the initial analyses, or predicted within the first three days of the forecast. The poorer performance in zonal situations reflects not only phase errors of travelling disturbances, but also errors (which have a systematic element) in the cyclone tracks and in the rate of filling of cyclones. The systematic errors appear to inhibit the formation of blocking later in the forecast period and give rise to a tendency for the cyclonic activity on the western side of a ridge or block to break down that feature. An indication of this has been discussed in respect of the forecast from 26 November, 1983.

Situations in which the development of blocking is accurately predicted may also be used to examine the mechanisms and interactions involved, and the features of the forecasting system which are of crucial importance. This can be achieved by controlled numerical experimentation, and case studies examining a range of sensitivities to such features as orography, model resolution and parameterizations have been reported by Bengtsson (1981), Ji and Tibaldi (1983), Tibaldi and Buzzi (1983) and Tibaldi and Ji (1983).

6. DEVELOPMENTS IN PREDICTIVE SKILL

Miyakoda et al. (1972) suggested that the results of their first comprehensive trial of 2-week predictions, using a hemispheric model, might be taken as a benchmark for future comparisons. Their ensemble-mean anomaly correlations of 500 mb height for the extratropical Northern Hemisphere based on 12 January cases taken from the years 1964 to 1969 are presented in Fig. 12, together with corresponding results from the operational ECMWF forecasts for the December to February period for the years 1982/83 and 1983/84. Included in the ECMWF results is the month of January 1983 which we have already shown to have been one of high predictability with the then operational grid-point model, but results with the new spectral model and envelope orography for January 1984 are only slightly poorer than achieved in the operational forecasts for the previous January, and Fig. 12 thus appears to be reasonably representative of current ECMWF skill for the winter months.

Although a number of other qualifying remarks could be made about the comparison presented in Fig. 12, it is nevertheless clear that a substantial improvement has taken place over the past 10-15 years in our ability to predict at least the larger scales of motion over the extratropical Northern Hemisphere. The improvement is confirmed by comparison of limited-area verifications of operational forecasts over Europe (Bengtsson, personal communication). Figure 12 shows a large increase in accuracy in the very early part of the forecast, suggesting an important contribution from a more accurate specification of the initial state. Overall, the more recent forecasts reach the 0.8 level of anomaly correlation almost 2½ days later, and they also exhibit a somewhat smaller error growth beyond this time, as measured either by anomaly correlation or by standard deviation. Differences between the two sets of forecasts amount to 2.8 days at the 0.6 level and 3.7 days at the 0.4 level of anomaly correlation. The mean standard deviation of the 500 mb height error reaches a value of about 130 m by day 10 of the

ECMWF forecasts; this level of error was reached soon after day 6 in the experiments of Miyakoda et al..

Improvements in predictive skill can also be identified within the shorter period over which operational forecasting has been carried out at ECMWF. This may be seen in Fig. 13, which shows times at which monthly-mean anomaly correlations of 500 mb and 1000 mb height reach the value of 0.6 for the two extratropical hemispheres, for the period from January, 1980 to July, 1984. Annual running means are also plotted. Results for both hemispheres exhibit generally higher correlations at 500 mb than at 1000 mb, and the variability from month to month noted previously for Northern Hemisphere. They indicate overall improvements of about ½ day in the 'useful predictability' attained by the forecasting system, with somewhat larger improvements at 1000 than at 500 mb. Such advances are of course not unique to the ECMWF system; Lange and Hellsten (1984) present results showing distinct improvements to have been achieved at the 3-day range by a number of operational forecasting centres.

Corresponding results for the tropical belt are presented in Fig. 14, in this case using the forecast range at which the absolute correlation of vector wind falls below 0.75 to indicate the improvements brought about in the forecasting system. The level of accuracy of wind forecasts achieved initially at the one-day range is seen to be reached at about day 1½ at 850 mb, and improvement is also evident at 200 mb. This indication is consistent with a reduction in root-mean square errors, which at 850 mb have fallen (in the annual mean) from a maximum of 4.9 ms^{-1} to a value of 4.2 ms^{-1} in the 48-hour forecasts, and from 3.7 to 3.2 ms^{-1} at the 24 hour range. Falls at 200 mb have been from 11.8 to 9.8 ms^{-1} at the 3-day range and from 10.5 to 8.5 ms^{-1} in the 2-day forecasts. Although the overall level of forecast skill in the tropics

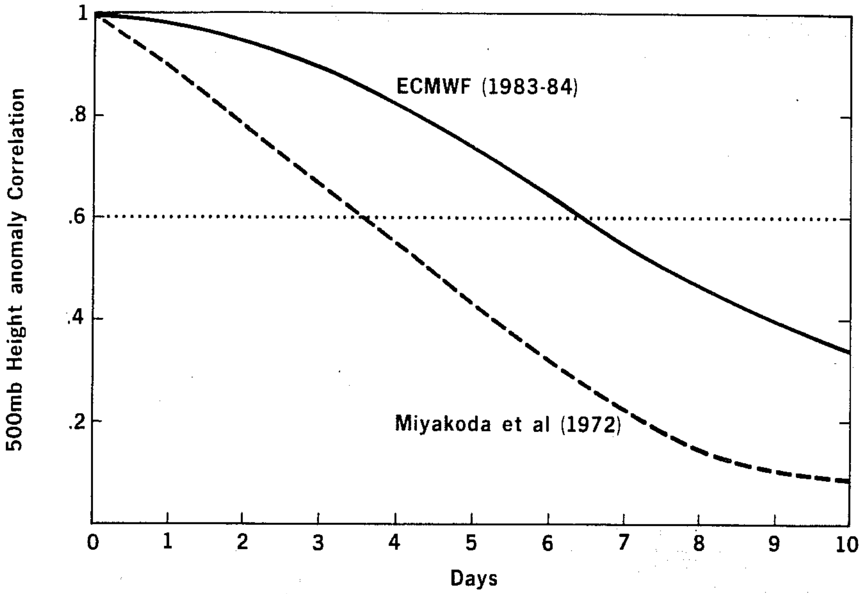
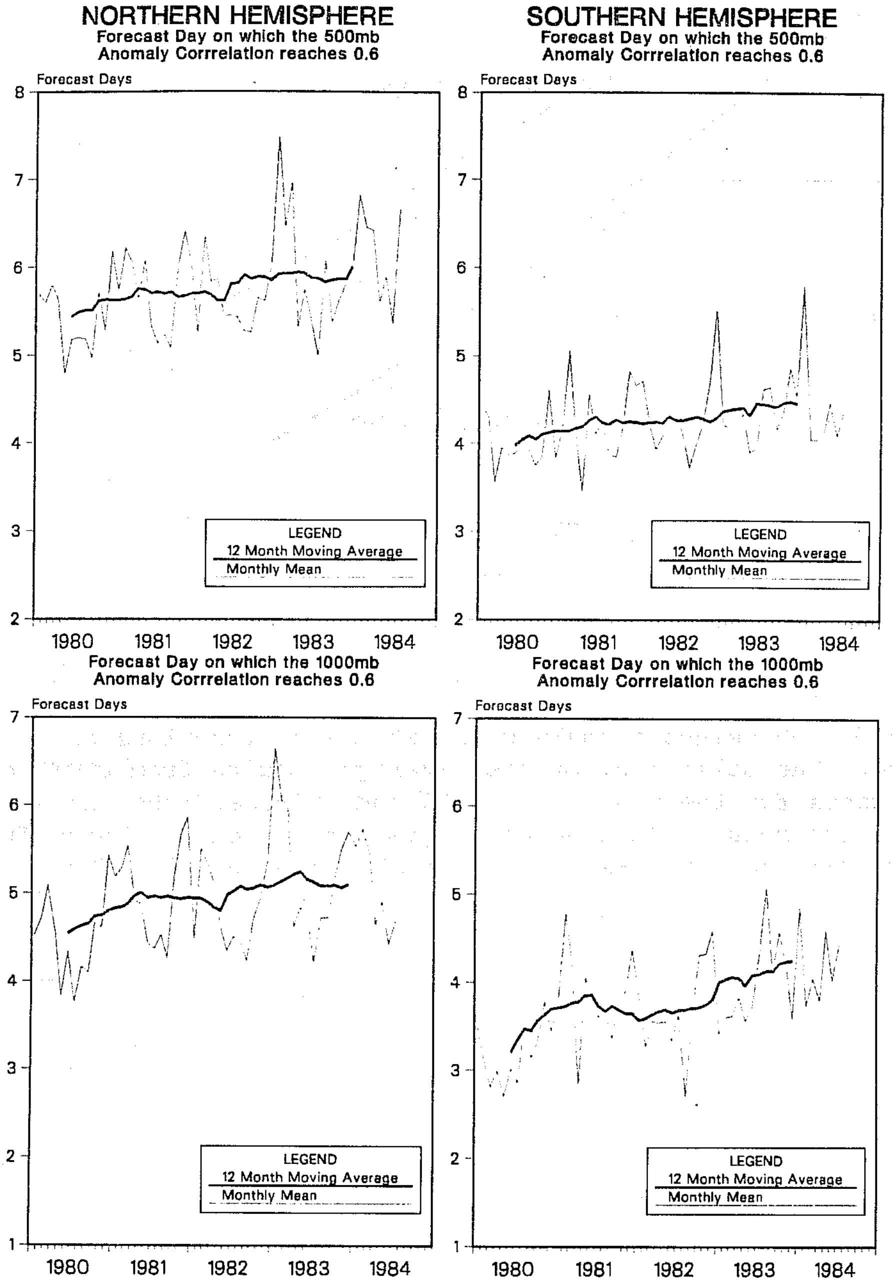


Fig. 12 Mean 500 mb height anomaly correlations as functions of forecast range. The solid line denotes average results from ECMWF operational forecasts for the winters of 1983 and 1984, and the dashed curves shows the mean of 12 forecasts from January cases chosen from the years 1964 to 1969, as reported by Miyakoda et al. (1972).


NORTHERN HEMISPHERE

Forecast Day on which the 500mb Anomaly Corrrelation reaches 0.6

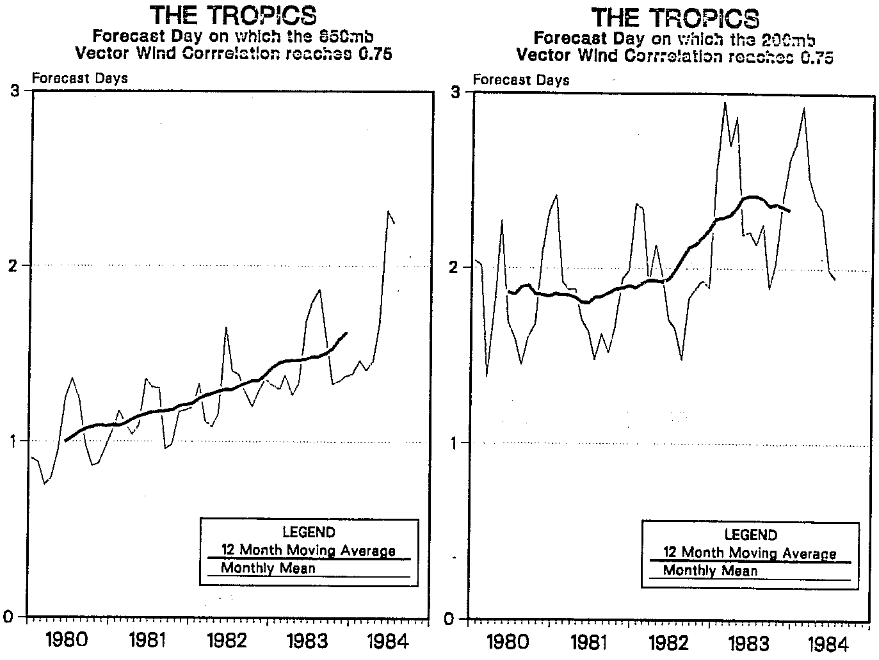
SOUTHERN HEMISPHERE

Forecast Day on which the 500mb Anomaly Corrrelation reaches 0.6

Forecast Day on which the 1000mb Anomaly Corrrelation reaches 0.6

Forecast Day on which the 1000mb Anomaly Corrrelation reaches 0.6

Fig. 13 The forecast ranges (days) at which monthly-mean height anomaly correlations reach a values of 0.6. Results are shown for the extratropical Northern (left) and Southern (right) Hemispheres, and for the 500 mb (upper) and 1000 mb (lower) levels. Both individual monthly values and 12-month running means are plotted.



THE TROPICS
Forecast Day on which the 850mb Vector Wind Corrrelation reaches 0.75

THE TROPICS
Forecast Day on which the 200mb Vector Wind Corrrelation reaches 0.75

Fig. 14 The forecast ranges (days) at which monthly-mean absolute correlations of 850 mb (left) and 200 mb (right) vector wind evaluated for the tropical belt fall below a value of 0.75.

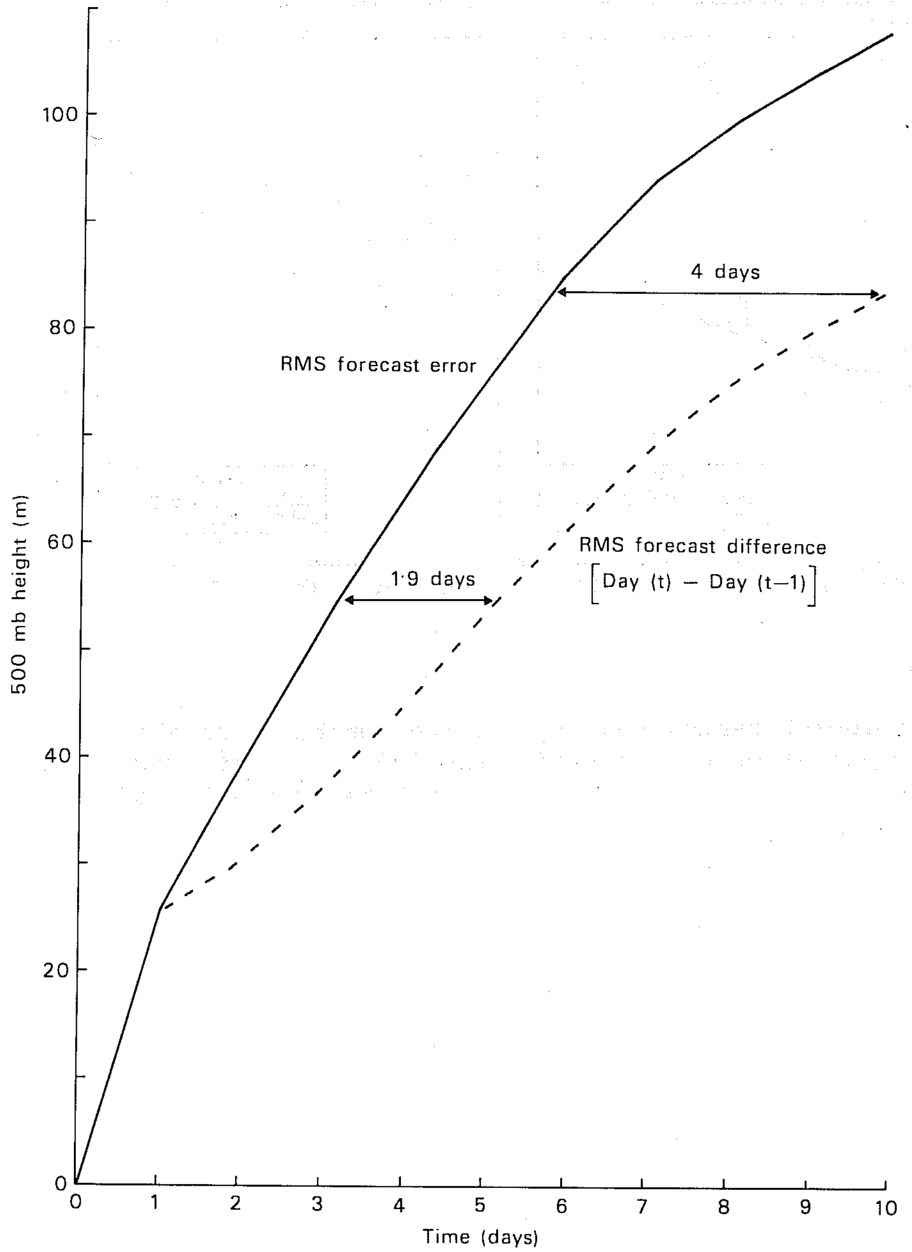


Fig. 15 The global root-mean-square error of 500 mb height (solid line) for all forecasts from 1 December 1980 to 10 March 1981, and the corresponding root-mean square difference between forecasts for a particular time starting from initial analyses one day apart, adapted from the work of Lorenz (1982).

remains far from satisfactory, it is encouraging that the performance of the forecasting system has shown itself to be sensitive to the various changes that have been made in the data assimilation and forecast model.

An indication of some of the potential for improving extratropical forecasts further has been derived recently by Lorenz (1982) from analysis of the performance of the ECMWF forecasting system. For a particular point in time the current analysis of the atmosphere and the one-day forecast from the day before provide two estimates of the actual atmospheric state. Thus the growth (as time t increses) of differences between a particular forecast for day t and the previous day's forecast for day (t+1) shows how initial, relatively small differences amplify in time. If this growth in the model occurs at a rate similar to that of real differences in the atmosphere, then it indicates a limit to the accuracy with which a forecast may be made, assuming the one-day forecast error to remain unchanged. Specific results such as shown in Fig. 15 suggest a maximum possible improvement of about 2 days at the level of predictability currently reached at day 3, and about 4 days at the level now reached at day 6. Further improvements would result from increased accuracy of the one-day forecast and Lorenz argues that halving the one-day root-mean-square error should add 2 days to the range of predictability. To put such a reduction in context, present ECMWF results for winter indicate a one-day error which is 40% of the value reported by Miyakoda et al. (1972). Thus an average limit of deterministic predictability of the order of two weeks or more is indicated, a result in general agreement with earlier estimates discussed by Charney et al. (1966) and Smagorinsky (1969).

7. THE REPRESENTATION OF MONTHLY-MEAN ANOMALIES

Some results of relevance to the longer-range prediction of large-scale anomalies may be obtained by examining the extent to which observed monthly-mean deviations from climatology are present in the day-10 forecast fields. For example, Bengtsson and Simmons (1983) illustrate how the mean of all ECMWF forecasts for 10 days ahead produced daily during July 1980 succeeded in representing the temperature anomaly associated with a major heat-wave and drought over south eastern North America. Such an anomaly was, of course, present in the initial conditions for the predictions, but the forecast model (which used a number of climatological initial surface conditions) was evidently capable of maintaining the strength of the anomaly over at least a 10 day period.

A brief examination of the treatment of Northern Hemispheric 500 mb height anomalies during the winter months of 1980/81 and 1981/82 has been carried out and some objective verification of day-10 forecasts is presented in Fig. 16. Since mean anomalies computed for a period of about 30 days change little when the averaging period is shifted by 10 days, persistence forecasts score highly in Fig. 16, and the challenge for the numerical model is in the first place to maintain initial anomalies over the 10-day forecast period, even though the ultimate requirement for long-range prediction models is to forecast the change in anomaly pattern from one month (or season) to the next. The ECMWF forecasting system is evidently partially successful in maintaining anomalies, achieving an average correlation of about 0.5 when monthly means of day-10 forecast fields are verified. The lower mean of about 0.25 obtained from averaging the day-10 correlations of individual forecasts reflects the inaccurate prediction by day 10 of transient disturbances which do not directly influence the verification of monthly-mean fields.

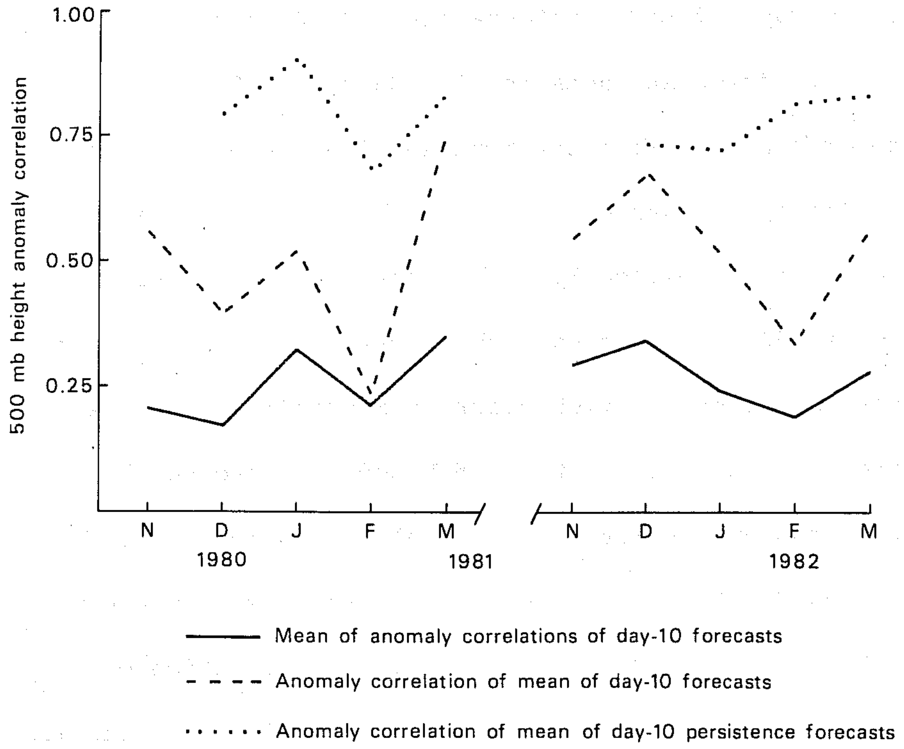
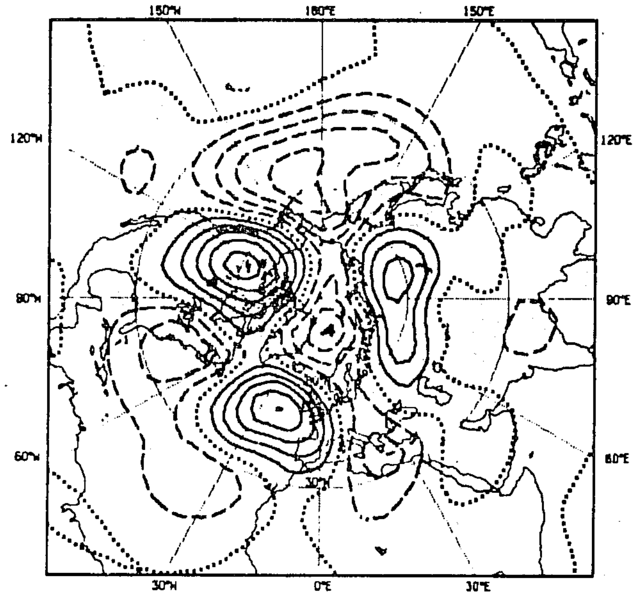


Fig. 16 Monthly-mean 500 mb height anomaly correlations for the extratropical Northern Hemisphere.

Inspection of individual cases indicates both anomalies which are well represented throughout the forecast period and anomalies which virtually disappear over 10 days of model integration. As an example, Fig. 17 presents analysed and day-10 forecast anomalies for January, 1981. It is clear that the anomaly resembling the Pacific/North American teleconnection pattern discussed by Wallace and Gutzler (1981) has been well-maintained by the forecasts, whereas the anomalously high pressure to the west of Europe has disappeared by day 10. In this case there may be an element of coincidence as the anomalously low pressure which occurred in reality over the Pacific matches a systematic tendency of the ECMWF model to produce lower than normal pressure in this region. However, the simulated pattern over North America does not correspond particularly with a systematic error pattern. Bengtsson and Simmons (1983) have noted the successful representation of low-level temperature anomalies for this case, and the sensitivity of the mean error implied by Fig. 17 to the representation of orography in the forecast model has been discussed by Tibaldi, 1985.

Another interesting example, for November 1981, is illustrated in Fig. 18. The predominant anomaly during this month has the appearance of a wavetrain over Europe and Asia, and the amplitude of this wavetrain is substantially weakened over all but its easternward limit in the mean of the day-10 forecasts. A wavetrain of smaller amplitude extending over North America from the Pacific to the Atlantic can also be seen to decay over the forecast period. Fig. 19 shows that almost all of the weakening of the monthly-mean anomaly pattern occurs in the second half of the forecast period. Further diagnosis would be required to determine the reason for this result, but it could, for example, be that the model misrepresents a localized remote forcing

ANALYSIS



FORECAST

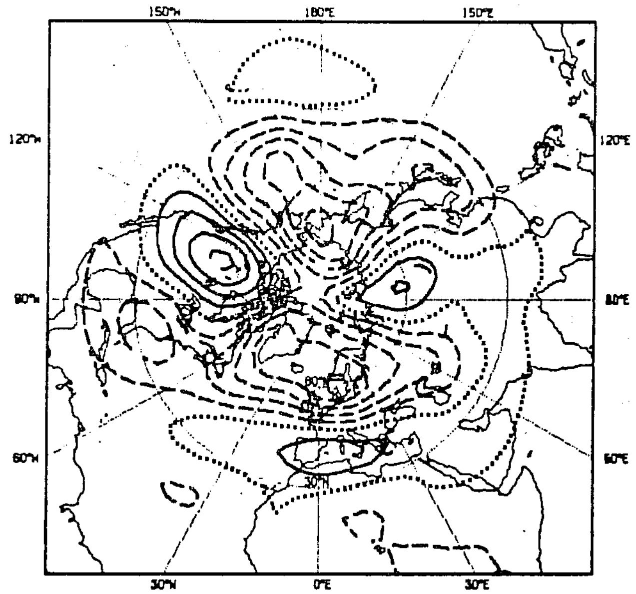
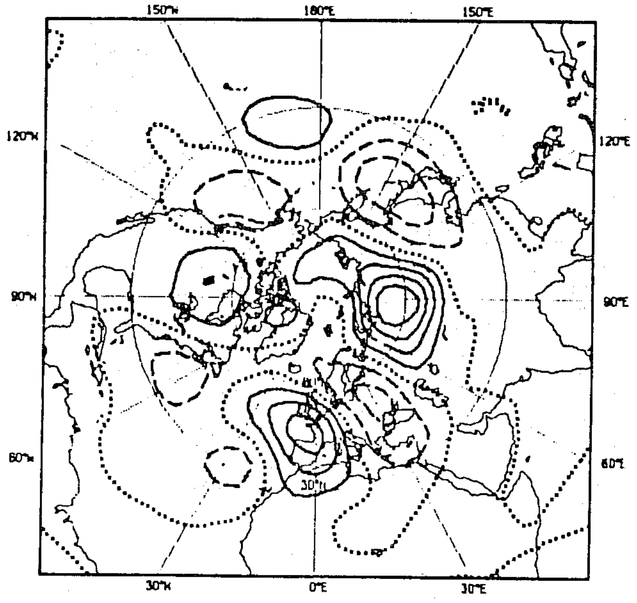


Fig. 17 The mean deviation of the 500 mb height from climatology for January 1981 (upper), and the corresponding anomaly computed from all day-10 forecasts verifying within this month (lower). The zero contour is dotted and negative contours dashed. The contour interval is 4 dam.

ANALYSIS



FORECAST

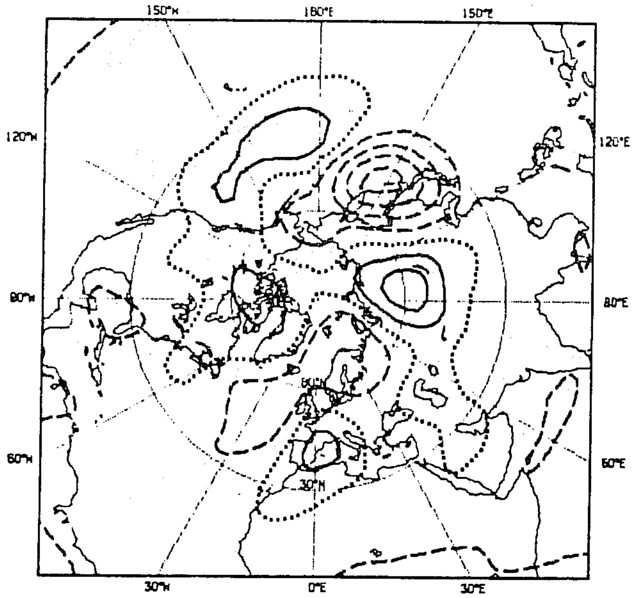


Fig. 18 As Fig. 17 but for November 1981.

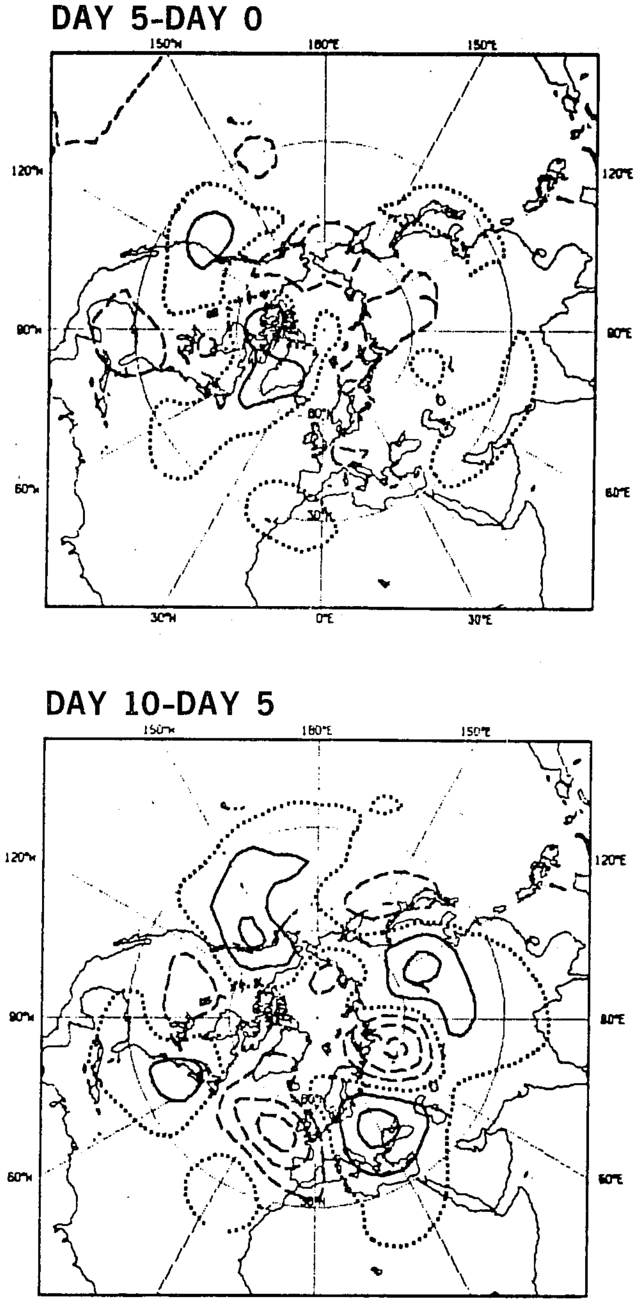


Fig. 19 The mean 500 mb height error of day-5 forecasts verifying in November 1981 and the mean difference betwen day-5 and day-10 forecasts verifying within this month. Contouring is as in Figs. 17 and 18.

or propagation which does not influence middle latitudes until day 5 or later. In any case, it would appear that the archived results from operational medium-range prediction form a valuable data source which may be diagnosed to further the understanding and numerical modelling of short-term climatic anomalies.

8. SYSTEMATIC MODEL ERRORS

Examination of monthly-mean maps such as discussed in the preceding section shows a tendency for certain errors to recur. These 'systematic errors' characteristically grow in amplitude throughout the forecast period, and their general similarity to errors in the model climatology revealed by integration over extended periods indicates that their growth represents a gradual drift from the climate of the atmosphere (as represented by the average of many initial states) to the climate of the model. The rate of this drift is found to vary from case to case, but the overall nature of the errors appears to be independent of the initial data. The errors can be recognized in substantial distortions of the flow pattern within individual forecasts towards the end of the 10-day range, and for longer-range prediction they may act to destroy predictability by, for example, causing an erroneous response to anomalous surface forcing which cannot be corrected in a simple statistical way.

A particularly important error of the ECMWF forecast model, and indeed of a number of other prediction and climate models (as discussed, for example, by Wallace et al., 1983, or Simmons and Bengtsson, 1984), occurs in the representation of the large-scale quasi-stationary wave patterns of the extratropical hemispheres. At the surface, centres of erroneously low pressure are characteristically found over the eastern Atlantic and Pacific Oceans, with too zonal a flow over North America and Europe. Stationary wave patterns in the middle and upper troposphere are typically weakened in both hemispheres. Over Europe this appears as a southward shift of the jet stream over the northern part of the continent, with a related erroneous track of transient disturbances, and an eastward shift of the mean trough over the eastern Mediterranean. A manifestation of the latter within an individual forecast has been discussed earlier in relation to Figs. 10 and 11. Overall, the error pattern has an 'equivalent barotropic' structure suggestive of inadequate orographic or thermal forcing of the stationary waves, although a contribution from systematic deficiencies in the behaviour of mature baroclinic eddies cannot be ruled out. The interaction between transient eddies and the time-averaged non-zonal flow has been the subject of several recent diagnostic studies (e.g. Hoskins et al., 1983; Illari and Marshall, 1983; Lau and Holopainen, 1984), and this interaction is almost certainly biased in the ECMWF model by a tendency to exaggerate meridional phase tilts of baroclinic waves (Arpe, 1983). Again, the example shown in Figs. 10 and 11 illustrates this deficiency in the forecast trough over the eastern Atlantic, which appears to contribute to a spurious eastward extension of the zonal Atlantic jet.

One specific area of investigation has involved examination of the prescription of the model orography. Diagnostic and barotropic model studies reported by Wallace et al. (1983) have suggested that use of a grid-square mean orography significantly underestimates the orographic forcing of the synoptic and larger-scale flow in the ECMWF model. Prediction experiments, some of which are described by Wallace et al, have been carried out using a series of 'envelope' orographies formed by adding to the mean orography multiples of the standard deviation of the actual orography over the grid square, this being computed from a very high resolution data set. Some significant improvements in the accuracy of medium-range forecasts have been found, and the growth of some systematic forecast errors has been substantially reduced. Further discussion is given by Tibaldi in this volume.

Dependence of these errors on other aspects of the model formulation has also been identified. They tend to be less at lower horizontal resolution

(Cubasch, 1981), as reported earlier for another model by Manabe et al (1979), and have been found to be sensitive to the parameterization of boundary-layer turbulence and shallow convection. Some sensitivity to the representation of the stratosphere has been found in extended integrations, in agreement with experience elsewhere (e.g. Mechoso et al., 1982; Boville, 1984), although the impact on the 10-day time range has not been found to be large at the vertical resolution used operationally (Simmons and Strüfing, 1983).

9. CONCLUDING REMARKS

A summary has been given of the predictive skill attained by the ECMWF forecasting system. We have largely utilized an extensive body of results from routine objective verification, although conclusions are generally borne out by subjective synoptic assessment. Some aspects of the prediction of blocking have been discussed, and a superficial investigation of the treatment of anomalies on the monthly-mean time-scale by the forecast model has been reported. Indications of significant developments in forecast accuracy have been presented.

Forecasts for the extratropical Northern Hemisphere are generally of good quality for 3 to 4 days ahead, with useful indications of weather type given on average for a further 2 or 3 days. Substantial variations in accuracy occur on a time-scale of weeks and months, and determination of the reason for this is of importance, as it may lead either to an ability to give guidance as to the expected accuracy of a particular forecast, or to ways of improving less accurate forecasts. Predictability is typically poorer by about 1½ days in the more data-sparse Southern Hemisphere, and more substantially poorer in the tropical belt. Both data assimilation and numerical modelling for the tropics pose particular problems, but objective verification indicates that significant advances are being made.

A stage has been reached in the development of the ECMWF forecasting system at which a number of major changes are anticipated. An enhancement of computing power has made possible a forthcoming increase in model resolution, and several changes are expected in the parameterization, with the introduction of a representation of shallow convection and new treatments of long-wave radiation and the specification of clouds. Changes in the parameterization of deep convection may also be introduced. In addition, the data analysis is being recoded to allow more flexibility, particularly with respect to the specification of structure functions, the use of data and the various interpolations carried out. We are unable to give a reliable estimate of the net improvement in forecast accuracy which will result from these changes and await the monitoring of future operational performance with interest.

The treatment of blocking in the forecasts appears to pose no outstanding problem, and cases of high predictability involving blocking have been noted. Such problems as do occur appear to be related to general deficiencies of the forecast model. Some success in the maintenance of pronounced large-scale anomalies on the monthly time-scale has been discussed, but examples of the failure to represent such features, particularly in the second half of the forecast range, have also been presented. Systematic errors can seriously influence predictions later in the forecast period, and become more pronounced in integrations carried out beyond the 10-day range. Reduction of these errors is of importance for the improvement of medium-range forecasting, and must be a crucial element in the development of a capability for the longer-range prediction of anomalies.

Acknowledgements

Many staff members and visiting scientists of ECMWF have contributed to the work reported here in some way or another, and gratitude is expressed to them.

REFERENCES

Arpe, K., 1980: Confidence limits for verification and energetic studies. ECMWF Tech.Rep.No.18, 24pp.

Arpe, K., 1983: Diagnostic evaluation of analysis and forecasts: Climate of the model. Proceedings of 1982 ECMWF Seminar on Interpretation of Numerical Weather Prediction Products, 99-140.

Bengtsson, L., 1981: Numerical prediction of atmospheric blocking - a case study, Tellus, 33, 19-42.

Bengtsson, L. and A.J. Simmons, 1983: Medium range weather prediction - operational experience at ECMWF. Large-scale Dynamical Processes in the Atmosphere, eds. B.J. Hoskins and R.P. Pearce, Academic Press, pp.337-63.

Bengtsson, L., Kanamitsu, M., Kållberg, P. and Uppala, S. 1982: FGGE research activities at ECMWF. Bull.Am.Met.Soc., 63, 277-303.

Boville, B.A., 1984: The influence of the polar night jet on the tropospheric circulation in a GCM. J.Atmos.Sci., 41, 1132-1142.

Burridge, D.M., 1979: Some aspects of large scale numerical modelling of the atmosphere. Proceedings of ECMWF Seminar on Dynamical Meteorology and Numerical Weather Prediction, Vol.2, 1-78.

Burridge, D.M. and J. Haseler, 1977: A model for medium range weather forecasts - adiabatic formulation. ECMWF Tech.Rep.No.4, 46pp.

Charney, J.G., R.G. Fleagle, V.E. Lally, H. Riehl and D.Q. Wark, 1966: The feasibility of a global observation and analysis experiment. Bull.Amer.Met.Soc., 47, 200-20.

Cubasch, U., 1981: The performance of the ECMWF model in 50-day integrations. ECMWF Tech.Memo No.32, 74pp.

Geleyn, J.-F. and A. Hollingsworth, 1979: An economical analytical method for the computation of the interaction between scattering and line absorption of radiation. Beitr.Phys.Atmosp. 52, 1-16.

Girard, C. and Jarraud, M. 1982: Short and medium range forecast differences between a spectral and grid point model. An extensive quasi operational comparison. ECMWF Tech.Rep.No.32, pp.178.

Grønaas, S., 1983: Systematic errors and forecast quality of ECMWF forecasts in different large-scale flow patterns. Proceedings of the ECMWF Seminar on Interpretation of Numerical Weather Prediction Products, 161-206.

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

Hoskins, B.J., I. James and G.H. White, 1983: The shape, propagation and mean-flow interaction of large-scale weather systems', J.Atmos.Sci., 40, 1595-612.

Illari, L. and J.C. Marshall, 1983: On the interpretation of eddy fluxes during a blocking episode. J.Atmos.Sci., 40, 2232-2242.

Ji, L.R. and S. Tibaldi 1983 Numerical simulations of a case of blocking. The effects of orography and land-sea contrast. Mon.Wea.Rev., 111, 2068-2086

Kuo, H.L 1974: Further studies of the influence of cumulus convection on large-scale flow. J.Atmos.Sci., 31, p.1232-1240.

Lange, A. and E. Hellsten, 1984: Results of the WMO/CAS NWP Data Study and Intercomparison Project for forecasts for the Northern Hemisphere in 1983. WMO Geneva.

Lau, N.-C and E.O. Holopainen, 1984: Transient eddy forcing of the time-mean flow as identified by geopotential tendencies. J.Atmos.Sci., 41, 313-328.

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

Lorenz, E.N., 1975: Climate Predictability, GARP Publication Series No. 16, pp. 132-6, WMO, Geneva.

Lorenz,E.N., 1982: Atmospheric predictability experiments with a large numerical model. Tellus, 34, 505-513.

Louis, J.-F. 1979: A parametric model of vertical eddy fluxes in the atmosphere. Boundary-layer Meteorol., 17, 187-202.

Manabe, S., D.G. Hahn, and J.L. Holloway, 1979: Climate simulation with GFDL spectral models of the atmosphere: Effect of spectral truncation. GARP Publication Series No.22, 41-94, WMO, Geneva.

Mechoso, C.R., M.J. Suarez, K. Yamazaki, J.A. Spahr and A. Arakawa, 1982: A study of the sensitivity of numerical forecasts to an upper boundary in the lower stratosphere. Mon.Wea.Rev., 110, 1984-93.

Miyakoda, K., G.D. Hembree, R.F. Stricker and I. Shulman, 1972: Cumulative results of extended forecast experiments I: Model performance for winter cases. Mon.Wea.Rev., 100, 836-55.

Miyakoda, K., T. Gordon, R. Caverly, W.Stern, J.Sirutis and W.Bourke, 1983: Simulation of a blocking event in January 1977. Mon.Wea.Rev., 111, 846-69.

Nieminen, R., 1983: Operational verification of ECMWF forecast fields and results for 1980-1981. ECMWF Tech.Rep.No.36, 48pp.

Shukla, J., 1981: Dynamical predictability of monthly means, J.Atmos.Sci., 38, 2547-72.

Simmons, A.J. and L. Bengtsson, 1984: Atmospheric general circulation models. Their design and use for climate studies. The Global Climate, ed. J.T. Houghton, Cambridge University Press, 37-62.

Simmons, A.J. and Burridge, D.M., 1981: An energy and angular momentum conserving vertical finite difference scheme and hybrid vertical coordinate. Mon.Wea.Rev., 109, p.758-766.

Simmons, A.J. and Hoskins B.J. 1978: The life cycles of some nonlinear baroclinic waves. J.Atmos.Sci., 35, 414-432.

Simmons, A.J. and M. Jarraud, 1984: The design and performance of the new ECMWF operational model. Proceedings of 1983 ECMWF Seminar. To be published.

Simmons and R. Strüfing, 1983: Numerical forecasts of stratospheric warming events using a model with a hybrid vertical coordinate Q.J.R.Met.Soc., 109, 81-111.

Smagorinsky, J., 1969: Problems and promises of deterministic extended range forecasting. Bull.Amer.Met.Soc., 50, 286-311.

Temperton, C. and Williamson, D.L. 1981: Normal mode initialisation for a multi-level grid-point model, Part I : Linear Aspects. Mon. Wea. Rev., 109, pp.729-743.

Tibaldi, S., 1985: Envelope orography and maintenance of quasi-stationary waves in the ECMWF model. Workshop on Global Scale Anomalous Circulations in the Atmosphere and Blocking. IBM Italia, Rome, 27-31 August 1984.

Tibaldi, S. and A. Buzzi, 1983: Effects of orography on Mediterranean lee cyclogenesis and its relationship to European blocking. Tellus, 35A, 269-286.

Tibaldi, S. and L.R. Ji, 1983: On the effect of model resolution on numerical simulation of blocking. Tellus, 35A, 28-38.

Tiedtke, M., J-F. Geleyn, A.Hollingsworth, J-F.Louis, 1979: ECMWF model, parameterization of subgrid scale processes. ECMWF Tech.Rep.No.10, 46pp.

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.

Wallace, J.M., S.Tibaldi and A.J.Simmons, 1983: Reduction of systematic forecast errors in the ECMWF model through the introduction of an envelope orography. Quart.J.Roy.Met.Soc., 109, 683-717.

Williamson, D. L. and C. Temperton 1981 Normal mode initialisation for a multi- level grid-point model, Part II: Nonlinear aspects. Mon. Wea. Rev., 109, pp.744-757.