TECHNICAL MEMORANDUM

146

Mid latitude atmospheric prediction on time scales of 10 30 days

A. Hollingsworth, U. Cubasch, S. Tibaldi, Č. Branković, T.N. Palmer and L. Campbell

Research Department

June 1988

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

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

ABSTRACT

Theoretical views on the possibility of useful extended-range forecasts range from considerable optimism (in the tropics) to guarded optimism, and even pessimism (in mid-latitudes). Paradoxically, general circulation and weather prediction modelling has been much more successful in mid-latitudes than in the tropics. In recent years the production of isolated but remarkably successful extended-range mid-latitude forecasts at several modelling centres has stimulated a re-evaluation of the possibilities for useful mid-latitude forecasts in this range.

Current levels of skill in medium-range (2-10 day) and extended-range (10-30 day) forecasts in mid-latitudes are much affected by the ability to forecast persistent anomalies. Blocking action is recognised as a key forecast problem in both time ranges. Recent progress in the understanding of the phenomenon is discussed, before reviewing developments in the theory of predictability.

A few examples of extended-range forecasts are then discussed in detail. Some are considerable successes; others are partial successes. One example demonstrates that a pair of forecasts for 10-day means can diverge from each other quite rapidly, even though they start from closely spaced initial conditions. Moreover each of the two contradictory forecasts were successful in some regions, but the regions of success did not overlap.

Statistical verification results for a set of twelve cases in the winter 1985/6 indicate that the average level of skill in the extended range is often controlled by the skill in the medium range. Model and analysis improvements will be as important for extended range forecasts as they are for medium range forecasts. Some attention is given to the problem of climate drift in the models. Sustained attacks on this problem have provided much of the progress in simulation capability in recent years.

Beyond the range of deterministic predictability, the most useful forecasts are likely to be those which can identify the range of likely atmospheric evolutions, and the probability of each evolution. For this one needs a model capable of simulating the full range of atmospheric variability, one needs methods to generate a good sample of likely evolutions, and one needs a means of interpreting the results in a useful way. Current work in these areas is discussed. Our conclusion is that current results are encouraging and justify continued research in the area.

1. INTRODUCTION

In recent years there has been a quickening of interest in the possibility of producing useful mid-latitude forecasts in the range of 10-30 days. The interest is stimulated by an increase in our understanding of low-frequency variability in the atmosphere, by improvements in forecast skill in the medium range, by improvements in space-based observational technology, and by improvements in computer technology. Most workers in the field are agreed that forecasts of instantaneous states beyond about 14 days are likely to have little skill. There is hope that one may be able to forecast some properties of the time-mean state, or the distribution of possible mean states, beyond 10-14 days. In this review we consider the scientific problems that must be solved if we are to attain this long-term goal.

Predictability studies and observational studies both suggest that extended-range prediction may be more successful in the tropics than in mid-latitudes because of the great importance of the oceanic forcing in the tropics. However both medium-range predictions and general circulation simulations of the tropics show lower levels of skill than is found in mid-latitudes. Most reported work on prediction using numerical models has concentrated on mid-latitudes. This will change as progress is made in simulating the tropical atmosphere and ocean.

Theoretical and observational studies in mid-latitudes have provided a clearer description and understanding of the processes involved in the development and maintenance of long-lived persistent phenomena such as blocking. Modelling developments for operational medium-range (2-10 day) weather forecasting have led to a convergent evolution of forecast and general circulation models which can produce fairly accurate simulations of atmospheric flow both in the time-mean and transient components of the flow. Prediction experiments and predictability experiments with these models in the range beyond 10 days suggest that it may indeed be possible to produce useful forecasts in the 10-30 day range given reasonable progress in modelling and given current trends in the development of computer technology. There is feeling amongst many workers in the field that the area is ripe for sustained and concerted attack.

Views on the possibility of useful forecasts in the 10-30 day range vary from extreme optimism to extreme pessimism. The optimists are encouraged by the fact that every group that makes experiments in the area seems to find at least one spectacularly successful case. They are also encouraged by many of the linear and non-linear theoretical and/or observational studies. The pessimists feel that there can be useful predictive skill for at most one or two eddy turnover times and any linear forcing effect will be obliterated by the turbulence of the flow. In their view the eddy turnover time is a few days and so there can be no predictive skill beyond about six days and certainly no more than ten days. In mid-latitudes one must average the atmospheric data for at least a season to see clear evidence of effects of sea surface temperature anomalies (Wallace, 1987; Palmer, 1987). This suggests that the internally generated variability is substantially greater than the externally generated variability on time scales of 10-30 days.

New developments in our understanding of the low frequency behaviour of the atmosphere have led to a cautious edging together of the optimistic and the pessimistic views. Observational and theoretical studies are gradually uncovering the rich and complex nature of the processes which excite and maintain long-lived coherent blocking structures in the mid-latitude atmosphere. The effect of the 30-60 day oscillation on the tropical atmosphere is being steadily documented. The role of tropical SST anomalies in generating mid-latitude variations on the seasonal time-scale is the subject of considerable observational and theoretical work, while the role of the ocean in exciting and maintaining the much longer lived ENSO phenomena has been clearly established.

Since the topic for this paper is the predictability of the mid-latitude atmosphere on time scales up to a month, a key phenomenon to be forecast is mid-latitude blocking. The status of the theory of blocking and of the theory of mid-latitude predictability are reviewed in section 2. Examples of successful and unsuccessful monthly forecasts are discussed in section 3. Section 4 reviews the available, and limited, experimental material and discusses the average level of skill achieved to date. Section 5 discusses the importance of the climate drift problem, while section 6 discusses the statistical work necessary for the interpretation of monthly forecasts. We conclude in section 7 with some speculations on future developments.

From our present standpoint it seems that useable forecasts in this time range will have to be probabilistic in nature and may need to be couched in terms of weather regimes, giving the probability of transition between regimes as a function of space and time. We have a long way to go, but computer power is improving so rapidly that the tempo of research in this area will accelerate rapidly in the coming decade.

2. BLOCKING AND PREDICTABILITY

The problem of forecasting blocking action is a central problem in forecasting for mid-latitudes (e.g. Grønaas, 1982). There is a growing body of theoretical and observational evidence that blocking action is dominated by internal dynamics. Despite strenuous efforts, there is little consistent evidence that anomalous boundary forcing in the form of sea surface temperature (SST) anomalies plays more than a subsidiary role on these time scales. On longer seasonal or inter-annual time scales the role of SST and other forcing anomalies appears to be much more significant (Palmer, 1987).

2.1 Theories and Observations of Blocking

The mechanism of blocking has received considerable attention in recent years. A number of theories have been proposed which offer insight into the problem. The current surge of interest was pioneered by the work of Green (1977) and Egger (1981) who suggested that the phenomenon could be maintained by transient eddy effects, and by Charney and DeVore (1979) who suggested that blocking could be interpreted as a resonance phenomenon arising from the effect of orographic forcing on the flow. A few years earlier, Stern (1975) developed a class of isolated and exact solutions of the barotropic vorticity equation, modons, whose relevance to blocks were discussed by McWilliams (1980). All three approaches to the problem have been followed vigorously and it now seems that there may be a role for all three theories.

After an initially enthusiastic reception, the resonance theories fell into temporary disfavour for a number of reasons. They required that blocking action be a hemispheric phenomenon (Charney and Strauss, 1980; Kallen, 1981). This result was not supported by synoptic experience nor by diagnostic work, (Dole and Gordon, 1983), which showed that blocking action occurred independently in the two main centres of action, the East Atlantic and the East Pacific. The resonance theory of Charney and DeVore also suggested that the two stable steady states occurring in the theory should lead to bi-modal distribution in the amplitudes of the zonal flow and of the wave components of the flow. No evidence for bi-modality in the zonal flow has been found to date, Hansen (1986) and Hansen and Sutera (1986). Modifications of the resonance theory to take account of some non-linear effects (Trevisan and Buzzi, 1980; Malguzzi and Speranza, 1981; Reinhold and Pierrehumbert, 1982; Speranza, 1986; Benzi et al, 1986a) suggest that the resonance mechanism might lead to bi-modal behaviour in the wave components but not in the zonal flow.

Sutera (1986) and Benzi et al (1986b) found that a simple indicator of the amplitude of the waves in the large scale flow has, in fact, a bi-modal distribution. There are suggestions that Sutera's two modes explain a substantial part of the variance of the flow in the 10-30 range, but this is as yet unverified.

A quite different view of blocking is presented in the work of Green (1977), Austin (1980), Mahlman (1979), Hoskins et al (1983), Illari (1984), Illari and Marshall (1983), Shutts (1983, 1986) and Hoskins et al (1985). These studies have emphasised the fact that the slowly varying flow field constituting the block can strain the high-frequency transients in such a way as to maintain the slowly varying field against dissipation. The discussion is most persuasively presented in terms of isentropic distributions of potential vorticity (IPV), which is conserved in adiabatic flow. The papers just cited include several case-studies which demonstrate very graphically how the straining of the eddies by the block leads to repeated refuelling of the anti-cyclonic part of the block with low vorticity air from the south, while the cyclonic part of the block is repeatedly replenished by high vorticity air from the north (Shutts 1986). In quantitative terms the effect is vividly depicted in terms of the anisotropy of the eddies using the anisotropy-vectors introduced by Hoskins et al (1983) and used to good effect by Shutts (1986).

The relevance of modons to blocking has been discussed in several papers since the work of McWilliams (1980). The existence of closed contours of stream-function (ψ) and vorticity (q) imply different functional relations between ψ and q inside and outside the closed streamlines. Observational evidence of this effect has been produced by Illari and Marshall (1983).

More recently, Haines and Marshall (1987) have used linearised and non-linear calculations to study the interaction of a modon and a field of transients. They produce empirical support for the result of Pierrehumbert and Malguzzi (1984) that the stability of the modon in a forced dissipative flow requires only a balance between forcing and dissipation of a parcel's vorticity, averaged over closed lagrangian circuits (closed streamlines) rather than strict balance at each instant. Haines and Marshall also show that the eddy straining phenomenon found in observational studies is precisely the mechanism by which the modon maintains itself against dissipation in their model. Their results show that the interpretation of a modon as a local resonance phenomenon is quite reasonable, but they do not discuss what processes might lead to the initial formation of a modon. Growth as a local resonant response to forcing is a reasonable possibility. At this point it may be permissible to invoke the global Charney-DeVore mechanism to explain the initial formation, at least for the Pacific cases, with the local non-linear resonance taking over as the amplitude increases. The experiments suggest that the presence of horizontally propagating Rossby waves can substantially modify the generation of the modon.

2.2 Medium Range Predictability

There has been considerable progress in medium range forecasts over the last ten to fifteen years. Over Europe in winter nowadays, the 6-day forecast is as good as the 2-day forecast was in the early 1970s (Bengtsson, 1985). Lorenz (1982) used statistics on global forecast verifications, and on the growth rate of the spread of forecasts started from analyses separated by 24h, to estimate the potential for improvement in the medium range. Using ECMWF forecasts for the boreal winter of 1981/82, he estimated that model improvements alone could extend the then current 7-day level of skill to 10 days or beyond. Fig 1 shows Northern Hemisphere rms skill and spread scores for winter 1980/81 (using Lorenz's data) and equivalent data from 1985/86. Lorenz argues that the potential for improvement is given by the horizontal distance between the skill and spread scores, assuming no improvement in the skill of the day one forecast. The results in Fig 1 show that the potential for improvements arising from model changes alone is almost as large in 1985/86 as it was in 1980/81, even though there have been substantial gains in forecast accuracy over the period. The reason is that the both model and analysis upgrades have helped to improve the 1-day forecasts.

In the period beyond 4 to 5 days, operational medium range forecasts have wide variations in skill. Fig 2 shows, for June 1985-May 1986, the length of time for which forecasts for Europe had an anomaly correlation larger than 0.6, which is judged to be a limit of useful predictive skill (Hollingsworth et al, 1980). The running 10-day mean of the daily curve is also shown for each quarter. The figure illustrates that forecast skill varies substantially from day to day, but also on much longer time scales. The remarkable jump in forecast skill for Europe in mid-November 1985 is discussed in section 3; the even larger jump in late January 1986 occurred because the forecasts failed to predict a major European block until 3 days before the onset and thereafter did a good job of forecasting its maintenance.

The time evolution of the cumulative distribution of the Northern Hemisphere anomaly correlation scores for the periods Oct-Dec 1984 and 1985 are illustrated in Fig 3. In both years there is a substantial spread in the scores. The marked improvement in the scores between 1984 and 1985 is due to the introduction of a much improved model in May 1985 (Jarraud et al, 1985; Jarraud, 1986; Tiedtke et al, 1987). The improvement affects not just the median of the scores but also the spread. In 1984, the top quartile of forecasts had an anomaly correlation of 0.3 or better at day 10, while in 1985 the top quartile had a score of 0.5 or better. These forecasts were of good quality at day 10, but they were of little practical value since the operational forecaster had no way of knowing which forecasts would be good and which bad. The ability to identify, a-priori, the forecasts good enough to fall in the best quartile or tercile would be of great practical value.

However much we improve the mean level of performance, there will always be a forecast range beyond which the forecasts become unreliable (Tennekes et al, 1986). There is therefore a growing demand for a-priori estimates of forecast skill in the medium range. Most effort at forecast centres has been put into model and analysis development, in the hope of reaching the goals which theory says are achievable for deterministic forecasts. Efforts to identify the skilful cases a priori, and to estimate the most likely range of atmospheric evolutions, will become matters of major concern in the medium range as well as in the extended range. Preliminary explorations of this new area (Grønaas, 1985; Palmer and Tibaldi, 1986) suggest there is scope for much useful work.

2.3 Extended Range Predictability

Predictability studies in the 10-30 day range have concentrated on the possibility of forecasting the time mean state, rather than on the predictability of particular phenomena. In a pioneering study, Shukla (1981) used the growth of white noise perturbations of the the wind-field to estimate the potential predictability of 10-day means. He concluded that time-mean states in mid-latitudes are potentially more predictable than instantaneous states. In a companion paper (Shukla 1984) he examined the effect of forcing from slowly varying boundary conditions on atmospheric predictability. He concluded that an ability to predict the evolution of the boundary conditions would enhance atmospheric predictability, most notably in the tropics. His general conclusions about the importance of boundary forcing for the tropics are supported by the recent study of Chervin (1986). In mid-latitudes however, Arpe et al (1985) have criticised his methodology, while Tribbia (1986) in a rather similar study reaches much more pessimistic conclusions. On the other hand, Shukla (1985) argues that the conclusions of his 1981 study are conservative and that the potential for predictive skill in mid-latitudes may be larger than he first estimated. Since most of the conclusions of the predictability studies are dependent on the realism of the models employed, many workers have felt that it may be more valuable to determine the currently attainable level of skill.

3. EXAMPLES OF EXTENDED RANGE FORECASTS

Extended range forecasting with numerical models is being explored in a number of institutes, notably the U.K. Meteorological Office (UKMO), the Geophysical Fluid Dynamics Laboratory, Princeton (GFDL), the European Centre for Medium Range Weather Forecasts (ECMWF), the U.S. National Meteorological Centre (NMC), the Goddard Laboratory for Atmospheres, Greenbelt (GLA), the Japan Meteorological Agency (JMA), the University of Maryland (COLA), and the Canadian Climate Centre (CCC) amongst others. The UKMO has had an active experimental programme for about 10 years (Gilchrist, 1977, 1986), while GFDL has been active for almost the same length of time. The other institutes have recently undertaken experimental programmes in the area, partly because of the interest generated by the pioneers, and partly because extended range integrations have become an essential means of diagnosing the climate drift problems of medium range models.

A number of these institutes have begun a long series of experiments running over several years, to develop an extensive database of results for a wide variety of cases. Many of the institutes are coordinating both their experimental strategies and the cases used for experimentation. This should provide valuable estimates of the representativeness and reliability of the results.

In this paper we can only review the set of results which have been submitted for publication. Many new results may be expected in the next year or two. We review two successful forecasts and a pair of contradictory forecasts, (each of which was correct in one region and wrong in another), before discussing the general level of performance in the next section. It seems that most institutes working in the area have at least one striking success. These successes are tantalising because they seem to be very sensitive to the details of the starting data and, to a lesser extent, to the details of the model used. So far the successes have been difficult to reproduce with another institute's model, despite determined efforts in that direction.

3.1 Examples of Successful Forecasts

a) GFDL forecast from Jan 1 1977

Miyakoda and his collaborators at GFDL have made extensive studies of the extreme weather over North America in January 1977. They investigated the sensitivity of monthly forcasts of the event to numerical formulations, to physical parameterisations, to surface boundary conditions, to model resolution and to starting analyses. They found (Miyakoda et al, 1983) that their most advanced and expensive model showed a remarkable level of skill in reproducing the extreme blocking pattern over North America. Their forecast for the time mean 850mb temperature for days 10-30 is shown in Fig 4, together with the verifying observed field and the climatological mean field. The forecast is an impressive success, and was the major stimulus to the current interest in the area.

The forecast quality in Miyakoda's experiments was extremely sensitive to the numerical technique used (grid-point versus spectral), to the physical parameterisations used (Mellor-Yamada formulation versus a much simpler formulation for the boundary layer) and to the starting date (the forecast from Jan 2 was much less successful than the forecast from Jan 1). Attempts have been made to reproduce his forecast with other models (Strüfing, 1982), but with no success.

b) ECMWF forecast from Jan 17 1984

The first tantalising success at ECMWF was the 30-day forecast from January 17 1984 reported by Jarraud (1986) and Molteni et al (1986). The starting analysis is shown in Fig 5a, the forecasts from day 6 through day 26 are shown at four day intervals in Fig 5b-g in the right hand panels, while the verifying analyses are shown on the left. Initially the flow over the Atlantic is quite zonal, with a trough over the Pacific. The large scale pattern has changed little by day 6, and the day 6 forecast is reasonably successful (Fig 5b). Between day 6 and 14 a weak ridge develops over western Europe and decays again. This feature is reasonably well forecast (Fig 5c-d). By day 18 the flow has become rather mobile over the Atlantic. The forecasts for the transient features over the Atlantic are wrong in the detailed timing of events, but correct in indicating the synoptic type (Fig 5e-f). Between day 22 and 26 (Fig 5f-g) a marked blocking high develops over western Europe again and is very successfully forecast. The 1000mb forecast for the surface developments in the Mediterranean between days 24 and 27 were remarkably accurate in the development and eastward movement of a Genoa cyclone, as may be judged from the 1000mb forecast and verifying analysis for day 26 shown in

Fig 6. Success in forecasting such a small-scale feature at this range obviously depends on a good forecast for the larger scales.

In the first 5-10 days the forecast performance is robust to small changes in the starting analysis, in the boundary conditions and in the the model parameterisations. Beyond day 10 the forecast is very sensitive to any or all of these aspects of the forecasting system.

Similarly successful forecasts have been produced at UKMO and at NMC (G. White, pers. comm.). The accumulation of these cases shows that the phenomenon of blocking can sometimes be forecast at extended range. We need a great deal more theoretical insight before we can understand when and why the forecasts work well, or work badly.

3.2 A Pair of Contradictory Half-Successes

As noted already in the discussion of Fig 2, forecast skill varies on many time scales and there can be large differences between the quality of medium range forecasts made from successive days. The medium range forecasts for Europe from November 15 and 16 1985 (called F1511 and F1611 hereafter) showed marked differences in skill, as shown by Fig 2. When continued to 31 and 30 days respectively, these forecasts showed the largest differences between successive forecasts that we have seen in a sample of 17 such pairs, one from each month between April 1985 and August 1986. The differences were large in the 10-day and 30-day means as well as in the daily maps. A study of these forecasts suggests that 10-day mean forecasts will, on occasion, diverge from each other at a rather fast rate.

Fig 7 shows ten-day mean maps for days 1-10, 6-15, 11-20, and 16-25 for the observed 500mb field and for the two forecast fields, all counting being reckoned from Nov. 16. The forecast mean maps for days 1-10 (Fig. 7a-c) are similar to each other and to the verifying analysis, with good agreement regarding the Aleutian and Atlantic ridges and for the low over Western Europe.

Important differences between the forecasts are already evident in the day 6-15 maps (Fig. 7d-f). The retrogression of the Atlantic ridge is more successful in F1611, which is why this forecast has a much better skill score for Europe than F1511. The same forecast, F1611, is much less successful than its companion in the treatment of the Pacific ridge, moving it westward and northward quite rapidly and incorrectly; the forecast from the 15th is much better in this regard.

By day 11-20 (Fig. 7g-i) the two forecasts are quite different from each other and each of them is wrong in some important respects. F1611 continues to treat the flow over the Atlantic better than F1511. In the Pacific however, the positions are reversed and F1511 is much more successful than F1611 in its treatment of ridge over Alaska.

The day 16-25 (Fig. 7j-l) means show little resemblance to each other or to the observed field. By this time too both forecasts show evidence of the loss of eddy energy which is discussed in the section on climate drift below. Thereafter the forecasts continue their separate developments. The differences between the forecasts are so large that even the 30 day means (not shown) are quite different.

This example illustrates some of the difficulties that must be faced if we are to produce useful extended range forecasts. The 10-day mean fields in the Pacific diverged from each other at a very fast rate. The earlier forecast F1511 was much more successful in treating the Pacific block than its companion, while the positions were reversed in the Atlantic. We need a deeper understanding of the initiation and maintenance of blocking, so as to understand the patent sensitivity to details of the initial data. A critical point seems to be whether the (possibly unpredictable) high-frequency transients make the lower frequencies unpredictable as suggested by Opsteegh and van den Dool (1979), or whether the lower frequencies can organise the high frequencies so as to reinforce existing blocks as discussed by Shutts (1986). At the very least we need to overcome the climate drift problem in forecast models, so that the correct climatology is preserved and the low frequency variability is well simulated. This latter is essential if probabilistic forecasts of blocking are to have any chance of success.

4. VERIFICATION RESULTS FROM EXPERIMENTS IN THE EXTENDED RANGE

Forecast verification at extended ranges is likely to be as difficult a problem as it is at shorter ranges. The verification method should be appropriate for the application for which the forecast is used. Wind, temperature and precipitation are parameters of general interest. The averaging period and the region of verification should both be chosen to suit the particular application. Since we are still at a very early stage in the development of dynamical extended range forecasting, we limit our discussion to verifications of the hemispheric height field using rms and anomaly correlation statistics. This has the advantage of highlighting important problems and it allows a degree of comparison between the three sets of results which are available to us, from UKMO, GFDL and ECMWF.

4.1 The ECMWF Experiment

Since we can get from our own experiments the set of statistics needed to make some general points, we start by discussing the ECMWF experiments, though these were by no means the first in point of time. In May 1985 a new operational model was introduced at ECMWF. It has a revised set of physical parameterisations (Tiedtke et al, 1987) and higher resolution (T106 compared to T63) than the old model (Jarraud et al, 1986). Given the importance of the climate drift problem for medium range forecasting (Hollingsworth et al, 1980; Bengtsson and Simmons, 1982) it was thought necessary to develop a comprehensive sampling of the climate drift properties of the new model. This is being done by generating pairs of 30-day forecasts from two successive days in mid-month from April 1985 to August 1986. Two model resolutions are used, the operational T106 resolution and the T63 resolution used prior to 1985. Both sets of integrations use the same physical parameterisations, and the same orography before the spectral truncation.

The experimental design provides the required sampling of the climate drift as a function of model resolution and season, and also provides an estimate of the rate of spread of forecasts in the extended range - an important consideration for predictability estimates.

Figure 8a shows the 500mb rms height verifications for the extended range T106 forecasts for the 6 pairs of cold season forecasts (October 1985- March 1986). Also shown are the rms errors of persistence forecasts and the expected rms errors of climatological forecasts, which provide cheap zero-cost controls. The rms spread between the pairs of successive forecasts is also shown. All forecast and analysis fields were truncated at total wave-number 40 before the verifications.

An indication of possible sampling problems may be had from comparison of these results with those in Fig 1 for one hundred 10-day forecasts in winter 1985/86. The forecast error curves agree quite well out to about 8 days, after which our small sample of extended range forecasts show slightly larger errors. There are larger discrepancies in the forecast spread curves; the successive forecasts in our small ensemble diverge from each other much more rapidly than they do in the larger sample of shorter forecasts. The results for the ensemble of medium range forecasts are based on 100 verifications and 99 measurements of spread, while the results for the ensemble of extended range forecasts are based on 12 verifications and 6 measurements of spread. Thus the set of extended range forecasts has obvious sampling problems which must be borne in mind in our later discussions of the possibility of improving forecast skill in the extended range. Application of an F-test to the measures of spread in the two samples at day 10 suggests that the differences are not significant at the 10% level.

4.2 Asymptotic Limits of Forecast Scores

The asymptotic levels of the four curves in Fig 8a convey information about important model characteristics. The rms error of the climatological forecast is the atmospheric variability of daily values within a given season, say S_1^a, where the 1 denotes a statistic for daily values. The asymptotic level of the persistence error for daily forecasts, P_1^a, is the expected value of the difference between two randomly chosen days from the same season and is given by

$$P_1^a = \sqrt{2}\ S_1^a$$

The asymptotic level of the forecast spread when all correlation between the forecasts is lost, say P_1^m, is equivalent to the expected value of the difference between two randomly chosen model states from the same season. We will sometimes call this the asymptotic level of model persistence and it is given by

$$P_1^m = \sqrt{2}\ S_1^m$$

where S_1^m is the variability of the model about its own mean state.

Let us define E as the asymptotic rms value of the difference between the mean atmospheric state and the mean model state, i.e. the rms of the mean climate drift for the variable of interest. Then F_1, the asymptotic level of the forecast error, is given by

$$F_1 = S_1^a \left(1 + \left(\frac{S_1^m}{S_1^a}\right)^2 + \left(\frac{E}{S_1^a}\right)^2\right)^{\frac{1}{2}}, \qquad (1)$$

as the correlation between observed and forecast anomalies vanishes.

If the mean climate drift were zero and if the model variability were the same as the atmospheric variability, then P_1^a, P_1^m, and F_1 would all have the same asymptotic level (Miyakoda et al, 1972). In Fig 8a the model persistence curve is about 15% lower than the true persistence curve, so model variability for daily values is about 15% lower than atmospheric variability.

We can estimate the magnitude of the climate drift contribution to Eq (1) as follows. If R_1 is the asymptotic limit of the spread of daily values as measured by anomaly correlation, then it is easily shown that

$$\frac{R_1}{(1 - R_1^2)^{\frac{1}{2}}} = \frac{E}{S_1^m} \qquad (2a)$$

so that if R_1 is less than 0.5 we have

$$R_1 \sim \frac{E}{S_1^m} \qquad (2b)$$

to a good approximation. Fig 8b shows the anomaly correlation scores for the forecasts, for persistence forecasts and for the spread of the forecasts (where one forecast of a pair is treated as a forecast for the other). From Fig 8b we see that R_1 is about 0.35. This suggests that the asymptotic value F_1 should be approximately .95 P_1^a, which is not too far from the observed value.

4.3 Scores for ECMWF 10-day mean forecasts

Figure 9a,b shows the rms error and anomaly correlation scores for forecasts of 10-day mean 500mb height for the same set of winter forecasts as Fig 8; the climatological, persistence and spread curves are also shown. The discussion of the asymptotic limits of the curves for daily values in Fig 8 also applies, mutatis mutandis, to the curves for 10-day means in Fig 9; we shall use the suffix x to denote the equivalent properties of 10 day mean scores. The relative difference between P_x^a and P_x^m, the asymptotic persistence levels for atmosphere and model are just as large in rms as they were for the daily values. The asymptotic value of the anomaly correlation R_x is larger than R_1, because the denominator on the right hand side of Eq (2a) or Eq (2b) has decreased while the numerator is essentially unchanged. Thus the contribution of mean error to forecast error is larger for 10-day means than for daily values.

a) Actual and Potential Predictability Times

Though the rms height error and anomaly correlation are very simple verification tools, they nevertheless have provided the basis for many discussions of useful and potential predictability times (Charney et al, 1966; Smagorinsky, 1969; Lorenz, 1982; Shukla, 1981). Many of the earlier discussions neglected the role of the mean climate drift and even some more recent discussions have neglected the differences between model variability and atmospheric variability.

The 'theoretical predictability time' has been defined as the time needed for the rms forecast spread to reach a given fraction of P_1^a, e.g. $.95P_1^a$ (Kalnay and Dalcher, 1986) and 6/7 P_1^a (Mansfield, 1986). Such a definition ignores the difference between model variability and atmospheric variability and from Figure 8 would draw the conclusion that successful monthly forecasting is already a reality. A definition of 'theoretical predictability time' based on model persistence, e.g. $.95P_1^m$, would be more sensible. An equivalent definition of 'useful predictability time' based on the time taken for the error to reach S_1^m would would make little difference to current estimates, because of the way the slope of the forecast error curve changes with forecast range. Equivalent definitions may also be given for predictability times of 10-day means in terms of S_x^m and P_x^m.

Definitions of theoretical and useful predictability times may also be based on the behaviour of the anomaly correlation. We may describe the time taken for the forecast to reach zero correlation as the theoretical predictability time. Synoptic experience with daily maps indicates that useful skill is lost when the correlation drops below 0.6 (Hollingsworth et al, 1980). For daily maps this usually happens at about the same time as the rms error reaches the level of the climatological standard deviation. Figs 8 and 9 suggest that the same is approximately true for time-averaged fields. If one wishes to work with anomaly correlation, one may define the times of useful and theoretical predictability as the times taken for the correlation to reach 0.6 and 0.0 respectively.

b) <u>Predictive Skill for Time-Averages in the ECMWF Experiments</u>

Lorenz (1982), Shukla (1981) and others have argued that time-averaged or space averaged properties of the atmosphere may be more predictable than instantaneous states. Fig 9 was prepared to test this proposition for 10-day means, although the 10-day mean of the model output is not necessarily the best model-derived predictor for 10-day means (Roads, 1986). The 10-day time averaging produces fields which are also spatially smooth as may be seen by comparing Figs 5 and 7.

Fig 9 shows that the error growth for the 10-day mean forecasts reaches the level S_x^a just a day or two after the forecast error for daily maps reaches S_1^a. Thus the time averaging of these forecasts has brought only a little extra in terms of useful predictive skill. The reason is that the spectrum of the height field is dominated by the larger scales, as is the spectrum of forecast error beyond day 5. Since both the rms and anomaly correlation scores are dominated by the larger, slowly varying, scales, the time averaging has little impact on the useful predictability time as defined using either measure.

The main features of the extended range experiments with the operational ECMWF model are as follows. Occasional very good forecasts have been made in the extended range in certain regions. We have as yet no way to estimate the skill in a given region a-priori; the question is discussed in section 6. The climate drift represents about 20% of the total forecast error variance in the extended range. The atmospheric variability is underestimated by 15-20% if we look at daily or 10-day mean variances. The quality of the medium range forecast is important for the average level of skill in the extended range. Time or spatial averaging of the height fields appears to have a small, though positive, effect on useful forecast skill.

4.4 The UKMO experiment

There has been an on-going experiment on extended range forecasting at the U.K. Meteorological Office for about a decade. The experimental philosophy and preliminary results are reported by Gilchrist (1977, 1981, 1986); the latest results are reported in Mansfield (1986). Mansfield focuses most attention on a set of nine pairs of forecasts, each pair run from successive days and the dates chosen at random between Dec 1 and Jan 15 in the winters of 1978/79 and 1979/80, so the experimental strategy used by Mansfield was also used at ECMWF. The model used is the 5-level hemispheric model described in Corby et al (1977). The starting analyses were operational analyses to 15 N, extrapolated to the equator. The forecasts may therefore be affected by the problems of specifying initial data and boundary conditions in the tropics, as discussed by Somerville (1980) and by Daley et al, (1981).

Fig 10 shows the results of the verification of the first 30 days of the daily forecasts in these cases. The daily forecasts lose useful skill somewhat earlier than in the ECMWF case. This is to be expected in view of the difference in the forecasting systems; there has been considerable progress in medium range forecasting since the development of the UKMO 5-level model. The asymptotic level of the persistence errors in Figs 8 and 10 are fairly similar, so the meteorological situations must be fairly normal in both sets of experiments. On average the daily scores suggest little forecast skill in the extended range.

A notable difference between the UKMO and ECMWF results is seen in the low rate of divergence of the forecasts with the 5-level model. This is one symptom of the climate drift problem discussed at length by Mansfield and is due to an underestimation of atmospheric variability in the model. The fact that the forecast error curve nevertheless exceeds the level of the persistence error suggest that the mean climate drift in the 5-level model must be of the same order of magnitude as the atmospheric variability and therefore substantially larger than the model's variability.

For time mean (15-day) results, Mansfield only shows the anomaly correlation (Fig 11), which indicates that even for the first 15-day mean the useful forecast skill is marginal at best. He comments that the best forecast, with his model, for the mean of the next 15 days may be the 15-day mean formed from the last 8 days of observations and the first 7 days of the dynamical forecast.

Comparing these results with the ECMWF results we see that the main differences in useful skill occur in the medium range, where the more recent model performs better. The older model has a larger climate drift and also a more serious underestimation of atmospheric variability.

Mansfield demonstrates that despite these short-comings his model can nevertheless reproduce some features of the climatology of blocking such as the longitudinal and seasonal variations of the occurrence of persistent features. Mansfield found that the spread between forecasts with starting time separated by one year decreases systematically with time. He used this result in an ingenious way to provide upper bounds on what we have called potential theoretical predictability. The results of his estimates (22 days) are rather different from what one might get from simple interpretations of Fig 10.

4.5 The GFDL experiment

Miyakoda et al (1986) present the results of forecasts with the GFDL N48 model with 'E4' physics for the Jan 1 of each year from 1977 through 1983, together with the case of Jan 16 1979. Their experimental design is different from the ECMWF and UKMO experiments already discussed. Instead of using forecasts from different successive starting dates they generated three forecasts for each case starting from analyses made by GFDL, NMC, and ECMWF.

Unfortunately, these authors do not provide the daily verification results. The verification results for 10-day means are presented in Fig 12. The verification is shown for the ensemble mean of the forecasts from the three different analyses, and is then averaged over the eight cases. The plot shows the actual error of a climatological forecast rather than the expected value as shown in the ECMWF and UKMO results. Both in terms of rms height error and anomaly correlation for height, the 10-day mean forecasts lose useful skill around day 7 to 8 and have the level of a persistence forecast between day 15 and 20. For the 20-day means of these forecasts, useful skill is lost between days 10 and 15 (not shown).

Climate drift contributes more to total error in the GFDL model than in the ECMWF model. Eqs 1 and 2, together with the results in Fig. 9a,b imply that E/S_x^a is at most 0.5 for the EC model, so that the mean error contributes at most 25% to the total error variance of the ECMWF 10-day mean 500mb height forecasts. Miyakoda et al. estimate that the equivalent figure for their 20-day mean forecasts is 64%.

As noted by Gilchrist (1977), Arpe (1982), Shukla (1981), and Molteni et al (1986), it may be advantageous to use the model's departure from its climatology as a forecast for the observed departure from the atmosphere's climatology. Such a strategy is equivalent to correcting the model's climatology. This is a legitimate forecasting procedure if one has independent information on the model's climate, apart from the forecasts being verified, as in Molteni et al (1986). In the case where no independent information on the model's climatology is available, then such a correction of the mean error gives some indication of the level of skill one might expect if the mean error could be eliminated.

Miyakoda et al (1986) found a dramatic impact on their forecast scores when they tried this device. The anomaly correlation of 10-day means increased from about 0.4 to about 0.6 at day 10. The rms error of the 10-day mean forecasts reached the level of a climatological forecast at about day 20 rather than at about day 10. The implications of these results for possible future skill are ambiguous. The rms error, after correction for the mean error, never exceeds the error of the climatological forecast. One must conclude that the GFDL model variability for 10-day means is about 20% lower than the observed value. If the GFDL model under-estimates the atmospheric variance by this amount, one should rely more on the anomaly correlation of the corrected forecasts, rather than on the rms error, for a pointer to what might happen with an improved model which had no climate drift. Since Miyakoda et al did not use independent information on the model's climate drift, this result can only be taken as suggestive rather than conclusive evidence that correction of the climate drift would lead to a useful forecast of the 10-day mean centred on day 10.

Miyakoda et al (1986) use the dispersion of the triplet of forecasts about their mean as a measure of the sensitivity of the forecasts to errors in the initial analyses. They comment that their scatter statistic reached the persistence level for only two out of eight cases at the end of the month, and that it reached the climate level for four out of eight cases. One cannot be precise without detailed calculations, but impression is that the dispersion in their model grows much more slowly than in the experiments with the ECMWF model, which were done for a quite different period.

To summarise, the medium range forecast skill of the ensemble mean of the triplets of forecasts from the GFDL model seems to be roughly the same as in the ECMWF experiments. The GFDL model has a substantial climate drift which contributes 64% to the total error variance of the forecasts for 20-day means. The GFDL model appears to underestimate the atmospheric variability by about the same amount as the ECMWF model. Elimination of the systematic errors leads to corrected forecasts which show useful skill for 10-day means centred on day 10, judged from the anomaly correlation results.

4.6 The Potential for 10-30 Day Forecasts

The results available to date do not permit a reliable estimate of the potential for useful forecast skill beyond day 10. The number of forecasts that have been made so far is small. It is hoped that the number of forecast experiments will increase rapidly in the near future. Estimates of future improvements using current prediction studies or predictability studies are affected by two aspects of the climate drift problem which are common, to a greater or lesser extent, to all models discussed here - the mean climate drift and the under-estimation of atmospheric variability.

This latter problem even inhibits one from using Lorenz's (1982) arguments to estimate the potential for improvements in medium range skill beyond day 5. From an optimistic viewpoint Fig 2 suggests considerable room for improvement expensive to repeat such studies for models with resolutions close to those used operationally. They would provide a useful documentation of many important aspects of model behaviour, but they might not give much help on how to quantify the change in the transient behaviour in the first month of a run. A multi-year integration with resolution at about T40 is an important first step (v. Storch et al, 1985).

It will be essential to take account of these aspects of climate drift in the development of probabilistic methods to provide forecasts beyond the range of deterministic predictability.

6. PROBABILISTIC FORECASTS FOR THE MEDIUM AND EXTENDED RANGE

It has been recognised at least since the work of Epstein (1969) and Gleeson (1970) that the uncertainty of a single forecast or of a group of forecasts is a proper subject for prediction. There is a growing demand from users of medium range forecasts for skill forecasts of some kind (Tennekes et al 1986). Studies of fluctuations in predictive skill are underway, in response to these demands (e.g. Palmer and Tibaldi, 1986). In the extended range it is likely that the most useful forecasts will be couched in probabilistic terms, giving a range of likely atmospheric evolutions with an indication of the likelihood of transition from one weather regime to another. The most practical method of providing such forecasts is through ensembles of forecasts starting from a group of initial states that are reasonably close to each other in some sense (Leith, 1974).

Current research centres on a number of problems: methods to generate the initial states, relationships between ensemble spread and forecast skill, and sensitivity of the results to the models used. A variety of methods have been proposed to generate the set of initial states: white noise perturbations in the mass or wind field (Leith,1974; Pitcher, 1977; Shukla, 1981), geostrophic white noise (Hollingsworth and Savijärvi, 1980), successive analyses separated by a short time (Hoffman and Kalnay, 1983), and perturbations generated by adding a proportion of randomly chosen earlier analyses (Murphy, 1986). The work of Daley (1981) suggests that deep geostrophic perturbations are most effective at inducing rapid spread of the forecasts in mid-latitudes.

Most idealised studies (Leith, 1974; Seidman, 1981; Hayashi, 1986; Hoffman and Kalnay, 1983; Murphy 1986) and real data studies (Pitcher, 1977; Hollingsworth and Savijärvi, 1980; Shukla, 1981; Molteni et al, 1986; Murphy, 1986) have concentrated on forecasts of the time mean state, using an implicit or explicit assumption that the statistics of both atmosphere and model are Gaussian. The possibility of clustering of forecast states has recently been examined by Palmer et al (1986a). Most of these studies have examined the relationship between forecast skill and ensemble spread, with encouraging theoretical results, and some encouraging results from the real-data experiments (Molteni et al, 1986; Murphy, 1986).

As an example we consider a recent experiment carried out at ECMWF using the T63 version of the operational model (Cubasch, Tibaldi and Brankovic, work in progress). A number of extended range forecasts from ensembles of 9 analyses made 6 hours apart and covering a 2 day period were run, following the Lagged Average Forecasting approach, suggested by Hoffman and Kalnay (1983) and discussed by Dalcher et al (1985). Fig 15 shows the variation with starting time of the skill of the 10-day mean forecasts centred on day 5, 10 and 15, all counting being reckoned from the latest starting time. The four examples were chosen from mid-month in May, September, December 1985 and March 1986. The skill of the 10-day mean centred on day 5 is relatively insensitive to starting time. Thereafter there is a strong sensitivity to the starting time, which is particularly marked in the May 1985 and March 1986 cases. For success in medium range forecasting one would like to see a more uniform, and higher skill for the 10-day mean centred on day 10. However, the sensitivity of the forecasts to the initial data gives hope that the ensembles can encompass the full range of possible atmospheric evolutions.

In a similar set of experiments (carried out using a T40 resolution version of the same model) to examine the relation between forecast skill and forecast spread, Molteni et al (1986) stratified the point-wise rms error of the ensemble mean as a function of the spread of the ensemble (Fig 16). An encouraging degree of correlation between low (high) error and low (high) spread is evident in their results.

The prospect of developing probabilistic forecasts based on ensembles of forecasts, in the range beyond the deterministic limit of predictability, must depend heavily on the ability of the model to represent reasonably well the range of atmospheric states and their probability distribution. Without such a capability the problem becomes very difficult, as one would have to allow for the biases in the model's distribution of states. For this reason the elimination of model bias in the mean fields and in the transients is just as critical for probabilistic forecasts at longer range as it is for deterministic forecasts at short and medium range.

error is an important contributor to analysis error (Bengtsson and Simmons, 1982), so improved models lead to improved analyses (Brankovic, 1986). It is noticeable in Fig 1 that recent gains in predictive skill in the medium range depend markedly on the improvement at day 1, where reductions in analysis error must contribute (Arpe et al, 1985). The magnitude of the effect is difficult to quantify, but is obviously important for esimating the gains to be made from improving the model.

5.2 Errors in the Transient Fields

Errors in the transient fields can be documented in terms of the evolution of the eddy kinetic energy of the forecast model. Fig 14 shows such a calculation for the ECMWF forecasts discussed above. It is evident that there is a serious change in the ratio of eddy to zonal kinetic energy in the course of the forecasts to thirty days. The drop in the eddy kinetic energy is independent of horizontal resolution differences between T63 and T106. Changes of similar magnitude have been noted by workers with other models.

The changes in eddy kinetic energy affect both the stationary component and the transient component of the flow. Because of the secular change in the mean state it is difficult to partition the changes in the transients into changes in low-frequency behaviour and changes in high-frequency behaviour. Klinker and Capaldo (1986) identified systematic changes in the high-frequency behaviour of the ECMWF model using ensembles of analyses and of day-3 and day-5 forecasts. Mansfield (1986) investigated the low frequency behaviour of his model by determining the frequency of (an analogue of) blocking action as a function of season and longitude. The data base for this calculation consisted of the 18 winter forecasts already mentioned, plus a further 57 forecasts distributed through the year. Even though the overall frequency of (the analogue of) blocking was much reduced, he found that the distribution of blocking events in longitude and season bore encouraging similarities to nature. Palmer (1987) has shown how the pattern of low-frequency variability is stongly affected by the errors in the mean field. Improvements in the one led to marked improvements in the other.

Multi-year runs of general circulation models have been used successfully for the study of the frequency, distribution and structure of blocking action (Manabe and Hahn, 1981; Lau, 1981; Volmer et al, 1983). These are intrinsically interesting and useful studies. It would be enormously expensive to repeat such studies for models with resolutions close to those used operationally. They would provide a useful documentation of many important aspects of model behaviour, but they might not give much help on how to quantify the change in the transient behaviour in the first month of a run. A multi-year integration with resolution at about T40 is an important first step (v. Storch et al, 1985).

It will be essential to take account of these aspects of climate drift in the development of probabilistic methods to provide forecasts beyond the range of deterministic predictability.

6. PROBABILISTIC FORECASTS FOR THE MEDIUM AND EXTENDED RANGE

It has been recognised at least since the work of Epstein (1969) and Gleeson (1970) that the uncertainty of a single forecast or of a group of forecasts is a proper subject for prediction. There is a growing demand from users of medium range forecasts for skill forecasts of some kind (Tennekes et al 1986). Studies of fluctuations in predictive skill are underway, in response to these demands (e.g. Palmer and Tibaldi, 1986). In the extended range it is likely that the most useful forecasts will be couched in probabilistic terms, giving a range of likely atmospheric evolutions with an indication of the likelihood of transition from one weather regime to another. The most practical method of providing such forecasts is through ensembles of forecasts starting from a group of initial states that are reasonably close to each other in some sense (Leith, 1974).

Current research centres on a number of problems: methods to generate the initial states, relationships between ensemble spread and forecast skill, and sensitivity of the results to the models used. A variety of methods have been proposed to generate the set of initial states: white noise perturbations in the mass or wind field (Leith,1974; Pitcher, 1977; Shukla, 1981), geostrophic white noise (Hollingsworth and Savijärvi, 1980), successive analyses separated by a short time (Hoffman and Kalnay, 1983), and perturbations generated by adding a proportion of randomly chosen earlier analyses (Murphy, 1986). The work of Daley (1981) suggests that deep geostrophic perturbations are most effective at inducing rapid spread of the forecasts in mid-latitudes.

Most idealised studies (Leith, 1974; Seidman, 1981; Hayashi, 1986; Hoffman and Kalnay, 1983; Murphy 1986) and real data studies (Pitcher, 1977; Hollingsworth and Savijärvi, 1980; Shukla, 1981; Molteni et al, 1986; Murphy, 1986) have concentrated on forecasts of the time mean state, using an implicit or explicit assumption that the statistics of both atmosphere and model are Gaussian. The possibility of clustering of forecast states has recently been examined by Palmer et al (1986a). Most of these studies have examined the relationship between forecast skill and ensemble spread, with encouraging theoretical results, and some encouraging results from the real-data experiments (Molteni et al, 1986; Murphy, 1986).

As an example we consider a recent experiment carried out at ECMWF using the T63 version of the operational model (Cubasch, Tibaldi and Brankovic, work in progress). A number of extended range forecasts from ensembles of 9 analyses made 6 hours apart and covering a 2 day period were run, following the Lagged Average Forecasting approach, suggested by Hoffman and Kalnay (1983) and discussed by Dalcher et al (1985). Fig 15 shows the variation with starting time of the skill of the 10-day mean forecasts centred on day 5, 10 and 15, all counting being reckoned from the latest starting time. The four examples were chosen from mid-month in May, September, December 1985 and March 1986. The skill of the 10-day mean centred on day 5 is relatively insensitive to starting time. Thereafter there is a strong sensitivity to the starting time, which is particularly marked in the May 1985 and March 1986 cases. For success in medium range forecasting one would like to see a more uniform, and higher skill for the 10-day mean centred on day 10. However, the sensitivity of the forecasts to the initial data gives hope that the ensembles can encompass the full range of possible atmospheric evolutions.

In a similar set of experiments (carried out using a T40 resolution version of the same model) to examine the relation between forecast skill and forecast spread, Molteni et al (1986) stratified the point-wise rms error of the ensemble mean as a function of the spread of the ensemble (Fig 16). An encouraging degree of correlation between low (high) error and low (high) spread is evident in their results.

The prospect of developing probabilistic forecasts based on ensembles of forecasts, in the range beyond the deterministic limit of predictability, must depend heavily on the ability of the model to represent reasonably well the range of atmospheric states and their probability distribution. Without such a capability the problem becomes very difficult, as one would have to allow for the biases in the model's distribution of states. For this reason the elimination of model bias in the mean fields and in the transients is just as critical for probabilistic forecasts at longer range as it is for deterministic forecasts at short and medium range.

Several recent workshops have been devoted to this topic as a whole. Current activity in the area is accelerating and it will profit from collaboration between the modelling and academic community. Work is underway to build a large database of extended range runs at several centres, using different models and analysis systems. The choice of cases is coordinated so as to extract the maximum benefit from the experiments. The runs will be of two types- regular monthly or quarterly runs at full operational resolution and ensembles of runs with clustered starting conditions, probably at somewhat lower than operational resolution. It is not absolutely clear at this stage that the best short or medium range forecast model is necessarily also the best model for extended range forecast studies. Miyakoda et al's (1983) result for the 1977 case suggests that one should use the very best models available. On the other hand, the experience of Molteni et al (1986), Palmer et al (1986a), and others shows that a great deal of valuable development and experimental work can be done at lower resolution. Extended range runs with operational models are essential for model development work. Since they would be made for that reason alone, it makes sense to verify them and use them as controls. The ensembles of runs at lower resolution will provide the raw material for studies of probabilistic forecasting and will provide experimental material to test the value of the statistical tools under development.

7. SUMMARY AND DISCUSSION

There is a great deal of scientific interest in the possibility of producing useful deterministic or probabilistic forecasts in the range beyond seven days. There are some theoretical indications that useful deterministic forecasts may be possible as far as 10-14 days ahead, given improvements in models and in the quality of the initial analyses. Beyond the limit of deterministic predictability, the available evidence suggests that useful probabilistic forecasts may be possible. Such an approach would use ensembles of forecasts and the skill would be dependent on the ability of models to represent the range of possible atmospheric states and the transitions between them. The isolated but remarkable 30-day forecasts that have been produced at several centres gives good grounds for expecting that such a capability is possible.

The central problem in both the medium and extended range is to forecast blocking action. Much more work is needed on diagnosis, on theory and on modelling of the phenomenon, before we can understand the ingredients necessary for forecast success.

The elimination of model bias is seen as a central problem for improving both deterministic and probabilistic forecasts. The progress of recent years in identifying physical processes (orography, gravity wave drag, shallow convection) which were not represented or seriously mis-represented in models gives grounds for expecting that the pace of progress will be maintained. The ultimate driving force comes from the operational forecasters who need to able to give their customers both a better product, and a sensible estimate of the quality of the product on a case dependent basis.

REFERENCES

Arpe, K., 1983: Diagnostic evaluation of analysis and forecasts: climate of the model. ECMWF Seminar/Workshop on Interpretation of Numerical Weather Prediction Products, 13-24 September 1982.

Arpe, K., A. Hollingsworth, M.S. Tracton, A.C. Lorence, S. Uppala, and P. Kållberg, 1985: The response of numerical weather prediction systems to FGGE level IIB data. Part II. Forecast verification in implications for predictability. Quart.J.R.Met.Soc., 111, 67-102.

Arpe, K. and E. Klinker, 1986: Systematic errors of the ECMWF operational forecasting model in mid-latitudes. Quart.J.R.Met.Soc., 112, 181-202.

Austin, J.F., 1980: The blocking of middle latitude westerly winds by planetary waves. Quart.J.R.Met.Soc., 106, 327-350.

Bengtsson, L., 1985: Medium-range forecasting at the ECMWF. Advances in Geophysics, Vol. 28, Ed. by S. Manabe, 3-54.

Bengtsson, L., and A. Lange, 1981: Results of the WMO/CAS numerical weather prediction data study and intercomparison project for forecasts for the northern hemisphere in 1979-80. WMO PSMP Report Series No.1, WMO, Geneva.

Bengtsson, L., and A.J. Simmons, 1982: Medium range weather prediction - operational experience at ECMWF. In "Large-Scale Dynamical Processes in the Atmosphere". (B.J. Hoskins and R.P. Pearce, eds.) pp.337-363. Academic Press, New York.

Benzi, R., A. Speranza and A. Sutera, 1986a: A minimal baroclinic model for the statistical properties of low-frequency variability. J.Atmos.Sci., 43, 2962-2967.

Benzi, R., P. Malguzzi, A. Speranza and A. Sutera, 1986b: The statistical properties of general atmospheric circulation: observational evidence and a minimal theory of bimodality. Quart.J.R.Met.Soc., 112, 661-674.

Bettge, T.W., 1983: A systematic error comparison between the ECMWF and NMC prediction models. Mon.Wea.Rev., 111, 2385-2389.

Boville, B.A., 1984: The influence of the polar night jet on the tropospheric circulation in a GCM. J.Atmos.Sci., 41, 1132-142.

Brankovic, C. 1986: Zonal diagnostics of the ECMWF 1984-85 operational analyses and forecasts. ECMWF Tech. Rep. 57, Reading, U.K., 72 pp.

Charney, J.G., R.G. Fleagle, H. Riehl, V.E. Lally and D.Q. Wark, 1966: The feasibility of a global observation and analysis experiment. Bull.Amer.Meteor.Soc., 47, 200-220.

Charney, J.G., and J.G. DeVore, 1979: Multiple flow equilibria and blocking. J.Atmos.Sci., 36, 1205-1216.

Charney, J.G., D.M. Straus, 1980: Form-drag instability, multiple equilibria and propagating planetary waves in baroclinic, orographically forced, planetary wave system. J.Atmos.Sci., 37, 1157-1175.

Chervin, R.M., 1986: Interannual variability and seasonal climate predictability.. J.Atmos.Sci., 43, 233-251.

Chouinard, C., M. Beland and N. McFarlane, 1986: A simple gravity wave drag parametrization for use in medium-range weather forecast models. Atmosphere-Ocean, 24, 91-110.

Corby, G.A., A. Gilchrist and P.R. Rowntree, 1977: United Kingdom Meteorological Office five-level general circulation model. Methods in Comp. Physics, 17, 67-110.

Daley, R., 1981: Predictability experiments with a baroclinic model. Atmosphere-Ocean, 19, 77-89.

Daley, R., J. Tribbia and D.L. Williamson, 1981: The excitation of large-scale free Rossby waves in numerical weather prediction. Mon.Wea.Rev., 109, 1836-1861.

Dalcher, A., E. Kalnay, R. Livezey and R.N. Hoffmann, 1985: Medium range lagged average forecasts. Preprints from the Ninth AMS Conference on probability and statistics in atmospheric sciences. Virginia Beach, October 1985, 130-136.

Dole, R.M. and N.D. Gordon, 1983: Persistent anomalies of the extratropical northern hemisphere wintertime circulation: geographical distribution and regional persistence characteristics. Mon.Wea.Rev., 111, 1567-1586.

Egger, J., 1981: Stochastically driven large-scale circulation with multiple equilibria. J.Atmos.Sci., 38, 2606-2618.

Epstein, E.S., 1969: Stochastic dynamic prediction. Tellus, 21, 739-759.

Fawcett, E.B., 1969: Systematic errors in operational baroclinic prognoses at the National Meteorological Center. Mon.Wea.Rev., 97, 670-682.

Gilchrist, A. 1977: An experiment on extended range prediction using a general circulation model and including the influence of sea surface temperature anomalies. Beitr.Phys.Atmos., 50, 25-40.

Gilchrist, A., 1981: Long-range forecasting in the Meteorological Office: Part I. ECMWF Seminar on Problems and Prospects in Long and Medium Range Forecasting, 14-18 September 1981, 21-54.

Gilchrist, A., 1986: Long-range forecasting. Quart.J.R.Met.Soc., 112, 567-592.

Gleeson, T.A., 1970: Statistical-dynamical predictions. J.Appl.Meteor., 9, 333-344.

Green, J.S.A., 1977: The weather during July 1976: some dynamical considerations of the drought. Weather, 32, 120-128.

Grønaas, S., 1982: Systematic errors and forecast quality of ECMWF forecasts in different large-scale flow patterns. ECMWF Seminar/Workshop on Interpretation of Numerical Weather Prediction Products, 13-24 September 1982. ECMWF, Shinfield Park, Reading, UK., 161-206.

Grønaas, S. 1985: A pilot study on the prediction of medium range forecast quality. ECMWF Tech.Memo.No.119, 23 pp.

Haines, K. and J.M. Marshall, 1987: Eddy forced coherent structure as a prototype of atmospheric blocking, Quart.J.R.Met.Soc., to appear.

Hansen, A.R., 1986: Observational characteristics of atmospheric planetary waves with bimodal amplitude distribution. Advances in Geophysics, ed. by S. Manabe, Vol. 29. Academic Press.

Hansen, A.R. and A. Sutera, 1986: On the probability density distribution of planetary-scale atmospheric wave amplitude. J.Atmos.Sci., 43, 3250-3265.

Hayashi, Y., 1986: Statistical interpretations of ensemble-time mean predictability. J.Meteor.Soc.Japan., 64, 167-181.

Heckley, W.A., 1985: Systematic errors of the ECMWF operational forecasting model in the tropics. Quart.J.R.Met.Soc., 111, 709-738.

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

Hollingsworth, A. and H. Savijarvi, 1980: An experiment in Monte Carlo forecasting. WMO Symposium on Probabilistic and Statistical Methods in Weather Forecasting, Nice, 8-12 September 1980, 45-47.

Hollingsworth, A., K. Arpe, A.C. Lorenc, M.S. Tracton, G. Cats, S. Uppala and P. Kallberg, 1985: The response of numerical weather prediction systems to FGGE level II-b data. Part I. Analysis. Quart.J.Met.Soc., 111, 1-66.

Hoffman, R.N. and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. Tellus, 35A, 100-118.

Hoskins, B.J., I.N. James and G.H. White, 1983: The shape, propagation and mean-flow interaction of large-scale weather systems, J.Atmos.Sci., 40, 1595-1612.

Hoskins, B.J., M.E. McIntyre and A.W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. Quart.J.R.Met.Soc., 111, 877-946.

Illari, L., 1984: A diagnostic study of the potential vorticity in a warm blocking anticyclone. J. Atmos. Sci., 41, 3518-3526.

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

Jarraud, M., 1987: Sensitivity of the operational ECMWF model to horizontal resolution. ECMWF Workshop on Predictability in Medium and Extended Range, Reading, 17-19 March, 1986, 109-154.

Jarraud, M., A.J. Simmons and M. Kanamitsu, 1985: Development of the high resolution model. ECMWF Tech.Memo.107, 61 pp.

Kallen, E. 1981: Bifurcation properties of quasi-geostrophic, barotropic models and their relation to atmospheric blocking. Tellus, 34, 255-265.

Kalnay, E. and A. Dalcher, 1987: Forecasting forecast skill. Submitted to Mon.Wea.Rev. 115, 349-356.

Kalnay, E., R. Livezey, J. Pfaendtner and Y. Sud, 1987: GCM studies on medium and long range forecasting at GLA. ECMWF Workshop on Predictability in the Medium and Extended Range, Reading, 17-19 March, 1986, 389-414.

Klinker, E. and M. Capaldo, 1986: Systematic errors in the baroclinic waves of the ECMWF model. Tellus, 38A, 215-235.

Lange, A., and E. Hellsten, 1983: Results of the WMO/CAS NWP data study and intercomparison project for forecasts for the northern hemisphere in 1981-82. WMO PSMP Report Series No.2, WMO, Geneva.

Lange, A., and E. Hellsten, 1985: Results on the WMO/CAS NWP data study and intercomparison project for forecasts for the northern hemisphere in 1984. WMO PSMP Report Series No.16, WMO, Geneva.

Lau, N.-C., 1981: A diagnostic study of recurrent meteorological anomalies appearing in a 15-year simulation with a GFDL general circulation model. Mon.Wea.Rev., 109, 2287-2311.

Leith, C.E., 1974: Theoretical skill of Monte Carlo forecasts. Mon.Wea.Rev., 102, 409-418.

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

Mahlman, J.D., 1979: Structure and interpretation of blocking anticyclones as simulated in a GFDL general circulation model. In Proceedings of the Thirteenth Stanstead Seminar (T.Warn, ed.) McGill Univ., Publ. in Met., 123, 70-76.

Malguzzi, P. and A. Speranza, 1981: Local multiple equilibria and regional atmospheric blocking. J.Atmos.Sci., 38, 1939-1948.

Manabe, S., and D. Hahn, 1981: Simulation of atmospheric variability. Mon.Wea.Rev., 109, 2260-2286.

Mansfield, D.A., 1986: The skill of dynamical long-range forecasts, including the effect of sea surface temperature anomalies, Quart.J.R.Met.Soc., 112, 1145-1176.

McWilliams, J.C. 1980: An application of equivalent modons to atmospheric blocking. Dyn.Atmos.Oceans, 5, 43-66.

Miller, M.J., and T.N. Palmer, 1987: Orographic gravity-wave drag: its parameterisation and influence in general circulation and numerical weather prediction models. ECMWF Seminar/Workshop on Observation, Theory and Modelling of Orographic Effects, 15-20 September 1986, Vol. 1. ECMWF, Reading, U.K. 283-333.

Miyakoda, K., G.D. Hembree, R.F. Strickler and I. Shulman, 1972: Cumulative results of extended forecast experiments. 1. Model performance for winter cases. Mon.Wea.Rev., 100, 836-855.

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, 846-869.

Miyakoda, K., Sirutus and T. Knutson, 1987: Experimental 30-day forecasting at GFDL. Proceedings of ECMWF Workshop on Predictability in the Medium and Extended Range, Reading, UK, 17-19 March, 1986, 17-50.

Molteni, F., U. Cubasch and S. Tibaldi, 1986: 30- and 60-day forecast experiments with the ECMWF spectral models. ECMWF Workshop on Predictability in the Medium and Extended Range, Reading, 17-19 March, 1986, 51-108.

Opsteegh, J.D. and H.M. van den Dool, 1979: A diagnostic study of the time-mean atmosphere over northwestern Europe during winter. J.Atmos.Sci., 36, 1862-1879.

Palmer, T.N. and S. Tibaldi, 1986: Forecast skill and predictability. ECMWF Tech.Memo.No.127, 40 pp.

Palmer, T.N., 1987: Modelling low-frequency variability of the atmosphere. In "Atmospheric and Oceanic Variability" ed. H. Cattle. Pub. Roy.Met.Soc., Bracknell, 75-104.

Palmer, T.N., J.M. Murphy and J.A. Owen, 1987: Results from experimental extended range forecasts of the UK Meteorological Office. ECMWF Workshop on Predictability in the Medium and Extended Range, Reading, 17-19 March, 1986, 155-182.

Palmer, T.N., G.J. Shutts and R. Swinbank, 1986b: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization. Quart.J.R.Met.Soc., 112, 1001-1039.

Pierrehumbert, R.T. and P. Malguzzi, 1984: Forced coherent structures and local multiple equilibria in a barotropic atmosphere. J.Atmos.Sci., 41, 246-257.

Pitcher, E.J., 1977: Application of stochastic dynamic prediction to real data. J.Atmos.Sci., 34, 3-21.

Ramanathan, V., E.J. Pitcher, R.C. Malone and M.L. Blackmon, 1983: The response of a spectral general circulation model to refinements in radiative processes. J.Atmos.Sci., 40, 605-630.

Reinhold, B.B. and R.T. Pierrehumbert, 1982: Dynamics of weather regimes: quasi-stationary waves and blocking. Mon.Wea.Rev., 110, 1105-1145.

Ritter, B., 1984: The impact of an alternative treatment of infrared radiation on the performance of the ECMWF forecast model. Proceedings of the IAMAP International Radiation Symposium, pub Deepak Publ, ed. G. Fiocco, 277-280.

Roads, J.O., 1986: Forecasts of time averages with a numerical weather prediction model. J.Atmos.Sci., 43, 871-892.

Seidman, A.N., 1981: Averaging techniques in long-range weather forecasting. Mon.Wea.Rev., 109, 1367-1379.

Shukla, J., 1981: Dynamical predictability of monthly means. J.Atmos.Sci., 38, 2547-2572.

Shukla, J., 1984: Predictability of time averages. Part II. The influence of the boundary forcings. In Problems and prospects in long and medium range weather forecasting. (D.M. Burridge and E. Kallen, eds), 155-206. Springer-Verlag, Berlin and New York.

Shukla, J., 1985: Predictability. Advances in Geophysics, 28, Academic Press, 87-121.

Shutts, G.J., 1983: Parameterization of travelling weather systems in a simple model of large-scale atmospheric flow. Advances in Geophysics, 25, Theory of Climate, Academic Press, 117-169.

Shutts, G.J., 1986: A case study of eddy forcing during an Atlantic blocking episode. Advances in Geophysics, 29, 135-164.

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

Speranza, A., 1986: Deterministic and statistical properties of northern hemisphere middle latitude circulation: minimal theoretical models. Advances in Geophysics. 29, Academic Press, 199-266.

Sommerville, R.C.J., 1980: Tropical influences on the predictability of ultralong waves. J.Atmos.Sci., 37, 1141-1156.

Stern, M.E., 1975: Minimal properties of planetary eddies. J.Mar.Res., 33, 1-13.

Storch, E. von, E. Roeckner and U. Cubasch, 1985: Intercomparison of extended range January simulations with General Circulaion Models: Statistical Assessment of Ensemble Properties. Beitr.Phys.Atm., 58, 477-497.

Strüfing, R. 1982 Some comparisons between linear and non-linear horizontal diffusion schemes for the ECMWF grid-point model. ECMWF Tech Memo 60, 42 pp.

Sutera, A., 1986: Probability density distribution of large scale atmospheric flow. Advances in Geophysics, 29, 319-338.

Tennekes, H., A.P.M. Baede and J.D. Opsteegh, 1987: Forecasting forecast skill. ECMWF Workshop on Predictability in the Medium and Extended Range, 17-19 March, 1986. Reading, UK, 277-302.

Tibaldi, S., 1986: Envelope orography and maintenance of the quasi-stationary circulation in the ECMWF global models. Advances in Geophysics, 29, 339-374.

Tiedtke, M. and J. Slingo, 1985: Development of the operational parameterization scheme. ECMWF Tech.Memo.No.108, 38 pp.

Tiedtke, M., W.A. Heckley, J. Slingo, A. Simmons, M. Jarraud, G. Sommeria, 1987: Tropical forecasts at ECMWF: Impact of revised physics and increased horizontal resolution. Submitted to Quart.J.R.Met.Soc.

Trevisan, A. and A. Buzzi, 1980: Stationary response of barotropic weakly nonlinear Rossby waves to quasi-resonant orographic forcing. J.Atmos.Sci., 37, 947-957.

Tribbia, J.J., 1986: The predictability of monthly mean teleconnection patterns. To appear in the "Proceedings in the Pontifical Academy Study Week on "Persistent Meteo-Oceanographic Anomalies and Teleconnections". Rome 23-24 September 1986, Pout.Ac.Sc.Scripta Vazia.

Volmer, J.-P., M. Deque and M. Jarraud, 1983: Large-scale fluctuations in a long range integration of the ECMWF spectral model. Tellus, 35A, 173-188.

Wallace, J.M. and J.K. Woessner, 1981: An analysis of forecast error in the NMC hemispheric primitive equation model. Mon.Wea.Rev., 109, 2444-2449.

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.R.Met.Soc., 109, 683-717.

Wallace, J.M. and J. Quanrong, 1987: On the observed structure of the interannual variability of the atmosphere-ocean climate system. Atmospheric and Oceanic Variability, ed. H. Cattle. Roy.Met.Soc., Bracknell, 17-44.

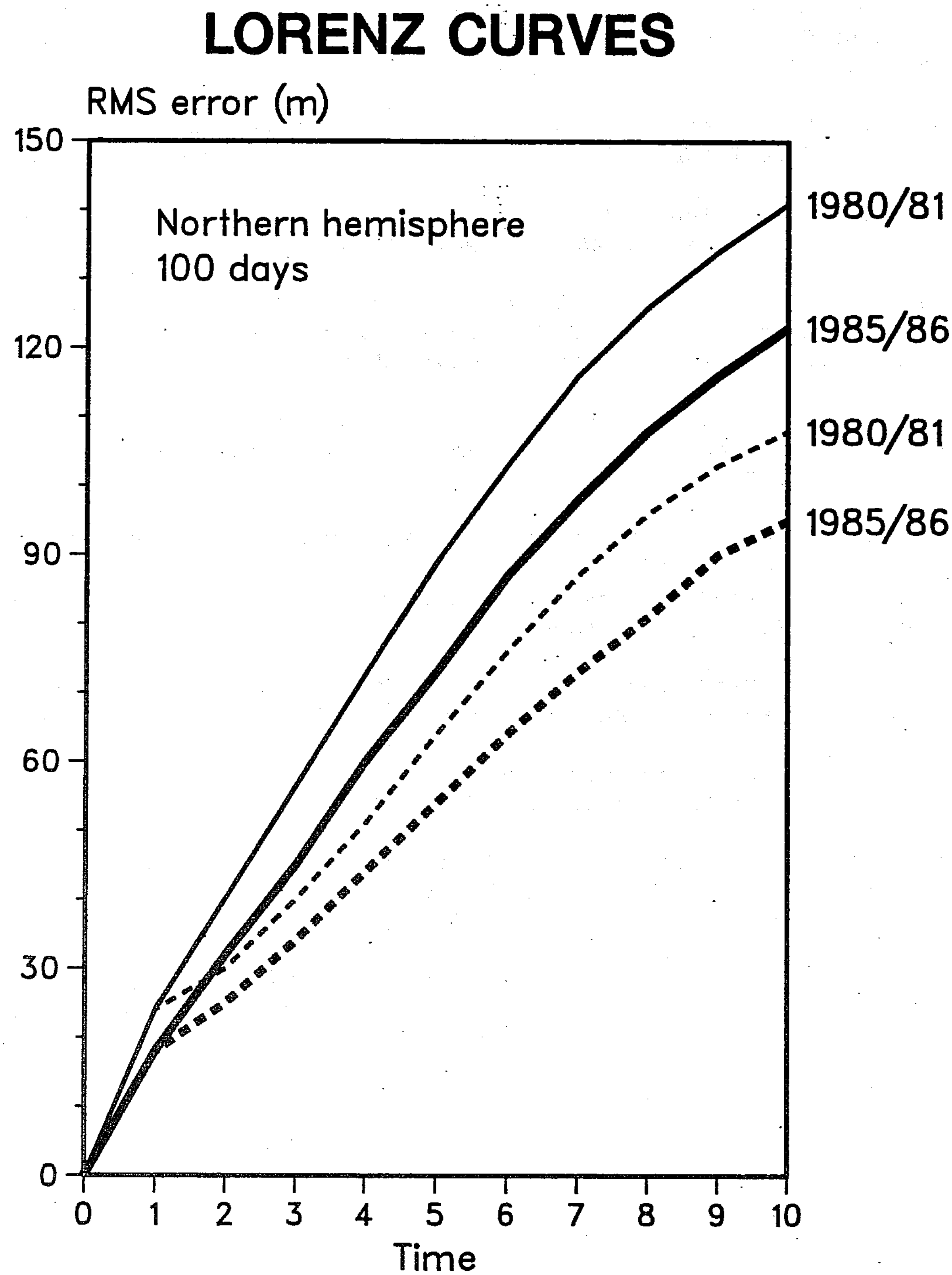


Fig 1 Northern Hemisphere rms 500mb height forecast errors for 100 10-day forecasts from 1 Dec 1980 (thin solid line, 1980/81), and the corresponding errors for 100 forecasts from Dec 1 1985 (thick solid line, 1985/86). Also shown are the spread of successive forecasts for the first of these winters (thin dashed line, 1980-81) and for the second winter (thick dashed line 1985/86).

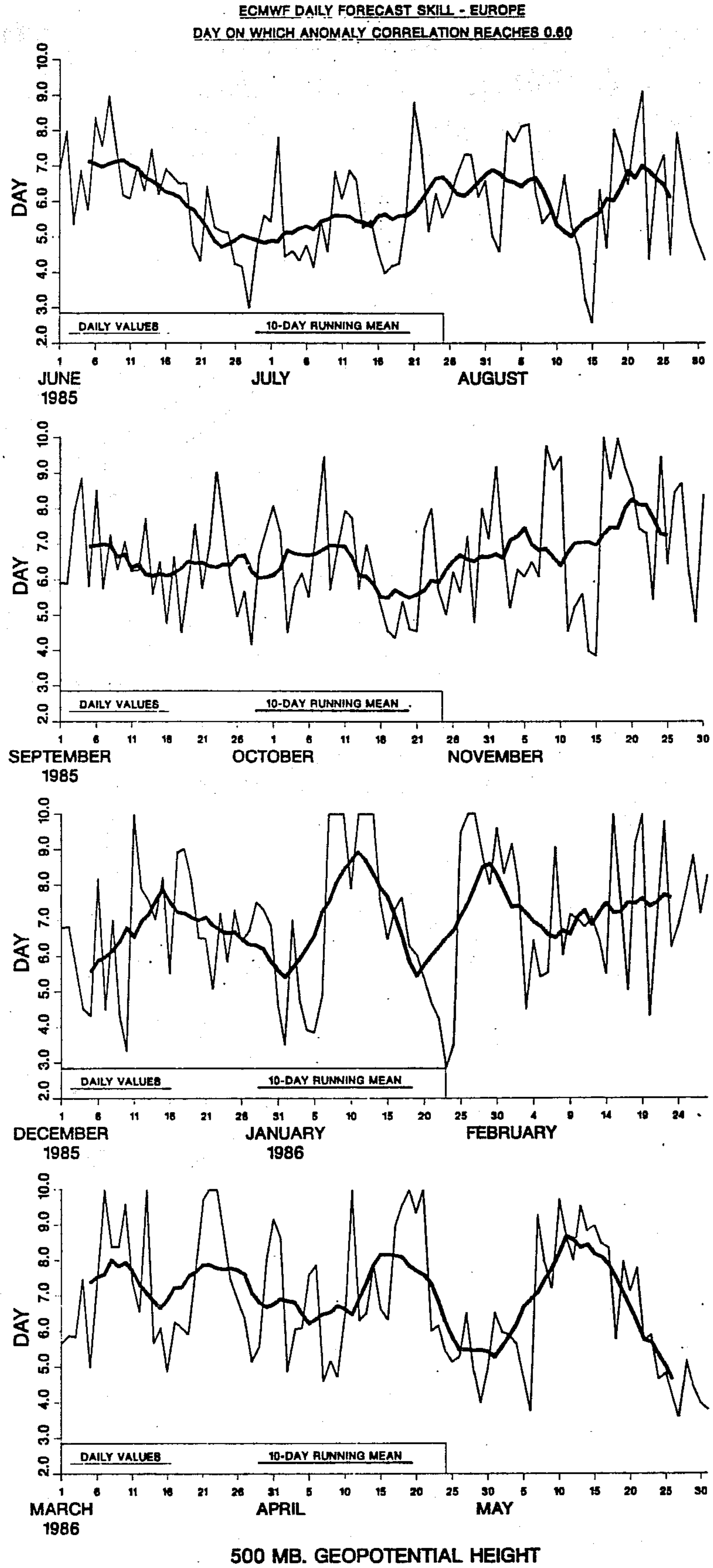


Fig 2 Quarterly time series of the day on which the anomaly correlation for Europe reached 0.6, for ECMWF forecasts between June 1985 and May 1986. The 10-day running mean of the score is also shown for each quarter.

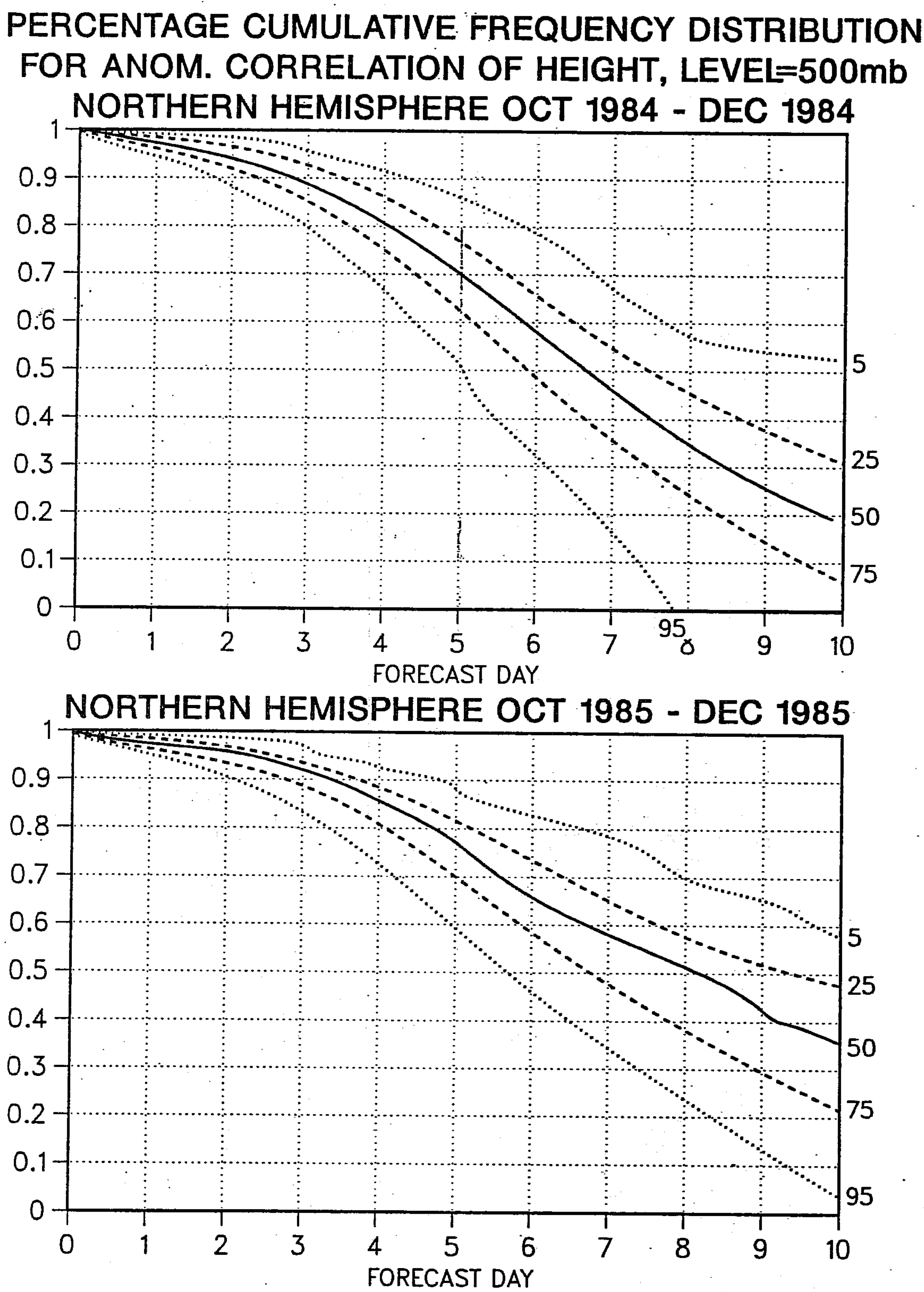
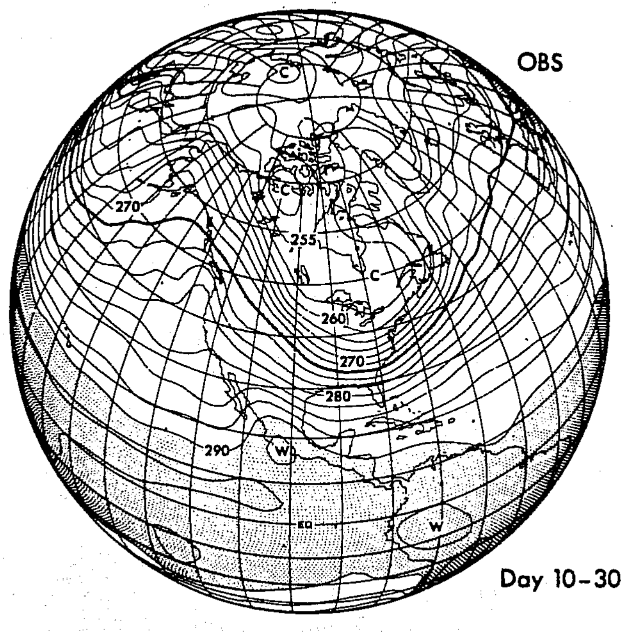


Fig 3 Dependence on forecast range of the cumulative probability distribution of Northern Hemisphere anomaly correlation scores for operational ECMWF forecasts for a) Oct-Dec 1984, b) Oct-Dec 1985.



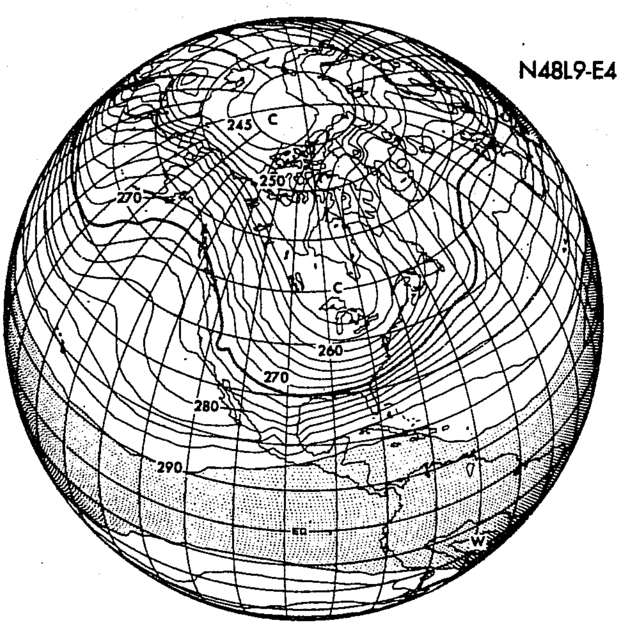
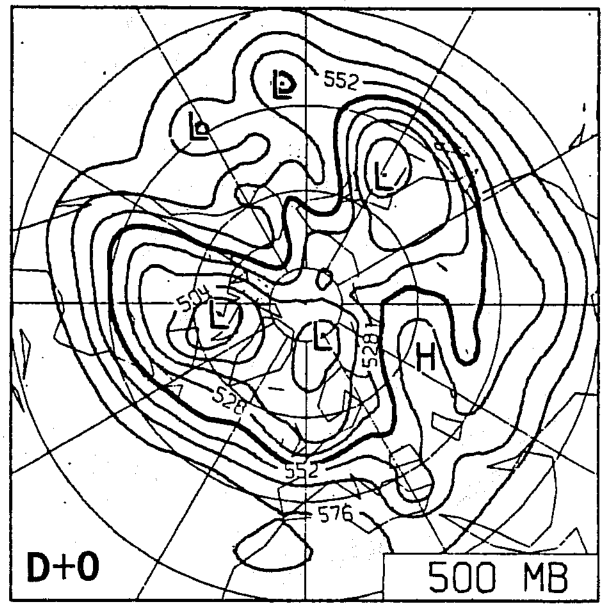


Fig 4 The lower panel shows the 20-day mean forecast 850mb temperature centred on Jan 20 1977 for a GFDL forecast made from Jan 1 1977; the verifying observed state is shown in the upper panel, (courtesy of Miyakoda et al 1983).

ANALYSIS

a)



17 January 1984

Fig 5 a) Northern Hemisphere 500mb analysis for 1200UT on 17 Jan 1984. Panels b)-g) show the observed (left) and forecast (right) 500mb height fields at 4 day intervals starting at day 6 (i.e. days 6,10,14,18,22,26) of the ECMWF forecast starting from 1200UT on Jan 17 1984.

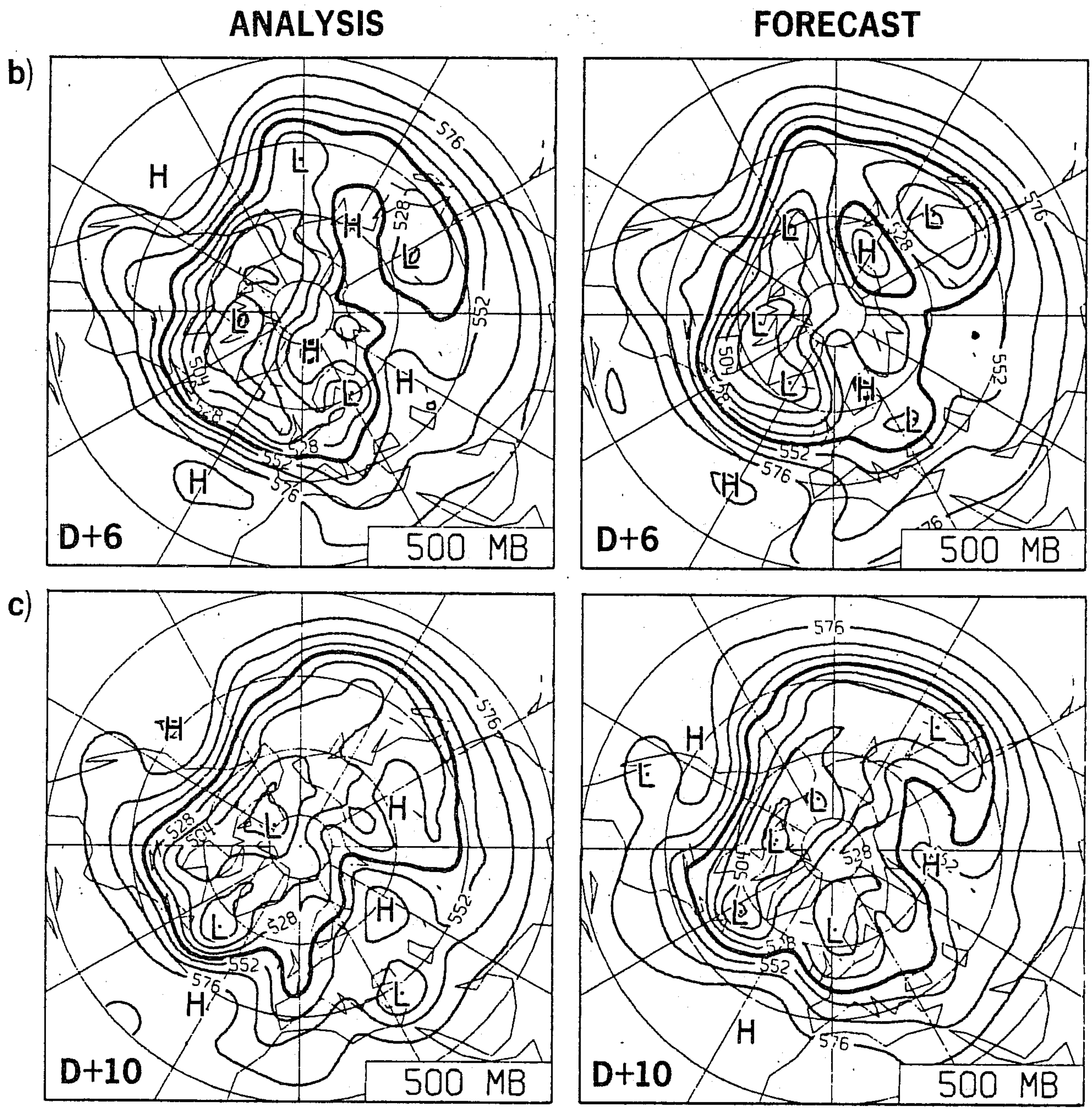


Fig 5 cont

ANALYSIS FORECAST

d)

e)

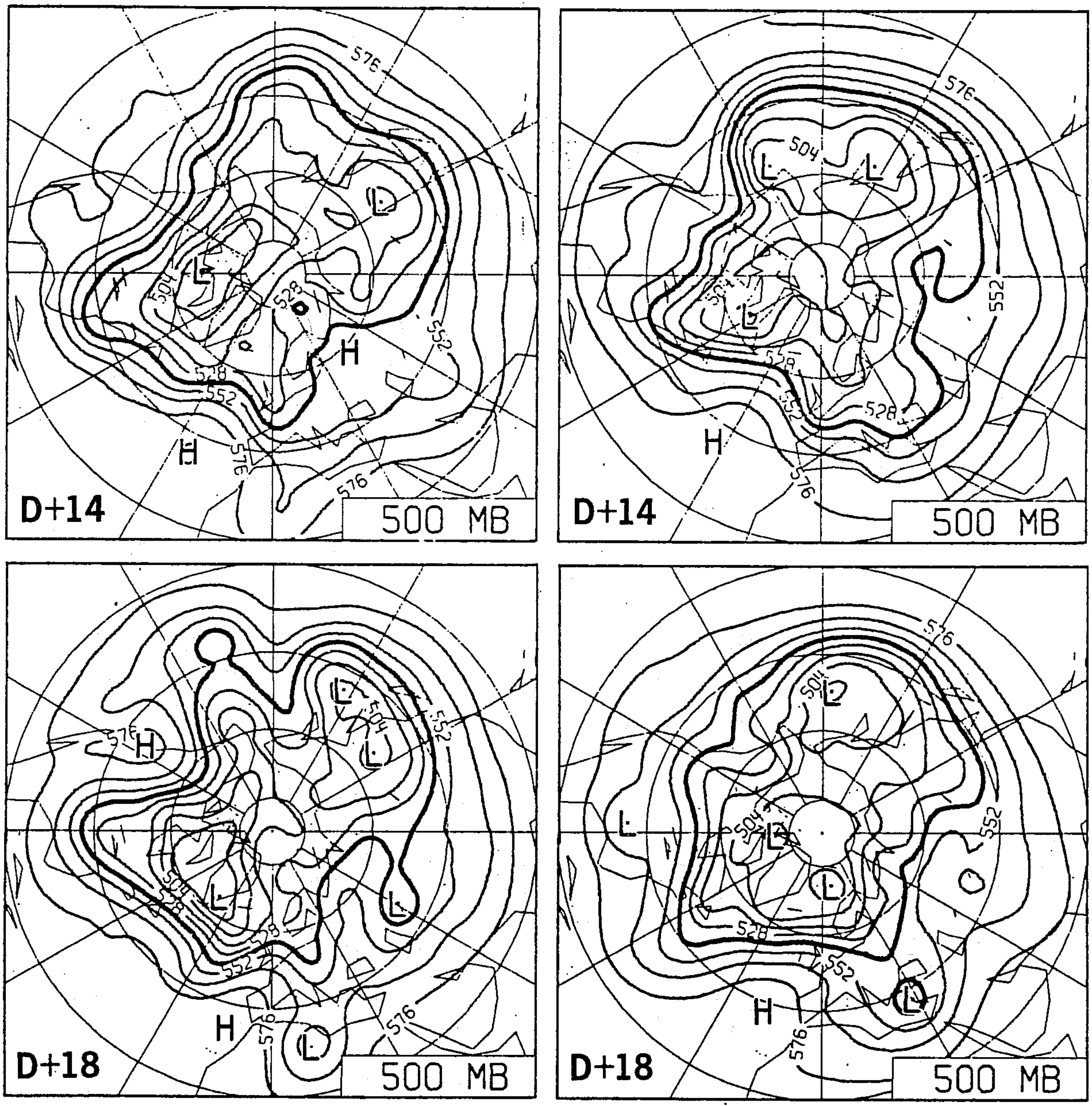


Fig 5 cont

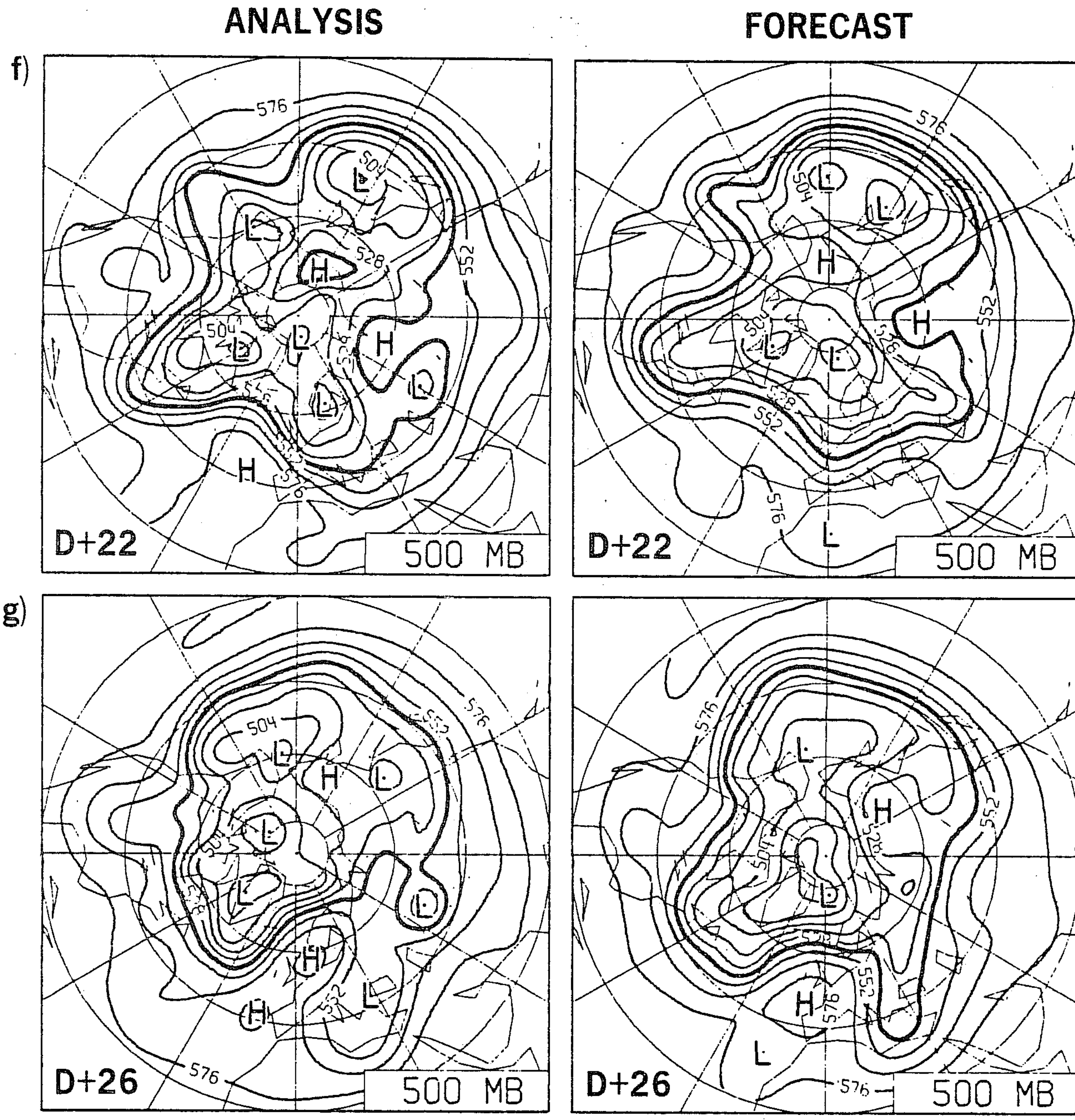

ANALYSIS
FORECAST
f)
D+22
500 MB
D+22
500 MB
g)
D+26
500 MB
D+26
500 MB

Fig 5 cont

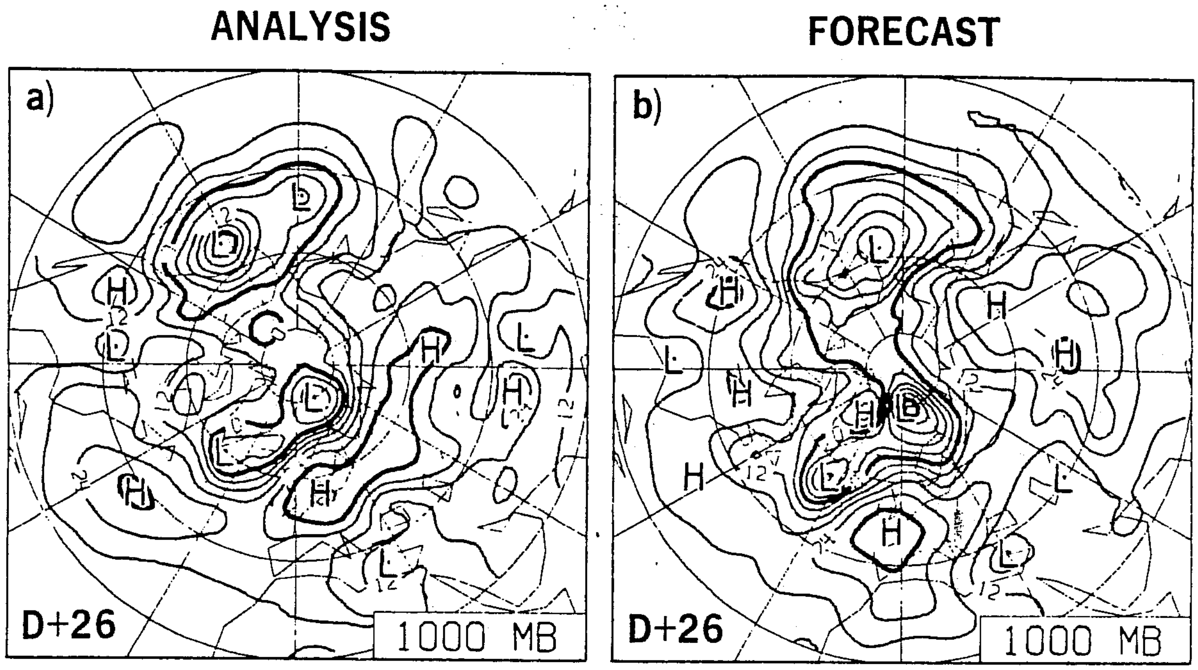


Fig 6 The observed (a) and forecast (b) 1000mb field at day 26 of the forecast from 1200UT, Jan 17 1984.

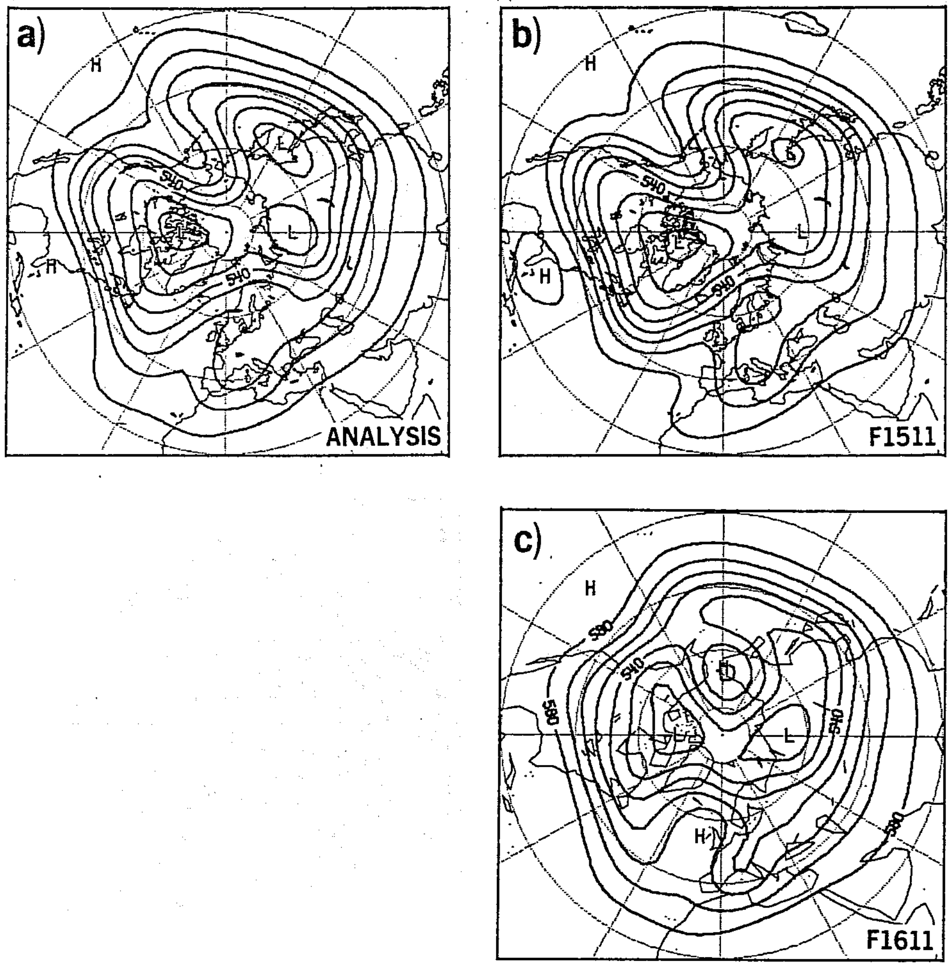


Fig 7 Time evolution of observed (top left) and forecast 10-day mean fields at 500mb in forecasts from 1200UT Nov 15 1985 (top right F1511) and from 1200UT Nov 16 1985 (bottom, F1611) for days 1-10 (a-c), days 6-15 (d-f), days 11-20, (g-i) and days 16-25 (j-l).

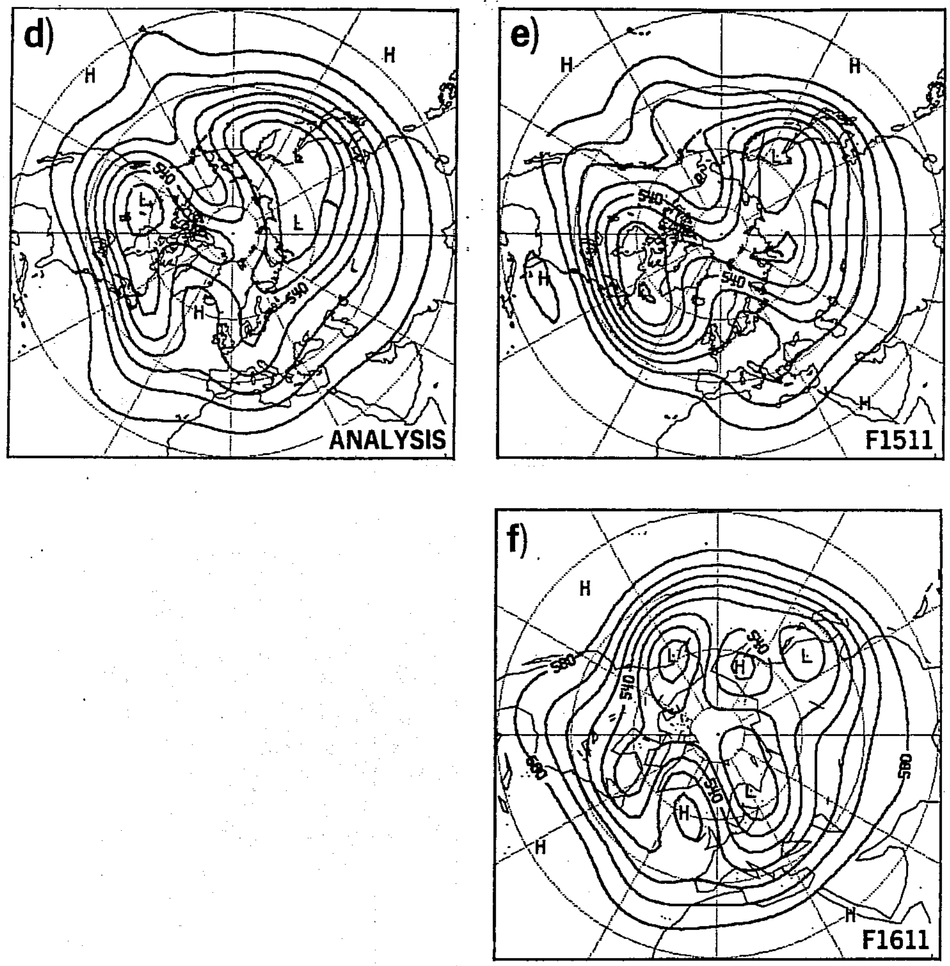


Fig 7 cont

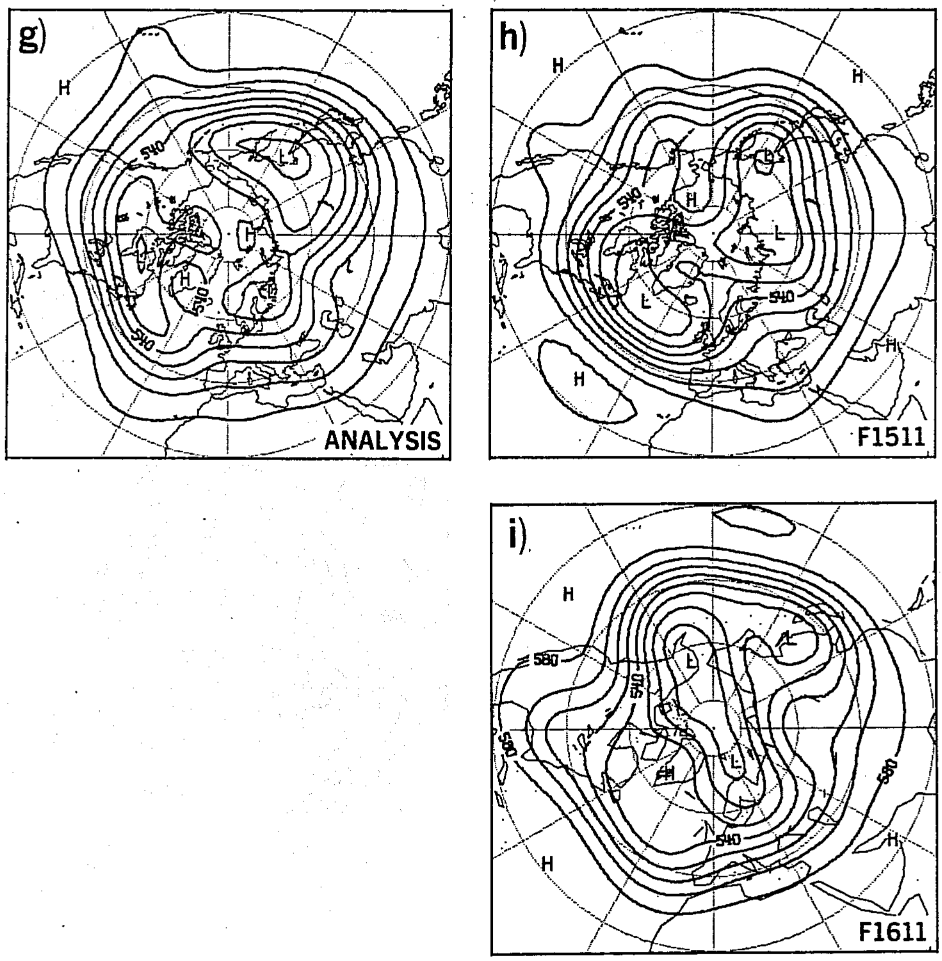


Fig 7 cont

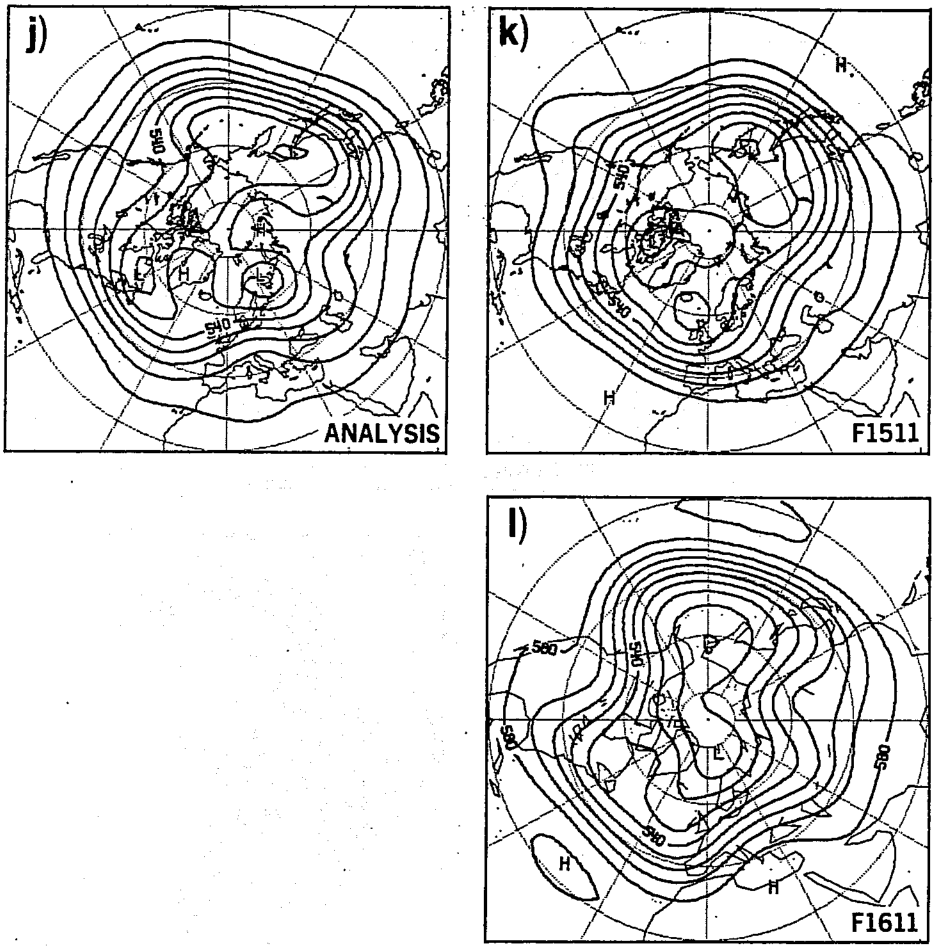


Fig 7 cont

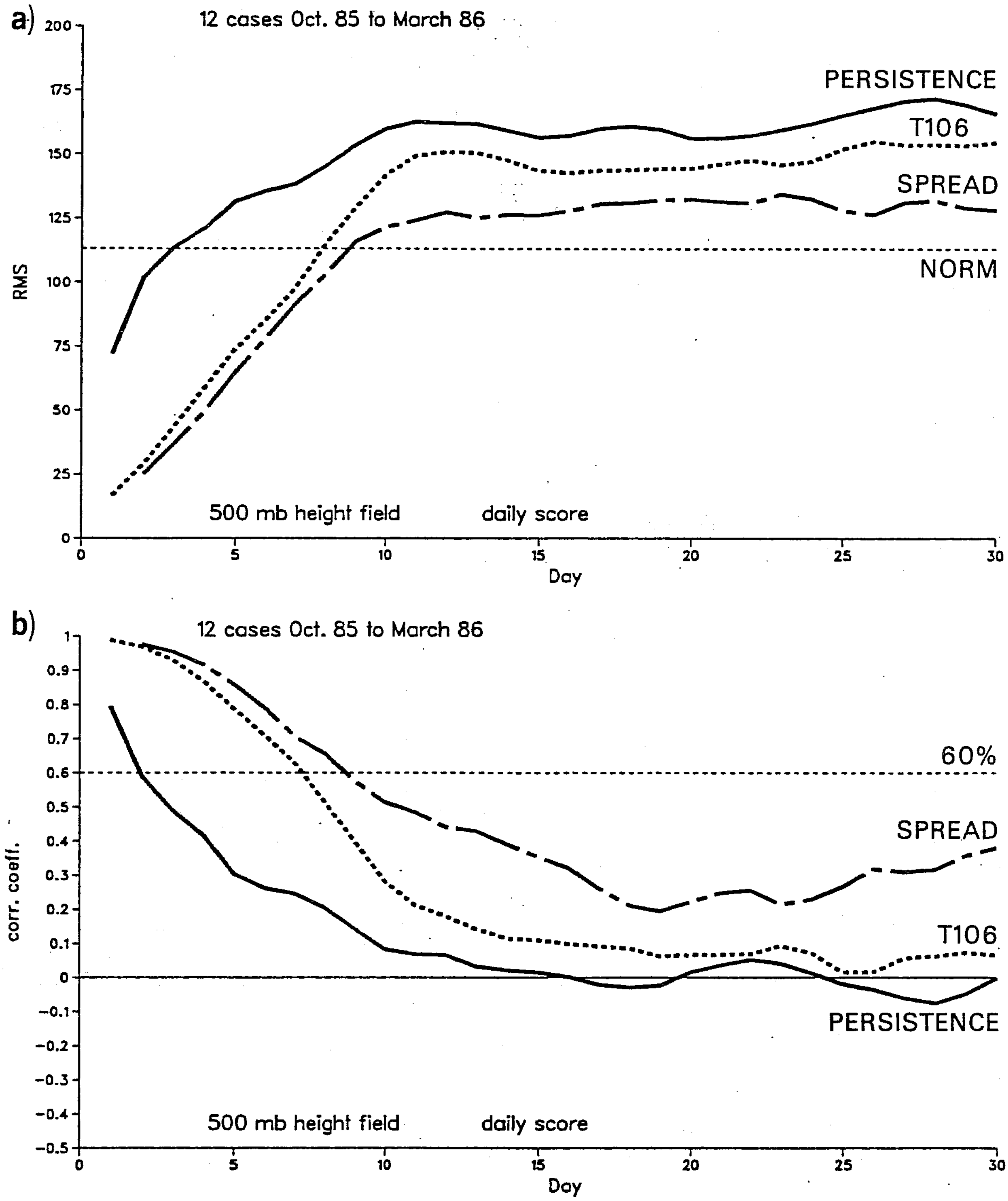


Fig 8 Daily verification results for six pairs of 30 day forecasts run from 15 and 16 of each month from Oct 1985 to March 1986 with the ECMWF T106 model (for the Northern Hemisphere 22.5-86.5N). The verification parameter is 500mb height, and the scores are a) rms error and b) anomaly correlation. The curves show the persistence error (solid), the forecast error (dots) and the forecast spread (pecked). Also shown is the expected error for a climatological forecast (norm, a) and the 60% level (b).

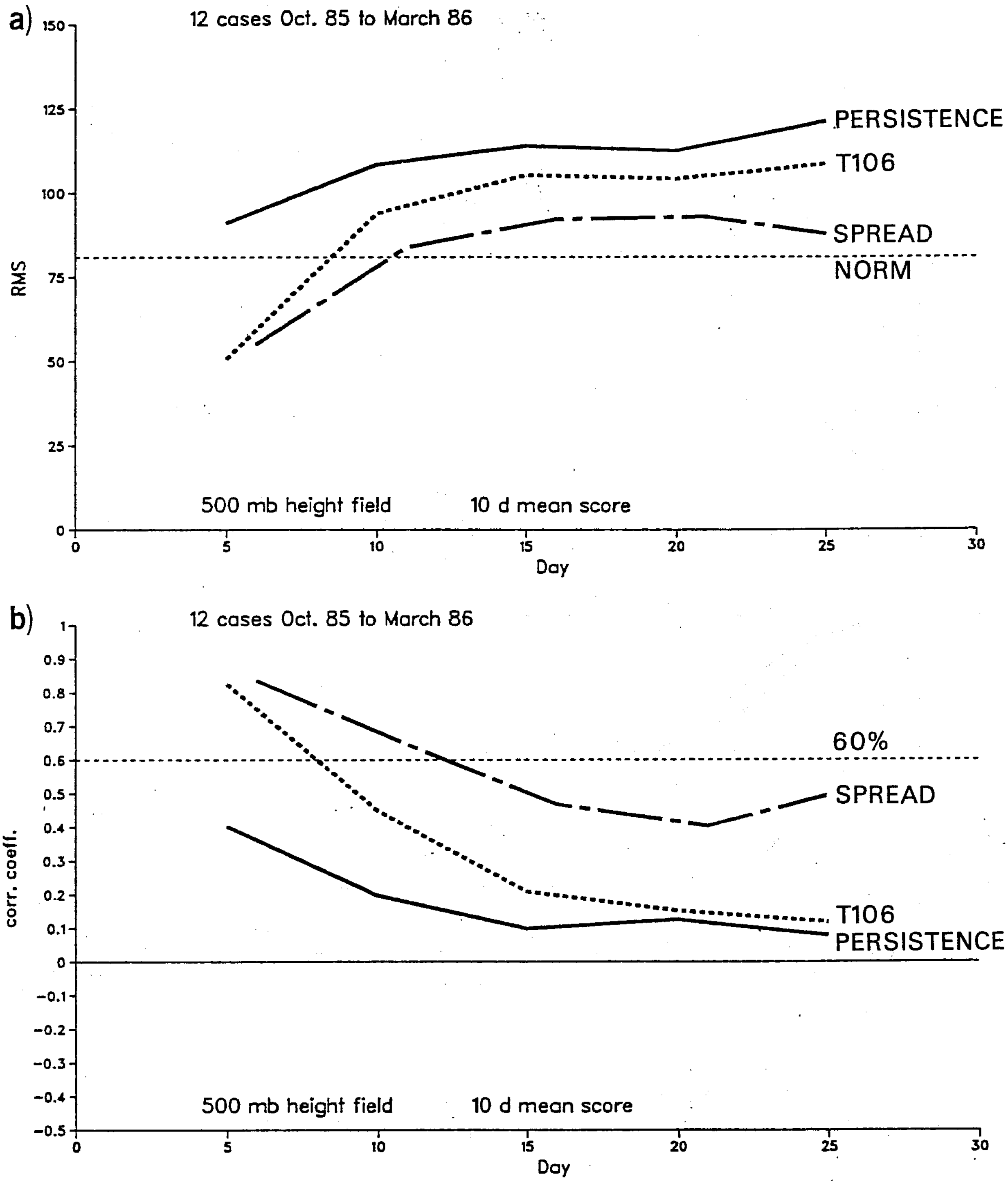


Fig 9 As Fig 8 for 10-day mean 500mb height forecasts with the ECMWF T106 model.

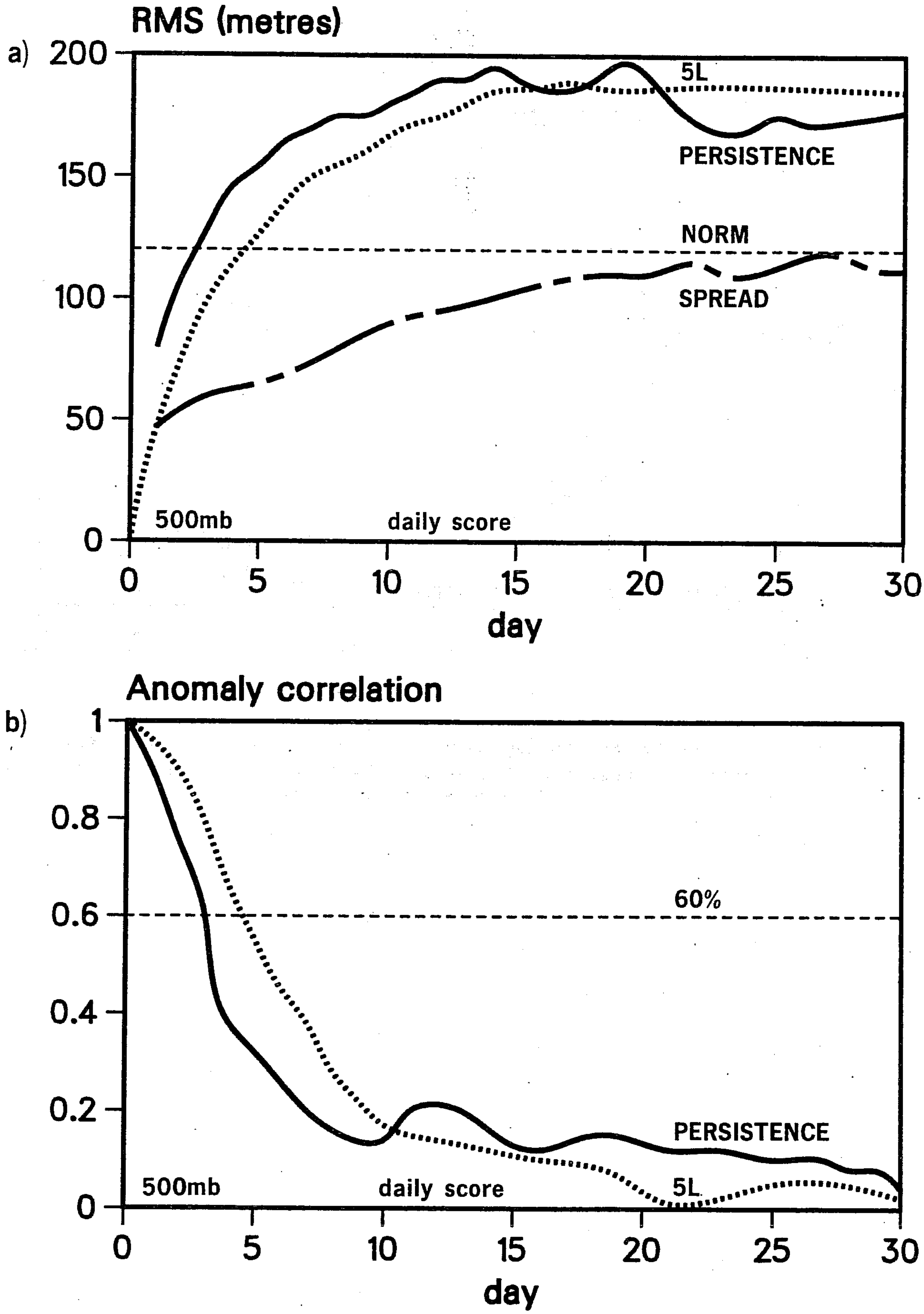


Fig 10 As Fig 8 for the daily scores of the nine pairs of forecasts reported by Mansfield (1986) with the UKMO 5-level model. The spread statistics in anomaly correlation are unavailable.

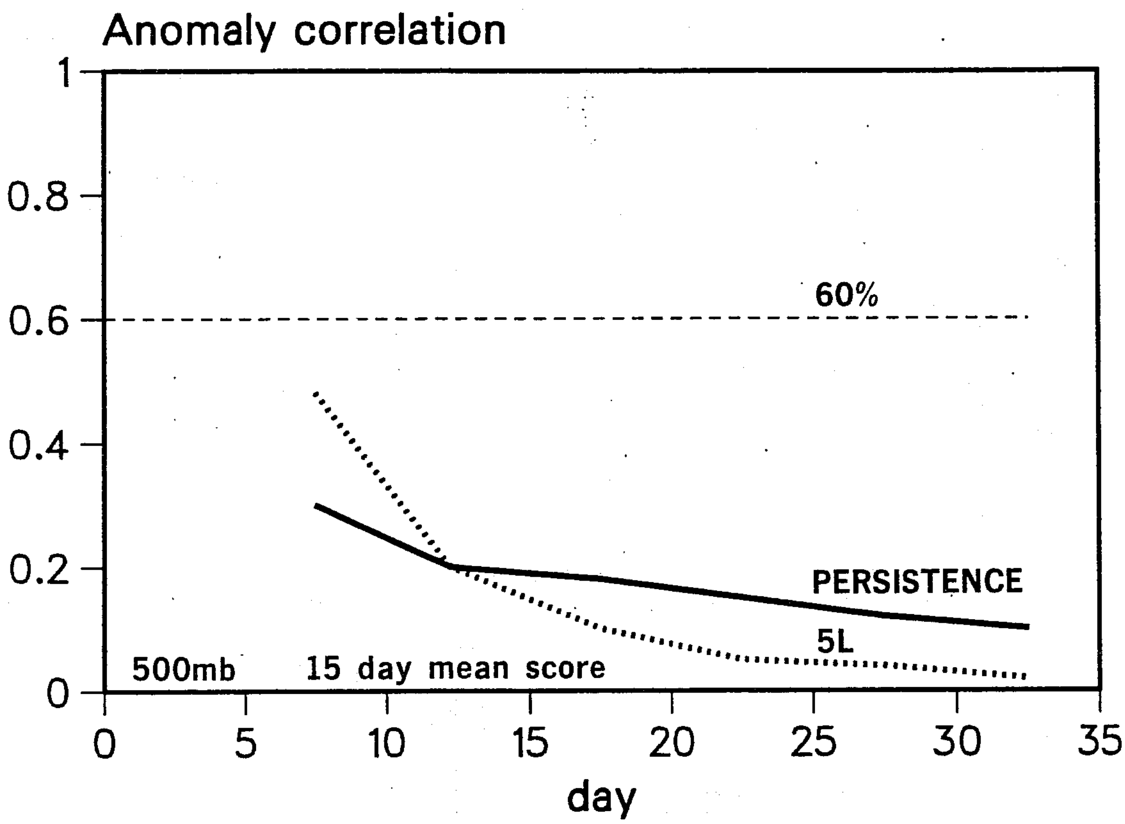


Fig 11 As Fig 8b for the 15-day mean scores of the nine pairs of forecasts reported by Mansfield (1986) with the UKMO 5-level model. The spread statistics in anomaly correlation are unavailable.

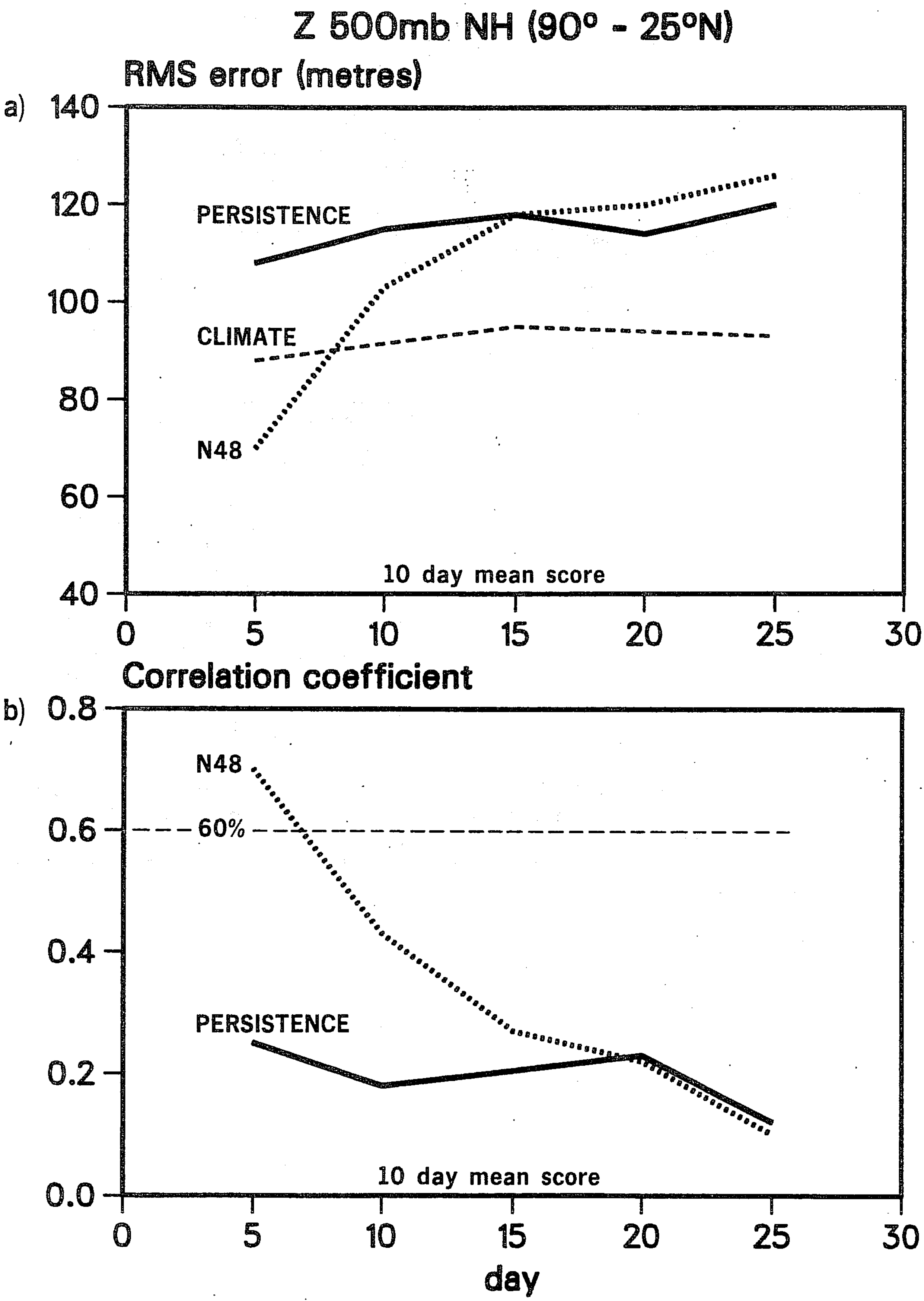


Fig 12 As Fig 9 for the scores of the 10-day mean of the average forecasts resulting from the eight triplets of forecasts reported by Miyakoda et al (1986) with the GFDL model. The spread statistics are unavailable. The dashed line in panel a indicates the actual errors of climatological forecasts.

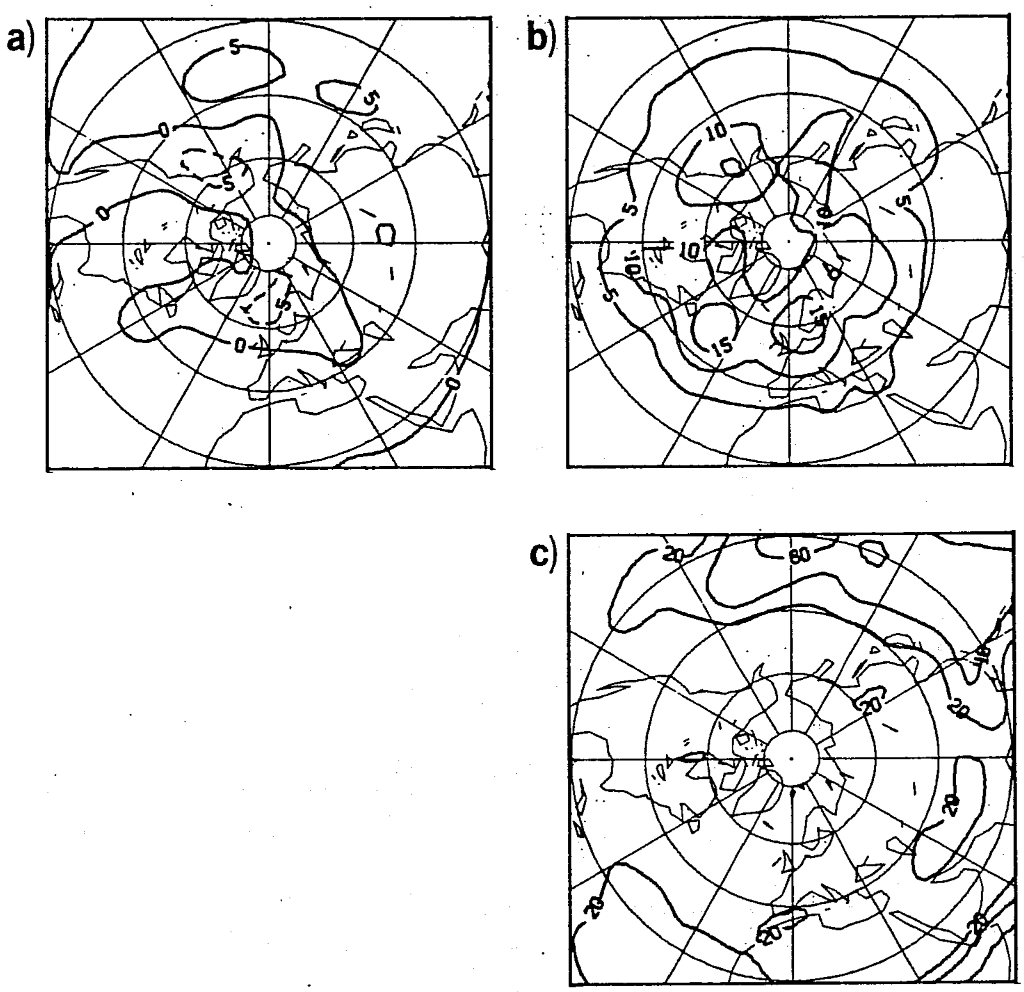


Fig 13 The mean error (a), the rms error (b), and the ratio of mean error variance to total error variance (c), averaged over days 1-10 for the 6 pairs of winter forecasts with the ECMWF T106 model discussed in the text. The corresponding results for the day 16-26 forecasts are shown in d) e) and f). The variable is 500mb height and the contour interval is 5 dam (a,b,d,e centre), and 20% (c,f).

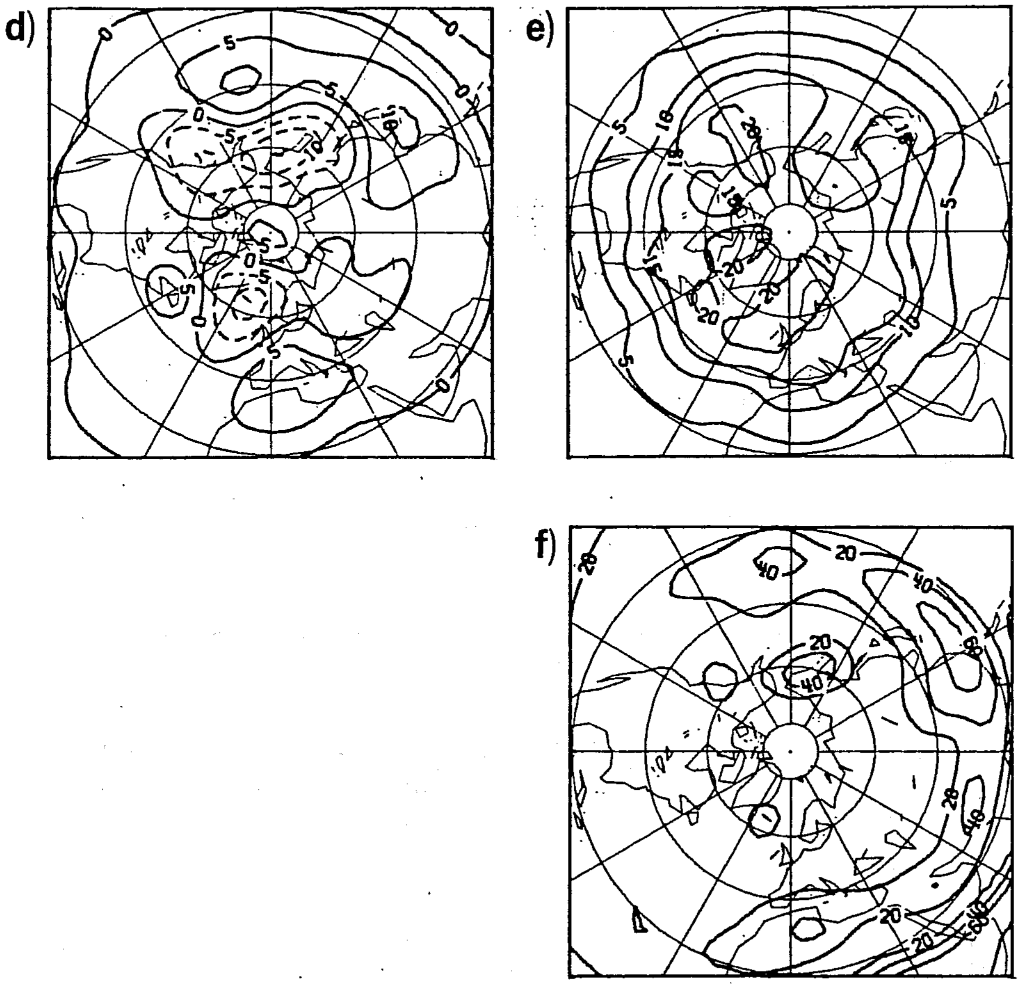


Fig 13 cont

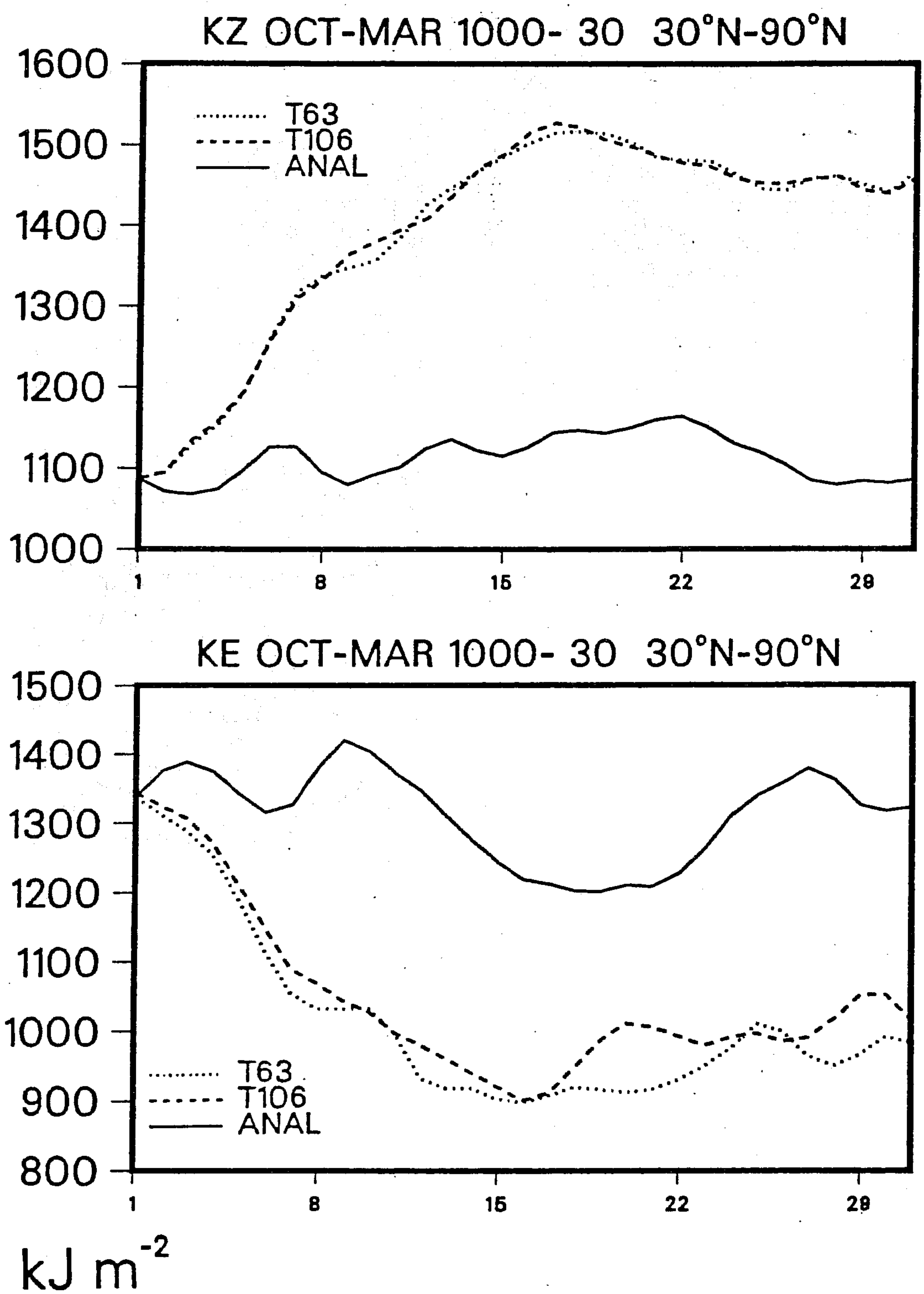
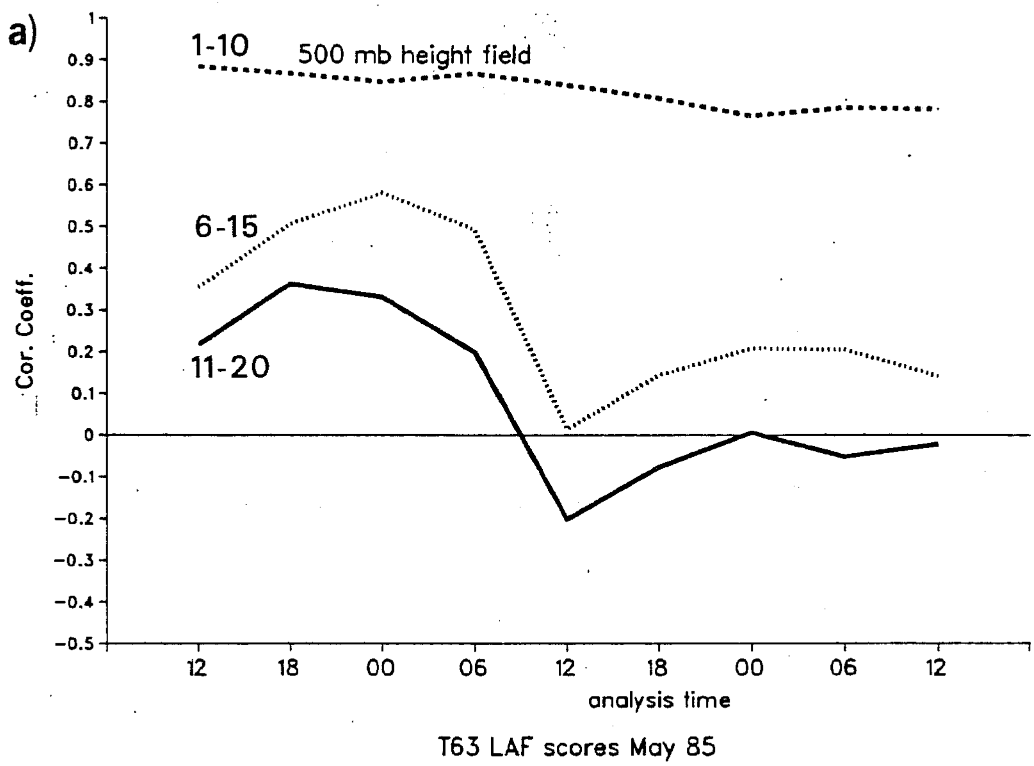


Fig 14 Evolution in time of zonal (top) and eddy (bottom) kinetic energy for the atmosphere (solid) and the sets of ECMWF T106 forecasts (dashed) and T63 forecasts (dots) discussed in the text. The unit is kJ/m*m



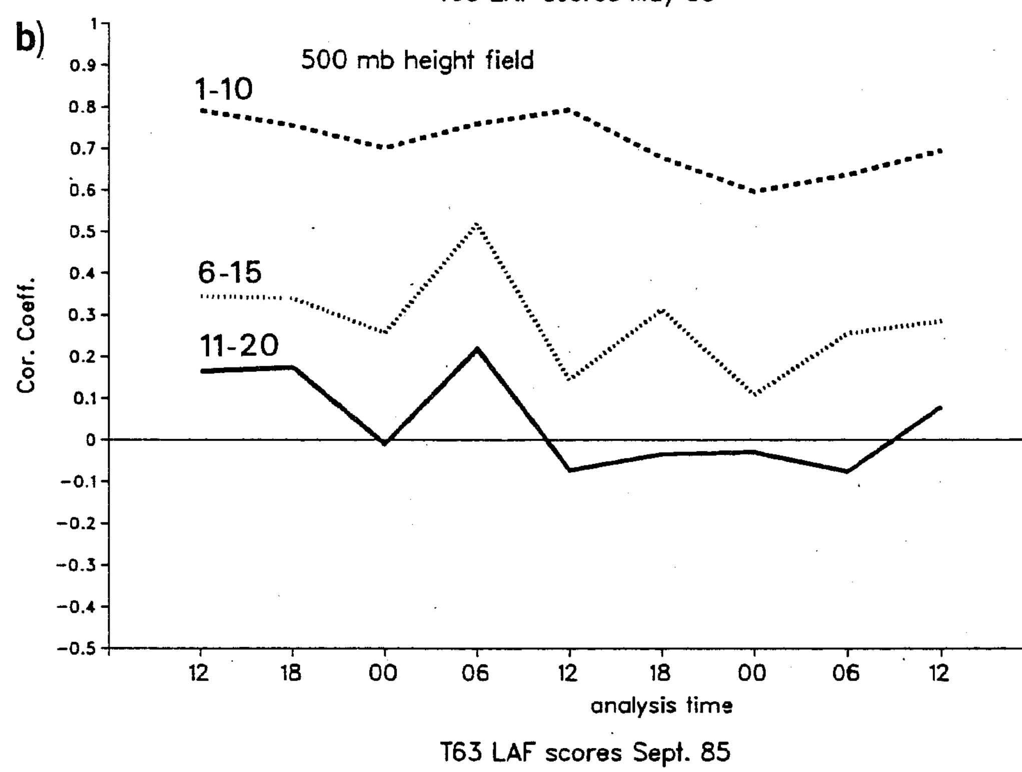
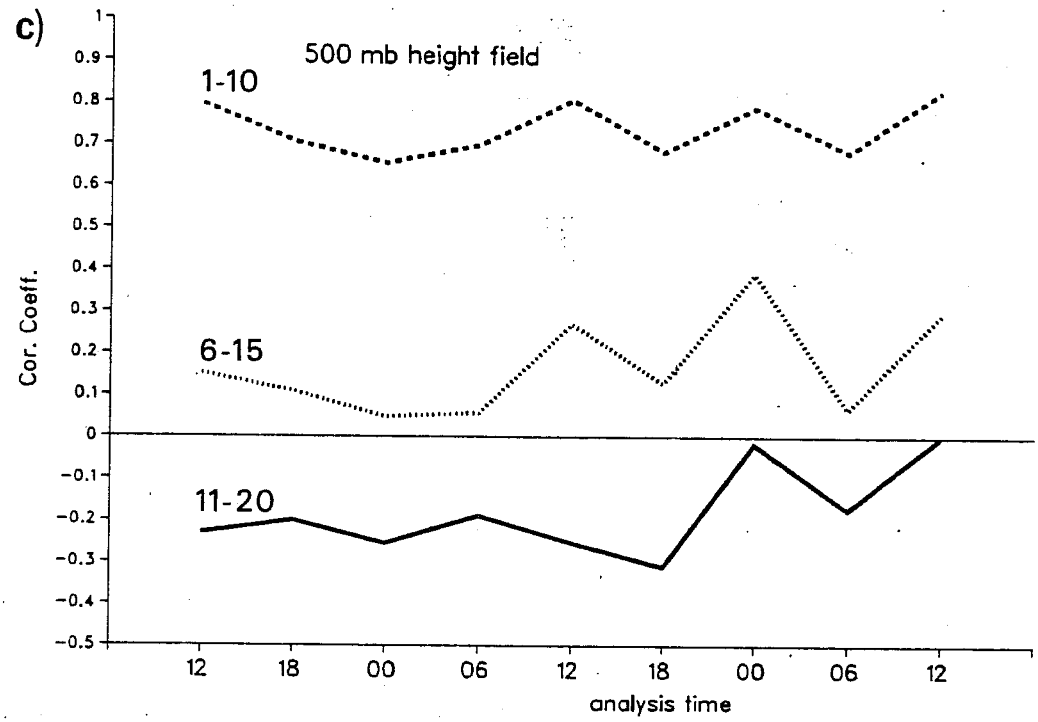
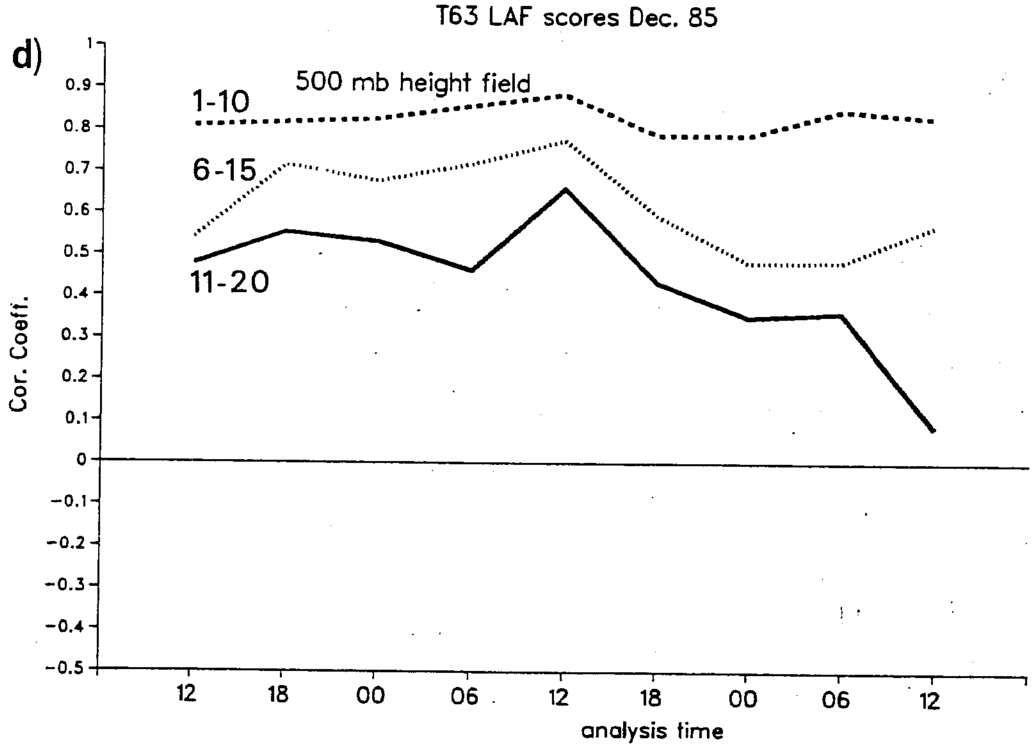


Fig 15 Anomaly correlation scores from four ensembles of T63 integrations with the ECMWF T63 model, for a) May 1985 b) September 1985, c) December 1985, d) March 1986. Within each ensemble, the starting times were separated by 6 hours. The horizontal axis indicates the validity of the starting analysis in hours before the last analysis of the set (nine starting times in all). The curves show the anomaly correlation scores for the 10-day mean fields for days 1-10 (heavy dots), days 6-15 (light dots) and days 11-20 (solid). The variable is 500mb height.



T63 LAF scores Dec. 85



T63 LAF scores March 86

Fig 15 cont

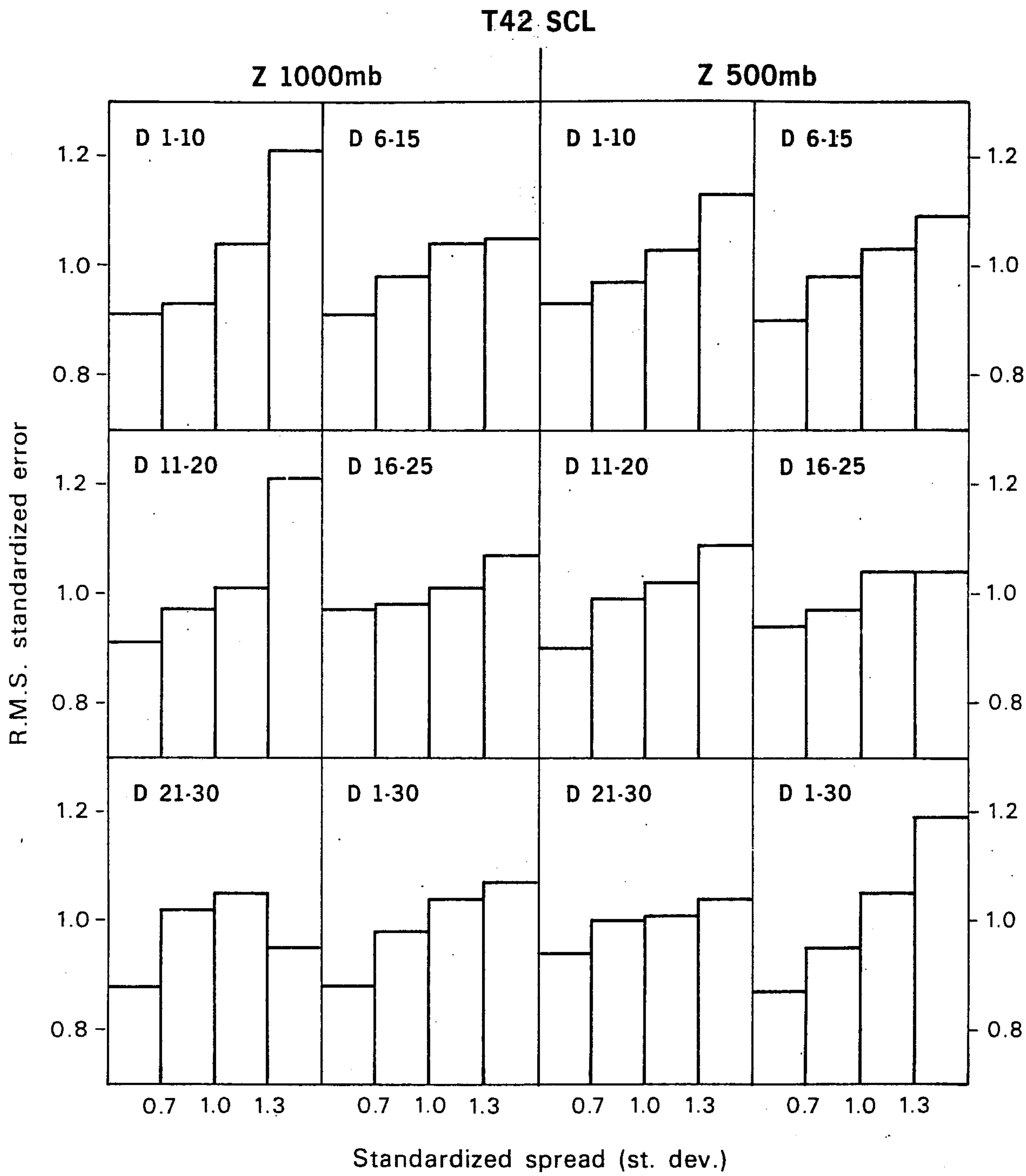


Fig 16 RMS forecast error (normalised by the persistence error) plotted against forecast spread (also normalised by persistence), binned in four categories, for the ensembles of T63 and T42 integrations described in Molteni et al (1986). The results are shown for 1000 mb height (left) and 500 mb height (right) for 5 overlapping 10 day mean forecasts, and for the 30-day mean forecast (from Molteni et al, 1986).