PREDICTABILITY IN THE MEDIUM RANGE AND BEYOND

T.N. Palmer, C. Brankovic, F. Molteni, R. Mureau
European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading

1. INTRODUCTION

In the 1960s, a quantitative definition of 'medium range' was actively sought. A variety of techniques were used to define the limit beyond which the day-to-day evolution of mid-latitude weather became unpredictable, and ranged from a few days to several weeks. Today it is generally accepted that this limit is about two weeks. However, as many studies have found, there are cases when forecast models do indeed show skill several weeks into the future; conversely, there are situations when they lose skill in a few days. In this sense, the definition of medium range is not invariant. One of the crucial tasks of predictability research in present day medium-range prediction is to develop tools for forecasting this variable predictability on a day-to-day basis.

In this paper, we review some of the earlier work using general circulation models (GCMs) to define the limits of predictability, and try to draw some lessons for modern day research on Monte Carlo forecasting. Diagnostic studies of the dependence of forecast skill on weather regime, and the use of these types of diagnostics for statistical forecasts of forecast skill, are discussed. Dynamical studies of the relationship between predictability and instability will be made, emphasising the role of non-normal mode growth.

We review some existing techniques for Monte Carlo prediction, and methods of 'postprocessing' results from multiple integrations. However, it is suggested that an optimal strategy for producing initial Monte Carlo perturbations can be designed using eigenmodes of the symmetric part of the evolution equation operator. Some examples of these eigenmodes are shown.

2. THE LIMIT OF DETERMINISTIC PREDICTABILITY

As characterised by low-order models, the atmosphere is a chaotic system, with finite predictability. However, almost since their inception, quantitative evaluation of the limit of deterministic predictability has been sought using state-of-the-art comprehensive GCMs. The first serious attempt at this was reported by Charney et al. (1966). Three GCMs were run until they had attained their statistically stationary climate. A random state from each of these was climates was then perturbed, and, for each model, a twin integration run until the RMS difference between the two integrations was equal to the RMS difference between two randomly chosen states of the model's climate. The perturbations were either spatially

uncorrelated, or consisted of a sinusoidal component. Their amplitude was considered representative of analysis uncertainties. An example of the growth in RMS difference for the model considered by Charney et al. (1966) to be most realsistic is shown in Fig. 1.

It can be seen that in the first few days, the perturbation decreases in amplitude, then increases with an initial doubling time of about 5 days, asymptoting at about day 30. On the basis of this result, prediction several weeks into the future appears to lie within the limit of deterministic predictability.

On the other hand, Fig. 2 shows two examples of a similar calculation using the ECMWF T106 operational model. It can be seen that the RMS difference increases monotonically with a doubling time of just over two days, asymptoting at about two weeks. This limit of two weeks is generally accepted to be the limit of deterministic predictability.

Why are the predictions from the two models so different? There are two principal reasons. Firstly, the current operational model has a much finer resolution than the early GCMs, and is able to simulate more accurately the important synoptic scale instabilities in the atmopshere. As shown in Tibaldi et al. (1989), the level of simulated extratropical eddy kinetic energy increases monotonically with resolution from T21 to T106.

Secondly, the effective perturbations used in the ECMWF calculations are quite different to those used in the Charney et al. experiments. In particular, Fig. 2 shows the RMS spread between two model integrations initialised from operational analyses one day apart. Hence the perturbation at day 0 is equal to the day one forecast error of the first forecast. The important difference between the two techniques is that in Fig. 2, the effective perturbations are meteorologically balanced, since the forecasts themselves start from initialized datasets. On the other hand, the Charney et al perturbations are not balanced, and the initial decrease of perturbation amplitude corresponds to the process of (geostrophic) adjustment. In modern parlance, the slow manifold has the character of a strange attractor, and the perturbed integration in the Charney et al study was initially attracted onto the slow manifold, before the generic divergence of the attractor could be realized.

It can be mentioned that some of the shortcomings of the integrations in the Charney et al. report were recognised at the time (e.g. Robinson, 1967). As a result, a number of calculations of predictability time were made using idealized turbulence models. These critically depended on the assumed spectral distribution of turbulent energy in the inertial subranges.

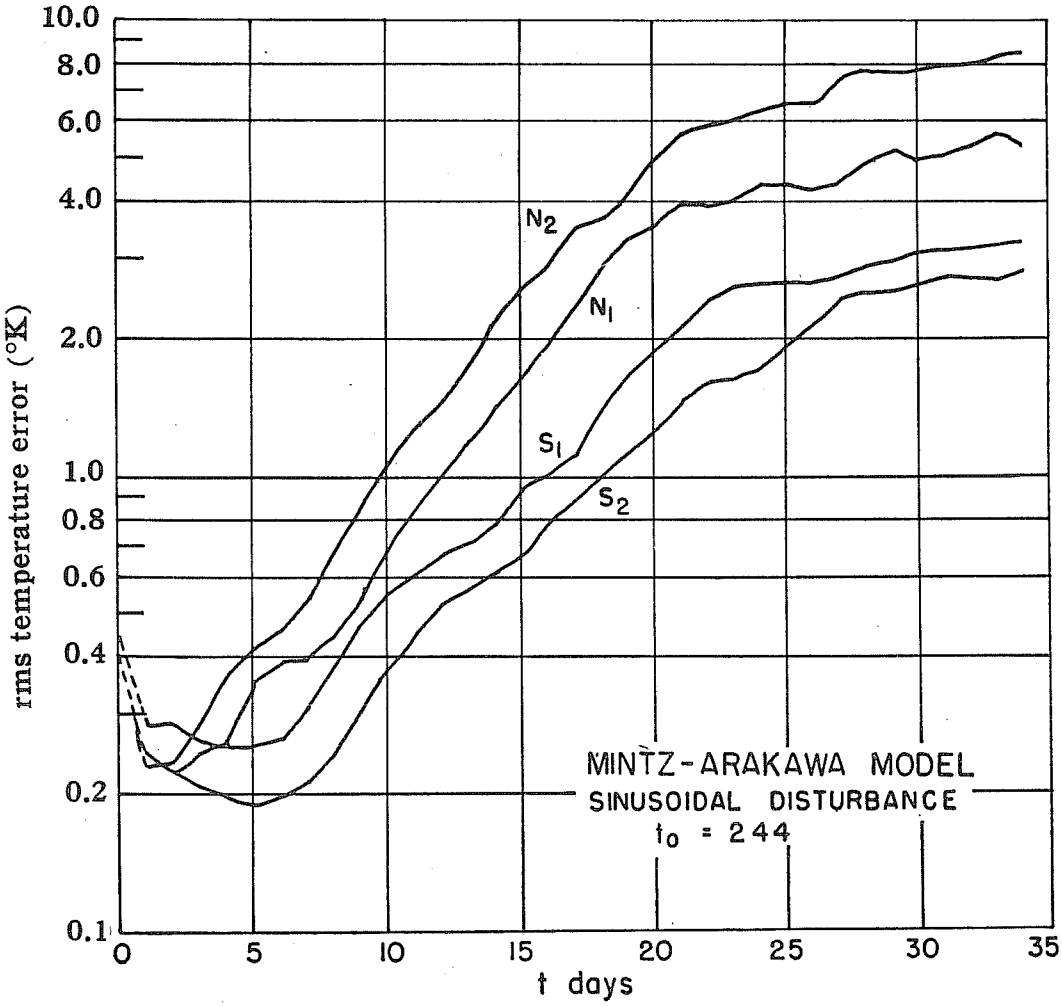

Fig. 1 Growth of RMS temperature perturbations in the two-level Mintz-Arakawa GCM. From Charney et al. (1966).

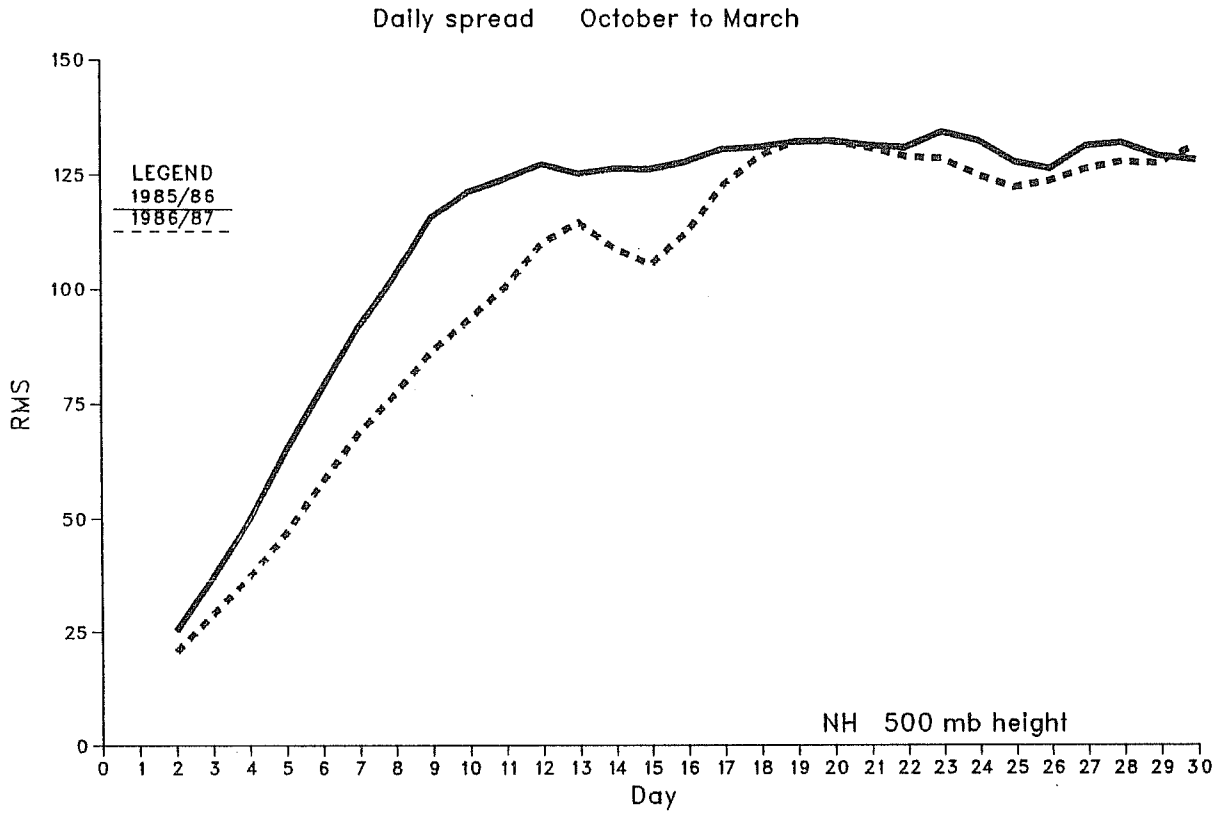

Fig. 2 The ensemble-mean Northern Hemisphere 500mb height RMS difference between 30 day forecast twins. Each twin comprises two forecasts initialized 24 hrs apart. Solid line October 86 to March 87. Dashed line October 85 to March 86.

3. FORECASTING PREDICTABILITY - PREDICTING FORECAST SKILL

Predictability is neither a spatial nor temporal invariant of the atmosphere; it fluctuates with geographical position, with time of year, with weather regime. In general it is as much a variable of the atmosphere as is wind, temperature and humidity. A predictability limit of two weeks merely represents a 'climatological' average estimate; there will be occasions when it is longer than this limit, other times when it is shorter.

If one is to treat predictability just as any other atmospheric variable, can it be predicted using numerical weather prediction (NWP) models? For medium range prediction, Tennekes et al. (1987) have argued that no forecast is complete without a forecast of predictability, and therefore, by implication, a prediction of forecast skill.

Since predictability is based on the notion of the dispersal of an ensemble of initial states which are in some sense close to each other, a forecast of predictability requires some sort of extension of current NWP practice. Attempts to formulate the equations of motion in terms of moments of some infinite ensemble operator (Epstein, 1969) lead to problems of closure. On the other if one calculates the ensemble operator by a finite number of integrations of a deterministic model, one trades closure problems for sampling problems. Nevertheless, this latter, Monte Carlo, approach to forecasting predictability is a practicable in an operational NWP environment, given adequate computational facilities.

However, within operational forecast centres, there is a wealth of data which can be used to find aproximate predictors of forecast skill. Predictors taken from ECMWF operational forecasts have been discussed by Palmer and Tibaldi (1988). Examples include the consistency between today's and yesterday's forecast, and large-scale forecast flow patterns.

The dependence of medium and extended-range forecast skill on one particular forecast flow pattern, the Pacific/North American (PNA) mode, has been discussed in some detail by Palmer (1988). This mode of variability emerges naturally from a regression analysis, as the pattern most strongly correlated with ECMWF forecast skill (at the end of the forecast period), particularly over North America, but also over the whole hemisphere (see Fig. 3). In particular, when the PNA index (forecast or analysed) is negative (anomalous ridge over the north-east Pacific), forecast skill tends to be poorer than average.

In order to interpret this result physically, it was hypothesised in Palmer (1988) that the state with negative PNA index is more barotropically unstable than the state with positive PNA index. This hyothesis was tested in a barotropic model choosing, as two basic states, one with positive PNA index, and one with negative PNA index. An identical localized initial

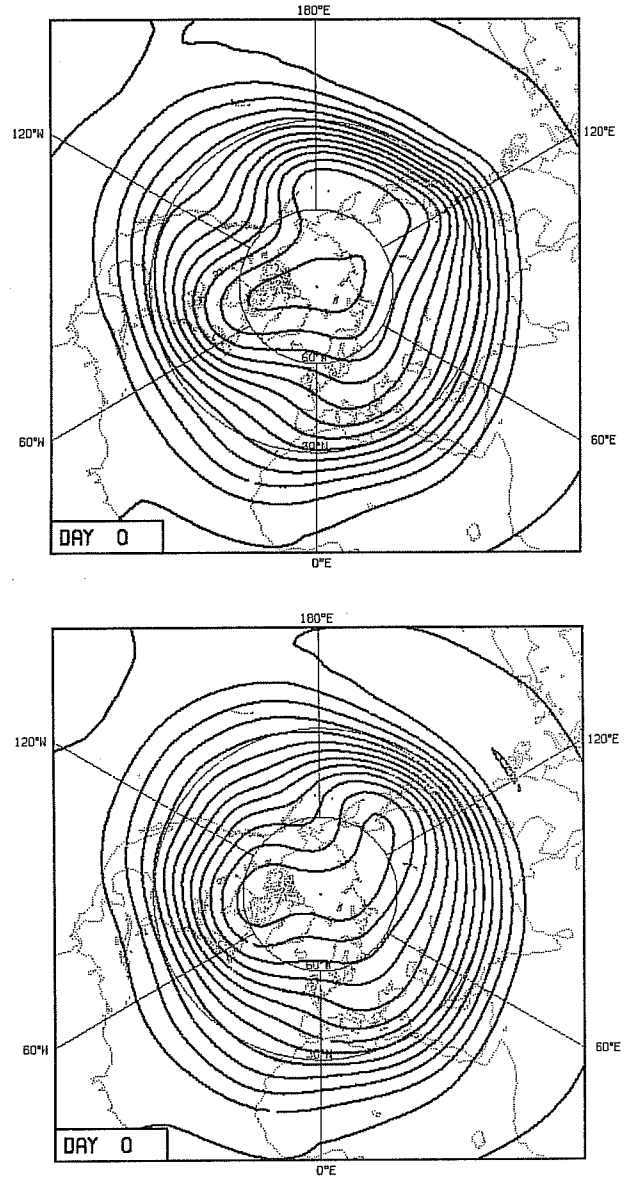

Fig. 3 300mb streamfunction basic states associated with a) skilful b) unskilful day 9 ECMWF forecasts, derived from a statistical analysis of 500 ECMWF wintertime forecasts (see Palmer, 1988, for details).

perturbation was added upstream of the PNA mode to each basic state, and the barotropic model was integrated for 30 days. The evolution of the initial perturbation is shown in Fig. 4. It is clear that by the end of the integration, the perturbation has grown to much larger amplitude on the negative PNA state.

One conundrum which we mention now, and resolve later, is that the growth of the perturbations on these two basic states cannot be resolved purely in terms of conventional normal mode analysis. In particular, it has been found that the most unstable normal mode for the positive PNA (more predictable) basic state has an e-folding time of 7.81 days. By contrast, the most unstable normal mode for the negative PNA (less predictable) basic state has an e-folding time of 8.96 days. On the basis of the Lyapunov exponents the less predictable basic state is more stable!

Based on predictors such as flow regime and forecast consistency, taken from the two most recent forecasts (together with a short cut-off 0z analysis to give an estimate of short-range forecast error), combined using linear regression analysis, a simple practical forecast of operational forecast skill can be made. During the winter of 1988/89, an experimental forecast of forecast skill was issued in real time to the ECMWF Member States. Details of the scheme and results are given in Molteni and Palmer (1989). (It can be mentioned that the flow pattern predictor used was based on the tendency of the forecast flow projected onto a rotated EOF which describes a global version of the PNA pattern index flow during the 10-day period.) Overall results are shown in Fig. 5, in terms of the correlation between predicted and actual skill scores (500mb height RMS error and anomaly correlation coefficient). It can be seen that skill is generally limited, particularly over Europe, and towards the end of the forecast period. On the other hand, a correlation of around 0.4 at day 7 over North America suggests that the scheme is able to predict some aspects of forecast skill variability.

4. PREDICTABILITY, PROJECTIBILITY AND THE ADJOINT EIGENMODES

In this section we shall persue further, dynamical reasons for the growth of perturbations on the positive and negative PNA basic states shown in Fig. 3, and try to resolve the 'conundrum' posed, that the growth of the perturbations in Fig 4 appeared to be quite independent of the e-folding times of the normal modes. Much of this analysis is taken from the PhD thesis of Zhang (1988). As discussed below, an analysis of the generalized instability characteristics of basic states such as these is of relevance to the Monte Carlo programme discussed below.

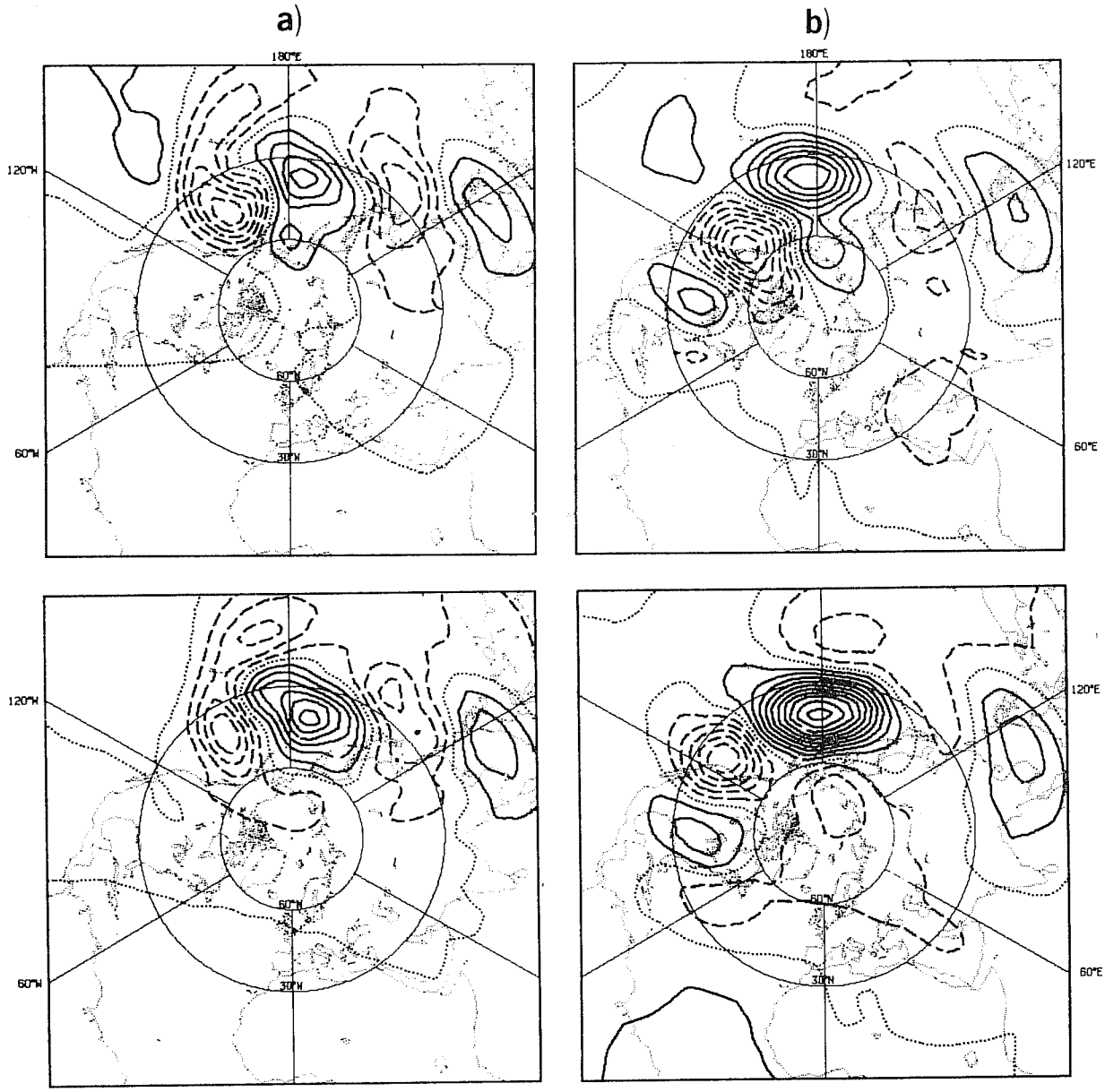

Fig. 4 Perturbation streamfunction from integrations of a barotropic model with the two basic states shown in Fig. 3, and identical perturbations positioned at 30°N, 120°E. The top diagrams are for basic state in Fig. 3a, the bottom diagram is for basic state in Fig. 3b (see Palmer, 1988, for more details.

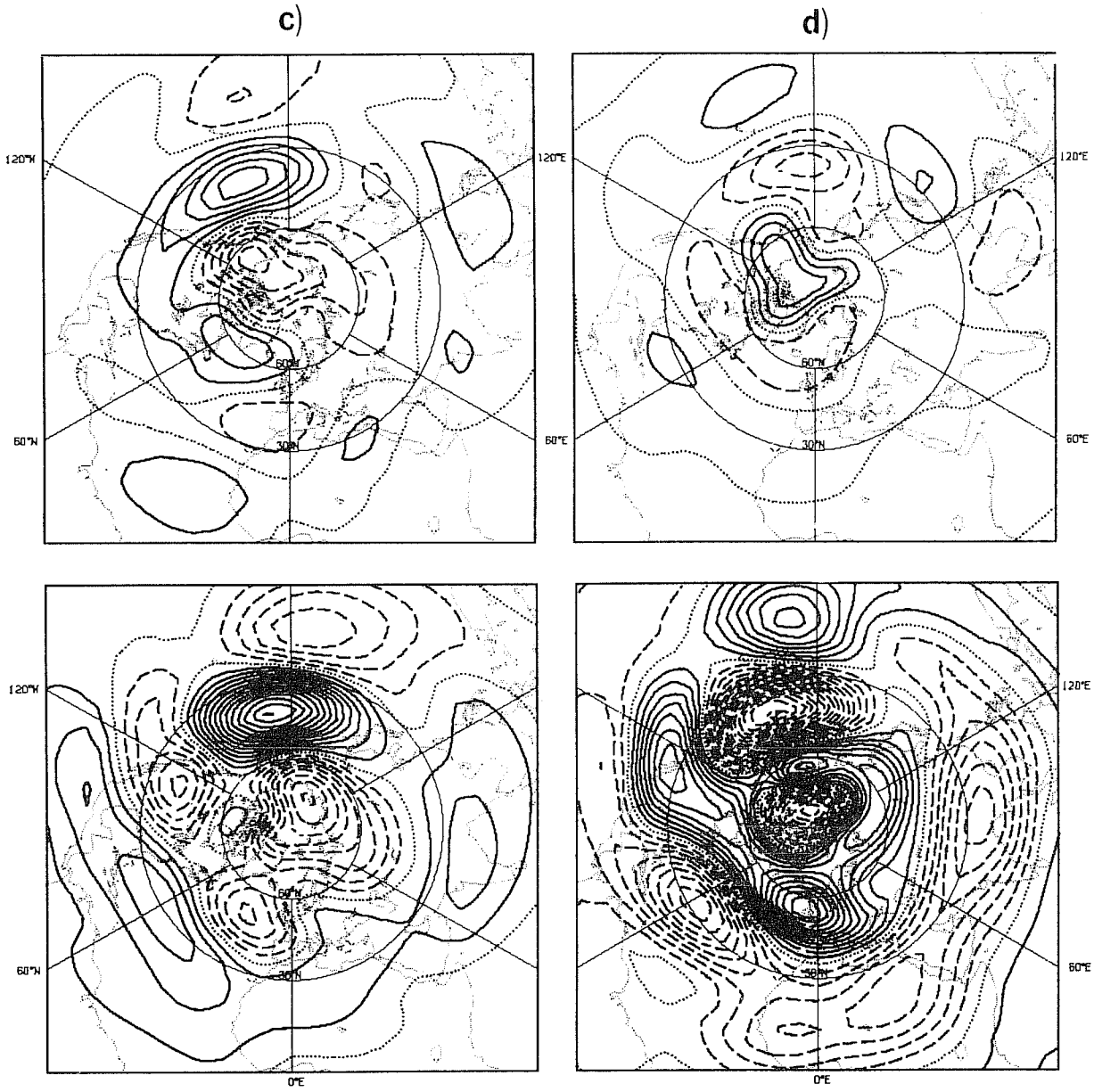

Fig. 4 continued

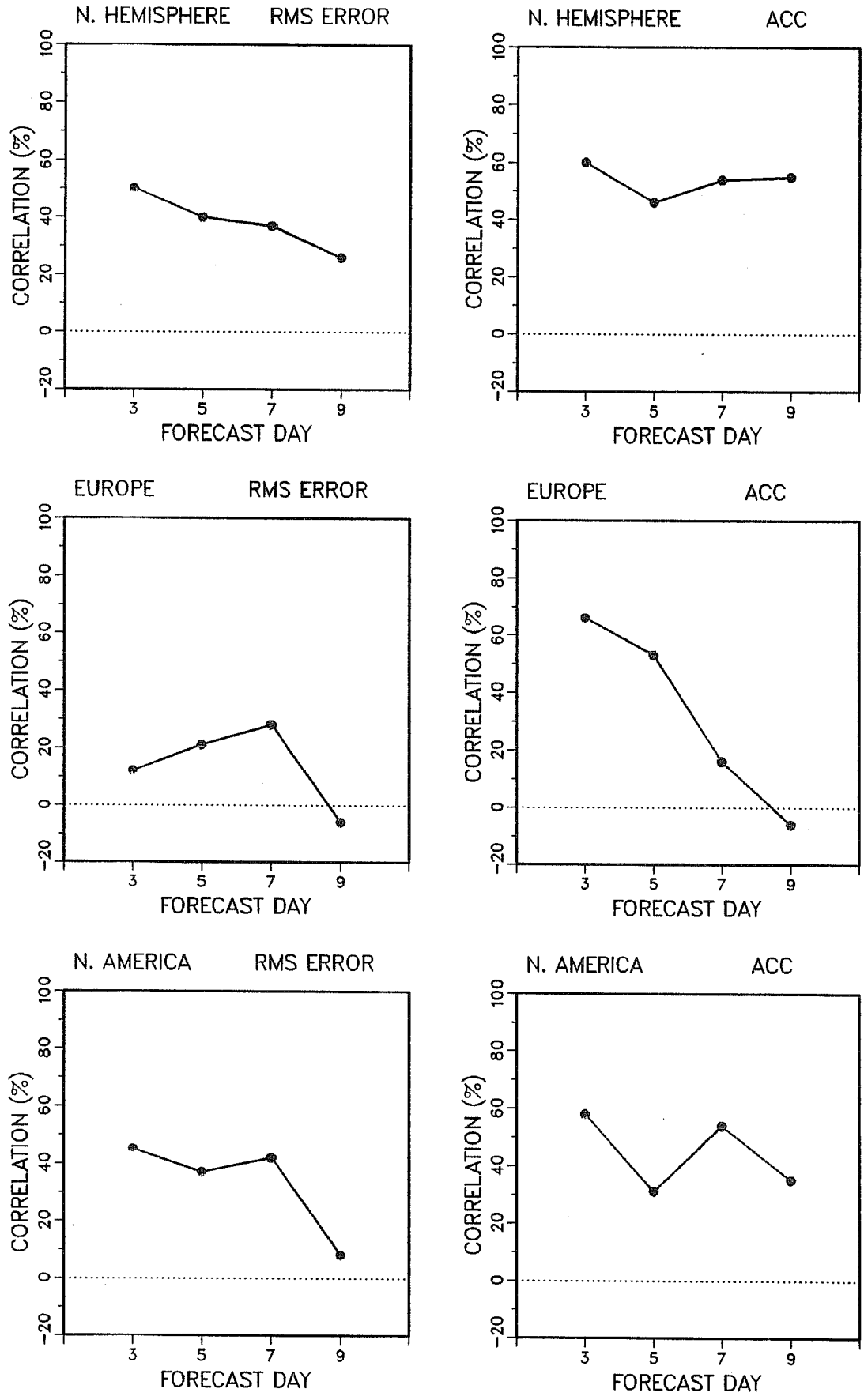

Fig. 5 Results of the quasi-operational forecasts of forecast skill for winter 1988/89.

Let us write the barotropic vorticity equation as

$$\frac{d}{dt}\,\zeta = L\,\zeta \qquad (1)$$

The operator L* is defined as the adjoint operator of L if

$$<\zeta,\ L\,\xi> = <L^*\,\zeta, \xi> \qquad (2)$$

where <..> is a suitable inner product. In the following, we assume $<\zeta\ \zeta>$ defines the global energy.

Now suppose that (λ_j) is the ensemble of discrete eigenvalues of L, and (ν_j) is that of L*, i.e.

$$L(E_j) = \lambda_j E_j$$

$$L^*(E^*_j) = \nu_j E^*_j$$

then, from eq. (2),

$$(\lambda_i - \nu_j^*) <E^*_i, E_j>=0$$

This fundamental bi-orthogonality relationship states that for any eigenvector and adjoint eigenvector that are not orthogonal, the eigenvalues must be a complex conjugate pair, i.e.

$$\lambda_j = \nu_j^*$$

Now, expanding an arbitrary vector G

$$G = \sum_j C_j E_j$$

we have

$$C_j = <E^*_j, G>/ <E^*_j, E_j>$$

Now write

$$\gamma_j = \langle E^*_j, E_j \rangle^{-1}$$

and choose a normalization of the eigenvectors and adjoint eigevectors (see Zhang, 1988 for details) so that γ_j is real.

The quantity γ_j is referred to by Zhang as the projectibility of the j'th eigenmode. At some future time t, we have

$$G(t) = \sum_j \langle E^*_j, G \rangle \gamma_j e^{\lambda_j t}$$

Hence, the initial growth of a perturbation depends as much on its projection onto adjoint modes with small e-folding time, as its projection onto adjoint modes with large projectibility. If a mode has large projectibility, this means it is almost degenerate with another mode. Large initial amplification then can be achieved if the initial conditions can be expressed in terms of the difference between these two almost degenerate modes.

For the positive PNA basic state shown in Fig. 3 (associated with high skill), Zhang(1988) has shown that the first unstable mode has a projectibility of 5.32. For the negative PNA basic state shown in Fig. 3, the most unstable mode has a projectibility of 19.3. The projections of the initial perturbations used in Fig. 4 onto the real and imaginary part of the most unstable mode for a) the positive PNA basic state, and b) the negative PNA basic state, are shown in Fig. 6. In the first 30 days, the amplitude of projection is very much larger for the negative PNA basic state. After about 90 days the different e-folding times of the two basic states becomes more relevant, and the positive PNA basic state has larger amplitude.

As well as having relevance to the diagnosis of the barotropic model integrations in the last section, this analysis suggests a strategy for choosing the fastest growing perturbations for Monte Carlo forecasts. This will be discussed in more detail in section 6.

5. TIME-LAGGED ENSEMBLE FORECASTING

The statistical prediction of forecast skill described in section 3 is clearly limited in scope, yet provides a base level of skill with which to compare more ambitious dynamical schemes. As indicated in section 2, dynamical predictability using Monte Carlo technique is certainly practicable within an operational environment. As discussed (cf the results of Charney et

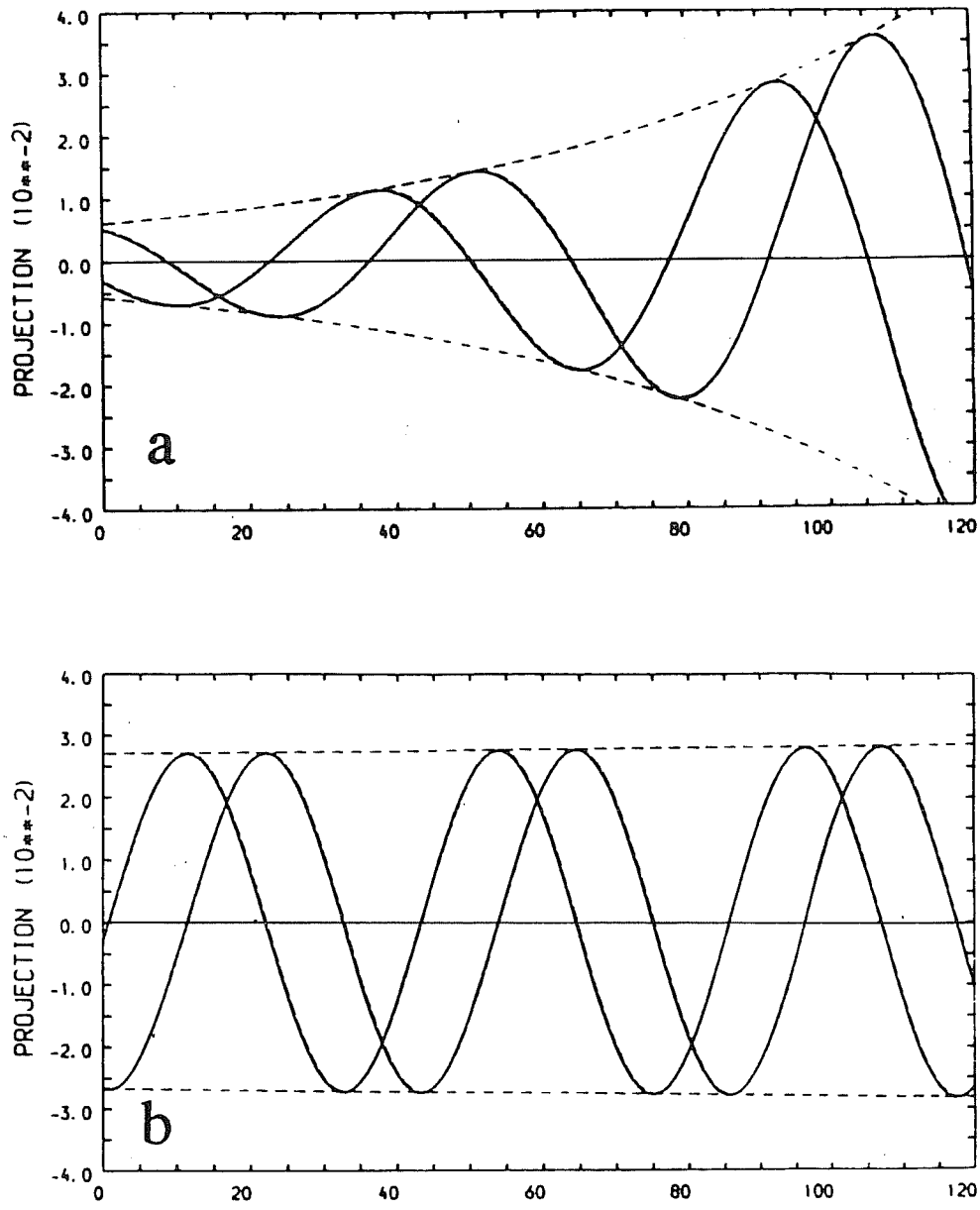

Fig. 6 The evolution of the projections onto the real and imaginary part of the most unstable mode for the basic state a) of Fig. 3a, b) of Fig. 3b). From Zhang, 1988.

al., 1966), estimates of predictability are just as prone to the effects of model systematic error as is the prediction of any other variable. Moreover, as with the initial analyses themselves, it is important that the perturbations be meteorologically balanced.

A simple and convenient way (in an operational environment) of effectively obtaining an ensemble of balanced initial states, is through the time-lagged technique advocated by Hoffman and Kalnay (1983). In the current operational analysis/forecast system at ECMWF, where analyses are produced at 6-hour intervals, an n-member ensemble at time t=0 would be composed of the operational analysis at t=0, the 6-hour forecast initialsed from the analysis at t=-6 hr, and so on to the 6(n-1)-hour forecast from the analysis at t=-6(n-1) hr. The effective perturbations at t=0 are the forecast errors at t=0.

In this section, we study the problem of predicting extended range forecast skill using the time lagged technique. See Brankovic et al. (1989) for more details. Following earlier studies we recognise that at least 10 integrations are necessary to form a reasonable sized ensemble (we address the question of what the least upper bound might be in the next section). With present computational constraints, integration of 10-member ensembles would cause practical difficulties if made at the currently operational T106 resolution. However, results from Tibaldi et al. (1989) have indicated that the extratropical tropospheric extended-range performance of the ECMWF model is not significantly worse at T63 resolution than at T106 resolution. The time-lagged ensembles described in this paper have therefore been integrated at this lower resolution.

5.1 Ensemble-mean skill versus the skill of the control forecast

As is well known, the RMS error of the ensemble-mean forecast is trivially lower than the mean RMS error of the individual forecasts. However, from a practical point of view, it is more important to consider whether the ensemble-mean forecast is superior to the latest member of the ensemble, the control forecast. When compared against a single forecast, it is no longer inevitable that the ensemble-mean forecast will be superior, even in terms of RMS error.

Fig. 7 gives a summary of 10-day mean ensemble mean forecast skill for 500mb height (Fig. 7a), and 850mb temperature (Fig. 7b). The dashed line shows the percentage of ensemble mean forecasts which are more skilful than their controls in terms of both RMS error and anomaly correlation coefficent (ACC). The dotted line gives the percentage of ensemble mean forecasts which beat a persistence forecast (RMS error and ACC) and also beat a forecast of climate.

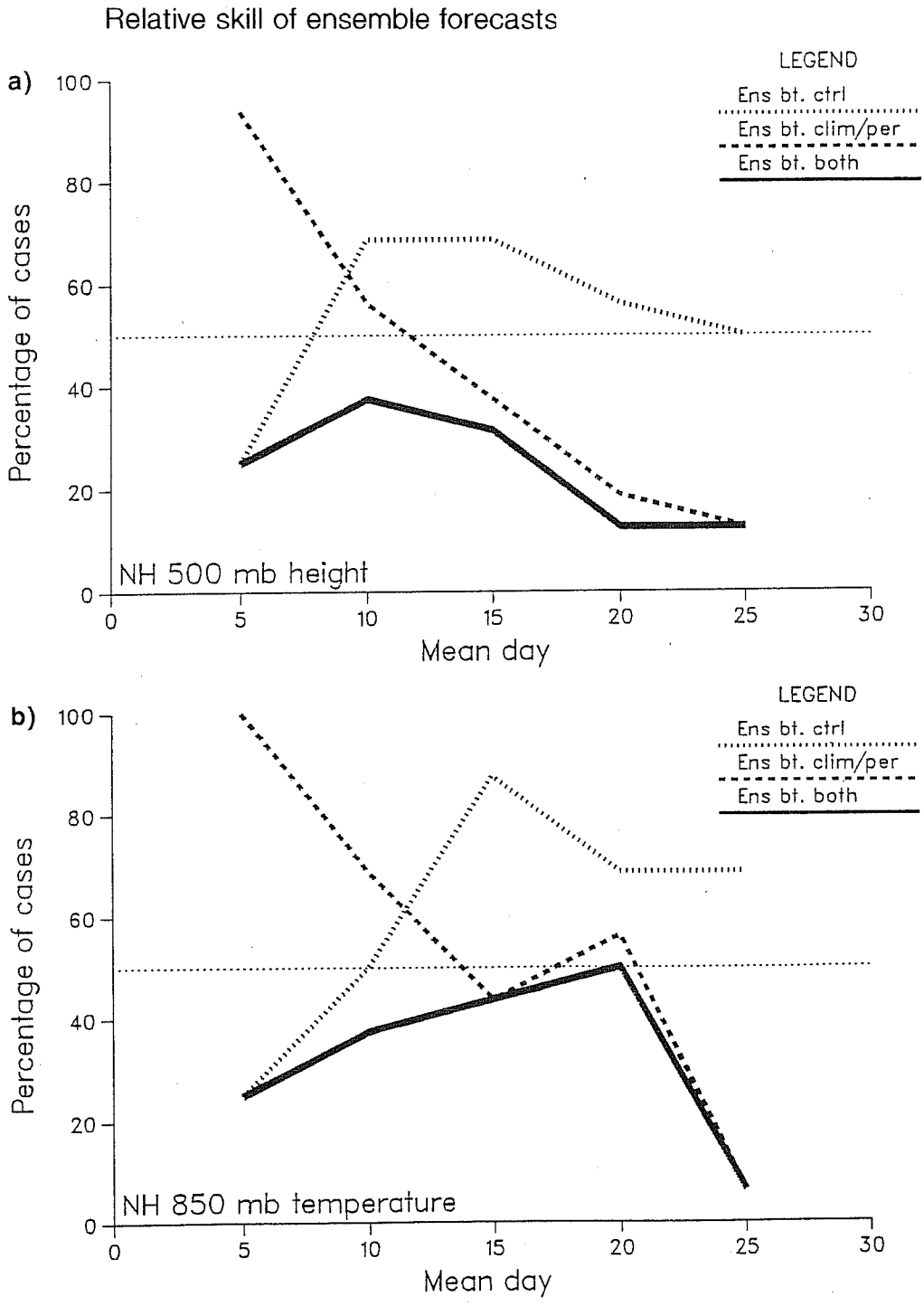

Fig. 7 10-day mean ensemble mean forecast skill for 500mb height (a), and 850mb temperature (b). The dashed line shows the percentage of ensemble mean forecasts which are more skilful than their controls in terms of both RMS error and anomaly correlation coefficient (ACC). The dotted line gives the percentage of ensemble mean forecasts which beat a persistence forecast (RMS error and ACC) and also beat a forecast of climate.

The solid line gives the percentage of forecasts which beat both their controls, and persistence and climate. For reference, a 50% 'breakeven' line is also shown. Mean day 5 corresponds to days 1-10 averaged; mean day 10 corresponds to days 6-15 averaged, and so on.

The solid line gives the percentage of forecasts which beat both their controls, and persistence and climate. For the sake of argument we suppose that extended range prediction becomes viable when the solid line exceeds 50%.

In the medium range all ensemble mean forecasts are better than the 'no cost' forecasts, however, few of the ensemble mean forecasts are superior than their controls. This does not mean that there is no benefit to ensemble forecasting in the first 10 days; rather that some weighting to reduce the impact of earlier members of the ensemble should be made. On the other hand, it may be preferable to employ some other technique to generate initial ensembles, so that all initial perturbations are, a priori, equally likely.

The maximum skill of the ensemble mean (in terms of quantities defined above) occurs at days 6-15 for 500mb height, and days 16-25 for 850mb temperature. For these time ranges the percentages are around 30-50%, and hence fail to consistently achieve our supposed 'breakeven' line.

5.2 Relationship between spread and skill

Scatter diagrams of NH 500 mb height and 850 mb temperature ensemble RMS standard deviation against ensemble-mean RMS error are shown in Brankovic et al., (1989). For days 11-20 some correlation exists for the transition season forecasts of 500 mb height, but none for the winter forecasts. For days 21-30, there is a correlation between spread and skill taking forecasts of 500 mb height from all seasons into account, but none for forecasts within a season. The correlation that exists in the full sample clearly reflects the impact of the annual cycle on spread and skill of 500 mb heights.

Within the winter season in particular, there is no apparent correlation between RMS spread and skill at any time within the forecast period. Hence, we conclude that in our imperfect model environment, hemispheric-scale spread is not a good predictor of hemispheric scale skill. These results are not improved by focussing on more regional areas.

5.3 Phase space trajectories of forecast ensemble

As discussed in section 5.2, the RMS spread of the winter ensembles was not well correlated with their skill. In this section we attempt to study the evolution of ensemble dispersion in more detail for two particular ensemble forecasts.

The technique we use in this section is disussed in more detail in Brankovic et al, 1989. Specifically we show below the time evolution of an ensemble as trajectories on a low dimensional phase-space cross-section. In our case, for each ensemble, the axes of the

phase space trajectories are coefficients of the first two EOFs of the 5-day mean forecst 500 mb height fields (from days 1-5 to days 26-30 inclusive). These cross-sections can be thought of as optimal in the sense that, of all two- dimensional cuts through the phase space of the ensemble-mean forecast, they capture the most variability of 5-day mean 500 mb height within each forecast ensemble.

The trajectories of the January 1986, and February 1986 ensemble forecasts are shown in Figs 8 and 9. The January forecast was the poorest of the whole set of forecast ensembles; from a synoptic point of view intense Euro/Atlantic blocking developed half way through the forecast period. By contrast, the February forecast ensemble was the most skilful of the set. The synoptic development of the January and February ensembles, and the principal EOFs of the forecast flow, are shown in Brankovic et al., 1989.

In Figs 9 and 10, the trajectories of the individual forecasts are numbered by the arrow head. The forecast from the earliest analysis is numbered 1, the forecast from the latest (ie most recent) analysis is numbered 9. The trajectory of the projection of the verifying analysis onto these EOFs is also shown in these diagrams (trajectory with open arrow head).

The scatter of the phase space trajectories towards the end of the two 30 day forecasts do in fact give an indication of the relative skill of the two forecasts. However, this is not the case for the first half. It could be argued that this failure reflects a fundamental inability of the model to simulate blocking activity. On the other hand, it is possible that none of the initial perturbations introduced by the time-lagged forecasting technique was sufficient to cause at least some of the members of the ensemble to develop into blocking patterns.

5.4 Cluster analysis of ensemble forecasts

From a practical point of view, one potentially important advantage of the ensemble technique is the ability to give probabilities of possible alternative developments. If an extreme event is predicted in just one member of the ensemble, it would be associated with a small probability. Nevertheless, this may be valuable information for a user. Information about possible extreme events would not be available from the ensemble-mean forecast.

The technique we use is the application of a Ward hierarchical clustering algorithm to the EOF coefficients of each forecast ensemble (see Anderberg, 1973). For the calculations shown here we use as many EOFs as is necessary to explain about 80% of the variance of 500 mb height within the forecast ensemble. The cluster algorithm is applied separately to

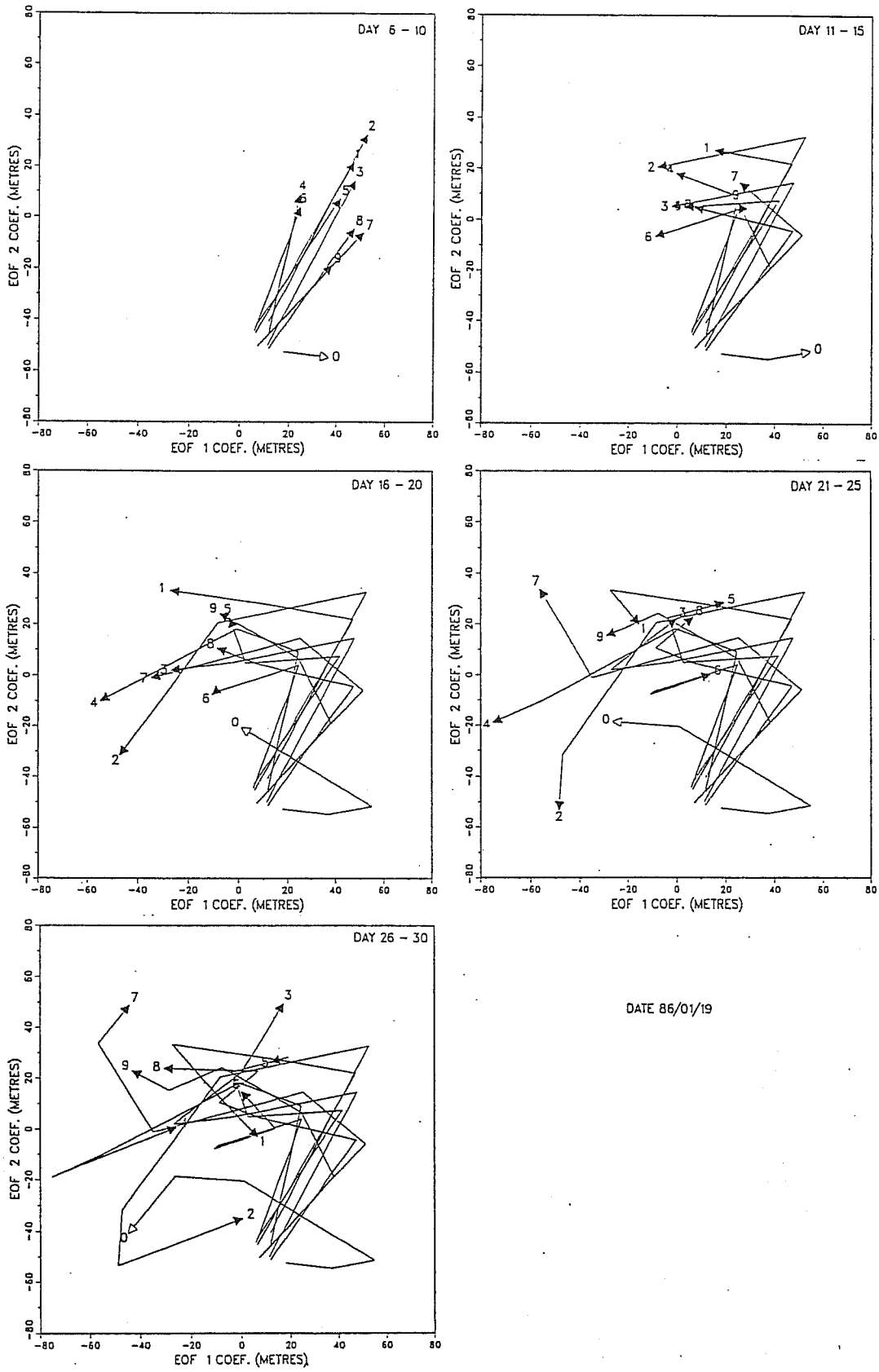

Fig. 8 Phase space trajectories of the 5-day average Northern Hemisphere 500 mb heights in the plane defined by the first two EOFs of the January 1986 ensemble. 1 to 9 denote individual forecasts. Zero and open arrow denote verifying analysis.

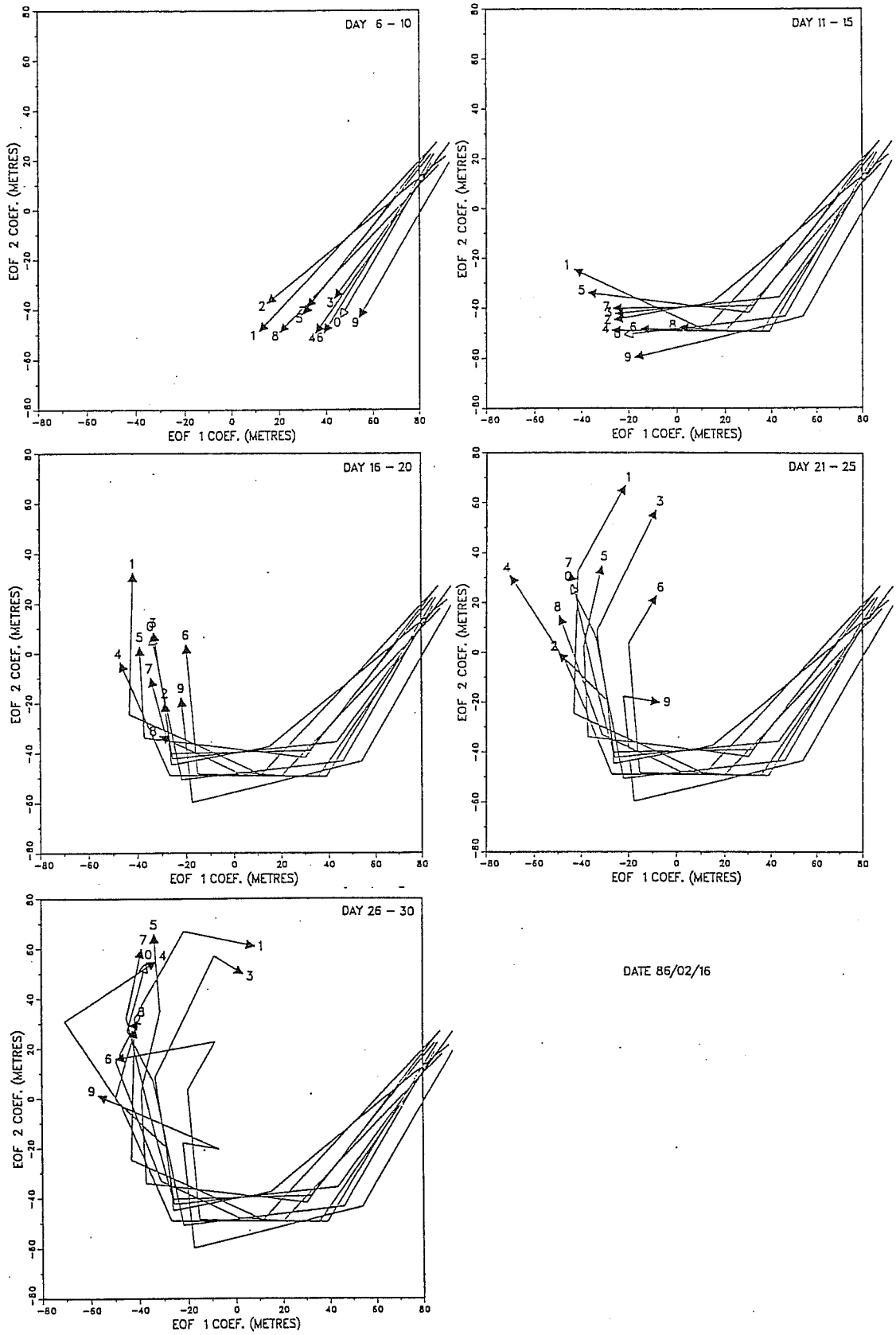

Fig. 9 As Fig.8 but for February 1986 ensemble.

Table 1.

11-20 Date	Cluster A	Cluster B	Cluster C	Ensemble mean
851116	(2,3,4,6,9) .51	(5,7,8) .76	(1) .09	.66
851215	(1,2,3,4) -.16	(5,6,8) -.24	(7,9) .06	-.15
860119	(1,3,5,6,8,9) -.28	(2,4) -.43	(7) -.44	-.36
860216	(1,3,4,5) .58	(7,8,9) .60	(2,6) .62	.64
860316	(1,3,6,7,8) .46	(2,5) .55	(4,9) .27	.49
860914	(1,3,4,5,6) .29	(2,7,8) .58	(9) -.10	.44
861214	(2,4,5,8,9) .18	(1,6,7) .23	(3) .19	.22
871213	(3,6,7,8,9) .35	(2,4,5) .31	(1) .04	.32

Northern Hemisphere 500 mb height anomaly correlation coefficients for three clusters and ensemble-mean: days 11-20.

the EOF coefficients for 10-day mean periods, i.e. for days 1-10, 11-20 and 21-30. For each 10-day mean the algorithm works in nine separate steps. In the first step it finds those two members of the forecast ensemble which are the most similar in the sense that their RMS difference is minimized. These two forecasts are then merged to form a single sub-ensemble (cluster). In the next step, the algorithm merges either two more forecasts, or any of the remaining forecasts with the cluster obtained in the first step, with the objective of finding the combination in which the internal variance of the new cluster is minimized. Hence the nine steps of the algorithm have a tree-like structure; at each step the number of clusters is reduced by one. In the last step, all nine members of the ensemble are merged to form the ensemble mean. At an earlier step n, $1<n<9$, probabilities could be assigned to each cluster according to the number of individual members of the ensemble that have been merged into that ensemble.

It is interesting to note that for the first 10 days the cluster algorithm merges forecasts from adjacent analyses i.e. these are most similar to each other. By the third 10-day mean, however, forecasts from non-adjacent analyses have been merged by the third step. This demonstrates quite clearly that in the extended range, there is no obvious way to 'weight' individual members of the ensemble as a means of reflecting the time-lagged technique used to construct the initial perturbations.

In Table 1 we give the anomaly correlation scores at days 11-20, of forecast Northern Hemisphere 500 mb heights for the centroids of the three clusters obtained at step 7 of the clustering analysis technique outlined above. We show the set of five winter ensemble forecasts, supplemented by the forecasts from November 1985, March 1986 and September 1986, deemed to be 'skilful' in section 4. The first three columns give the scores calculated for three forecast clusters; most populated cluster always being in column A, and the individual forecasts comprising each cluster are shown in brackets. We discuss here the stage in cluster analysis algorithm when only three clusters are defined, because it has been found that, on average, with three clusters about 50% of the variance of the full ensemble is explained (100% being explained by nine 'clusters', i.e. all individual forecasts). In the fourth column, the score for the single cluster (ensemble- meaan) is shown.

For days 11-20, in 7 out of 8 ensembles studied, at least one of the three clusters is superior to the ensemble mean. However, for days 11-20, in only 2 of the cases considered in Table 1, is the most skilful cluster also the most densely populated cluster. For days 21-30 in 5 cases the most skilful cluster was the most dense; but in three of these the density was equal to at least one other cluster (November 1985, December 1986 and December 1987). Consistent with earlier results above, this suggests a problem of sampling.

6. AN ALTERNATIVE MONTE CARLO TECHNIQUE.

For medium-range forecasting, the unweighted time-lagged technique is not optimal, as initial effective perturbations from the earlier analysis times may be quite unrealistic. Weighting would minimize the impact of these earlier forecasts. On the other hand, it may be desirable to find techniques to perturb the initial data in such a way that each perturbed integration is, a priori, equally likely.

A simple technique to produce such perturbations is underway at ECMWF, and some preliminary integrations made. The initial perturbations satisfy the following criteria:

a) they are meteorologically balanced

b) they are representative of short-range forecast errors

c) they are representative of the large-scale flow of the initial state

d) each perturbation is spatially uncorrelated with the others

e) each perturbation has the same amplitude in the northern hemisphere

This is achieved by the following procedure:

a) Retrieve the archived 6hr forecast error of the ECMWF model for the days immediately preceding the initial date

b) Apply a Gram-Schmidt orthonormalization procedure to the 500mb height error fields so that each has an RMS amplitude of 10m between 30°N and 80°N, and each is independent of the others.

c) For each 6 hr forecast error field, scale each variable by the coefficients derived from the Gram-Schmidt procedure.

d) Add (or subtract) the scaled field to the initial analysis to produce the ensemble of initial states.

Below is an example of a 10-day 24 member (T63) forecast ensemble created in this way, integrated from initial conditions from 0z 2/12/88. Fig. 10 shows the skill of the forecast ensemble over Europe. Note that after about day 6, a sizeable sub-ensemble of forecasts are significantly more skilful than the control forecast (dotted line).

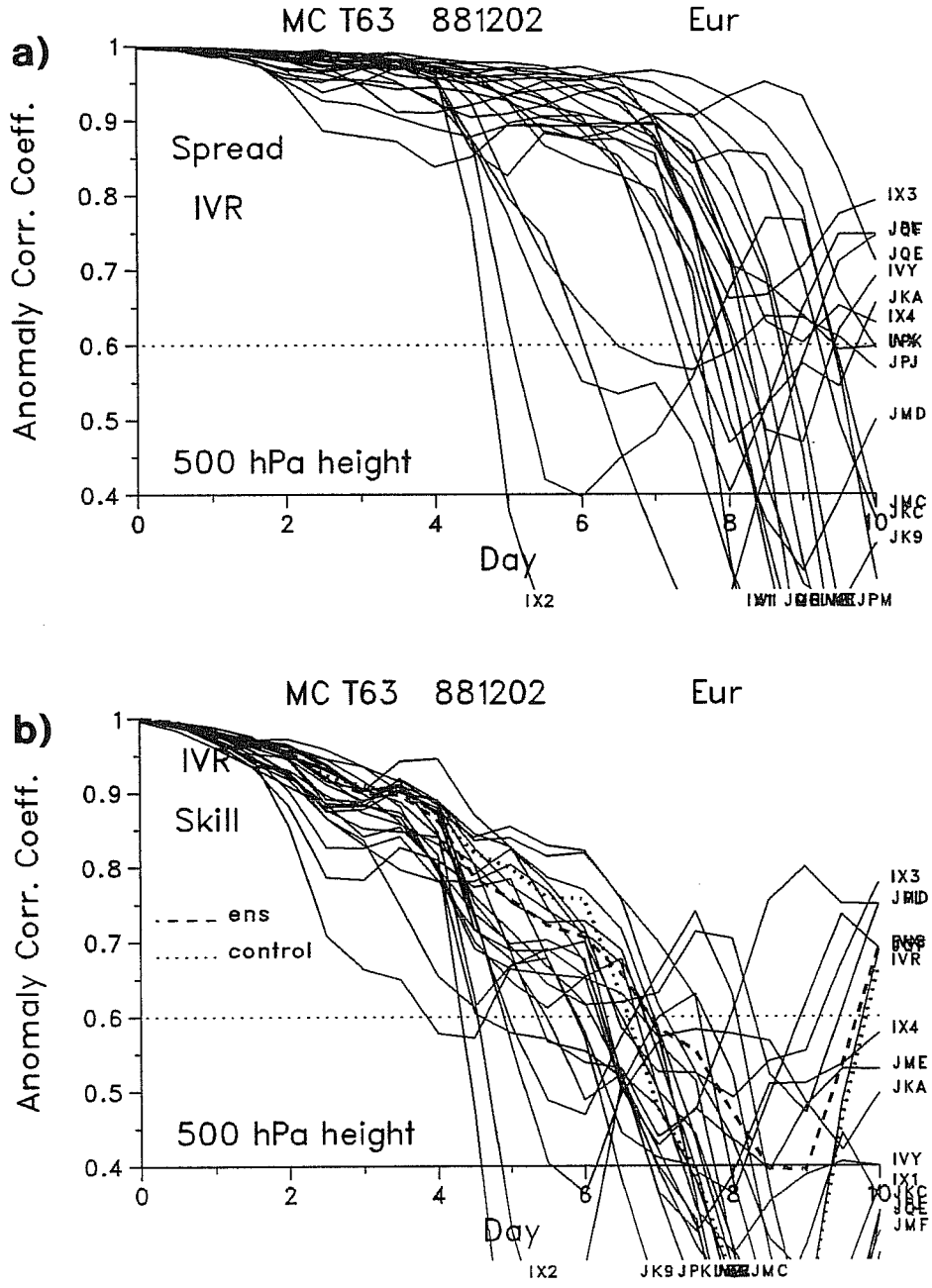

Fig. 10 a) spread b) skill (500mb height anomaly correlation) over Europe of 24 member forecast ensemble created using the technique described in section 6. Dotted/dashed line in b shows skill of control/ ensemble mean forecast respectively.

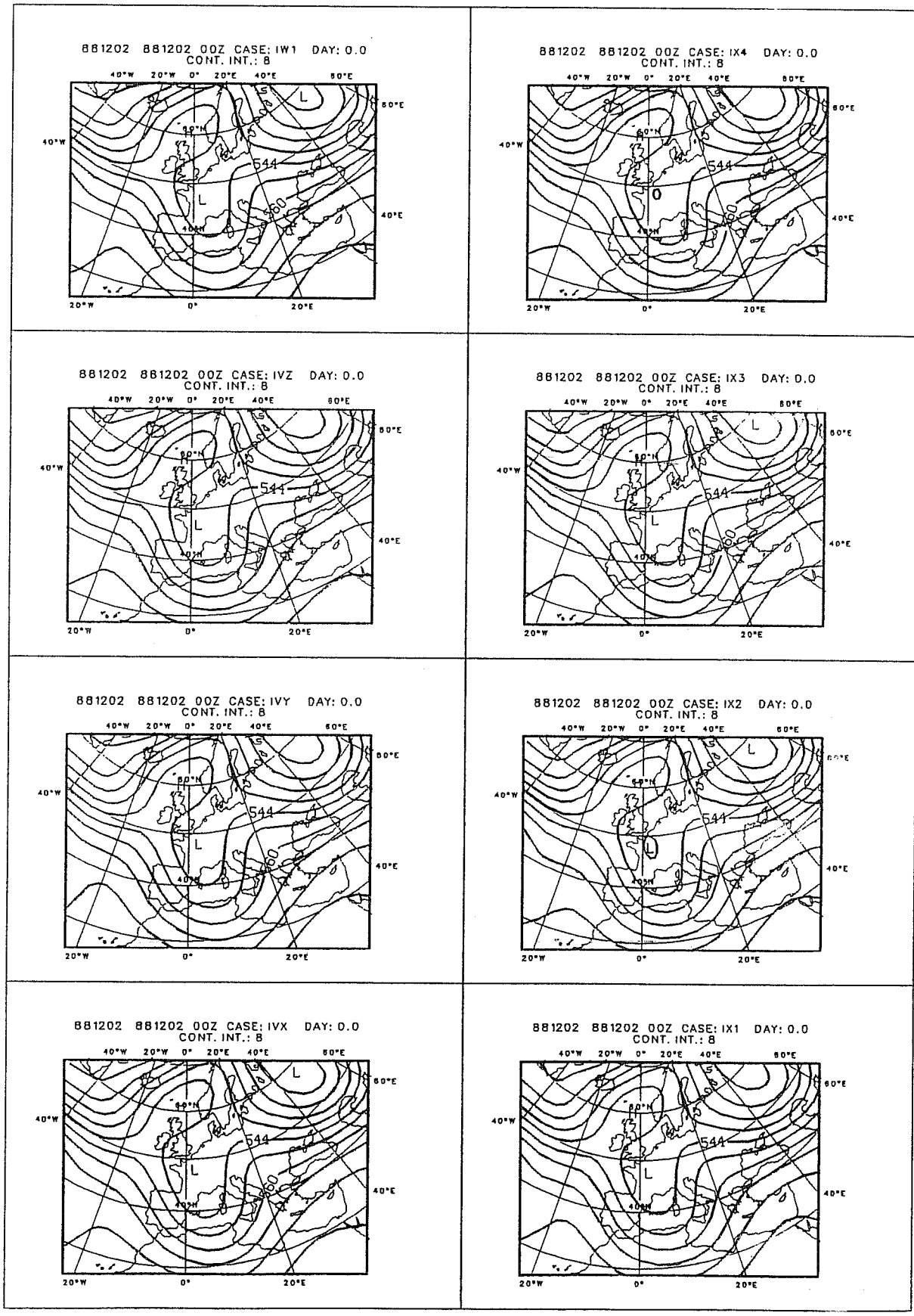

Fig. 11 Day 0 500mb height of first 8 members of forecast ensemble from 0Z 2/12/88.

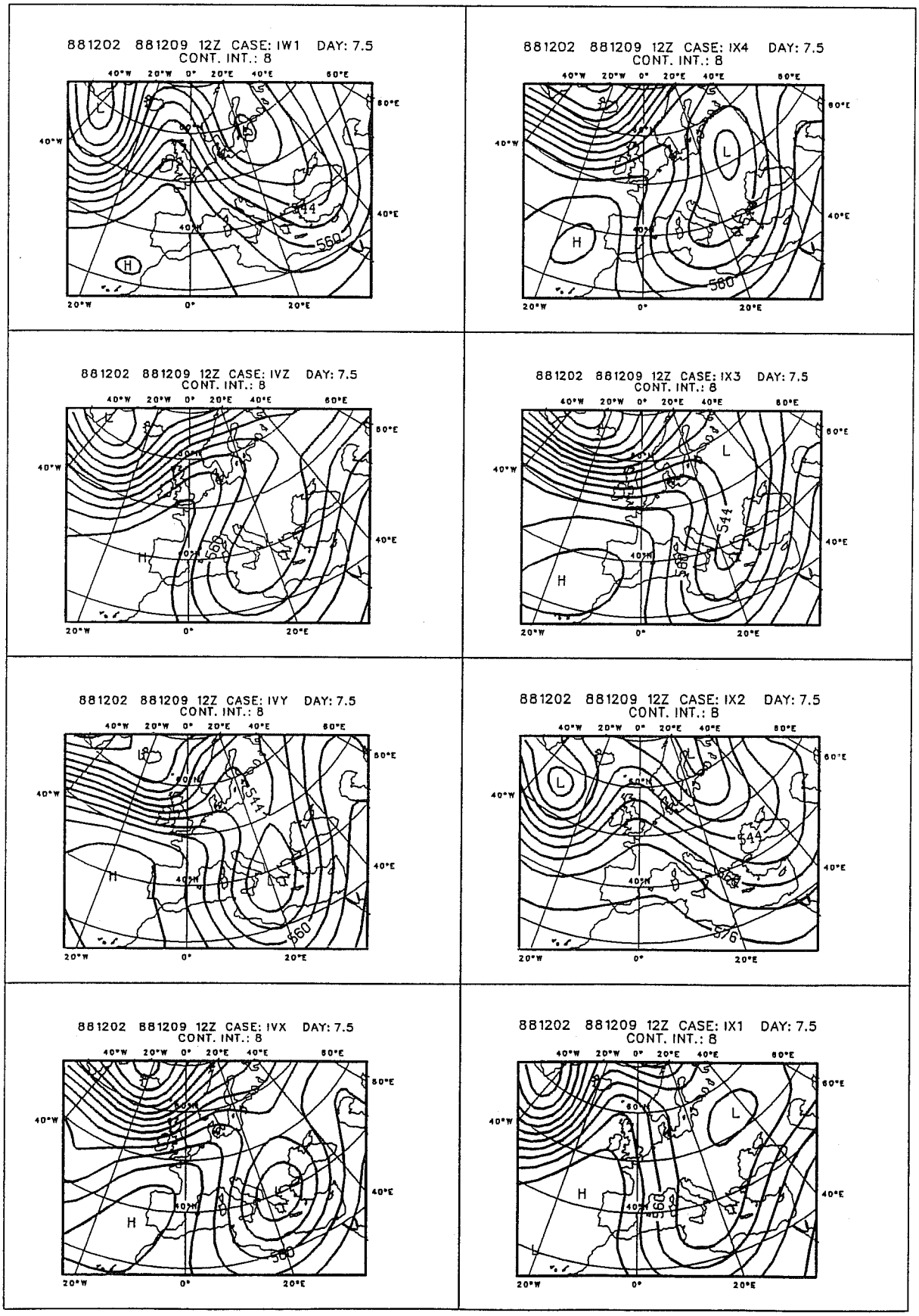

Fig. 12 Day 7.5 500mb height of first 8 members of forecast ensemble from 0Z 2/12/88.

To give an synoptic impression of the spread of the ensemble we show in Figs. 11 and 12, the day 0 and day 7.5 500mb height fields over Europe for the first eight members of the ensemble over Europe. In this particular case, most of the members of the ensemble develop significant troughing over southern Europe by the second week of the forecast. This could therefore be forecast with some confidence. By contrast, over the UK, the flow types are quite different, and less confidence can be attributed to a weather forecast for the second week.

At present, a number of techniques are being used to analyse forecast ensembles such as these, including the cluster analysis techniques discussed in section 5. All these techniques are essentially probabilistic.

7. USE OF MODES OF INSTABILITY FOR MONTE CARLO FORECASTING.

The technique described above for choosing initial perturbations has no knowledge of the most linearly unstable phase-space directions on the slow manifold. Yet, in section 4 we discussed the possibility that model skill was indeed related to large-scale modes of instability. Let us represent the energy of a perturbation at time t by

$$E(t)=\langle\zeta(t), \zeta(t)\rangle$$

Let R be the mapping from the tangent space at $t=t_0$, to the tangent space at $t=t_1$, then

$$E(t_1) = \langle R\,\zeta(t_0), R\,\zeta(t_0)\rangle$$

$$= \langle R^*R\,\zeta(t_0), \zeta(t_0)\rangle$$

For given initial perturbation $\zeta(t_0)$, $E(t_1)$ is maximised if $\zeta(t_0)$ is coincident with the dominant eigenmode of the (self-adjoint) product operator R*R (LaCarra and Talagrand, 1988; Farrell, 1988).

Now let $t_1=t_0+\delta t$, then from equation (1)

$$R^*R = (L\,\delta t + 1)^*(L\,\delta t + 1) = 1 + (L + L^*)\,\delta t + O(\delta t^2)$$

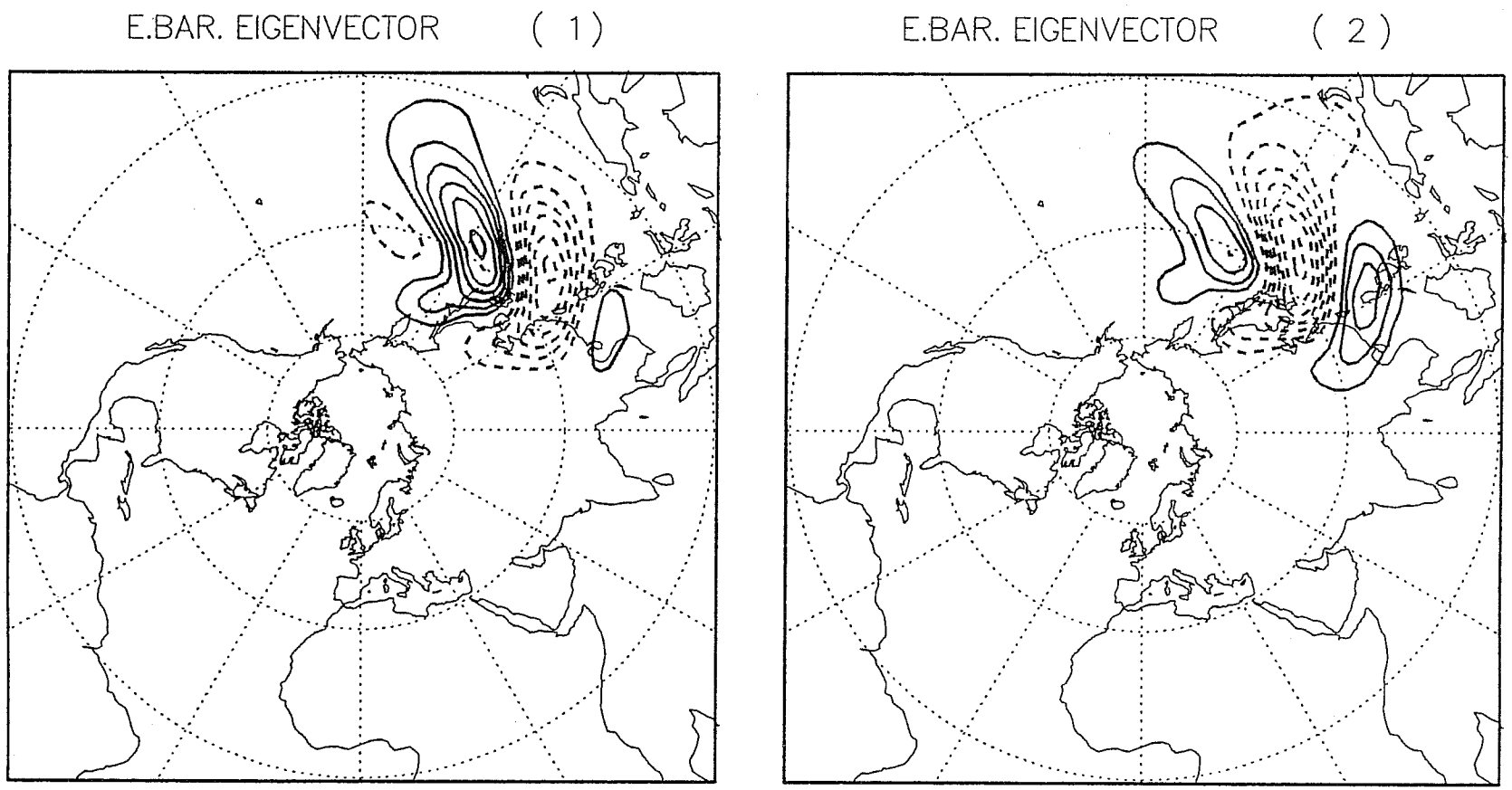

Fig. 13 First eigenmode of (L+L*) for equivalent barotropic climatological basic state.

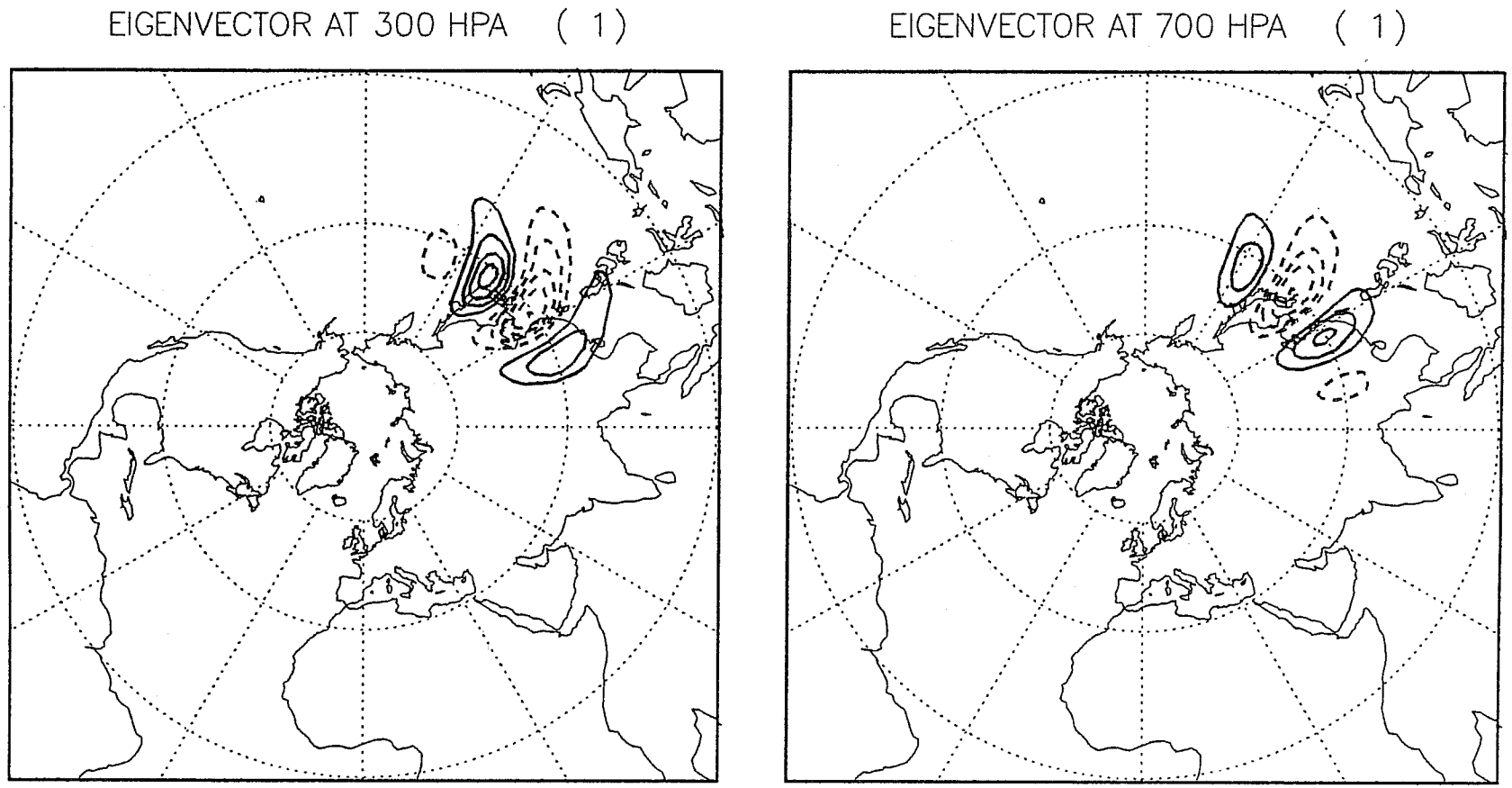

Fig. 14 First eigenmode of (L+L*) for baroclinic climatological basic state.

Hence the initial increase in perturbation energy is maximized if it is coincident with the dominant eigenvector of L + L*, the symmetric part of L.

Figs 15, 16 show some calculations of the dominant eigenmodes of L+L* for an equivalent barotropic and baroclinic two-level quasigeostrophic zonally varying basic state (respectively). For the barotropic basic state, the e-folding time is 1.1 days (compared with 10.1 days for the dominant normal mode); for the baroclinic basic state, the e-folding time is .65 days (compared with 6.4 days for the dominant normal mode). It is interesting to note that the structure of the optimum barotropic and baroclinic modes is quite similar.

Now in principal it would be possible to use the modes of the symmetric part of L to calculate initial perturbations for the Monte Carlo forecast-choose the N most unstable modes where N either represents the maximum size the computational constraints will permit, or, more satisfactorily, the number of significantly unstable modes. As such these modes may not necessarily have the structure of analysis errors, however by choosing the inner product to be spatially inhomogeneous, and represent a weighting in areas of large analysis uncertainty (eg over the oceans), more realistic perturbations can be attained by this technique.

Future development of this technique can be further enhanced using the full tangent model proposed for 4 dimensional data assimilation.

8. CONCLUSIONS

With projected upgrades in computer power, it will become technologically feasible to run Monte Carlo forecasts operationally in a few years. With estimates of optimal mode growth, it appears that a strategy for choosing the initial pertubations can be formulated. Finally, development of a probabilistic analysis of forecast flow fields will allow a synthesis of the ensemble forecasts to be given to the user. It can therefore be anticipated that there will be a significant change in the perception of the medium range forecast as a purely deterministic prediction. As a result it is hoped that the perceived skill of the medium range forecast will improve significantly.

References

Anderberg, M.R., 1973: Cluster analysis for applications. Academic Press, New York, 359 pp.

Brankovic C., T.N.Palmer, F.Molteni, S.Tibaldi, and U.Cubasch, 1989: Extended-range predictions with ECMWF models: III A study with time-lagged ensembles. Submitted to Q. J. R. Meteor Soc.

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. Am. Meteorol. Soc., 47, 200-220.

Epstein, E.S., 1969: Stochastic Dynamic Perdiction Tellus, 21, 739-759.

Farrell, B., 1988: Optimal excitation of neutral Rossby waves. J. Atmos. Sci., 45, 164-172.

Ferranti, L., T.N.Palmer, F.Molteni and E.Klinker, 1989: Tropical, extratropical interaction associated with the 30-60 day oscillation, and its impact on medium and extended range predictability. Submitted to J. Atmos. Sci.

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

Hoskins, B.J. and P.D.Sardeshmukh, 1987: A diagnostic study of the dynamics of the northern hemisphere winter of 1985-86. Quart. J. R. Meteor. Soc., 113, 759-778.

LaCarra, J.-F. and Talagrand, O., 1988: Short-range evolution of small perturbations in a barotropic model. Tellus, 40A, 81-95.

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

Palmer, T.N., 1988: Medium and extended range predictability and stability of the Pacific North American mode. Quart. J. R. Meteor. Soc., 114, 691-714.

Palmer, T.N. and S. Tibaldi, 1988: On the prediction of forecast skill. Mon.Wea.Rev., 116, 2453-2472.

Palmer, T.N., C.Brankovic, F.Molteni and S.Tibaldi, 1989: Extended range predictions with ECMWF models: I Interannual variability of operational model integrations. Submitted to Quart. J. R. Meteor. Soc.

Robinson, G.D.,1967: Presidential Address Q.J.R.Meteorol. Soc., 43, 409-418.

Tennekes, H., A.P.M. Baede and J.D. Opsteegh, 1987: ECMWF workshop on predictability in the medium and extended range. ECMWF, Shinfield Park, Reading.

Tibaldi, S., T.N.Palmer, C.Brankovic, and U.Cubasch, 1989: Extended-range predictions with ECMWF models: II Influence of model resolution on systematic error and extended range forecast skill. Submitted to Quart. J. R. Meteorol. Soc.

Zhang Z., 1988: The linear study of zonally asymmetric barotropic flows. PhD thesis, University of Reading.