ENSEMBLE PREDICTION USING DYNAMICALLY-CONDITIONED PERTURBATIONS

R. Mureau, F. Molteni, T.N. Palmer
ECMWF, Shinfield Park, Reading, UK

Abstract

We apply the technique developed in the companion paper by *Molteni and Palmer* (1992) as a means of providing dynamically-conditioned perturbations for ensemble forecasting with a primitive equation model. Four wintertime initial states are chosen, three at random, and one because of substantial development in the large-scale flow within four days, which the control forecast completely missed. A set of singular vectors are created using a quasi-geostrophic model linearised about basic states taken from data close to the chosen initial dates. These are interpolated onto the primitive equation model grid, and used as perturbations to the initial state. An ensemble forecast is made from the perturbed initial states. The dispersion of this ensemble is compared, for each date, with that from a second ensemble with initial perturbations constructed from 6-hour forecast errors. Throughout the forecast period, it is found that the amplitude of the perturbations is noticeably larger using the singular vectors. The dispersion of the ensembles using the forecast-error perturbations did not indicate that the control forecast from the case with substantial development was likely to be poor. By contrast, the (envelope) dispersion of the ensemble using the singular vectors was notably larger for this case than the other three. A number of members of this ensemble were particularly skilful in predicting weather-related elements of the flow, such as low-level temperature change. It is found that the evolution of perturbations which are initially localised over the western Pacific, or western Atlantic, develop blocking-like structures several days later, over the eastern oceans. The development of these structures is highly nonlinear.

The growth of the singular vector perturbations was not as large in the primitive equation model as in the quasi-geostrophic model, probably due to interpolation problems exacerbated by inconsistent orographic representation. Further work to overcome these, and other problems is indicated.

1. INTRODUCTION

Whilst operational numerical weather prediction (NWP) models make skilful predictions into the second week of the forecast about a third of the time, their practical value at this range is limited by day-to-day variability in skill. For example, within a typical month, the skill of a series of day 7 forecasts can vary from 80% anomaly correlation with the verifying analysis to complete lack of skill. Such variability undermines the usefulness of NWP, even though the mean skill over such a period may be significant.

The formulation of numerical models will continue to be improved. The accuracy of the initial state will also be increased, both through more advanced data analysis schemes, and through more complete data coverage. However, even when the effects of model imperfections are truly negligible and when the state of the whole atmosphere can be measured with high uniform accuracy, variability in forecast

skill associated with amplification of inevitable uncertainties in the initial state by flow-dependent instabilities will be sufficiently large to restrict the useful range of the forecast. *Palmer and Tibaldi*, 1988 argue that much of the observed variability in current forecast skill is associated with such flow-dependent instability. Flow-dependent estimates of the amplification of analysis uncertainties by such instabilities, ie estimates of the predictability of the flow, could therefore provide a key to the enhancing the practical usefulness of NWP in the coming years.

Predictability has been studied by stochastic-dynamic modelling (*Epstein*, 1969), ensemble prediction (*Leith*, 1974) and statistical techniques (*Kalnay and Dalcher*, 1987). In the stochastic-dynamic approach, prognostic equations for higher moments of the atmospheric probability distribution function are integrated forward in time. In the ensemble technique, a deterministic NWP forecast is made several times from a set of perturbed initial states, each of which is, in principle, consistent with the available data. The model-generated ensemble dispersion is taken to be a measure of forecast predictability. In the statistical method, predictors of forecast skill include the skill of earlier forecasts, the spread between consecutive forecasts and between current forecasts from different operational centres, and large-scale flow patterns.

Formulation of stochastic-dynamic equations is hampered both by the absence of physically-based closure schemes (at least for the range in which error growth is nonlinear), and by the need for massive computational power in determining individual components of the higher-order moments. Recently, progress has been made using adiabatic direct and adjoint tangent models to integrate an initial estimate of the analysis error covariance (*Veyre*, 1992). This is equivalent to stochastic-dynamic prediction in the range in which error growth is linear, probably within the first three days of integration, as discussed below. Computational problems are overcome by considering predictions for specific local regions.

The ensemble technique is, in principle, not restricted by an assumption of linear or adiabatic error growth, and is therefore a potentially useful tool for estimating predictability in the medium and extended range. However, problems of closure and computation cost are traded for problems of sampling, since the size of the ensemble is necessarily much smaller than the number of degrees of freedom available to perturb the initial conditions.

The statistical, 'no-cost', technique provides a basic measure of skill against which dynamically-based schemes will ultimately have to be assessed. Research to date indicates useful skill at 3 days, but insufficient skill at day 5 to be of operational value (*Molteni and Palmer*, 1991).

For the ensemble method, addition of random noise to an initial state (closest to the original meaning of the Monte Carlo method) is not viable, and perturbations which do not project onto meteorologically-balanced modes will be largely removed by initialization (*Hollingsworth*, 1980). However, balance is not in itself a sufficiently strong constraint to determine the structure of viable perturbations. Current NWP models have $M=O(10^6)$ balanced degrees of freedom, very much larger than a practicable maximum ensemble size N_e (say between 10 and 100). On the other hand, the number of degrees of freedom N_s required to explain most of the variance of the synoptic-scale flow is much smaller than M, perhaps on the order of 10-100. ($M>>N_s$ because of atmospheric nonlinearity, the requirement to represent accurately topographic and other physical forcing effects, and because of the requirement of assimilating local observations without large interpolation error.) It is very unlikely that N_e randomly chosen perturbations from the M balanced degrees of freedom will have any significant projection onto the N_s synoptic degrees of freedom.

The choice of initial perturbations for an ensemble forecast could be based on a subset of the synoptic degrees of freedom which are themselves energetically active and therefore affect forecast skill. Our identification of the energetic degrees of freedom in relation to the large-scale flow pertaining at the time of the initial analysis is guided by recent theoretical developments.

For many years it was widely believed that the energetically-active synoptic-scale disturbances arise primarily as exponentially-growing modal instabilities of the large-scale background flow. *Farrell* (1990, 1991), however, argues from both observational and theoretical grounds that the variance associated with synoptic disturbances is associated with non-modal energy transfer from the background flow associated with stochastic forcing, which could arise e.g. from a nonlinear upscale energy cascade. A quantitative description of these non-modal disturbances has been addressed in the preceding companion paper (*Molteni and Palmer*, 1992; hereafter MP; see also *Lorenz*, 1965, *Lacarra and Talagrand*, 1988), using singular vectors (SVs) calculated from the of the linearisation of a 3-level quasi-geostrophic (QG) model. On the synoptic timescale, the variance explained by the SVs would be very much larger than that of the exponentially-growing modal instabilities.

According to these results, analysis errors which project significantly onto these SVs could generate substantial forecast error. Conversely, analysis error which projects onto weakly growing or decaying phase-space directions would not be expected to provide a substantial contribution to synoptic error growth. One needs an a-priori knowledge of the likely projection amplitude of analysis error onto the energetic phase-space directions. Estimates of analysis error variance fields by optimum interpolation schemes provide some useful information on such projection amplitudes.

In this paper we describe a number of ensemble integrations using a selection of the fastest growing SVs as initial perturbations. By exciting the most energetically growing disturbances in the linear error growth range, it may also be possible to capture possible regime transitions of the large-scale flow in the range where error growth is predominantly nonlinear. The use of such perturbations to the initial analysis should minimise occurrences (*Brankovic et al.*, 1990) where the ensemble spread was small, yet all members of the ensemble failed to capture the development of the observed flow. An ensemble forecast which produces an overly optimistic forecast of predictability (through small ensemble dispersion) is less practically useful than one which errs on the side of caution (by maximising possible ensemble dispersion).

The purpose of this paper is not to compare the use of optimal SV perturbations (SVPs) with another more conventional technique, and to demonstrate the potential of the dynamically-conditioned perturbations for ensemble prediction. We study how these SVPs, interpolated onto a T63 19 level grid, evolve when integrated with a primitive equation model. The effectiveness of the SVPs will be compared with a second independent set of (balanced) perturbations with a more random character, which were based on 6 hour forecast errors of the ECMWF operational model of the days prior to the chosen initial state.

2. OUTLINE OF PAPER AND DESCRIPTION OF EXPERIMENTS

After some introductory material in this section, a description of the perturbations selected for the ensemble case studies is given in section 3, based either on optimal SVPs, or a set of perturbations derived from previous 6-hour forecast errors (hereafter FEPs). An evaluation of perturbation growth in these ensemble forecasts and on their verification is given in sections 4 to 6.

Ensemble forecasts were integrated from four different initial dates. Each ensemble comprised a control forecast and a number of perturbed forecasts which were run from the initial state of the control with the selected perturbations both added to, and subtracted from, this initial state (it is assumed that there is no a priori information concerning the sign of any one perturbation). A subset of the 20 fastest growing SVs was selected from the full set as calculated with the T21L3 QG model (see MP). Ensembles with the FEPs consist of 24 members each. In both sets, the perturbations were orthogonalised and scaled at all levels and all gridpoints such that the height field deviation at 500 hPa between 30N and 80N was 10 m RMS.

The initial dates, the size of the ensembles and the length of the forecasts are given in Table 1. For each case we have one 10-day ensemble forecast with the FEPs, and one 5 day ensemble forecast with

the SVPs (except for one SVP ensemble which consisted of 10-day forecasts). All forecasts were started from 00 UTC, and were carried out with a T63 version (cycle 36) of the ECMWF operational model.

Table 1

1	2 December 1988	00Z	SVP FEP	5 days 10 days	40 members 24 members
2	17 January 1989	00Z	SVP FEP	10 days 10 days	36 members 24 members
3	27 January 1989	00Z	SVP FEP	5 days 10 days	40 members* 24 members
4	1 March 1989	00Z	SVP FEP	5 days 10 days	40 members 24 members

* : 2 members integrated to 10 days

In addition, nonlinear ensemble integrations of the QG model were made using the SVPs. In order that the nonlinear QG integrations were consistent with the T63 integrations the following forcing was added to the QG equations of motion. Firstly, for the control integration of the QG model, a linear relaxation term was added to the model equations, which forced the model towards the T63 control integration. The tendencies from this relaxation were then stored and used as prescribed forcing terms in the perturbed QG integrations.

Three of the four cases were randomly chosen. They are characterized by fairly typical flow developments as listed below. One of the four cases (27 January, 1989) was chosen because of its strongly anomalous character, and because the model had great difficulties in capturing the correct flow development.

2 December 1988

Initially blocked flow exists over the European area. This weakens within the first few days, after which a trough develops over this area. Over the Pacific the flow changes from zonal to meridional within the first 5 days.

17 January 1989

Initially, the flow is zonal throughout the entire hemisphere, and remains zonal over the Pacific. Over

the Atlantic and Europe transients begin to disturb the flow at day 3, after which a very weak ridge develops.

27 January 1989

This case has meridional flow throughout the forecast period. Initially, a ridge occurs over Europe and a trough over the Western USA. At day 3 the ridge over the European area intensifies moderately at day 3, whilst an intense ridge develops over the Eastern Pacific. Over the Western USA a very severe cold air outbreak occurred: surface temperatures dropped in 5 days from values above 0°C to -30°C. The control forecast underestimated significantly this drop in temperature. This period will be discussed in more detail below.

1 March 1989

The initial flow is meridional with troughs over Northern Europe, the Eastern Pacific and the American continent, and more intense ridges over Eastern Europe and the Central Pacific. The flow patterns are fairly persistent, though the amplitude varies a little over the forecast range with smaller amplitude at day 3 recovering at day 5.

3. GENERATION OF PERTURBATIONS

In this section we detail the construction of both SVPs and FEPs. Both sets of perturbations give rise to ensembles of initial states that are in close meteorological balance.

3.1 The SVPs

The SVs were generated using the T21L3 QG model described in MP, using techniques outlined in that paper. Let A denote the (integral) resolvent operator associated with the linearised QG model, mapping the initial tangent space to a forecast tangent space. The optimal SV is, by definition equal to the singular vector with largest singular value, corresponding to the eigenvector of the operator A^*A with largest eigenvalue. Here '*' denotes the adjoint operation defined with respect to a global kinetic energy norm (integrated over the three levels of the model). For simplicity, a 5-day time mean of analysed fields around the initial date was taken as the basic state.

SVs were calculated for optimisation times of 12, 24, 48 and 96 hours (MP). For the nonlinear T63 Monte Carlo integrations described below, it was decided to use a 12 hour time interval for the following reasons:

1) In order to minimise possible inconsistencies between the QG model and the T63 model (particularly with respect to orographic specification), it was decided to keep the optimisation

period as short as reasonable. Possibly related to this we noted that fewer of the initial states perturbed with 12 hour SVPs suffered any significant changes when passed through the T63 model's initialisation scheme, compared with longer timescale SVPs.

2) Many of the 12 hour SVs were quite spatially isolated, which made analysis of their amplification and dispersion relatively easy.

3) The linearised assumption was most unequivocally appropriate for a 12 hour optimisation time.

4) As shown in MP, the difference in the QG model of a 12 hour optimal SV and a 2-day optimal SV, both evaluated at 2 days, is not very great. This can probably be related to the fact that the QG assumption necessarily filters sub-synoptic (e.g. frontal) instabilities which would otherwise tend to be dominant on a 12-hour timescale.

In MP, the growth of optimal SVPs were shown for the 2 December, 1988 and 27 January, 1989 cases. Below we illustrate some of the other significant eigenvectors further down the spectrum for all four cases. The amplification factor over 12 hours for the fastest growing perturbations exceeds a factor of 2 for each initial condition (Table 2) as calculated in the quasi-geostrophic model. About 25% of the total spectrum of (1449) eigenvectors represent growing structures (eigenvalues larger than 1), and about 50 had amplitude doubling times shorter than 2 days, typical of the growth of forecast errors. For practical reasons it was necessary to limit the size of our ensembles to the first 20 SVPs. Each was included with positive sign and negative sign.

Table 2

Global amplification factors in perturbation amplitude over 12 hours
of first 10 SVPs for the four initial states

	881202	890117	890127	890301
1	2.223	2.216	2.396	2.079
2	2.218	2.212	2.395	2.076
3	2.036	2.128	2.162	1.878
4	2.030	2.123	2.157	1.876
5	1.879	2.064	2.082	1.813
6	1.876	2.058	2.079	1.806
7	1.855	1.955	1.986	1.746
8	1.750	1.949	1.981	1.743
9	1.734	1.892	1.903	1.734
10	1.724	1.889	1.816	1.693

The SVPs were interpolated from the T21L3 QG model to the T63L19 primitive equation model. Quadratic interpolation was used to map the QG vorticity field onto the 19 levels. Temperature was then calculated from the thickness, derived from the QG streamfunction.

A selection of the 500 hPa height fields associated with 6 SVPs is shown in Fig. 1 for the initial state of 2 December 1988. The perturbations are ordered according to their global growth rates. SVPs 1 and 2 have a very localized structure over the Eastern-Asian Western Pacific jet entrance (and are in phase quadrature), confirming synoptic experience that the atmosphere is very unstable in this area. Higher SVPs, up to the tenth, have a weaker, less localized wavy pattern over the Asian-Pacific area. SVPs 11 and 12 are similar to SVPs 1 and 2, but with localized structures upstream of the West Atlantic jet entrance. Beyond SVP 12 we find perturbations spread out over larger parts of the hemisphere (e.g. SVP 15 is shown in Fig. 1). SVPs 9 and 10 (not shown) are located in the Southern Hemisphere.

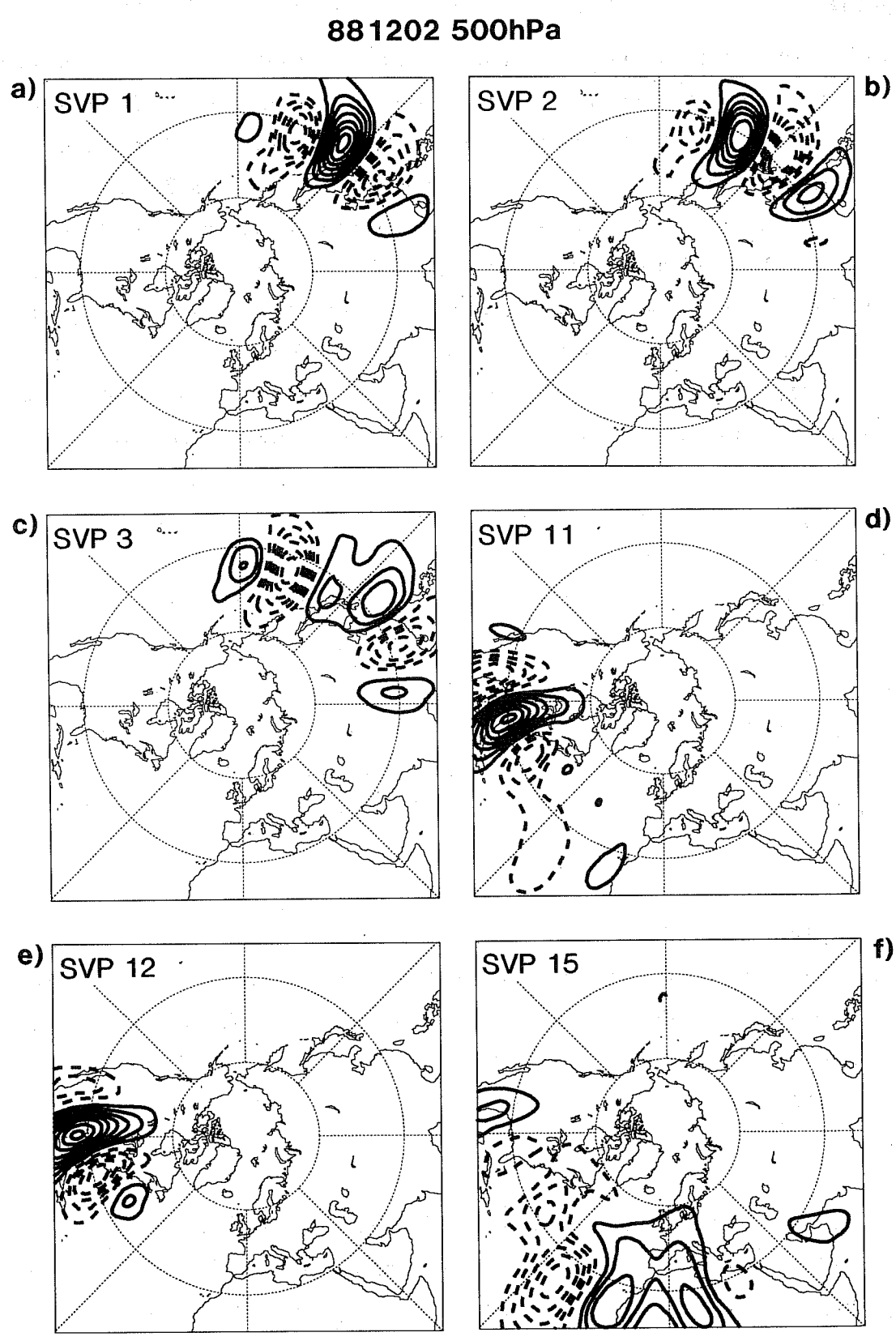

Fig. 1 Selected normalised SVPs optimised over 12 hours from 2 December, 1988. 500 hPa geopotential height, contour interval 10 m. a) 1st SVP b) 2nd SVP c) 3rd SVP d) 11th SVP e) 12th SVP f) 15th SVP.

It is possible to make a similar partitioning for the perturbations calculated for the other 3 initial states. The partitioning is broadly the same, but there are some differences of detail. This can be more clearly seen at the 200 hPa level. As shown in Fig. 2, the fastest growing SVP, for instance, is not always found in the same area; it is sometimes positioned near the West Pacific (case 1 and 3) or, shifted westward, over Asia (case 2 and 4). In all cases, the relative location of the fastest growing SVP can be broadly related to a downstream maximum in the jet strength. This is illustrated in Fig. 3 where the 200 hPa maps of the vector wind and isotachs are shown. The dots on the map show the perturbation height maxima and minima of the fastest growing SVPs, some of them are numbered according to the position of the eigenvalue in the spectrum. For example, the position of the first SVP in the 2nd case is located upstream of a region of 60 m/s winds between 50 and 90E.

For higher eigenvectors, comparison between the 4 cases is difficult. The localized 11th and 12th SVPs over the Atlantic jet entrance (Fig 1), for instance, are found in the second case as SVPs 15 and 16 and in the fourth case as SVPs 12 and 13. The third basic state did not have such perturbations among the first twenty. The reason for this appears to be related to the fact that the Pacific jet in this case is so strong compared with the Atlantic jet that the Pacific modes dominate the spectrum in the first twenty. (This creates possible sampling problems when using SVPs in global ensemble prediction, since the dispersion of an ensemble created from the first 20 SVPs could be very small e.g. over an area such as the Euro/Atlantic region in the first days of the forecast period. This will be discussed in a future paper.)

3.2 The FEPs

The ECMWF operational analysis is made every 6 hours using as first guess, a 6-hour forecast from the earlier analysis. The difference between the first guess and the final analysis gives a partial estimate of possible analysis error, at least in regions where there are validating data. All analyses and 6-hour forecasts are routinely archived (at 00, 06, 12, 18 UTC) for all model levels and all model variables.

A linear combination was taken of the most recent first guess errors (with respect to the initial date of the ensemble forecast). Two problems had to be overcome. First of all the observational data density differs from one analysis time to the next (associated in particular with satellite data coverage) causing a potential bias in the first guess error. This bias was reduced by combining the first guess errors of two successive analysis times. Secondly, to avoid using strongly correlated error patterns and to ensure that all perturbations had the same amplitude, the perturbations were orthonormalised by applying a Gram-Schmidt procedure such that the mean deviation at 500 hPa was 10 m RMS. As

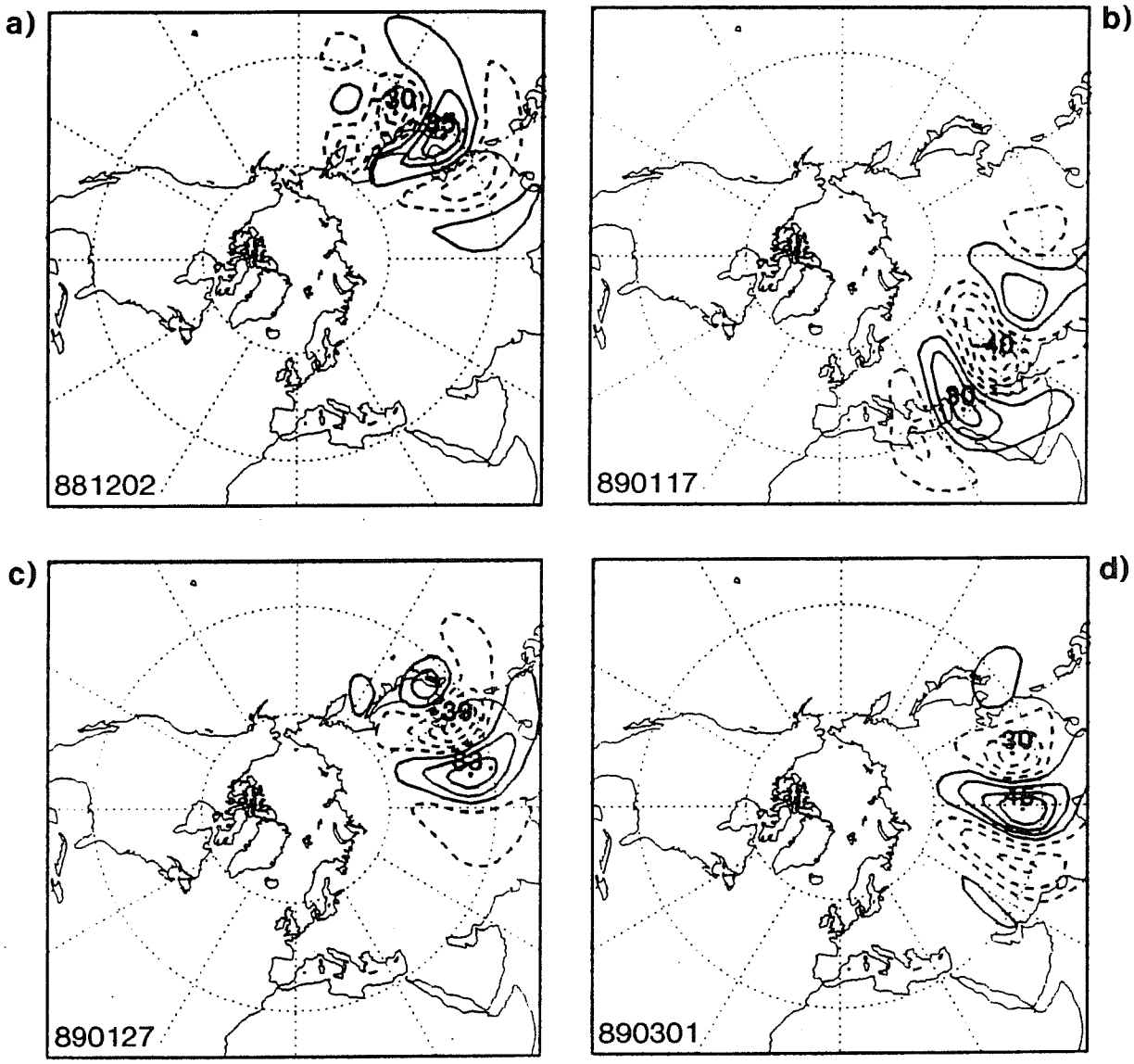

Fig. 2 Fastest growing SVP, optimised over 12 hours from a) 2 December, 1988 b) 17 January, 1988 c) 27 January, 1989 d) 1 March, 1989. 200 hPa geopotential height, contour interval 10 m.

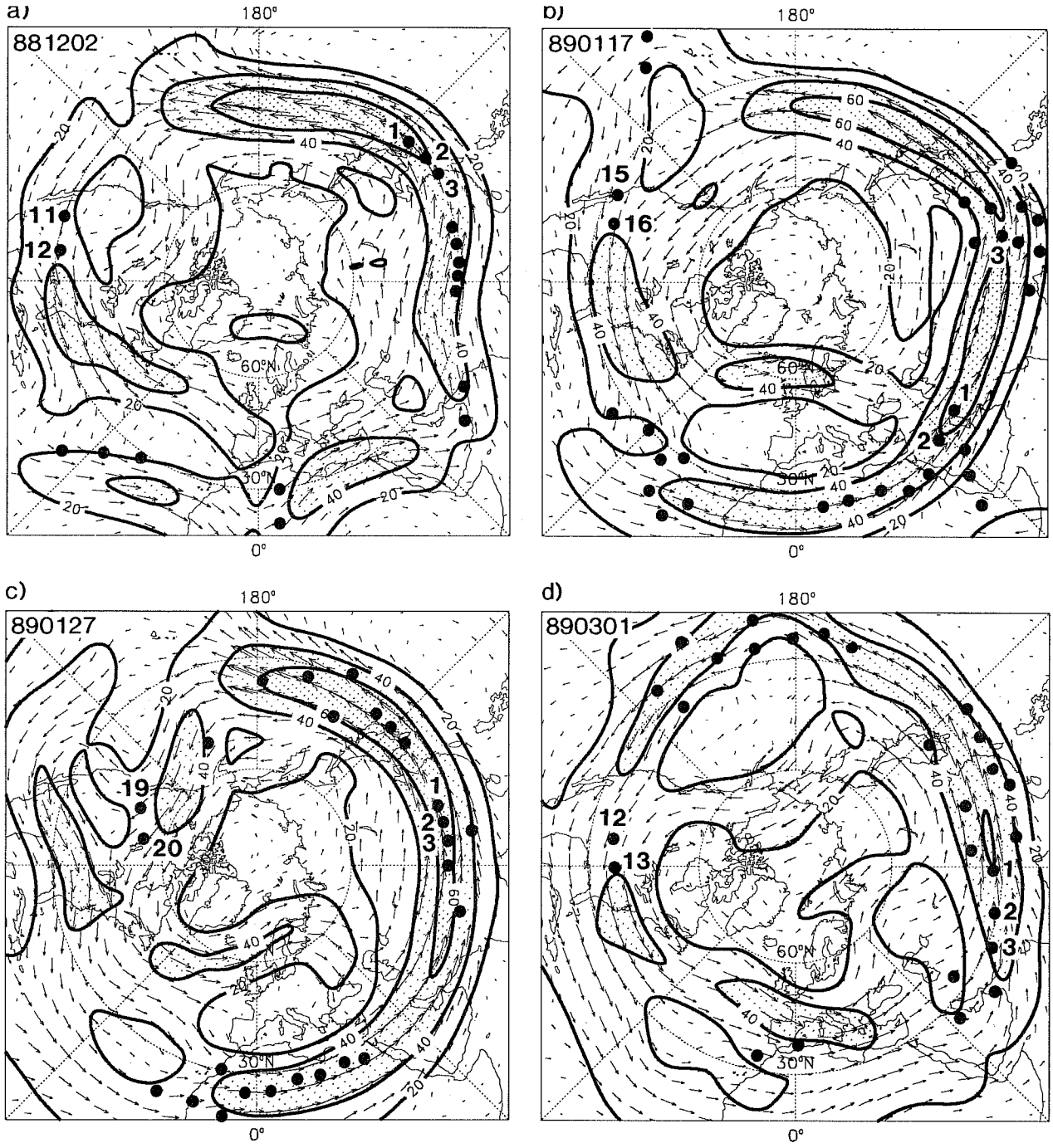

Fig. 3 200 hPa wind vectors and isotachs, together with positions of maxima in the 500 hPa height associated with the principal SVPs for a) 2 December, 1988 b) 17 January, 1988 c) 27 January, 1989 d) 1 March, 1989.

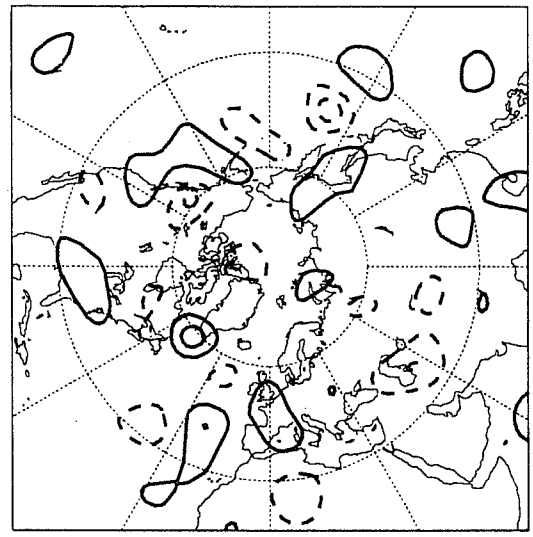

Fig. 4 500 hPa height field associated with typical FEP (contour interval: 10m).

it turned out, the corrections required to orthogonalise the perturbations were in general very small, suggesting that the successive forecast errors were, in fact, very nearly independent. Note that the perturbations thus created are in balance as long as one applies the technique to all model variables and all levels. For small-amplitude perturbations, the total fields will also be in balance. A typical example of the 500 hPa height field associated with an FEP is shown in Fig. 4.

4. ENSEMBLE SPREAD WITH RESPECT TO CONTROL FORECAST

Fig. 5 shows the 500 hPa RMS differences between control and perturbed forecasts, averaged between 30N and 80N, for the FEP and SVP ensembles for each of the four initial dates.

Within the forecast period, the growth of the spread of the optimal SVPs is much larger than that of the FEPs. Whilst the smallest spread values are similar for both sets of perturbations (e.g. about 20 m RMS at day 5), the extreme spread is about 50 to 60 m for the FEPs as compared to 80 to 140 m for the SVPs.

In Fig. 6 the amplification factors of the northern hemisphere (NH) mid-latitude 500 hPa height of the individual SVPs after 12 hours is shown for all four cases. The top set of solid lines show the SVPs as they develop in the nonlinear integrations in the QG model. The bottom set of dotted lines show the SVPs as they develop in the T63 model. Only the development of SVPs with one sign (defined as "positive") are shown. Where perturbations are missing, they had most of their amplitude in the Southern Hemisphere. Fig 6 shows that the 12 hour development of the SVPs in the QG and T63 models correlate reasonably well. Moreover (not shown) perturbations associated with the larger eigenvalues generally have largest amplitude throughout the T63 model integration. However, the growth in the T63 model is somewhat weaker than in the QG model. This may be attributable to problems associated with the vertical interpolation of the perturbations from one model to the other or to problems associated with the different representation of orography in the two models.

An example of the development of 2 positive and negative SVPs in the T63 model is given in Fig. 7; the first SVP for the 27 January, 1989 case (Pacific perturbation), and the 12th SVP for the 2 December, 1988 case (Atlantic perturbation). The panels show the development of the 500 hPa height field at days 0, 1, 2, 3, 4 and 5, with contour interval of 20 m for days 0, 1, and 2, and 50 m for days 3, 4, 5. By comparing the development of perturbations started from the positive and negative perturbations, the linearity of the growth can be examined. Initially, up to day 1.5, the growth is linear: positive and negative perturbations grow almost exactly in phase. Non-linearity first becomes apparent at about day 1.5 and the development of the positive and negative mode becomes

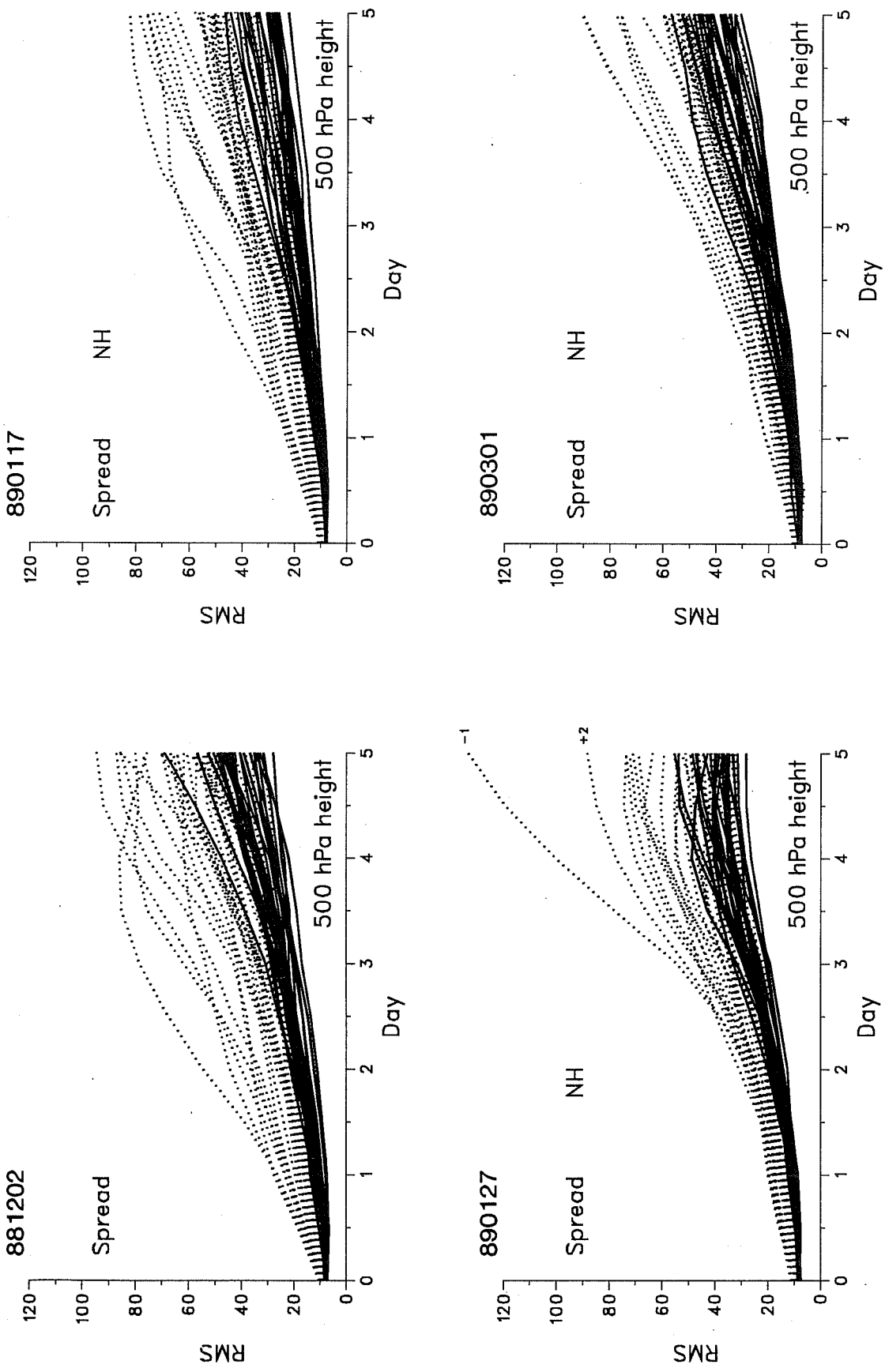

Fig. 5 500 hPa RMS height difference between perturbed and control integrations. Dotted lines - SVPs, full lines - FEPs.

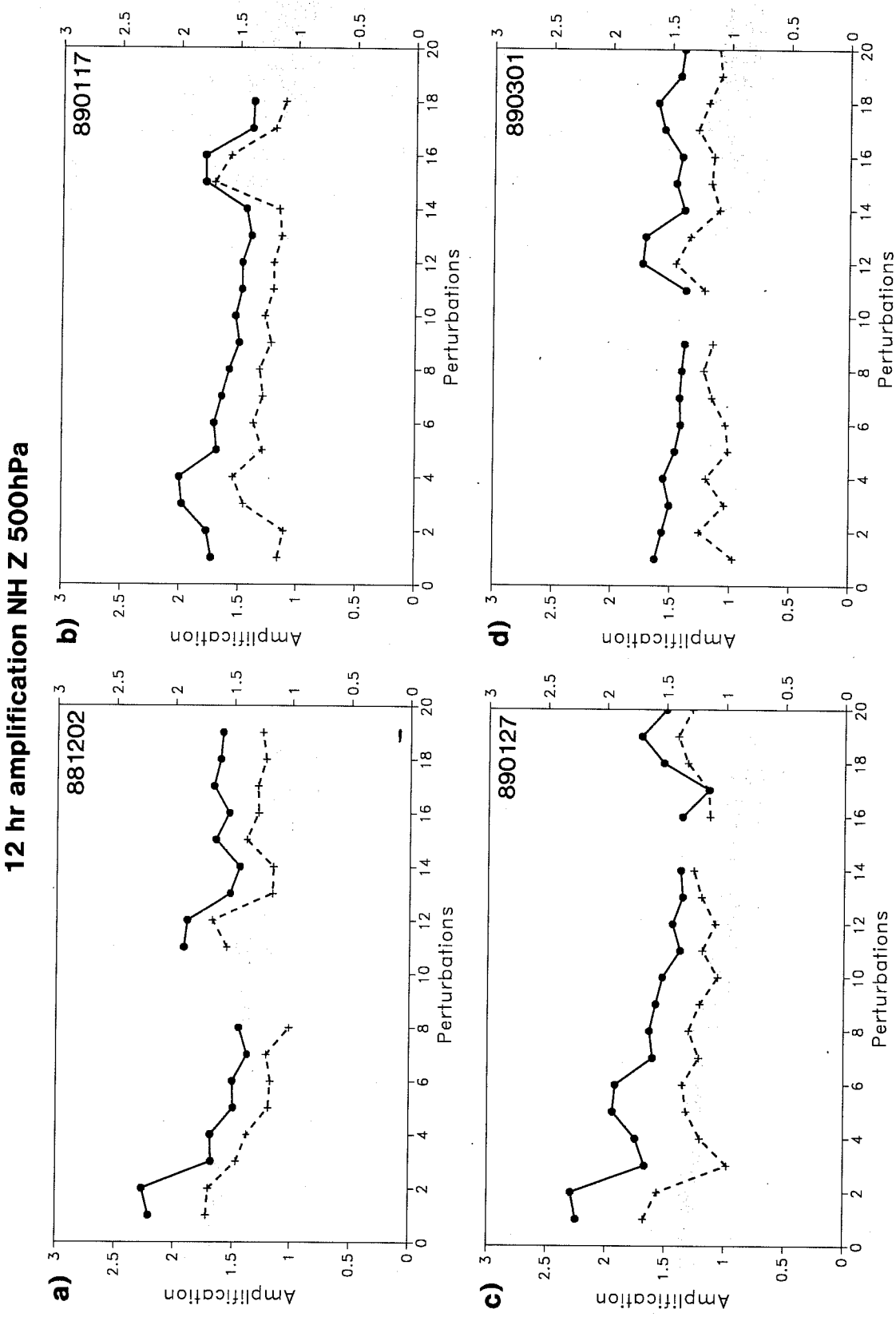

Fig. 6 Amplification of mid-latitude 500 hPa height of the individual perturbations after 12 hours for a) 2 December, 1988 b) 17 January, 1988 c) 27 January, 1989 d) 1 March, 1989. The top set of curves marked by the dots shows the SVPs as they develop in the nonlinear integrations in the QG model, the bottom set of curves marked by the crosses show the SVPs as they develop in the T63 model. Where perturbations are missing, they had most of their amplitude in the Southern Hemisphere.

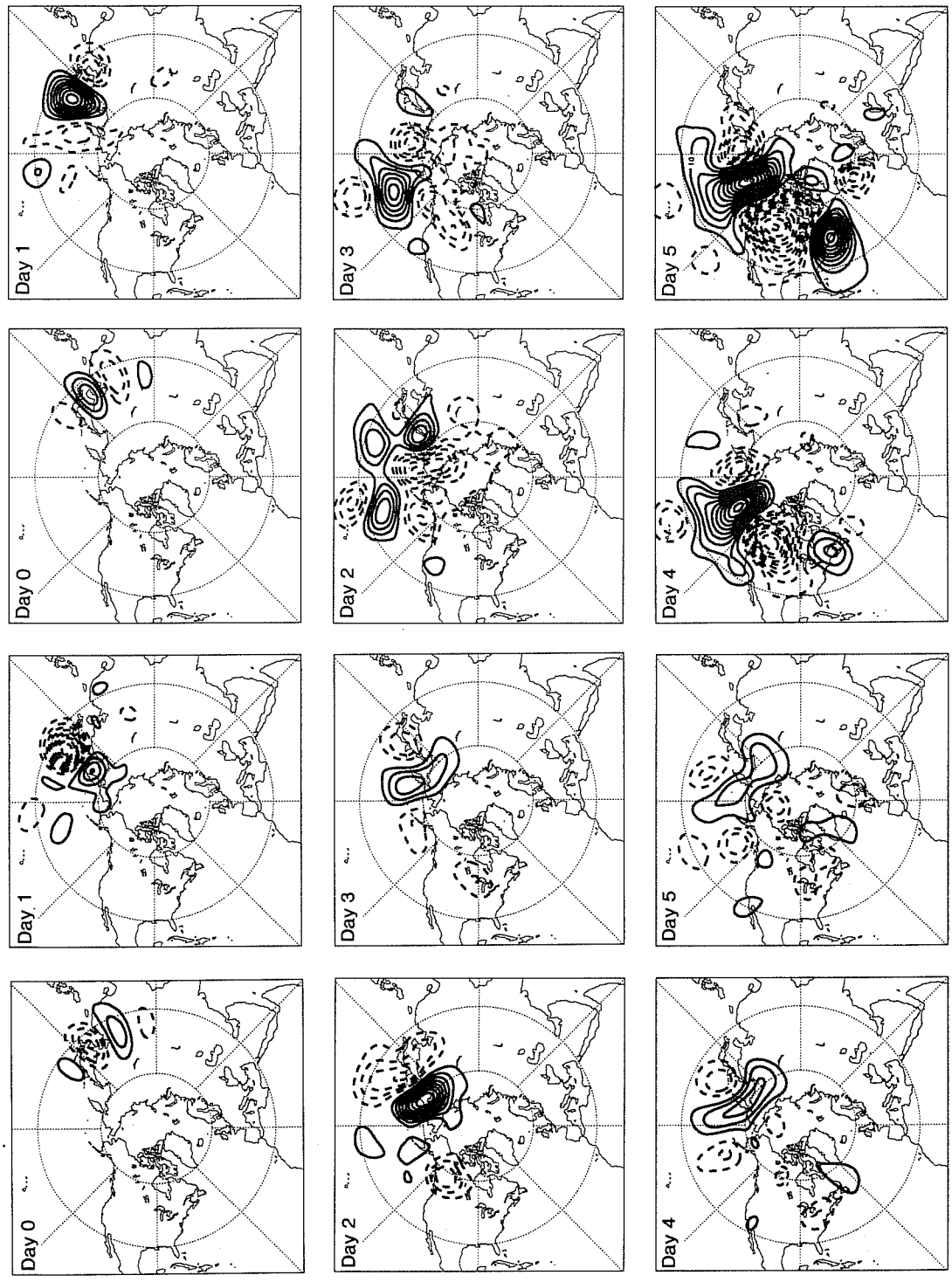

Fig. 7 Dispersion and amplification in the T63 model of 500 hPa height associated with SVPs a) SVP +1 from 27 January, 1989 b) SVP -1 from 27 January, 1989 c) SVP +12 from 2 December, 1988 d) SVP -12 from 2 December, 1988. Days 0, 1, 2 (contour interval 20 m), 3, 4, 5 (contour 50 m).

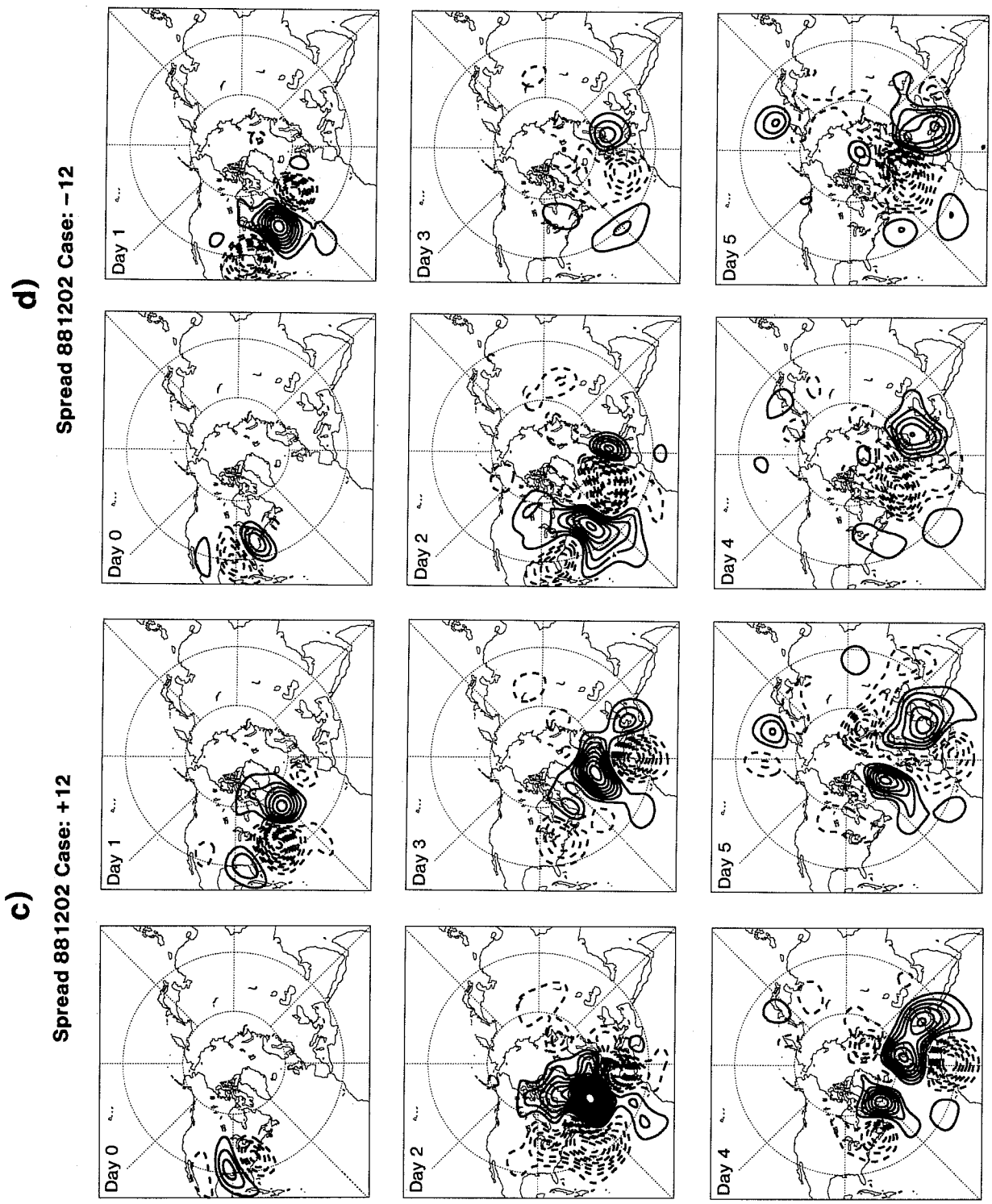

Fig. 7 Continued.

increasingly out of phase, to such a degree that in some areas the patterns at day 4.5 correlate positively. Because many of the SVPs modes are so spatially localised, the initial perturbations discussed here have an amplitude up to about 50 m, which may be a factor of two or more larger than likely analysis error. When the amplitude of the perturbations was scaled by a factor of 5, perturbation growth was linear up to day 3. The sensitivity of error growth to perturbation amplitude is discussed further below.

At about day 4 in Fig 7 there is considerable similarity between perturbations -1 and +12, and blocking patterns in the Pacific and Euro/Atlantic areas respectively. Note the strong dipole shape of the Euro/Atlantic block in particular. In both cases, the development is nonlinear. The resemblance with blocking patterns confirms our hope that individual members of the ensemble can simulate blocking activity when such development is possible from a given initial state, yet not predicted by the control integration.

In Fig. 8 we show a longitude height cross section of the development of SVP -1 for the 27 January, 1989, averaged between 30 and 50N at hours 0,12,24,36,48,72,96 and 120. The perturbation has pronounced westward tilt with height, though, as discussed in MP, it is gaining energy through barotropic conversions as well. Development between day 1 and day 3 is complex, and as discussed above, is associated with nonlinear development. The amplification of the positive anomaly centre around 160W between days 2 and 3 becomes a quasi-stationary feature which largely amplifies in situ.

5. ENSEMBLE DISPERSION AND FORECAST SKILL

The primary goal of this paper is to explore the possibility of developing an ensemble forecast technique in which initial perturbations are dynamically conditioned. We have not yet addressed the problem of how to constrain the initial perturbations according to inhomogeneities in data density. This is an essential component of a viable operational ensemble scheme, and will be dealt with in a future paper. Therefore, we only make a simple exploration of the relationships between the spread of the ensemble and the skill of the control integrations. In any case, a study of that kind based on just four ensembles cannot be definitive.

In Fig. 9 the RMS errors of the 4 control forecasts are shown at 12, 24, 72, 96 and 120 hours (NH mid-latitudes, 500 hPa height). Along the axes the 4 cases are labelled in chronological order. The very large forecast error of the third case (27 January, 1989) stands out very clearly beyond day 3. (Indeed this case was chosen for this reason). In view of the limited sample, we concentrate on just one question: was it possible to anticipate this poor skill from the spread of the ensemble?

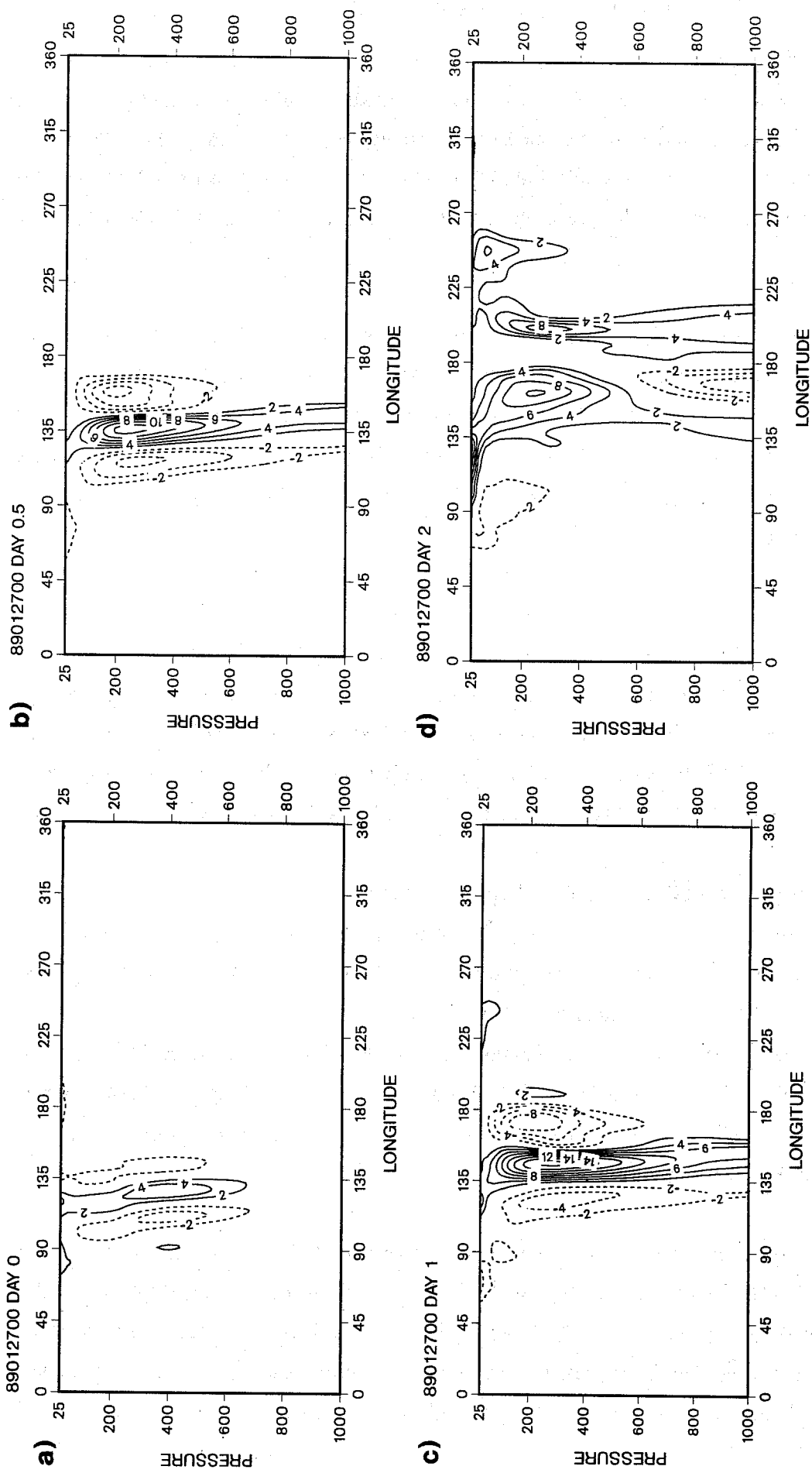

Fig. 8 Longitude-height cross section of geopotential height (averaged between 30N and 50N) of the dispersion and amplification in the T63 model associated with SVP -1 from 27 January, 1989. a) day 0, b) 12 hours, c) 24 hours, d) 48 hours, e) 72 hours, f) 96 hours, g) 120 hours. Contour interval 20 m.

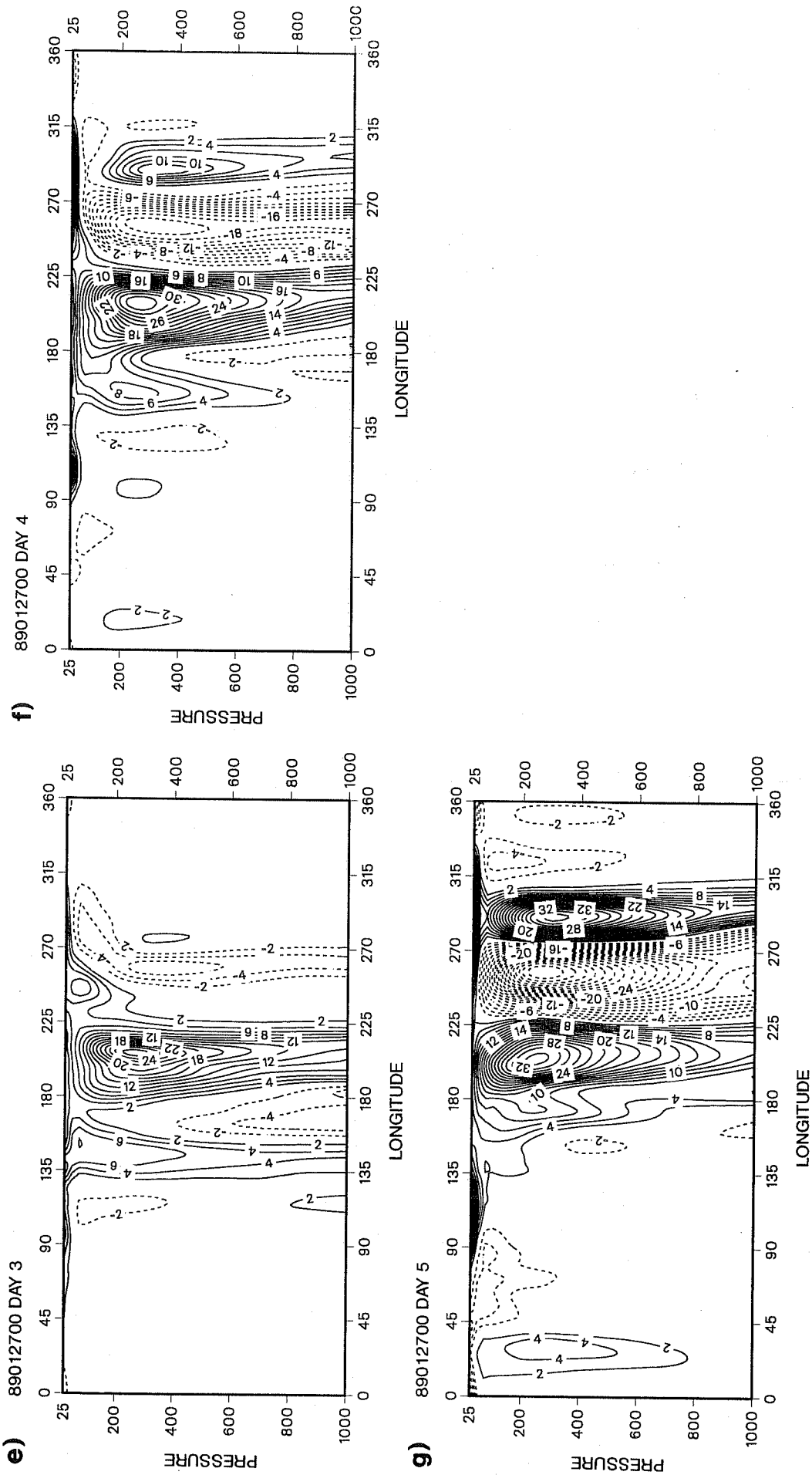

Fig. 8 Continued.

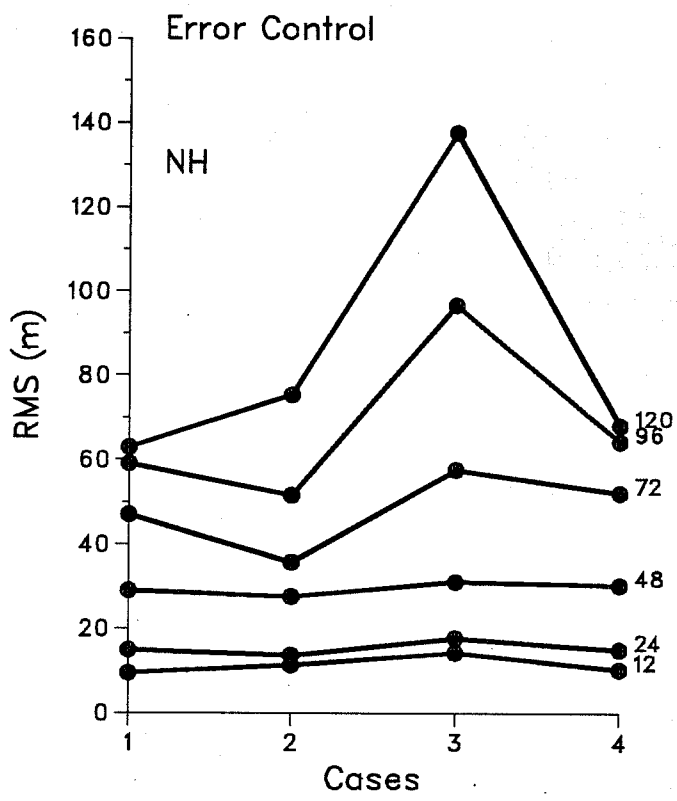

Fig. 9 NH 500 hPa RMS height error of the control integration of the four T63 control forecasts.

The spread of the three different ensembles (the SVP ensemble using the QG nonlinear model, the SVP T63 ensemble, and the FEP T63 ensemble) is compared against the error of the control. Two estimates of spread are tested: the mean and maximum RMS deviation from the control within the ensemble.

In the top panel of Fig. 10 is shown the spread due to the FEPs in the T63 model. In the middle and bottom panels is shown the spread from the SVPs, integrated in the T63 model and in the QG model, respectively. In all three panels, the mean spread is found on the left hand side and the maximum spread on the right hand side. In all cases, the spread is expressed as a multiple of its initial value.

The FEPs exhibit little variability from case to case, indicating little discrimination in forecasting predictability. The SVPs in the T63 model, on the other hand, are well able to anticipate the case of poor skill, but only if the extreme spread is used as an indicator. In the QG model, where growth is very rapid, both the mean and the extreme spread of the third case start to dominate within one day. The reason that the T63 model appears to perform poorly compared with the nonlinear QG integrations is probably associated either with the interpolation of the SVPs onto the 19-level grid or with different representations of orography.

The anomaly correlation skill of individual members of the FEP, SVP and QG ensembles are shown in Fig. 11 at day 5 of the forecast. The skill of the control forecasts are shown by dots joined by solid lines. The skill of the FEP ensemble members closely follows that of the control, whilst there is a substantial variation in skill of the SVPs. In all four cases an SVP ensemble members, note that in all four cases the best forecast has an anomaly correlation score of 0.8. A number of the SVP forecasts are also rather poor, but since individual initial perturbations have not been constrained by a priori knowledge of analysis error this is not surprising.

6. INTEGRATIONS FROM 27 JANUARY 1989

A major concern in formulating a strategy for ensemble forecasting is well illustrated by comparing the effect of the FEPs with the SVPs for the 27 January, 1989 case. It was clear from Fig. 10 that the FEPs generate insufficient ensemble dispersion to indicate that the control integration was likely to be particularly unskilful.

Several forecasts in the SVP ensemble deviated quite substantially from the control. In particular there are 3 individual forecasts with NH anomaly correlation skill over 0.6: forecasts from SVPs -1,+2 and +4. In order to study these further, the control and forecast from SVP -1 have been continued to day 10.

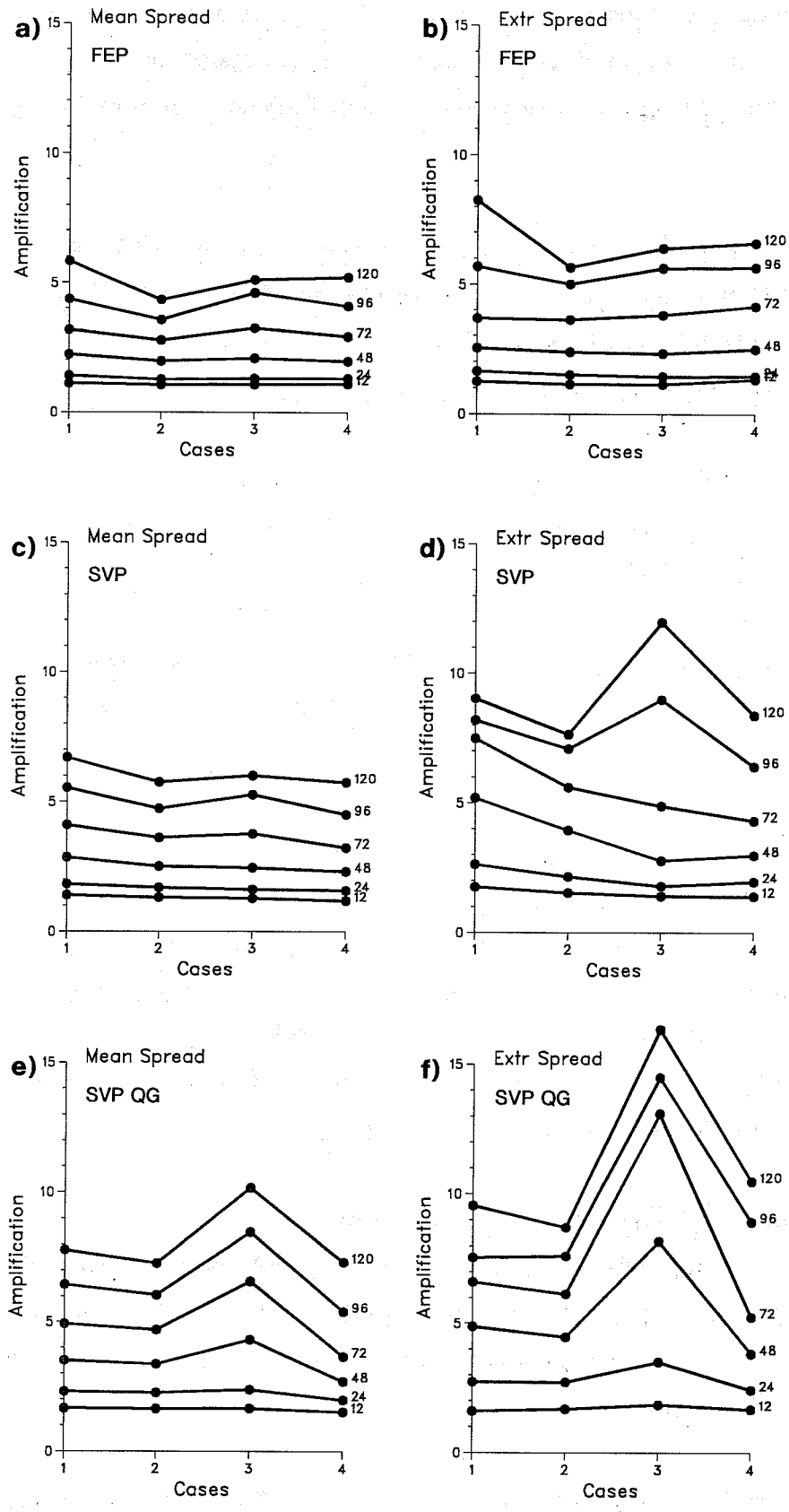

Fig. 10 Spread of 500 hPa RMS height a) FEPs, mean spread b) FEPs, extreme spread, c) SVPs, mean spread d) SVPs, extreme spread. e) SVPs, QG model, mean spread f) SVPs, QG model, extreme spread.

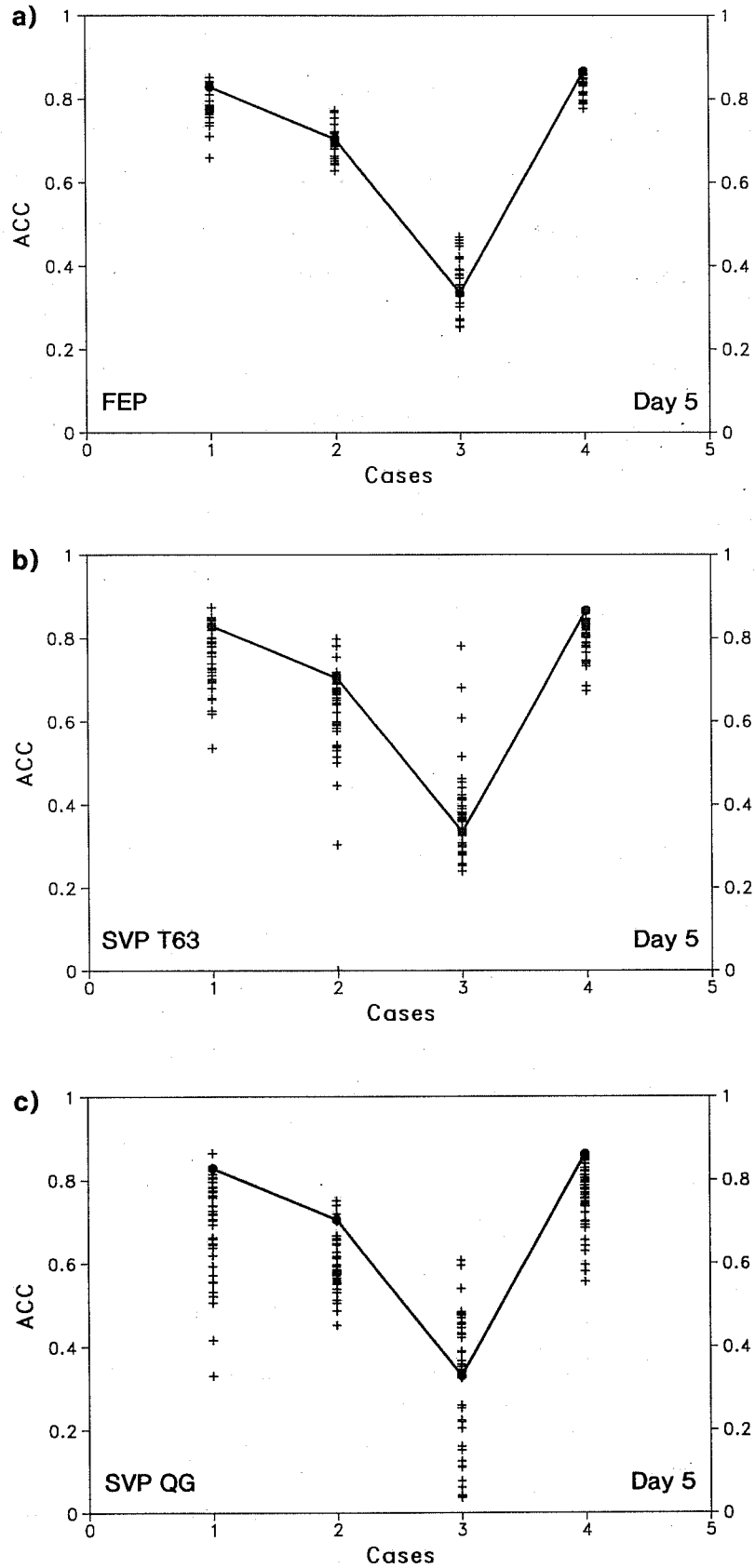

Fig. 11 Day 5 skill (anomaly correlation coefficient of 500 hPa height over Northern Hemisphere) of control forecast and ensemble members for a) FEP ensemble b) SVP ensemble c) QG ensemble.

500 hPa height fields for the verifying analysis, control and forecast using SVP -1, are shown in Fig. 12 for days 5 and 8. The flow is characterized by a transition to a strongly meridional flow over the Eastern Pacific. The verifying analyses show substantial ridging over the north Pacific, beginning on 31 January. The control forecast misses this development and its skill falls below zero within five days. Over the USA a very broad deep trough developed by day 3 covering the entire USA, whereas the control forecast generates 2 separate troughs, one centred over eastern Canada the other over the eastern Pacific. The forecast with SVP -1 is clearly more skilful than the control.

Using the primitive equation model, it is possible to study ensemble forecasts of direct weather-related variables such as precipitation and low-level temperature. This issue will be the focus of future studies. However, as an example, we show in Fig. 13 the low-level (850 hPa) temperature tendency (ie difference from initial conditions) at days 5 and 8, for the verifying analysis, control forecast, and forecast from SVP -1. The cold-air outbreak over western north America can be clearly seen, with a drop in temperature over 5 days of up to 50 K. The control forecast underestimates this by a factor of over 2, whilst in the SVP -1 integration, the maximum drop is in excess of 45 K. The ensemble forecast using the FEPs would not have indicated any probability for a drop in temperature this large.

The high skill of the SVP -1 forecast suggests that the errors in the control forecast over the USA could have been caused by analysis errors over the Eastern Asian-Western Pacific region. However, the implied magnitude of such errors 40-50 m is probably too large (although errors of that magnitude can occur as analysis experiments with the ECMWF model have indicated).

The sensitivity of the skill of the forecast from initial conditions perturbed with SVP -1 but with different amplitudes was tested by rerunning the SVP 1 forecast up to day 10, with initial amplitudes varying from +10 m to -10 m. The resulting anomaly correlation coefficient over the Pacific area is shown in Fig. 14). Reducing the amplitude of the negative perturbation from 10 m to 6 m RMS has little impact on the skill of the forecast. There is a remarkable and highly nonlinear impact on skill when the amplitude is reduced to lower values. The "discontinuity" occurs between day 3 and 4 of the forecast, again indicating that this represents the limit of applicability of linearization. There is little sensitivity of skill to the amplitude of the positive perturbation.

7. CONCLUSIONS

The analysis in this paper has been motivated by the recognition that amplification of uncertainties in the analysed state of the atmosphere by flow-dependent instabilities can lead to considerable variability in the skill of a numerical forecast, limiting the potential usefulness of the numerical weather prediction. As models continue to improve, and data coverage becomes more uniform in time,

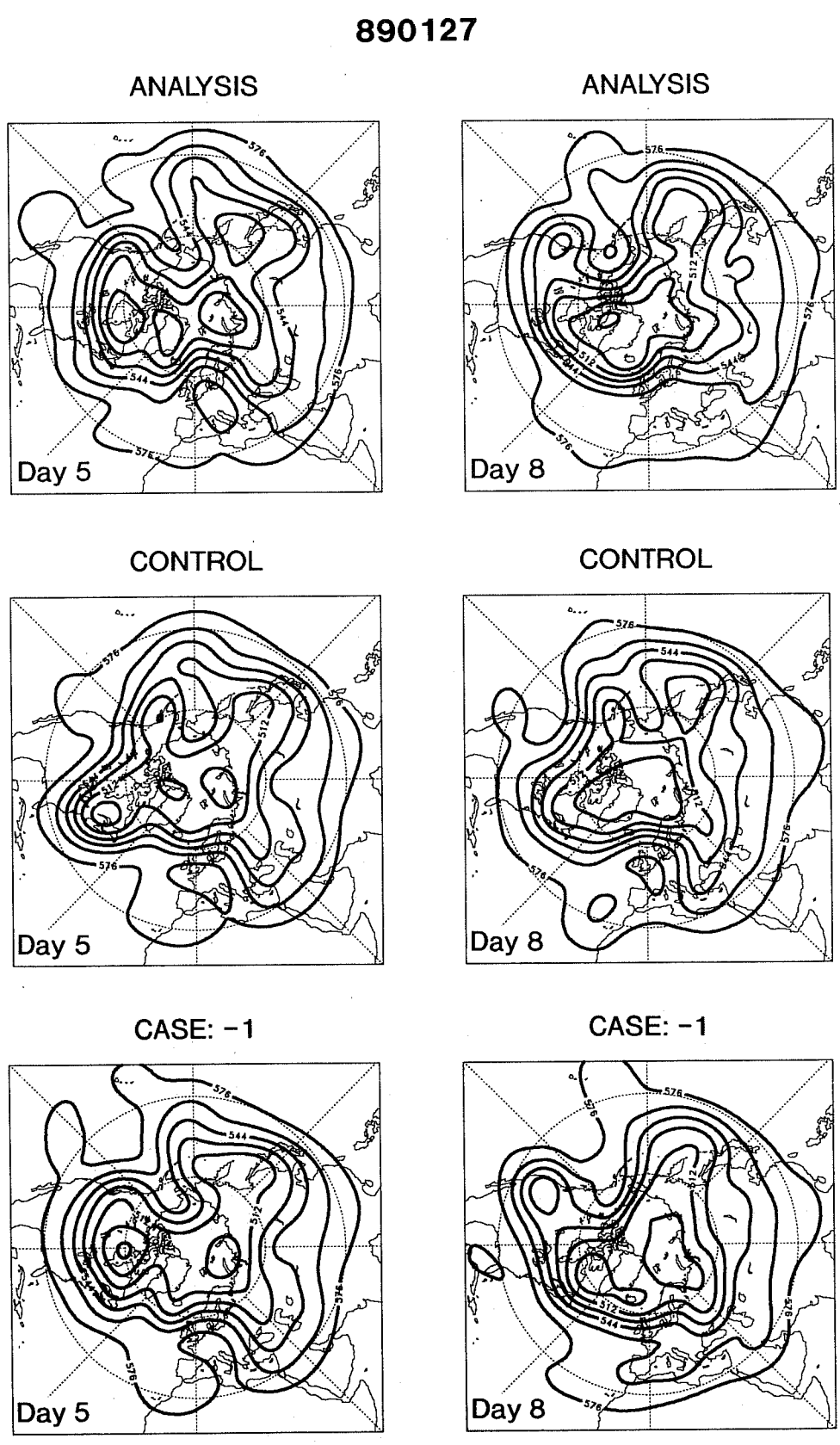

Fig. 12 500 hPa height forecasts (days 5 and 8) from 27 January, 1989 . Top row -verifying analysis. Middle row control forecast. Bottom row - forecast with SVP -1. Contour interval 16dam.

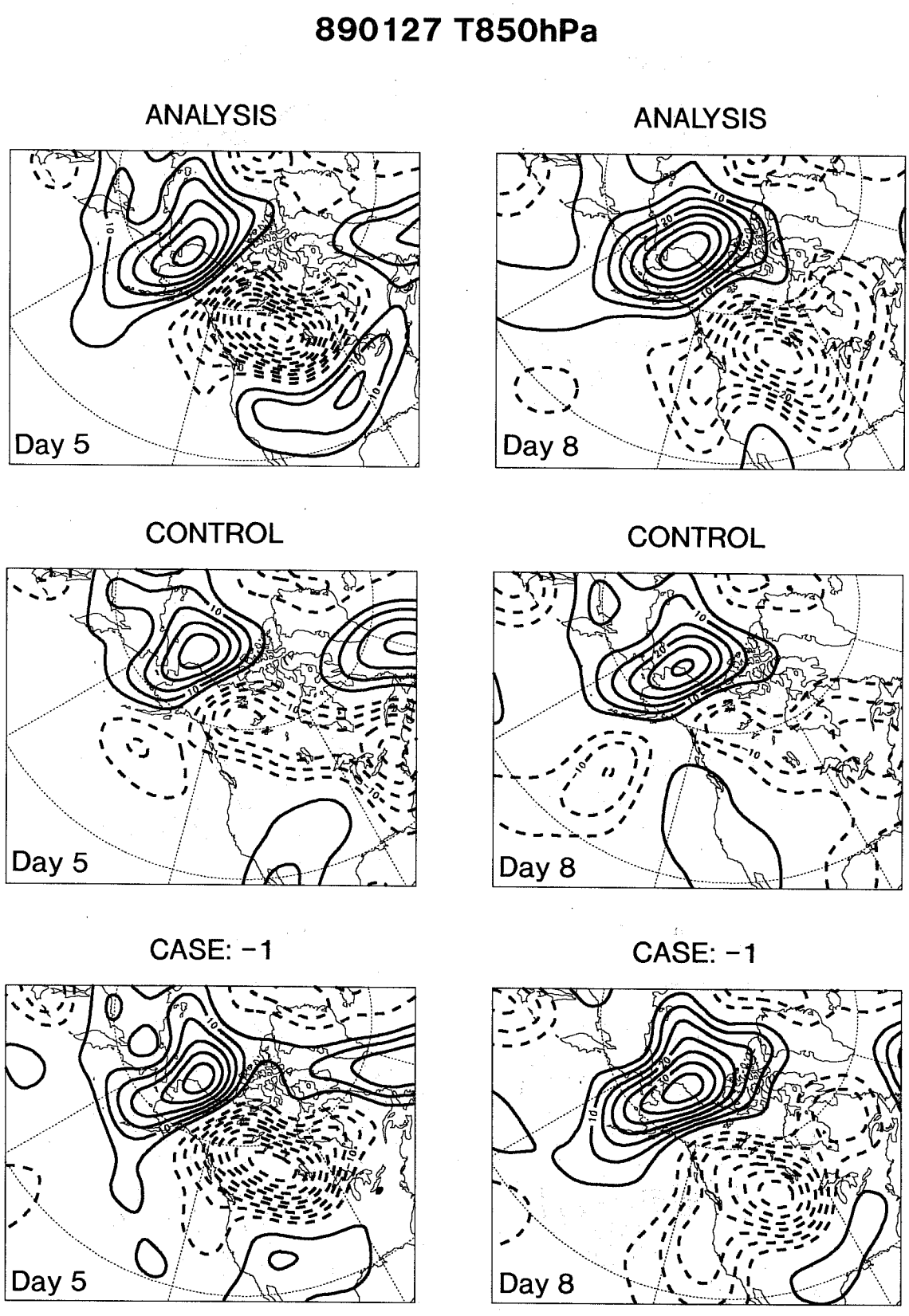

Fig. 13 As Fig. 12 but for 850 hPa temperature tendency (difference from initial value) over northeast Pacific and adjacent landmass. Contour interval 5K.

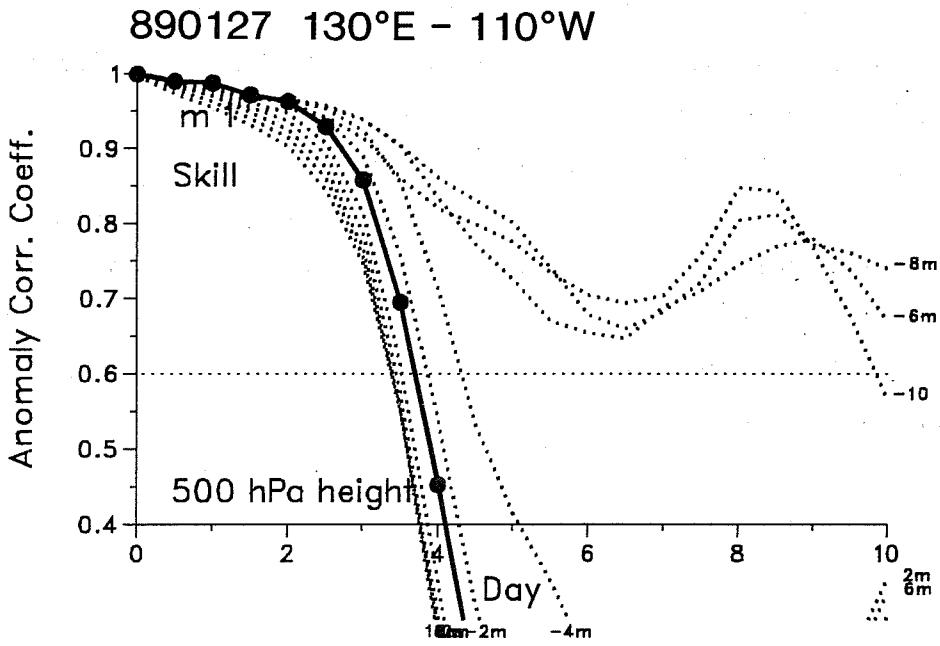

Fig. 14 Skill (anomaly correlation coefficient of 500 hPa height over north Pacific area) for 27 January, 1989 case. Control forecast shown as heavy line. Dotted lines are all for perturbations with SVP 1 but varying amplitude (given by numbers at the end of each skill curve).

variations in the predictability of the flow will become the dominant source of variability in predictive skill.

The use of ensemble forecasts is a practical technique to estimate the predictability of the flow. An advantage over other methods is that the ensembles are made with a numerical model in which nonlinearities due both to advection and unresolved physical processes are represented as accurately as possible. The ability to model the nonlinear development of analysis error becomes important beyond about day 3 of the forecast. However, the principal difficulty in implementing an ensemble forecast technique in practice is associated with sampling; the number of balanced degrees of freedom available to perturb the initial state vastly exceeds the largest practicable size of an ensemble. In practice this means that ensemble dispersion using randomly chosen perturbations are likely on average to underestimate predictability.

In addressing this sampling problem, we have taken the view that the synoptic disturbances whose predictability we are trying to forecast are in some sense associated with instabilities of a larger-scale background flow. Moreover, we have been guided by recent dynamical developments, promoted principally by *Lacarra and Talagrand* (1988) and *Farrell* (1990,1991), that such instabilities are not of the "classical" exponentially-growing normal mode type, rather they arise through energy transfer from the background flow by stochastically-forced non-modal disturbances. On this basis we have chosen our perturbations to the initial state from the fastest growing singular vectors, calculated using a T21 three level quasi-geostrophic model. These perturbations should capture as accurately as possible the dynamically-important uncertainties in the analysis of the synoptic disturbances at the initial time of the forecast. We hope that by exciting these particular disturbances, possible nonlinear regime transitions can be captured by some members of the ensemble. Thus a range of scenarios for the development of the large-scale flow can be forecast for the medium range.

We chose four initial dates for study. The control forecast from one of these dates was exceptionally poor, as the development of a large amplitude ridge over the north east Pacific around day 5 of the forecast was missed. The control forecasts from the other three dates were of typical quality. For each case, three ensembles were performed. The first used earlier operational 6-hour forecast errors, suitably orthonormalised, as initial perturbations. These were integrated using a T63 version of the ECMWF operational model. The second and third ensembles used the fastest-growing quasi-geostrophic singular vectors as initial perturbations, integrated in the T63 and quasi-geostrophic models respectively.

Comparing the two T63 sets of ensemble forecasts, the forecast divergence was much larger with the optimal non-modal disturbances than with the forecast error perturbations. However, neither was as large as in the integrations of the quasi-geostrophic model itself, suggesting that the interpolation procedure for the singular vectors had some deficiencies.

Using the singular vectors, the T63 ensemble was able to flag the exceptionally poor control forecast - indeed several members of the ensemble were able to capture correctly the development of the ridge. By contrast, none of the members of the corresponding ensemble with forecast error perturbations were able to predict the Pacific ridge. The correct development of the ridge depended on the nonlinear growth of the singular vector perturbations of the initial analysis. This result justifies the use of the ensemble technique over other purely linear methods for medium-range predictability research.

Further developments of the technique are required before operational implementation can be envisaged. Most importantly, the singular-vector perturbations must be further constrained by likely observation error structure. One must ensure that the RMS amplitude of the perturbations is consistent with the RMS amplitude of analysis error in any one geographical location. More generally, weighting with a full estimate of analysis error covariance would build into the optimal instabilities constraints due to analysis error correlations. Finally, since many of the individual singular vectors are spatially isolated, linear combinations of individual eigenvectors are required to give representative hemispheric-scale analysis error. Experimentation is currently under way with hemispheric-scale singular-vector perturbations constrained by analysis error.

A second potential problem needs to be addressed before the technique can be considered viable for operational forecasting. As noted above, the potential of the QG non-modal perturbations is not fully realised in the primitive equation model because of relative inconsistencies in vertical resolution and orography. Research is under way to calculate the singular vectors directly from the linearised version of the primitive equation model used for the nonlinear ensemble integrations.

Results from these current studies will be reported in due course.

9. REFERENCES

Brankovic, C., T.N. Palmer, F. Molteni, S. Tibaldi and U. Cubasch, 1990: Extended-range predictions with ECMWF models: Time-lagged ensemble forecasting. Q.J.R.Meteorol.Soc., 116, 867-912.

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

Farrell, B.F., 1990: Small error dynamics and the predictability of atmospheric flows. J.Atmos.Sci., 47, 2409-2416.

Farrell, B.F., 1991: Stochastic perturbations in shear flow. Extended abstract. Proceedings of eighth conference of American Meteorological Society on atmospheric and oceanic waves and stability, pages 64-66. American Meteorological Society. Massachusetts pp 417.

Hollingsworth, A., 1980: An experiment in Monte Carlo forecasting procedure. ECMWF workshop on stochastic dynamic forecasting. ECMWF, 1980, 99pp.

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

Lacarra, J.-F. and O.Talagrand, 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.

Lorenz, E.N., 1965: A study of the predictability of a 28-variable atmospheric model. Tellus, 17, 321-333.

Molteni, F. and T.N. Palmer, 1991: A real-time scheme for the prediction of forecast skill. Mon. Wea.Rev., 119, 1088-1097.

Molteni, F. and T.N. Palmer, 1992: Predictability and non-modal finite-time instability of the northern winter circulation. (MP). ECMWF Workshop on Predictability.

Palmer, T.N. and S. Tibaldi, 1988: On the prediction of forecast skill. Mon.Wea.Rev., 116, 2453-2480.

Veyre, P., 1992: Direct prediction of error variances by the tangent linear model: a way to forecast uncertainty in the short range? ECMWF Workshop on Predictability.