Tellus (2004), 56A, 485–500
Printed in UK. All rights reserved

TELLUS

Ensemble prediction of blocking regime transitions

By JORGEN S. FREDERIKSEN[1]*, MARK A. COLLIER[1] and ANDREW B. WATKINS[2],
[1]*CSIRO Atmospheric Research, PMB 1, Aspendale, Victoria 3195, Australia;* [2]*National Climate Centre, Australian Bureau of Meteorology*

(Manuscript received 22 October 2003; in final form 7 April 2004)

ABSTRACT

The skill of ensemble prediction during blocking regime transitions is examined for Northern Hemisphere flows within two atmospheric models. An ensemble prediction scheme based on fast growing perturbations has been implemented for conformal-cubic and spectral general circulation models. The methodology uses a breeding method, based on an implicit linearization of the models, in which perturbations likened to leading Lyapunov vectors are obtained and used to perturb the initial conditions. Detailed comparisons of the skill of ensemble mean forecasts with control forecasts have been carried out for Northern Hemisphere initial conditions in October and November 1979. A particular focus has been the variability of forecast skill during the development, maturation and decay of the large-scale blocking dipoles that occurred in the major blocking regions over Europe, over the Gulf of Alaska, over the North Atlantic and as well over North America. On average, the ensemble mean forecast performs better than the control forecast for forecast times longer than three or four days. Average error growth curves in the two models are quite similar with the conformal-cubic model ensemble generally performing slightly better than for the lower-resolution spectral model. Both ensemble and control forecasts initiated twice daily exhibit considerable variability in forecast skill that is shown to be related to instability regimes of particular synoptic events. At a given forecast lead time, errors tend to be larger for forecasts validating when blocks are developing or decaying and smaller for mature blocks. The spread of ensemble member forecasts has been studied and related to likely forecast skill. Comparison of results from the conformal-cubic and spectral models initialized with two different analysis data sets has allowed the determination of the robustness of our findings relating error growth to instability regimes.

1. Introduction

The inescapable uncertainty in specifying atmospheric initial states, due to chaos, means that weather prediction should be regarded as the problem of forecasting the probability density function of states in phase space or equivalently of determining the hierarchy of moments or cumulants of meteorological variables. Leith (1974) employed a second-order closure model in the context of two-dimensional turbulence to examine the growth of error spectra and their convergence towards the climatological variance associated with a typical atmospheric spectrum. His study was for the idealized problem of isotropic turbulence with zero mean fields. More recently, computationally tractable closures have been derived for general inhomogeneous barotropic flows over topography; these closures describe the time evolution of the mean fields and covariance and response function matrices (Frederiksen, 1999; O'Kane and Frederiksen, 2004). Currently, however, the only feasible way of predicting the statistics of future atmospheric states using multilevel primitive equation general circulation models is by running an ensemble of forecasts from perturbed initial conditions.

It is generally agreed that the perturbed analyses should represent the fast growing errors that are present in initial conditions and that may be selectively bred by the data assimilation scheme (Toth and Kalnay 1993). There has been less agreement on how these perturbations should be characterized. At the European Centre for Medium-Range Weather Forecasts (ECMWF), rotated singular vectors (SVs) in the total energy norm are used as the initial ensemble perturbations (Molteni et al., 1996; Palmer et al., 1998). At the National Center for Environmental Prediction (NCEP), an iterative breeding method is used to generate ensemble perturbations that have been linked to leading Lyapunov vectors (LVs; Toth and Kalnay, 1993, 1997; Noone and Simmonds, 1998; Kalnay, 2003). Houtekamer et al. (1996) take the view that because the data assimilation scheme selectively breeds fast growing errors, one can use multiple analysis cycles to form ensemble perturbations.

Frederiksen and Bell (1990) argued that superpositions of fast growing normal modes would approximately characterize the instability and error growth in atmospheric flows over several days. A successful ensemble prediction scheme based on

*Corresponding author.
e-mail: jorgen.frederiksen@csiro.au

normal modes was implemented by Anderson (1996) for a low-order dynamical system. Frederiksen (2000) and Wei and Frederiksen (2004) showed that initial errors converge over time towards structures that can be represented by low-order expansions in terms of the leading finite-time normal modes (FTNMs), which are the norm-independent eigenvectors of the tangent linear propagator. This suggested that it would therefore seem natural to be able to efficiently represent analysis error fields by a superposition of the leading FTNMs. Szunyogh et al. (1997) and Palmer et al. (1998) have also discussed the likely roles of different dynamical vectors in representing analysis errors. Comparative studies of the performance of ensemble prediction schemes based on SVs and bred perturbations have been carried out by Houtekamer and Derome (1995), Bontemps and Källén (1999), Hamil et al. (2000), Wei and Toth (2003) and Wang and Bishop (2003).

In this paper we examine ensemble prediction during blocking regime transitions using an iterative breeding method for ensemble prediction within two numerical weather prediction models: the Commonwealth Scientific and Industrial Research Organization (CSIRO) conformal-cubic model (McGregor and Dix, 2001) and the Bureau of Meteorology Research Centre (BMRC) spectral model (Hart et al., 1990; Frederiksen et al., 1996). Our primary aim is to examine the variability of forecast skill in different synoptic situations, focusing on the differences in skill between strong zonal flow situations and when mature blocks are present and as well on variability associated with the growth and decay phases of blocking.

In both hemispheres, numerical weather forecasts during the onset and decay of blocking frequently suffer from rapid loss of predictability (Bengtsson, 1981; Noar, 1983; Kimoto et al., 1992; Anderson, 1993; Tibaldi et al., 1995; Molteni et al., 1996; Watson and Colucci, 2002). Forecast errors during blocking transitions are also limiting factors in producing successful medium and extended range forecasts (Colucci and Baumhefner, 1992). Thus, examining the ensemble predictability during blocking regime transitions is of considerable practical importance for both short-term and extended range forecasts.

It has been shown in a number of theoretical and numerical studies that instability processes of the large-scale flow play major roles in the development of blocking anomalies and in the growth of errors during blocking transitions (Frederiksen, 1978, 1983, 2000; Simmons et al., 1983; Borges and Hartmann, 1992; Anderson, 1996; de Pondeca et al., 1998a,b; Li et al., 1999). Forecast errors are determined both by rapid dynamical growth of analysis errors and by model deficiencies (Reynolds et al., 1994; Frederiksen et al., 1996; Toth and Kalnay, 1997). In order to establish that the findings of our study are robust and are a reflection of the different instability regimes, depending on the synoptic situations, we compare results using two differently formulated general circulation models (GCMs), the CSIRO and BMRC GCMs, and with two different analysis data sets for the same times. The CSIRO model is initialized with the NCEP reanalyses and the BMRC model is initialized with the ECMWF analyses. Our aim is to show that, despite using quite different models verified against different analysis, the forecasts have similar variability in forecast skill which is determined by the instability regimes pertinent to the different synoptic situations. We focus on the sequence of dramatic blocking events that occurred over Scandinavia, over the Gulf of Alaska, over the North Atlantic and over North America during October and November 1979.

The plan of this paper is as follows. In Section 2, we present brief descriptions of the CSIRO conformal-cubic and BMRC spectral models which are used in the control and ensemble predictions. In Section 3, we present a summary of the ensemble prediction methodology used and the relation between the breeding method and four-dimensional data assimilation schemes. In Section 4, we describe the synoptic situations associated with the major blocking events over Scandinavia, over the Gulf of Alaska, over the North Atlantic and over North America that occurred during October and November 1979. The vertical and horizontal structures of the bred ensemble perturbations in the CSIRO and BMRC numerical weather prediction models are examined in Section 5.

In Section 6, we examine the growth of ensemble and control 500-hPa Northern Hemisphere zonal velocity errors (averaged between 20° and 90°N). We study these errors for forecasts in the two models started each 12 h between 0000 UTC on 11 October and 1200 UTC on 21 November 1979. We also calculate averages over this time interval and present similar diagnostics for 500-hPa geopotential heights. In Section 7 we focus on error growth of ensemble and control forecasts in the Atlantic sector between 0° and 60°W and relate the root-mean-square (rms) errors to the synoptic situations in this region and in the regions upstream and downstream. In Section 8, we examine in more detail ensemble mean forecasts of 500-hPa geopotential heights, spread of the ensemble members and error growth during the development of the four major blocking events. Our conclusions are detailed in Section 9.

2. GCM details

The forecast experiments performed in this paper have been carried out with two numerical weather forecast models developed at CSIRO and BMRC. The CSIRO GCM is a semi-Lagrangian model which uses a novel conformal-cubic grid; in all our simulations we use a *C*48 conformal-cubic grid, as in fig. 2 of McGregor and Dix (2001), with an average resolution of about 200 km, and we employ 18 vertical sigma levels. The model incorporates a comprehensive list of physical parametrizations as described by McGregor and Dix (2001) including representations for long-wave and short-wave radiation, interactive cloud distributions, cumulus and shallow convection schemes and evaporation of rainfall. It includes a stability-dependent boundary layer and vertical diffusion, Smagorinsky deformation-based horizontal diffusion (Smagorinksy et al., 1965) and a vegetation-canopy scheme with prognostics for soil temperatures and soil moisture.

The BMRC GCM is a spectral model employing spherical harmonic basis functions. In all simulations we use an $R31$ resolution, in which the rhomboidal truncation wavenumber is 31, and we use nine vertical sigma levels. The $R31$ resolution corresponds approximately to a grid spacing of 330 km at mid-latitudes. The BMRC model also includes a comprehensive list of physical parametrizations as described by Hart et al. (1990) and Frederiksen et al. (1996). These cover most of the processes described for the CSIRO GCM but are, in general, formulated somewhat differently. In particular, the BMRC model employs a Laplacian horizontal diffusion with a diffusion coefficient of 2.5×10^5 m^2 s^{-1} and is applied only to waves with a total wavenumber $n \geq 31$. The current version of the BMRC model has fixed zonally averaged cloud amounts.

3. Ensemble prediction methodology

As discussed in the introduction, a number of successful methods of generating perturbations for ensemble predictions have been proposed and applied both in research and operational modes. There appears to be consensus that, in order to obtain improved forecasts with ensemble prediction methods, it is necessary to use perturbations which have similar structures to the fast growing errors present in initial analyses. Toth and Kalnay (1993) argue that current four-dimensional (space–time) data assimilation schemes, although far superior to using persistence or climatology to fill data-poor regions, cause the breeding of fast growing perturbations. Typically, in an operational data assimilation cycle, a 6-h forecast is made from initial conditions determined by a previous analysis. The next analysis is then calculated as a weighted combination of the 6-h forecast and observations collected during a 6-h window centred on the analysis time, with the weighting determined geographically by the statistical reliability of the observations. Repetition of the process for subsequent analyses, four times every day, decreases the total error in the analysis but increases the ratio of the fast growing errors to the total error.

There is less agreement about what types of perturbations are most likely to efficiently represent analysis error (Szunyogh et al., 1997; Palmer et al., 1998; Frederiksen, 1997, 2000; Wei and Frederiksen, 2004). Superpositions of rotated SVs in the total energy norm are used to perturb the control analyses and form the initial ensembles at the ECMWF (Molteni et al., 1996; Palmer et al., 1998). At NCEP a ‘breeding’ method is used to generate the initial ensembles, and these bred perturbations have been linked to the leading LVs (Toth and Kalnay, 1993, 1997). Houtekamer et al. (1996) use multiple analysis cycles to generate their initial ensembles. Frederiksen (2000) has shown that initial errors contract on to a low-dimensional phase space spanned by the leading FTNMs, which are the norm-independent eigenvectors of the tangent linear propagator.

In this study we employ an iterative breeding method (Toth and Kalnay, 1993, 1997) primarily because it is a successful method which is relatively easy to implement and computationally relatively cheap. The version of the method that we employ is summarized in the schematic diagram in Fig. 1. It consists of a period of 10 d of breeding of perturbations starting at the beginning of October 1979 followed by a period of self-breeding and 10-d ensemble forecasts starting at 0000 UTC on 11 October and finishing at 1200 UTC on 21 November 1979.

During the breeding period, a specified small perturbation is added to the analysis and 12-h forecasts are performed from both the perturbed and unperturbed (control) initial conditions. The difference between the 12-h perturbed and control forecasts is then scaled to have the same root-mean-square (rms) magnitude as the initial perturbation and is added to the subsequent 12-h analysis. Again, 12-h forecasts are performed from both the new perturbed and unperturbed analysis and the process is repeated until the end of the 10-d breeding period. The breeding methodology corresponds essentially to an implicit linearization of the non-linear dynamics about the time-dependent analyses and would, with time, result in the bred perturbation converging to the leading LV, were it not for stochastic effects associated with the convective parametrizations in the numerical weather prediction models.

After the 10-d breeding period, 10-d forecasts are performed from both the perturbed and control analyses, as indicated in the schematic diagram. As well, 10-d forecasts are performed from the control analyses perturbed by bred perturbations with opposite signs. Twelve hours into the 10-d forecasts, the difference between the 12-h perturbed and control forecast is scaled to the standard rms magnitude and added to the subsequent 12-h analysis. Again, 10-d forecasts are performed from both the new perturbed and unperturbed analyses (and from the analysis perturbed by the bred perturbation with opposite sign) and the process is repeated until 1200 UTC on 21 November.

In practice, we perform eight separate breeding cycles starting from eight different perturbations over the first 10 d. The self-breeding is also performed for eight different perturbations with the ensemble forecast consisting of 16 forecasts employing the eight initial bred perturbations and eight initial identical perturbations with opposite signs. This ensures that the ensemble mean forecast starts from the same initial condition as the control. This method of using $2n$ paired perturbations of opposite signs from n independent breeding cycles appears to be superior to using $2n$ independent bred perturbations (with their mean subtracted from each perturbation) (Toth and Kalnay, 1997). As in the pioneering work of Leith (1974), our ensemble does not include the control. If the control were included in the ensemble, as in some operational forecasts, then mean square differences between the ensemble average and control errors would change at most by less than 6% and all of our conclusions would remain valid.

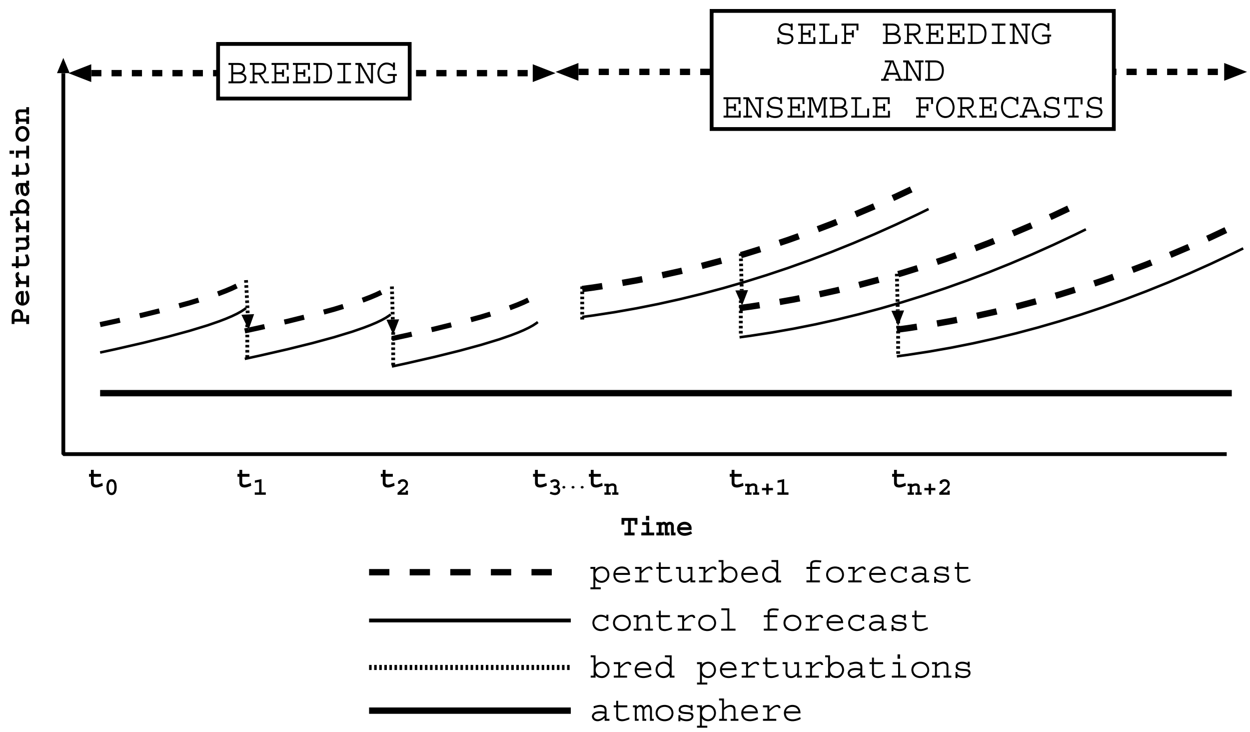

Fig 1. Schematic diagram showing the breeding cycle for generating perturbations followed by ensemble forecasts which maintain their own breeding cycle, known as self-breeding. See text for a detailed explanation.

4. Synoptic situations

We examine the performance of the ensemble prediction method described in Section 3 for initial conditions taken from the period between 0000 UTC on 11 October 1979 and 1200 UTC on 21 November 1979. We have focused on this interval because it was a time of large-scale Northern Hemisphere blocking in the major blocking regions over the North Atlantic Ocean and over Europe and over the Gulf of Alaska (Dole and Gordon, 1983; Lejenas and Okland, 1983; Lupo and Smith, 1995a,b; Wiedermann et al., 2002) and as well over North America. As this time interval coincides with the First Global Atmospheric Research Programme Global Experiment (FGGE), observations are of the highest quality. Ensemble and control forecasts have been commenced every 12 h during this time interval. This sequence of developing, mature and decaying blocks alternating with times of strong zonal flows seems an interesting period for studying the variability of predictability and for examining the differences between ensemble and control forecasts.

Some of the major blocks that occurred in October and November 1979 are shown in Fig. 2. The 500-hPa heights taken from NCEP analyses at 0000 UTC depict the mature phase of the large-scale Scandinavian block on 27 October (Fig. 2a), the Gulf of Alaska block on 8 November (Fig. 2b), the North Atlantic block on 16 November (Fig. 2c) and the North American block on 23 November (Fig. 2d). The European block started developing around 21 and 22 October over the UK and intensified to form the large-scale high–low dipole block on 27 October with the high centred over Scandinavia and the low over Spain, as depicted in Fig. 2a. These centres subsequently decayed and moved downstream in late October. In early November, the flow over the Gulf of Alaska became more diffluent with the blocking ridge developing rapidly around 5 November and forming the cut-off high–low dipole pair on 8 November, as shown in Fig. 2b. The Gulf of Alaska block weakened between 10 and 12 November and was followed by downstream development of a blocking high over the Atlantic Ocean around 13 and 14 November. The high–low dipole shown in Fig. 2c on 16 November subsequently decayed around 17 and 18 November. Around 20 November, ridging again occurred over the Gulf of Alaska leading to diffluent flow around a large-scale high–low dipole block over North America on 23 November (Fig. 2d). This block persisted until 26 November and then decayed.

5. Bred perturbations

The ensemble prediction methodology described in Section 3 results in the convergence of fairly arbitrary initial perturbations towards bred vectors, which have fairly similar large-scale features after about 10 d. This has been tested by computing pattern correlations between the different bred vectors, for both the CSIRO and BMRC models, all finishing on a given day. Bred vectors generated within different models, however, do differ

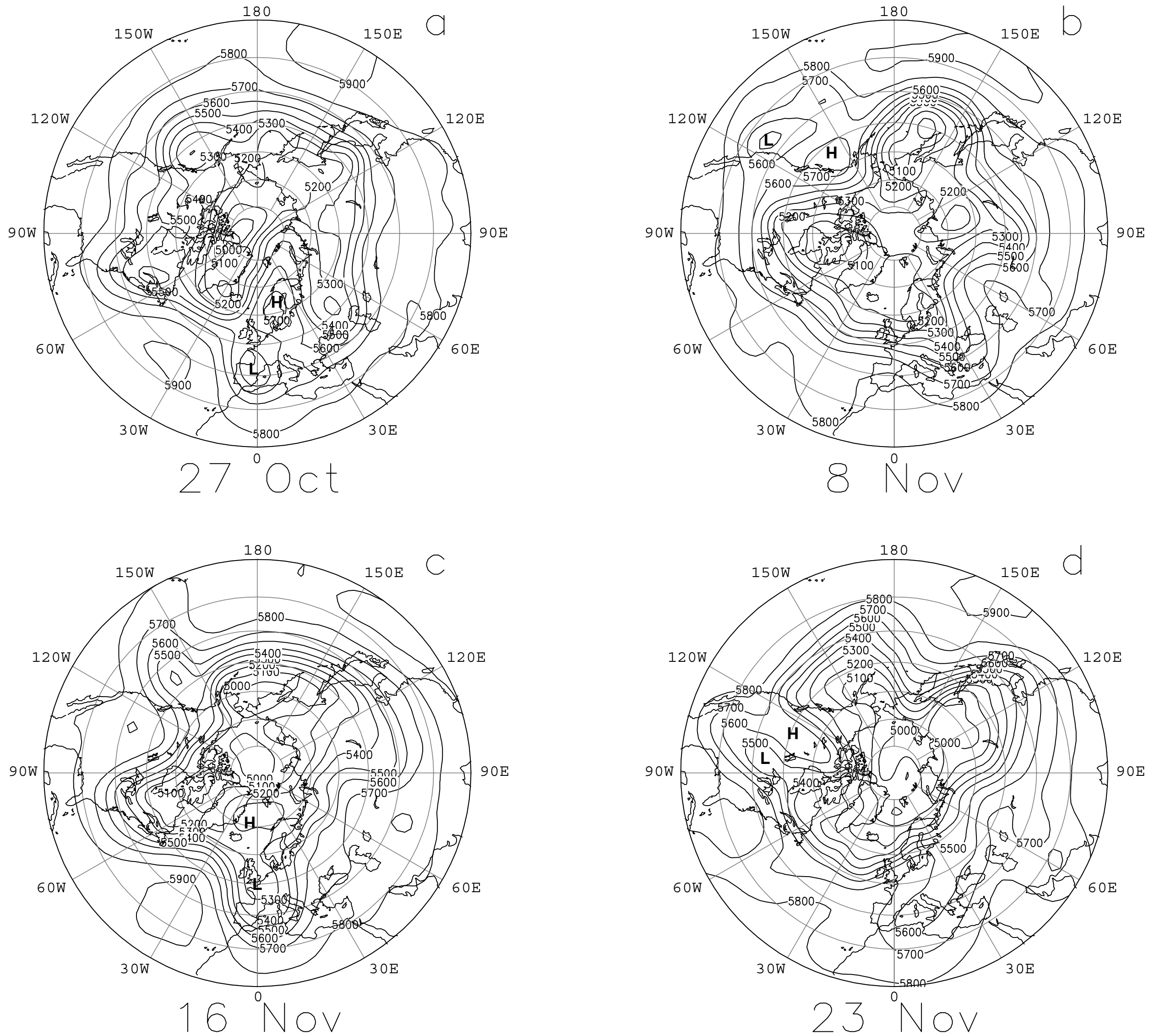

Fig 2. Northern Hemisphere 500-hPa geopotential heights (m) based on NCEP analyses for 0000 UTC on (a) 27 October, (b) 8 November, (c) 16 November and (d) 23 November 1979.

somewhat as discussed below. Within a given model, the main difference between bred vectors for breeding times longer than about 10 d is due to the stochastic element introduced by the convection schemes in the models as discussed by Toth and Kalnay (1993, 1997).

The vertical structures of geopotential heights for typical bred perturbations on days of particular interest in this study (27 October, 8 November and 16 November) are summarized in Fig. 3. It shows that for both the CSIRO and BMRC models the bred perturbations have fairly constant magnitudes throughout the troposphere with, in some cases, larger amplitudes at the surface. The structures of the perturbations are, in fact, largely equivalent barotropic throughout the upper troposphere and more baroclinic near the surface. Throughout this study, the Northern Hemisphere geopotential height bred perturbations are normalized to 15 m at 500 hPa. This has been found empirically to be nearly optimal for increasing the skill with ensemble forecasts (see also Toth and Kalnay, 1997).

As examples of the horizontal structure of 500-hPa geopotential heights for typical bred perturbations, we show in Figs. 4a and b the CSIRO and BMRC model bred perturbations on 16 November. Both perturbations, but in particular the CSIRO model perturbations, have high–low dipoles near the Greenwich meridian similar to the leading evolved SV and leading FTNM shown in figs. 8a and b of Frederiksen (2000). As discussed there, these structures in turn have essentially converged to the leading LV for the period 16 October to 16 November. Outside the key region of the North Atlantic where the large-scale blocking

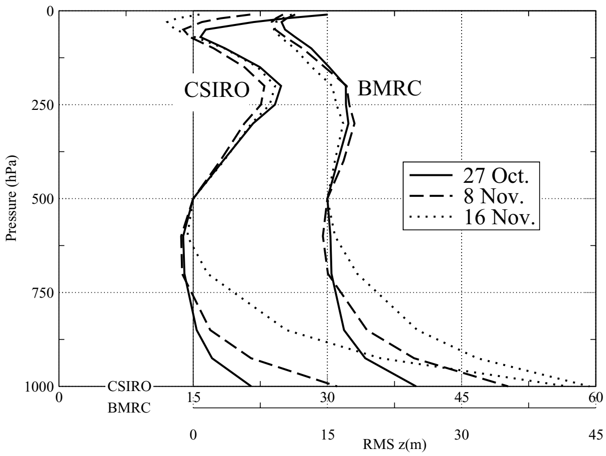

Fig 3. Vertical structures of Northern Hemisphere rms geopotential height (m) between 20° and 90°N for bred perturbations at 0000 UTC on 27 October, 8 November and 16 November 1979 for CSIRO and BMRC models.

occurs on 16 November (Fig. 2c), there are, however, differences in the structures of the bred perturbations in the two GCMs and in comparison with the dynamical vectors shown in figs. 8a and b of Frederiksen (2000).

It is clear from the results on 16 November, and from similar results on, for example, 27 October and 8 November (not shown), that the bred perturbations in both GCMs have major centres in the regions of dominant instability (as determined in the studies of Frederiksen, 1997, 1998, 2000). As well, they appear to mix in some of the non-leading modal structures, presumably due to non-linearity and the presence of stochastic effects due to convection.

6. Hemispheric ensemble mean forecast results

As discussed in Section 3, in all our ensemble forecasts we use 16 paired perturbations obtained by the breeding method with the two perturbations in each of the eight pairs having opposite signs. In our experiments, we have examined error growth in terms of 500-hPa geopotential height, zonal velocity, meridional velocity and temperature. In order to minimize differences in model forecast results and analyses, which have been derived based on other data assimilation systems and models, namely NCEP or ECMWF, we find it most useful to carry out some of our comparisons using fields such as the 500-hPa zonal wind. This field is directly carried in the prognostic equations of both the CSIRO and BMRC models. We also use the geopotential height but note that, because it is a derived field that involves an integration from the surface, it may be more influenced by differences in topography and surface processes between the models used here and the NCEP or ECMWF models.

In Fig. 5a we show for the CSIRO and BMRC models the 500-hPa Northern Hemisphere rms zonal wind errors (between

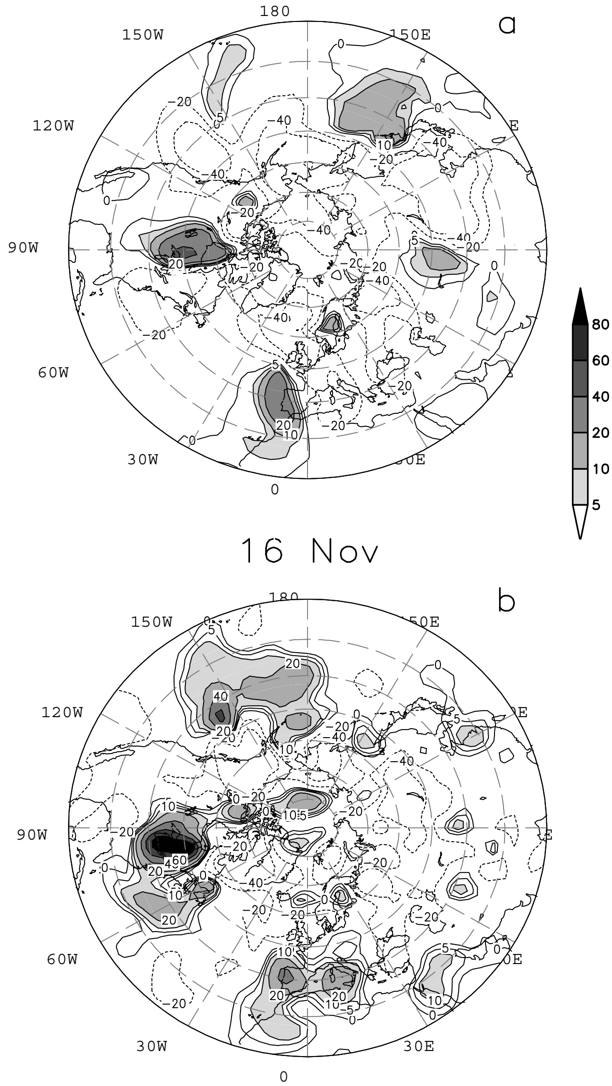

Fig 4. Horizontal structures of Northern Hemisphere 500-hPa geopotential height (m) for bred perturbations at 0000 UTC on 16 November 1979 for CSIRO (top) and BMRC (bottom) models.

20° and 90°N) for the ensemble mean and control forecasts averaged between 0000 UTC on 11 October and 1200 UTC on 21 November 1979. We note that for forecasts longer than about 3 d the average errors of the ensemble mean forecasts are lower than for the control forecasts. Quite similar results are obtained in terms of the 500-hPa geopotential height, as shown for the two models in Fig. 5b. For both BMRC and CSIRO models, 500-hPa geopotential height errors in m are typically 10 times 500-hPa zonal velocity errors in m s^{-1} for forecast times longer than 4 d. This relationship, which is evident from Figs. 5a and b for average forecast errors, also applies in good measure for errors on a particular day. Our results for average forecast errors are broadly comparable with the results of Molteni et al. (1996) for winter

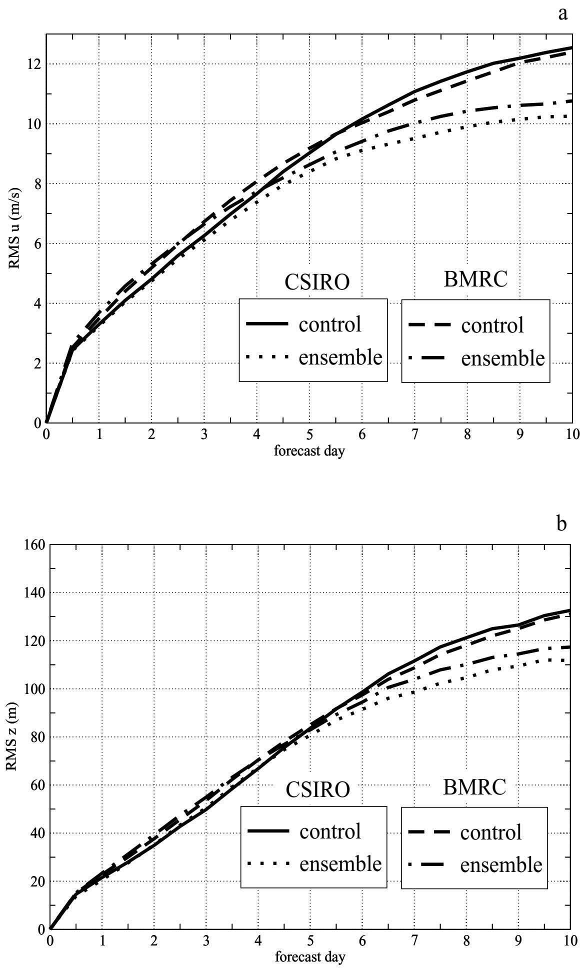

Fig 5. Northern Hemisphere rms errors of (a) 500-hPa zonal wind (m s^{-1}) and (b) 500-hPa geopotential height (m), averaged between 0000 UTC on 11 October and 1200 UTC on 21 November 1979, for ensemble mean and control forecasts with the CSIRO and BMRC models.

(their fig. 13a) in which they used ensemble perturbations based on rotated SVs in the total energy norm.

The average error growth results shown in Fig. 5a for the CSIRO and BMRC models have been determined from the twice daily 10-d forecast error results shown in Figs. 6 and 7 respectively. From Figs. 6 and 7 for the CSIRO and BMRC GCMs we see that, although the average of the ensemble mean forecasts is superior to the average of the control forecasts after a few days, there is considerable variability in the skill with the control forecast occasionally beating the ensemble mean as expected. We also note from Figs. 6a, 6b, 7a and 7b that there is also considerable variability in the skill of both the ensemble mean and control forecasts. We shall find that this variability can be related to particular synoptic events, particularly the development, maturation and decay of blocks, and that it is more distinct when the rms errors are calculated over smaller 60° longitude sectors (as in Section 6) that are particularly affected by the synoptic events. Even in the hemispherically (20°–90°N) averaged rms results shown in Figs. 6a, 6b, 7a and 7b we see some evidence of higher control and ensemble errors for forecasts ending around 25 October when the European block is forming and lower errors for forecast ending around 30 October, which covers the mature period of the European block. Similarly errors tend to be larger when the Gulf of Alaska block is developing around 5 November and smaller when it is in the mature phase around 8 November. Again, errors tend to be smaller when the North Atlantic block is mature, around 16 November, and when the North American block is mature, around 23 November, and are larger a few days earlier when these blocks are developing.

Despite the fact that the results in Figs. 6 and 7 have been obtained with different models (CSIRO and BMRC) initialized and validated against different data sets (NCEP and ECMWF), the Hovmoeller diagrams of 10-d forecast errors against analysis time show many similarities for the respective ensemble, control and control-ensemble. This similarity may be quantified by pattern correlations of the CSIRO and BMRC model Hovmoeller diagrams. These pattern correlations have the values of 0.996 for the ensemble errors, 0.994 for the control errors and 0.650 for the control-ensemble. If one subtracts the time mean errors (between 0000 UTC on 11 October and 1200 UTC on 21 November) from Figs. 6 and 7, then pattern correlations between the time mean anomalies are 0.628 for ensemble errors and 0.455 for control errors. These results indicate that not only is the average growth of errors similar in the CSIRO and BMRC models but so is the variability of error growth depending on the synoptic situation particularly for ensemble forecasts.

Next, we examine in more detail the skill and spread of ensemble forecasts focusing on periods relevant to the European and Gulf of Alaska blocks. The three panels in Fig. 8a show the hemisphere average error growth (north of 20°N) of the control and ensemble simulations for the 10-d period starting on 17 October for the CSIRO model (left panel), the corresponding error growth of the individual ensemble members, which provides a measure of the spread of the forecasts (central panel), and the error growth of the control and ensemble simulations for the BMRC model (right panel). The BMRC model (not shown) tends to have slightly less spread than the CSIRO model. For this particular case, there is good agreement between the CSIRO and BMRC errors for both the controls and ensembles. It tends to be the case that there is a general similarity in the improvement in forecasts with the ensemble methodology for both models although, as seen from Figs. 6 and 7, on occasions the two models produce qualitatively different results. One particular example is for the 10-d period starting on 22 October for which the ensemble forecast in the CSIRO model is significantly better than the control after 25 October, while in the BMRC model there is little difference between the two, with the control actually being slightly better than the ensemble in the last few days of the forecast (not shown).

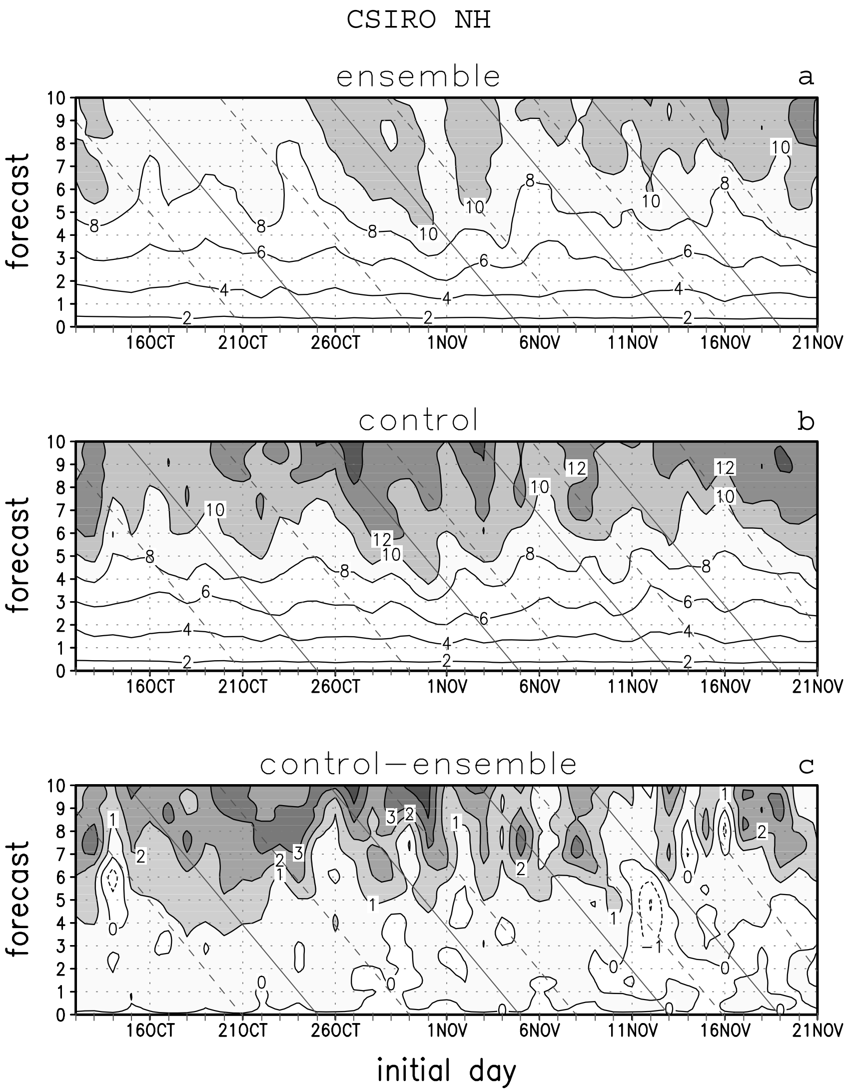

Fig 6. Northern Hemisphere rms errors of 500-hPa zonal wind (m s^{-1}) in 10-d forecasts initiated every 12 h between 0000 UTC on 11 November and 1200 UTC on 21 November 1979 and for ensemble, control and control-ensemble in the CSIRO model. The tick marks on both axes denote a day.

Figure 8b shows the same panels as in Fig. 8a but for the 10-d period starting on 31 October and covering the development and maturation of the Gulf of Alaska block and with the rms error calculations restricted to the sector 20° –90°N and 0° –60°W covering the North Atlantic region where the major storm track occurs. In Fig. 8b (left and right panels), we see relatively rapid error growth as the Gulf of Alaska block starts to form and this occurs in both the control and ensemble results, with the ensembles superior to the controls after 4 November. Particularly evident in these results for the North Atlantic sector is the recovery of the forecast accuracy once the block has formed and is maturing.

7. Forecast results for the Atlantic sector

As noted in Fig. 8b, examining the error growth of ensemble and control forecasts in particular sectors can highlight aspects of the relationship between developing and mature blocks and error growth. We have therefore carried out studies like those described in Section 5 but focusing on 60° longitude sectors of the Northern Hemisphere extratropics (20° –90°N) where the different blocks form. Although blocking appears to be primarily a local phenomenon (Lejenas and Okland, 1983; Lupo, 1997), the development of a block or the presence of a mature block in a given sector affects not only the error growth in the sector but tends to have a similar effect on sectors both upstream and downstream. For this reason, and because the Atlantic sector is a region for a major Northern Hemisphere storm track, we focus here on the sector 20°–90°N, 0°–60°W.

Figures 9 and 10 show the rms zonal wind errors of CSIRO and BMRC ensemble, control and control-ensemble 10-d forecasts started every 12 h for the Atlantic sector. These results should be compared with the corresponding hemispherically averaged results shown in Figs. 6 and 7. We note that in the Atlantic

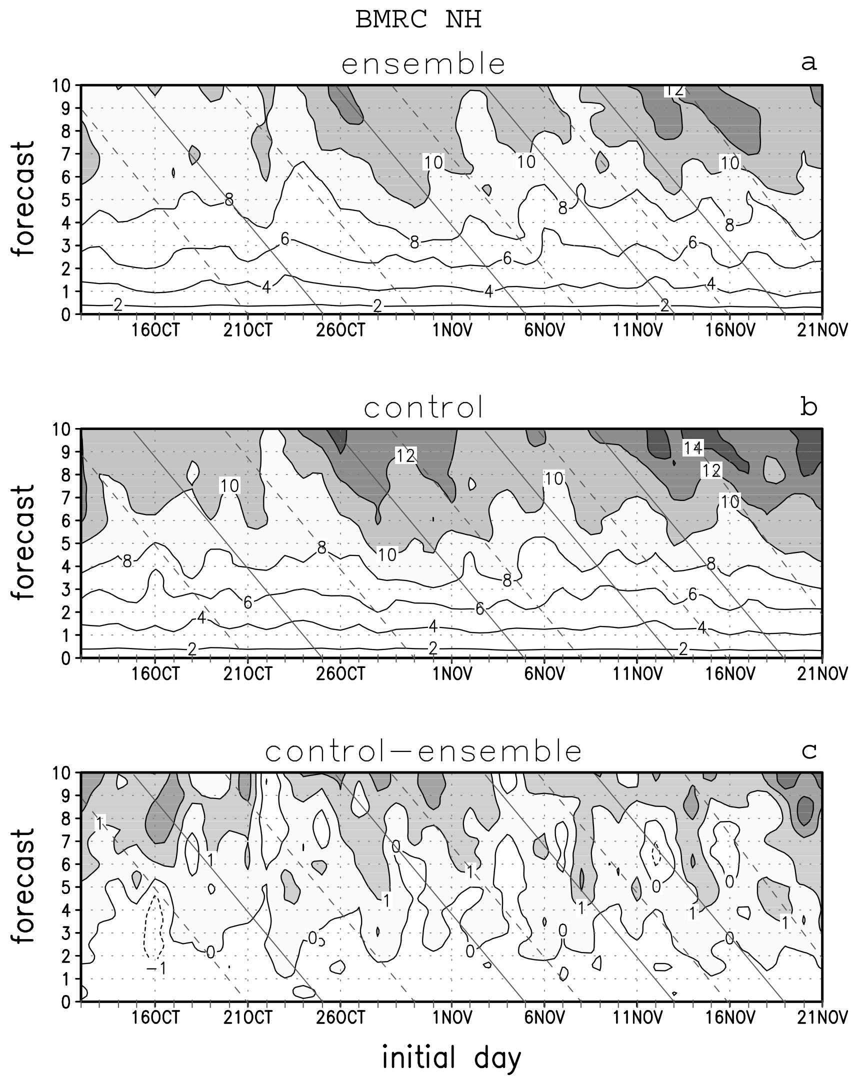

Fig 7. As in Fig. 6 for the BMRC model.

sector the variability of error amplitudes depending on the synoptic situation is considerably larger than for the hemispherically averaged results, as might be expected. The signatures of the life cycle of the four blocks summarized in Section 3 are particularly clear in the control and especially ensemble errors shown in Figs. 9 and 10. Forecasts on diagonals ending on 25 October, 5 November, 13 November and 19 November, when the blocks in the respective European, Gulf of Alaska, Atlantic and North American regions are rapidly developing, tend to have relatively large errors for a given forecast time. In contrast, when the blocks have reached their mature phase, on 30 October, 8 November and 23 November, forecast errors are somewhat smaller than average for a given forecast time. This is consistent with the notion that errors grow rapidly when dynamical development is rapid (Frederiksen and Bell, 1990) and are suppressed in the presence of large-scale equivalent barotropic waves such as mature blocks (Frederiksen, 1978; Colucci and Baumhefner, 1992). The less intense block on 16 November is reflected in the relatively smaller forecast errors around this time in Fig. 10 for the BMRC model but is less evident in Fig. 9 for the CSIRO model.

Again, we may quantify the similarities of the error growth shown in the Hovmoeller diagrams for the Atlantic sector in the CSIRO and BMRC models. These pattern correlations are 0.983 for ensemble errors, 0.980 for control errors and 0.412 for control-ensemble. The pattern correlations between anomalies from the time mean are 0.611 for ensemble errors and 0.540 for control errors. These results are quite comparable with the findings for the whole Northern Hemisphere extratropics discussed in Section 7.

8. Ensemble forecasts and error growth during block development

Next, we examine how forecast errors and forecast spread become structurally organized in particular geographical regions, focusing on cases of block development. Figure 11a shows the

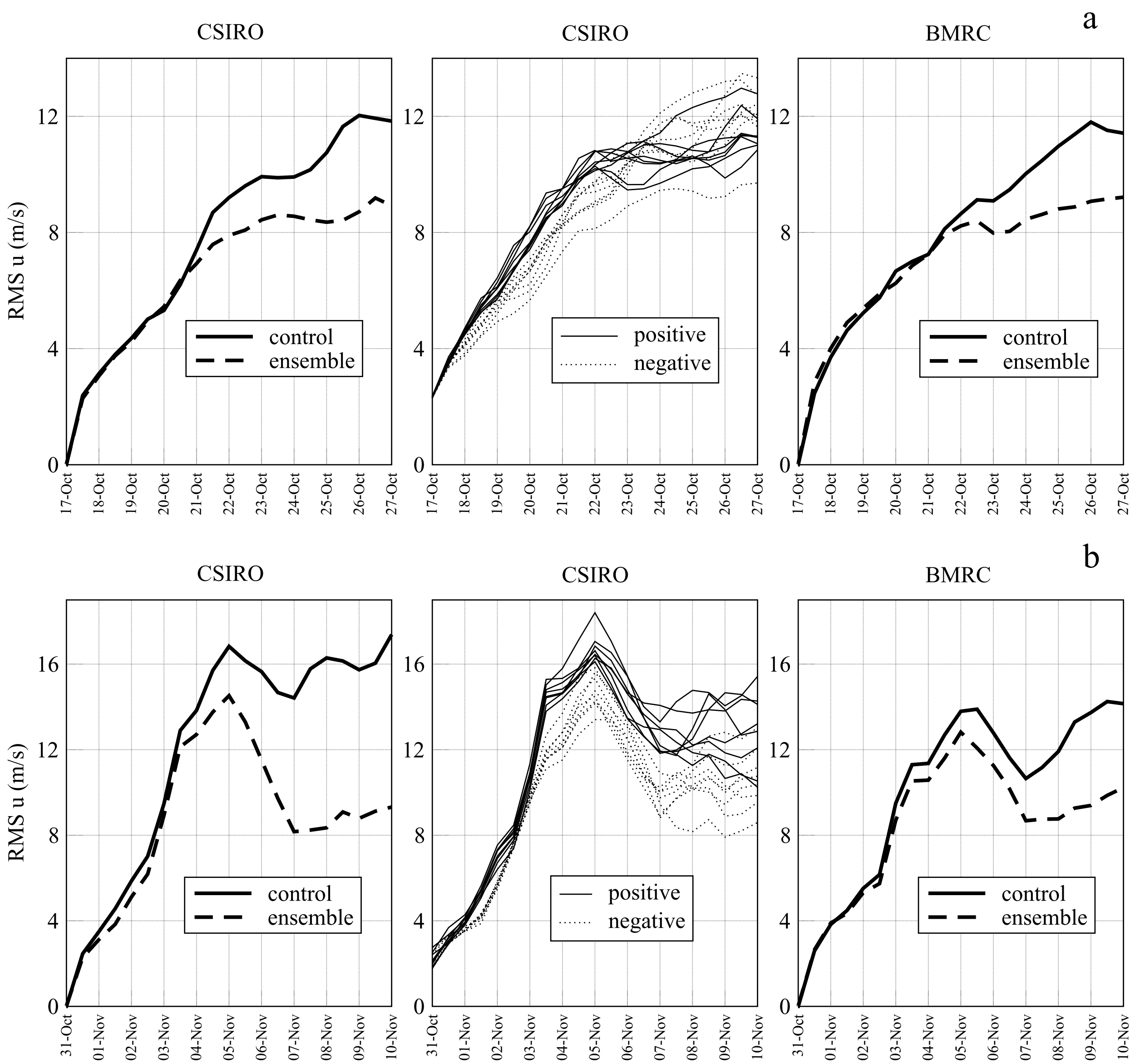

Fig 8. The rms errors of 500-hPa zonal wind (m s^{-1}) in 10-d ensemble mean and control forecasts in CSIRO (left panel) and BMRC (right panel) models initiated at 0000 UTC on (a) 17 October and averaged over the Northern Hemisphere north of 20°N and on (b) 31 October and averaged over the North Atlantic sector north of 20°N. Also shown in centre panels are corresponding error growth of individual members of the CSIRO ensembles.

ensemble mean forecast of 500-hPa geopotential height in the CSIRO GCM for forecasts starting on 24 October and valid for 27 October 1979. The 3-d forecast captures the block formation over Scandinavia but it is somewhat too weak. This may be seen from a comparison of Fig. 11a with Fig. 2a or from the error field of the 3-d CSIRO model ensemble forecast shown in Fig. 12a. Again, the low over Spain is too weak in the 3-d forecast but, as might be expected, is better captured in shorter time forecasts ending on 27 October (not shown). This is also the case for the structure and amplitude of the associated low over Asia Minor. These features of the 3-d forecasts, as seen by comparing Fig. 11a with the NCEP analyses in Fig. 2a, and from Fig. 12a, are reflected in the 3-d standard deviations of the 16 CSIRO model forecasts making up the ensembles.

The standard deviations measuring the spread of the individual forecasts in the ensembles are depicted by the shading in Fig. 11a. We note the quite large standard deviations associated with the 3-d forecast, with considerable spread in the region of the low over Asia Minor. Again, there is significant spread in the blocking region over Scandinavia and in the region of the low over Spain. As well, we notice the wave train of variable standard deviations extending across the major storm track over the

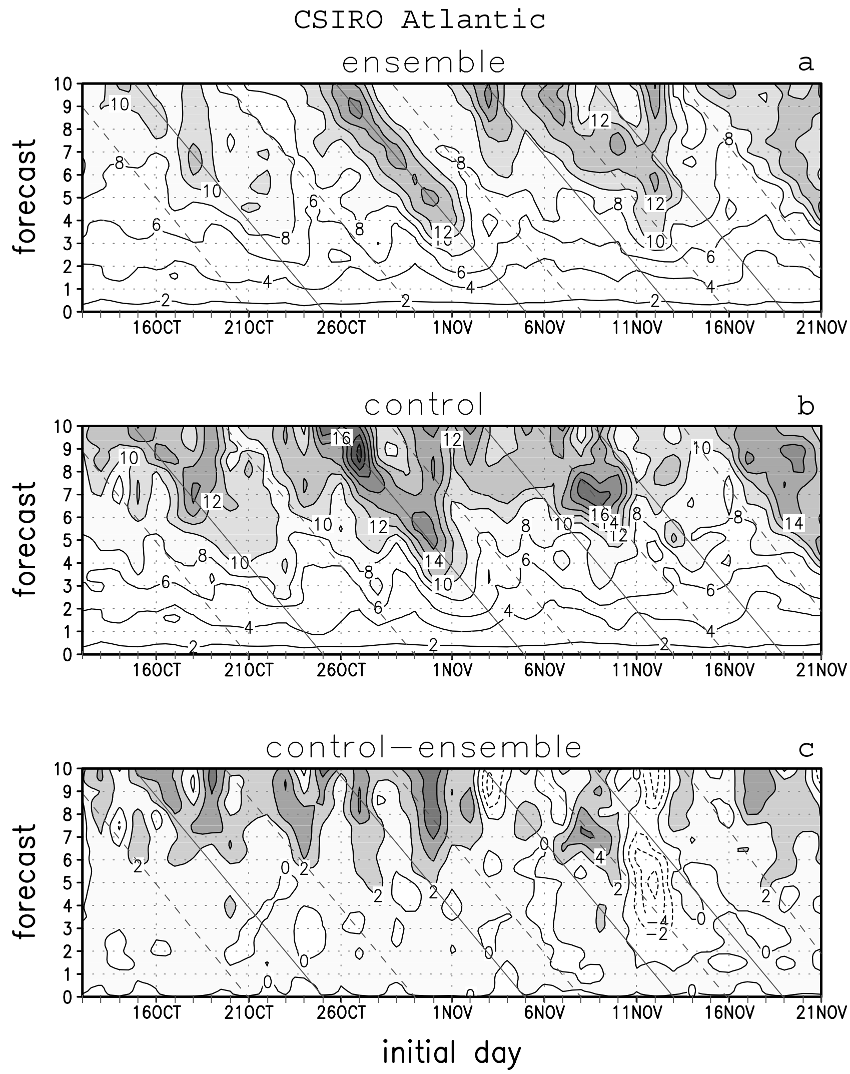

Fig 9. As in Fig. 6 but for the North Atlantic sector.

Pacific and across North America. This variability in the forecasts appears to be associated with the development of baroclinic transients in the major storm track. As expected, the spread associated with baroclinic transients is larger in the 3-d forecast but still quite distinct in the corresponding 1-d forecast (not shown). We also note the variability in the forecasts associated with the development of the lows in the Arctic region. These features of the spread of ensemble member forecasts are also seen in the error field in Fig. 12a indicating the close connection between the two.

We note from Figs. 12a and b that the 3-d error fields for the CSIRO model ensemble and the BMRC model ensemble have quite similar structures, particularly in the regions of large amplitude. This is also the case for the other time periods (Figs. 12c–f) shown in Fig. 12. Here, the BMRC model 3-d error fields are the differences between the BMRC ensembles and the ECMWF analysis. Again, the general relationships between ensemble spread and error structure found for the CSIRO model forecasts apply to the BMRC model forecasts for each of the three periods in Fig. 12.

Figure 11b shows the 3-d ensemble mean forecast of 500-hPa geopotential height in the CSIRO GCM valid on 8 November 1979. In comparison with the analysis in Fig. 2b and with a shorter 1-d forecast (not shown), the 3-d forecast of the blocking dipole over the Gulf of Alaska is of smaller scale and not yet fully cut-off from the westerly circulation. The standard deviations measuring the spread of the individual forecasts making up the ensemble are shown by the shading in Fig. 11b. For the 3-d forecast, the peak spread occurs just downstream of the blocking region in the Gulf of Alaska, with subsidiary maxima both upstream and downstream. The errors, shown in Fig. 12c at 3 d, and spread in the blocking region are consistent with the fact that the flow field over the Gulf of Alaska is a major region of fast growing instabilities, as discussed

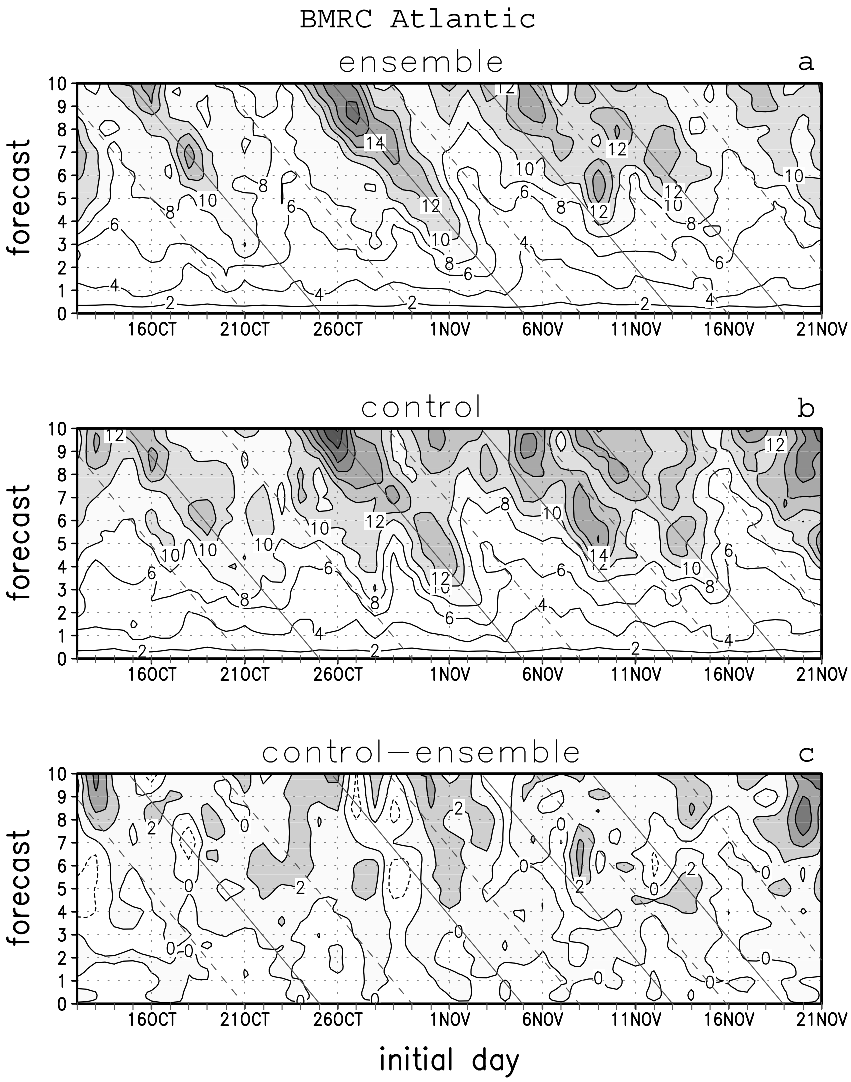

Fig 10. As in Fig. 7 but for the North Atlantic sector.

in detail by Frederiksen (1997, figs. 1 and 2). We note from Fig. 11b that there is also a wave train of forecast variability associated with the Pacific jet stream; this again appears to be due to baroclinic transients. There is also a region of significant spread associated with low intensification over Asia Minor and this is evident in both the 3-d and 1-d (not shown) forecasts. Here, too, there is a close correspondence between spread and error structure.

Figure 11c shows the 3-d ensemble mean forecast valid on 16 November 1979; again the 500-hPa height is depicted and the results are based on the CSIRO GCM forecast. The Atlantic block between Greenland and Scandinavia and the associated low over western Europe are well captured by a 1-d forecast (not shown) but more poorly represented in the 3-d forecast, as seen from Figs. 11b and 2c and the 3-d errors in Fig. 12e. The spread of individual forecasts has maxima in the region of the Atlantic high–low blocking dipole. These results for errors and spread are also consistent with the instability calculations of Frederiksen (1997, fig. 3), which showed that fast growing instabilities were associated with the development of the Atlantic block. Figure 11c shows wave trains of forecast variability from the east coast of Asia across the Pacific, which are associated with baroclinic transients in the major Pacific storm track. The close correspondence between spread and error structure is also evident for this period (Figs. 11c and 12e).

We have also examined the structural organization of forecast variability and errors associated with the American blocking dipole (Fig. 2d) that develops and matures around 22 to 23 November. Again 1-d and 3-d ensemble forecasts with the CSIRO GCM have maxima of forecast spread in and around the region of formation of the high–low dipole with generally larger values for the 3-d forecasts (not shown). The general relationships between errors, forecast variability and the formation of blocking dipoles shown in this section for the CSIRO GCM

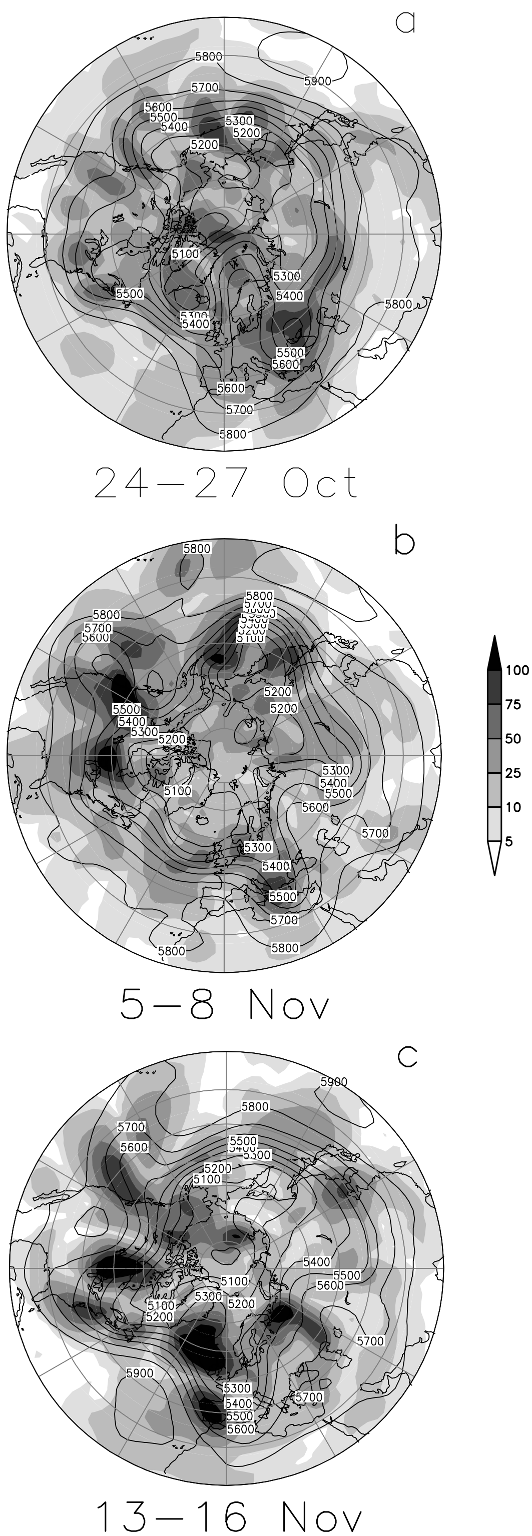

Fig 11. The 3-d ensemble mean forecasts in the CSIRO model of Northern Hemisphere 500-hPa geopotential height (m) valid on 27 October, 8 November and 16 November 1979. Also shown in shading is the spread of the 16 ensemble members measured by the standard deviation about the ensemble mean.

have also been found to be valid for forecasts based on the BMRC GCM using the ECMWF data set.

9. Conclusions

We have implemented and applied ensemble prediction methods based on fast growing perturbations within the CSIRO conformal-cubic and BMRC spectral general circulation models. The methodology employs an iterative procedure in which perturbations analogous to the leading (fastest growing) LV are obtained by a breeding method. We have used a procedure in which bred perturbations are generated over a 10-d period and thereafter ensemble (and control) forecasts are performed based on the bred perturbations and a subsequent self-breeding method. For the ensemble forecasts, which have been performed every 12 h and for 10 d ahead, we have used 16 paired perturbations of opposite signs generated from eight independent breeding cycles.

In this study we have examined the performance of the ensemble prediction method, in comparison with single control forecasts, for initial conditions taken from the period between 0000 UTC on 11 October 1979 and 1200 UTC on 21 November 1979. We have focused on Northern Hemisphere flow during this period because it was a time of development, maturation and decay of large-scale blocks in the major blocking regions over Europe, over the Gulf of Alaska, over the North Atlantic Ocean and also over North America. A major aim of our work has been to examine the variability of predictability depending on the particular synoptic situations, contrasting the growth, maturation and decay of blocks and alternating times of strong zonal flows. In order to establish that our findings are robust and are a reflection of the different instability regimes, depending on the synoptic situations, we have compared results from the two different CSIRO and BMRC models, employing two different analysis data sets from NCEP and ECMWF, respectively.

Within a given model, the eight pairs of bred perturbations have been found to have similar large-scale features, while the stochasticity associated with the convection schemes introduces some differences from the expected leading LV structure. The bred perturbations are found to have large amplitudes in the regions of major instabilities. In particular, during the growth of the Gulf of Alaska and North Atlantic blocks there are close similarities in the structures of the bred perturbations and the instabilities of Frederiksen (1997, 2000) in the blocking regions.

For both the CSIRO and BMRC models, on average the ensemble mean forecast is better than the control forecast for forecast times longer than 3 or 4 d. We have found that 500-hPa Northern Hemisphere rms errors of zonal wind and geopotential height (and meridional velocity and temperature), averaged between 0000 UTC on 11 October and 1200 UTC on 21 November 1979, are lower for the ensemble forecasts than for the control forecasts. Despite the different model formulations, the average error growth curves in the two models are quite similar.

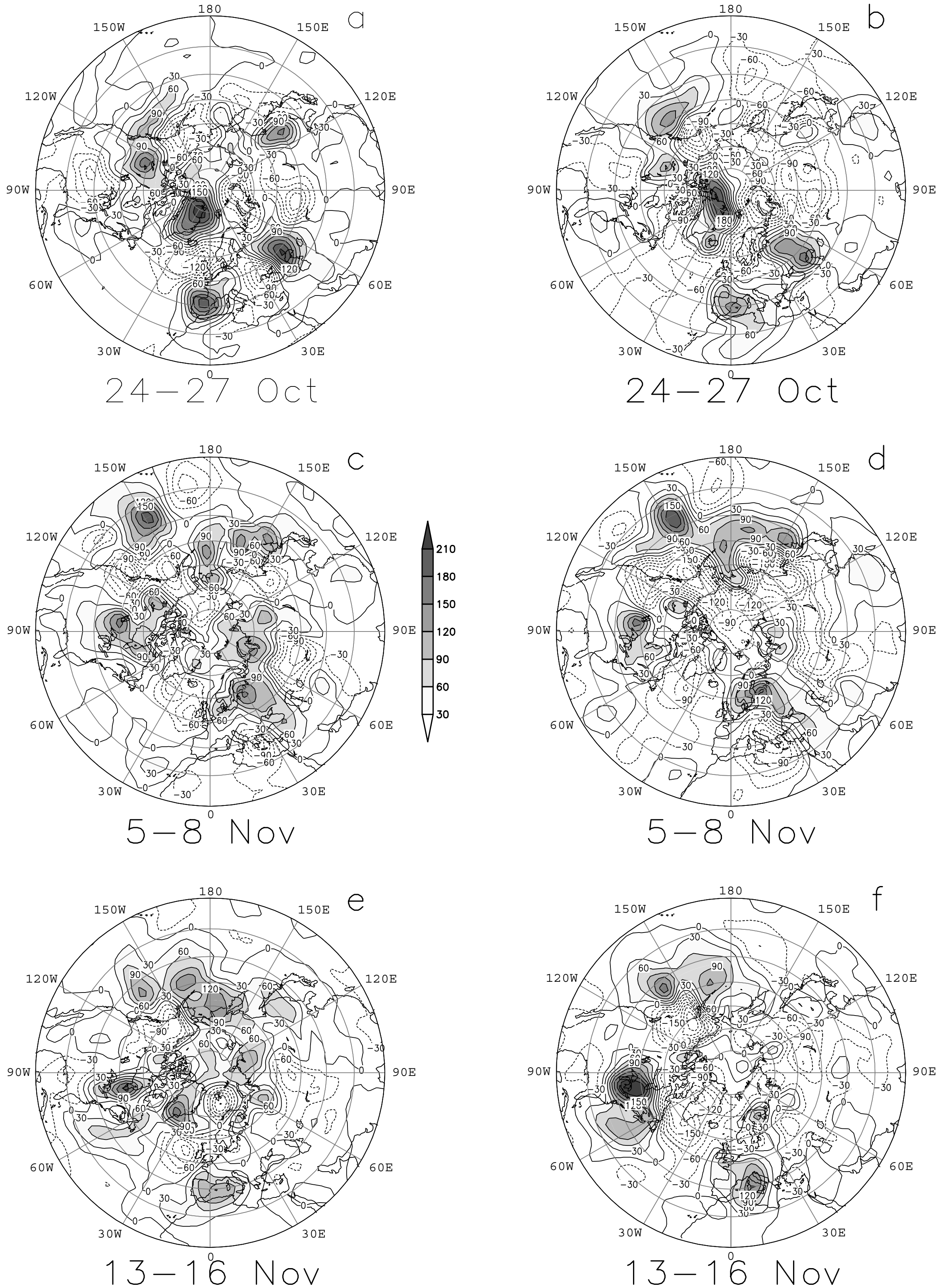

Fig 12. Horizontal structures of Northern Hemisphere 500-hPa geopotential height errors (m) in 3-d ensemble forecasts for the periods 24–27 October, 5–8 November and 13–16 November 1979 and for the CSIRO (left column) and BMRC (right column) models.

We have found that there is considerable variability in the skill of both ensemble and control forecasts initiated twice daily and that this variability is related to particular synoptic events. In particular, at a given forecast lead time, errors tend to be larger for forecasts validating when blocks are developing or decaying and smaller for mature blocks. We have found that this is true both for Northern Hemisphere average errors and for errors averaged over 60° longitude sectors of the particular blocking regions. Similar results are also found for 60° longitude sectors upstream and downstream of the blocking regions and, in particular, we have focused on the Atlantic sector between 0°–60°W. We note that these results are consistent with the notion that errors grow rapidly when dynamical development is rapid (Frederiksen and Bell, 1990) and are suppressed in the presence of large-scale equivalent barotropic waves such as mature blocks (Frederiksen, 1978; Colucci and Baumhefner, 1992).

We have also studied the spread of ensemble member forecasts which provides a measure of their consistency and likely skill. We have examined the variability of error growth associated with different ensemble members. This has been done for errors averaged over the Northern Hemisphere extratropics and averaged over the Atlantic sector, and also the geographical distribution of the spread has been analysed for the major blocking events. We have found that the variability of predictive skill is greatly influenced by rapidly developing systems such as baroclinic transients in the storm tracks and developing blocks. The rapidly growing errors and forecast spread in the regions of block development are consistent with the fast growing instabilities that have been shown to that occur in these areas (Frederiksen and Bell, 1990; Frederiksen 1997, 2000).

We have found that the use of an ensemble forecasting approach, based on an iterative breeding method, generally improves forecast skill including during blocking events. Furthermore, at time-scales beyond 4 d, it is possible to extend the skillful range of model forecasts by an extra half or one day by using ensemble mean prediction.

10. Acknowledgments

It is a pleasure to thank Martin Dix, Jack Katzfey and John McGregor for assistance in using the CSIRO conformal-cubic atmospheric circulation model and pre- and post-processing software. We wish to thank Mike Manton for making available the BMRC spectral atmospheric circulation model for this study.

References

Anderson, J. L. 1993. The climatology of blocking in a numerical forecast model. *J. Climate* **6**, 1041–1056.

Anderson, J. L. 1996. Selection of initial conditions for ensemble forecasts in a simple perfect model framework. *J. Atmos. Sci.* **53**, 22–36.

Bengtsson, L. 1981. Numerical prediction of atmospheric blocking: a case study. *Tellus* **33**, 19–24.

Bontemps, B. and Källén, E. 1999. Singular vectors and breeding modes: a comparitive study. Dynamic Meteorology Report No. 81, Department of Meteorology, Stockholm University, 38 pp.

Borges, M. D. and Hartmann, D. L. 1992. Barotropic instability and optimal perturbations of observed nonzonal flows. *J. Atmos. Sci.* **49**, 335–354.

Colucci, S. J. and Baumhefner, D. P. 1992. Initial weather regimes as predictors of numerical 30-day mean forecast accuracy. *J. Atmos. Sci.* **49**, 1652–1671.

Dole, R. M. and Gordon, N. D. 1983. Persistent anomalies of the extratropical Northern Hemisphere wintertime circulation. Geographical distribution and regional persistence characteristics. *Mon. Wea. Rev.* **111**, 1567–1586.

de Pondeca, M. S. F. A., Barcilon, A. and Zou, X. 1998a. An adjoint sensitivity study of the efficacy of modal and nonmodal perturbations in causing model block onset. *J. Atmos. Sci.* **55**, 2095–2118.

de Pondeca, M. S. F. A., Barcilon, A. and Zou, X. 1998b. The role of wave breaking, linear instability, and PV transports in model block onset. *J. Atmos. Sci.* **55**, 2852–2873.

Frederiksen, J. S. 1978. Instability of planetary waves and zonal flows in two-layer models on a sphere. *Q. J. R. Meteorol. Soc.* **104**, 841–872.

Frederiksen, J. S. 1982. A unified three-dimensional instability theory of the onset of blocking and cyclogensis. *J. Atmos. Sci.* **39**, 969–982.

Frederiksen, J. S. 1983. A unified three-dimensional instability theory of the onset of blocking and cyclogenesis. II Teleconnection Patterns. *J. Atmos. Sci.* **40**, 2593–2609.

Frederiksen, J. S. 1997. Adjoint sensitivity and finite-time normal mode disturbances during blocking. *J. Atmos. Sci.* **54**, 1144–1165.

Frederiksen, J. S. 1998. Precursors to blocking anomalies: the tangent linear and inverse problems. *J. Atmos. Sci.* **55** 2419–2436.

Frederiksen, J. S. 1999. Subgrid-scale parametrizations of eddy-topographic force, eddy viscosity and stochastic backscatter for flow over topography. *J. Atmos. Sci.* **56**, 1481–1494.

Frederiksen, J. S. 2000. Singular vectors, finite-time normal modes and error growth during blocking. *J. Atmos. Sci.* **57**, 312–333.

Frederiksen, J. S. and Bell, R. C. 1990. North Atlantic blocking during January 1979: linear theory. *Q. J. R. Meteorol. Soc.* **116**, 1289–1313.

Frederiksen, J. S., Dix, M. R. and Kepert, S. 1996. Systematic energy errors and the tendency toward canonical equilibrium in atmospheric circulation models. *J. Atmos. Sci.* **53**, 887–904.

Hamil, T. M., Snyder, C. and Morss, R. E. 2000. A comparision of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles. *Mon. Wea. Rev.* **128**, 1835–1851.

Hart, T. L., Bourke, W., McAvaney, B. J., Forgan, W. and McGregor, J. L. 1990. Atmospheric general circulation simulations with the BMRC global spectral model: the impact of revised physical parametrizations. *J. Climate* **3**, 436–459.

Houtekamer, P. L. and Derome, J. 1995. Methods of ensemble prediction. *Mon. Wea. Rev.* **123**, 2181–2196.

Houtekamer, P. L., Lefaivre, L., Derome, J. Ritchie, H. and Mitchell, H. L. 1996. A system simulation approach to ensemble prediction. *Mon. Wea. Rev.* **124**, 1225–1242.

Kalnay, E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, 341 pp.

Kimoto, M., Mukougawa, H. and Yoden, S. 1992. Medium-range forecast skill variation and blocking transition: a case study. *Mon. Wea. Rev.* **120** 1616–1627.

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

Lejenas, H. and Okland, H. 1983. Characteristics of Northern Hemisphere blocking as determined from a long time series of observational data. *Tellus* **35A**, 350–362.

Li, Z., Barcilon, A. and Navon, I. M. 1999. Study of block onset using sensitivity perturbations in climatological flows. *Mon. Wea. Rev.* **127**, 879–900.

Lupo, A. R. 1997. A diagnosis of two blocking events that occurred simultaneously over the mid-latitude Northern Hemisphere. *Mon. Wea. Rev.* **125**, 1801–1823.

Lupo, A. R. and Smith, P. J. 1995a. Climatological features of blocking anticyclones in the Northern Hemisphere. *Tellus* **47A**, 439–456.

Lupo, A. R. and Smith, P. J. 1995b. Planetary and synoptic-scale interactions during the lifecyle of a mid-latitude blocking anticyclone over the North Atlantic. *Tellus* **47A**, 575–596.

McGregor, J. L. and Dix, M. R. 2001. The CSIRO conformal-cubic atmospheric GCM. In: *IUTAM symposium on Advances in Mathematical Modelling of Atmosphere and Ocean Dynamics*, 298 pp (ed.P. F. Hodnett). Kluwer Academic, Dordrecht, 197–202.

Molteni, F., Buizza, R. Palmer, T. and Petroliagis, T. 1996. The ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc.* **122**, 73–119.

Noar, P. F. 1983. Numerical modelling of blocking with reference to June 1982. *Aust. Meteorol. Mag.* **31**, 37–49.

Noone, D. and Simmonds, I. 1998. Similarity of 'fast-growing perturbations' and an illustrative experiment with ensemble forecasting. *Aust. Meteorol. Mag.* **47**, 5–19.

O'Kane, T. J. and Frederiksen, J. S. 2004. The QDIA and regularized QDIA closures for inhomogeneous turbulence over topography. *J. Fluid Mech.* **504**, 133–165.

Palmer, T. N., Gelaro, R., Barkmeijer, J. and Buizza, R. 1998. Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.* **55**, 633–653.

Reynolds, C. A., Webster, P. J. and Kalnay, E. 1994. Random error growth in NMC's global forecasts. *Mon. Wea. Rev.* **122**, 1281–1305.

Simmons, A. J., Wallace, J. M. and Branstator, G. W. 1983. Barotropic wave propagations and instability, and atmospheric teleconnection patterns. *J. Atmos. Sci.* **40**, 1363–1392.

Smagorinsky, J., Manabe, S. and Holloway, J. L. 1965. Numerical results from a nine-level general circulation model of the atmosphere. *Mon. Wea. Rev.* **93**, 727–768.

Szunyogh, I., Kalnay, E. and Toth, Z. 1997. A comparison of Lyapunov and Optimal vectors in a low resolution GCM. *Tellus* **48A**, 200–227.

Tibaldi, S., Ruti, P. and Maruca, M. 1995. Operational predictability of winter blocking at ECMWF: an update. *Ann. Geophys.* **13**, 305–317.

Toth, Z. and Kalnay, E. 1993. Ensemble forecasting at NMC: the generation of perturbations. *Bull. Am. Meteorol. Soc.* **174**, 2317–2330.

Toth, Z. and Kalnay, E. 1997. Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.* **125**, 3297–3319.

Wang, X. and Bishop, C. H. 2003. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.* **60**, 1140–1158.

Watson, J. S. and Colucci, S. J. 2002. Evolution of ensemble predictions of blocking in the NCEP global spectral model. *Mon. Wea. Rev.* **130**, 3008–3102.

Wei, M. and Toth, Z. 2003. A new measure of ensemble performance: Perturbation versus error correlation analysis (PECA). *Mon. Wea. Rev.* **131**, 1549–1565.

Wei, M. and Frederiksen, J. S. 2004. Error growth and dynamical vectors during southern hemisphere blocking. *Non-linear Processes in Geophysics* **11**, 99–118.

Wiedermann, J. M., Lupo, A. R., Mokhov, I. I. and Tikhonova, E. A. 2002. The climatology of blocking anticyclones for the Northern and Southern Hemispheres: block intensity as a diagnostic. *J. Climate* **15**, 3459–3473.