Medium and extended range predictability and stability of the Pacific/North American mode

By T. N. PALMER

European Centre for Medium Range Weather Forecasts, Reading

(Received 14 April 1987; revised 22 October 1987)

SUMMARY

It is shown from an assessment of a small set of extended range forecasts from two centres, and from a much larger set of medium range forecasts from one centre, that variability in predictive skill is strongly related to fluctuations in the Pacific/North American (PNA) mode of low frequency variability. A hypothesis is put forward that this is associated with the dependence of large-scale instability of the forecast flow on the amplitude of the PNA mode. The hypothesis is tested in a barotropic model using as basic states, composite skillful and unskillful cases from the set of medium range forecasts and individual monthly mean fields. Results from the barotropic stability analysis suggest possible reasons for the asymmetric nature of the response of general circulation models to sea surface temperature anomalies of opposite sign, relevant to forecasting on monthly to seasonal timescales. Observational evidence for the stability hypothesis is also discussed.

1. INTRODUCTION

As our understanding of physical processes in the atmosphere has improved and as computer resources have increased, so numerical weather forecasts have become more skillful, particularly in the medium range. Despite this, such forecasts still show considerable variability in predictive skill. This variability occurs on all timescales, from daily to interannual. The ability to predict if a particular forecast is likely to be unusually skillful, or unskillful, could therefore have a substantial impact on the effective skill of a forecasting system, comparable with any change in model formulation.

In the last couple of years there has been an enormous upsurge of interest in the possibility of forecasting into the extended range. Whilst most centres have reported forecasts of exceptional extended range skill, these occur sufficiently infrequently to be of little practical use. A scheme to identify, a priori, such cases must be considered essential if extended range forecasting is to become a practical proposition.

The purpose of this article is to note that the skill of numerical forecasts, both in the medium and extended range, is strongly influenced by the amplitude of the Pacific/North American (PNA) mode of low frequency variability (Wallace and Gutzler 1981; Branston and Livezey 1987). A large number of forecasts are used to substantiate this, and a physical hypothesis is put forward. It is suggested that the growth of analysis or short range forecast errors is dependent on the barotropic stability of the forecast flow. The hypothesis is tested in a barotropic model, using as basic states, both composite skillful and unskillful forecasts from a set of 500 wintertime medium range forecasts, and from a set of monthly mean fields with high amplitude PNA index. It is found that the degree of instability is strongly dependent on the signed amplitude of the PNA mode. Observational evidence for the stability hypothesis is also given.

As well as being relevant to the problem of variability in the predictive skill of numerical weather prediction models, results from the barotropic stability analysis suggest reasons for the asymmetric nature of the response of general circulation models to sea surface temperature anomalies of opposite sign. This is of direct relevance to the forecasting problem on monthly to seasonal timescales.

691
2. Predictability and the PNA mode

(a) Extended range predictability

Results from a series of extended range experiments from two forecasting centres have recently appeared in the literature. Mansfield (1986) discusses the extended range skill of the Meteorological Office 5-layer model over 18 wintertime cases with initial conditions from December 1974 to December 1981. On average the anomaly correlation for these forecasts fell to zero by about day 19 (using 15-day-mean time averages). On the other hand, as Mansfield reports, some forecasts had exceptional skill (see Table 1). In particular, five out of the set of 18 forecasts had positive anomaly correlations at day 30: the initial conditions for these were 23 December 1975, 14 December 1976, 31 December 1976, 19 December 1980 and 29 December 1980. Mansfield notes that whilst the more skilful forecasts tend to be associated with greater than average persistence in the real data, persistence alone is not sufficient to explain the extended predictability.

| TABLE 1. Initial dates of a set of 18 50-day forecasts used by Mansfield (1986) to document the extended range skill of the Meteorological Office 5-layer model |
|---------------------------------|-----------------|-----------------|
| 14 December 1974               | 31 December 1976| 23 December 1979|
| 23 December 1975               | 31 December 1977| 29 December 1980|
| 23 December 1975               | 23 December 1978| 29 December 1980|
| 01 January 1976                | 29 December 1978| 23 December 1981|
| 14 December 1976               | 14 January 1979  | 29 December 1981|

Those initial dates in italic correspond to cases where the 15-day-mean forecast centred on day 30 showed positive skill.

Figure 1 shows the 200mb monthly mean observed streamfunction anomalies for the months of January 1977 and January 1981; the months for which Mansfield's forecasts showed positive skill, for all relevant initialization dates, at day 30. Over the Pacific and North America both have a similar pattern, with positive anomalies over the tropical Pacific and the Rockies, and negative anomalies near the Aleutian Islands, and over south-east U.S.A. This figure shows for each of these months an anomalous realization of the PNA mode of variability, see Fig. 2, and therefore raises the possibility that variability in the skill score of extended range forecasts may be linked to variability in the amplitude of a PNA-like mode.

Further support for this notion comes from a recent analysis of a set of eight extended range forecasts integrated with a GFDL model (Miyakoda et al. 1986). Table 2 shows the (bias-corrected) 30-day-mean scores together with the amplitudes of the observed (monthly mean) PNA index (as defined by Wallace and Gutzler). Five of the eight forecasts were for months common to Mansfield's set. It is interesting to note that, of these, Miyakoda et al.'s forecasts for January 1977 and January 1981 were the most successful, suggesting that the success of these exceptional cases may be model-independent. The streamfunction anomalies for January 1983, Miyakoda et al.'s most successful forecast period, are shown in Fig. 1(c). Again the strong PNA-like pattern is evident during the strong El Niño event (Quirez 1983). Furthermore, as Table 2 shows, there is a clear correlation over the full set of Miyakoda et al.'s integrations between forecast skill and the signed amplitude of the PNA mode, equal to -0.85.
Figure 1. Observed 200 mb streamfunction anomaly (×10^8 m^2 s^-1) for: (a) January 1977; (b) January 1981; and (c) January 1983. From U.S. National Meteorological Center Climate Analysis Center archives. Contour interval 6×10^8 m^2 s^-1.

Unfortunately, insufficient extended range forecast data are available to study more objectively such relationships between forecast skill and low frequency modes of variability. On the other hand, medium range forecasts have been run operationally for several years, and archived forecast data provide an excellent base with which to address this question.
Figure 2. Map showing teleconnectivity of 500 mb height from a base point over the east Pacific, and illustrating the PNA mode. From Wallace and Gutzler (1981).

<table>
<thead>
<tr>
<th>Initial date</th>
<th>Anomaly correlation</th>
<th>PNA index</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 January 1977</td>
<td>-81</td>
<td>150</td>
</tr>
<tr>
<td>01 January 1978</td>
<td>-68</td>
<td>120</td>
</tr>
<tr>
<td>01 January 1979</td>
<td>-42</td>
<td>20</td>
</tr>
<tr>
<td>16 January 1979</td>
<td>-57</td>
<td>-30</td>
</tr>
<tr>
<td>01 January 1980</td>
<td>-52</td>
<td>-10</td>
</tr>
<tr>
<td>01 January 1981</td>
<td>-71</td>
<td>210</td>
</tr>
<tr>
<td>01 January 1982</td>
<td>-49</td>
<td>-60</td>
</tr>
<tr>
<td>01 January 1983</td>
<td>-83</td>
<td>190</td>
</tr>
</tbody>
</table>

The correlation between data in the 2nd and 3rd columns is -85.

(b) Medium range predictability

Results from this section are part of an extensive study to devise a procedure for predicting the skill of medium range forecasts. Preliminary results have been given in Palmer and Tibaldi (1986, 1987).

The data base for this study is a set of 500 10-day wintertime forecasts using the ECMWF forecast system from 1980/1981 to 1984/1985 starting with the forecast for 1 December in each winter. In this paper, skill was defined in terms of the 500 mb height
r.m.s. error and calculated both on a regional and hemispheric basis. Each region is bounded by 60°N and 30°N and extends from longitudes 0 to 30°E, 30 to 60°E, . . . , 30°W to 0 (see Fig. 3).

Figure 3. Definition of 12 regions used in the regression analysis.

In order to study objectively possible relationships between forecast skill and configurations of the large-scale flow, the 500mb height data have been projected onto a set of basis functions. In principle, it would be possible to use, for example, a spherical harmonic basis. However, in an attempt to minimize the number of basis functions required to describe any significant relationships between large-scale flow variability and skill variability, it was decided to use a set \( \{E_i(x,y)\} \) of empirical orthogonal eigenfunctions (EOFs), defined over the whole hemisphere, from pentad-mean fields from 32 years of wintertime analyses (1952-84, taken from NMC and ECMWF archives). These EOFs were calculated separately on the zonal mean and on deviations from zonal symmetry. The first of these EOFs corresponds to variation in the hemispheric mean height. Then five ‘zonal’ EOFs were retained, explaining 99-96% of the total variance of the zonal mean fields, and 17 eddy EOFs, explaining 86-9% of the variance corresponding to fluctuations in the zonally varying component of the flow. Further details have been given by Molteni (1987).

A linear regression analysis was then performed using the regional or hemispheric skill as predictand, and EOF coefficients of the forecast flow as predictors. Let \( 1 \leq \alpha \leq 500 \) denote a daily index over the sample of forecast data. Furthermore, for day \( \alpha \), let \( e_{ja} \) denote the \( j \)th forecast EOF coefficient, \( 1 \leq j \leq 23 \). The hemispheric or regional score for that day will be denoted by \( S_{\alpha} \). We shall assume in the following that for fixed \( j \) \( \{e_{ja}\} \), \( \{S_{\alpha}\} \) have been normalized and standardized, i.e. if \( \langle \cdot \rangle \) denotes a mean over \( \alpha \),

\[
\langle S_{\alpha} \rangle = \langle e_{ja} \rangle = 0 \quad \text{and} \quad \langle S_{\alpha}^2 \rangle = \langle e_{ja}^2 \rangle = 1.
\]

In linear regression theory we find a set of weights \( a_j \) such that \( Q = \langle (S_{\alpha}a_{\alpha} - S_{\alpha})^2 \rangle \) is minimized. The condition \( \partial Q / \partial a_j = 0 \) straightforwardly gives these weights as \( a_j = \rho_j C^{-1}_{jj} \), where \( \rho_j = \langle S_{\alpha}e_{ja} \rangle \) and \( C_{jj} = \langle e_{ja}e_{ja} \rangle \). The quantity \( r_\alpha = \sum a_j e_{ja} \) is the regressed or ‘predicted’ skill score given a forecast with EOF coefficients \( e_{ja} \). The sample-mean correlation between the real and regressed skill score is \( \langle r_\alpha S_{\alpha} \rangle = \sum \rho_j C^{-1}_{jj} \rho_j = \rho \).
The so-called factor structure constants $\rho_i$ give the correlation between the scores $s_i$ and each EOF coefficient $e_{iq}$. Hence those EOF coefficients most strongly correlated with the scores will be emphasized in a forecast field with EOF coefficients equal to $\rho_i$. Alternatively, writing $\rho_i = a_i C_q$ then a forecast field with EOF coefficients equal to $\rho_i$ can be thought of as representative of the synoptic situations associated with the regression weights, taking climatological correlation between EOF coefficients into account. Forecasts with EOF coefficients proportional to $\rho_i$ are most strongly correlated with variations in forecast skill. The regressed skill score for a forecast with EOF coefficients $\rho_i$ is equal to $\Sigma e_{iq} \rho_i = \rho$. In nonstandardized form this would correspond to a score equal to $\rho \sigma + \mu$, where $\sigma$ is the standard deviation of the regional or hemispheric scores, and $\mu$ its average value.

The correlation, $\rho$, between the observed and regressed score for day 9 of the forecast exceeds .45 in only three regions; 1, 8 and 12, where $\rho = .48$, .46 and .46 respectively. (If a 5-day running mean is applied to the day-9 forecasts and day-9 skill scores, prior to performing the regression, the correlation, $\rho$, for these three areas increases to .62, .59 and .60 respectively.) It should be noted (see Fig. 3) that these regions are positioned close to the ends of the Pacific and Atlantic storm tracks.

The 500 mb height forecast field anomaly, $z(x, y) = \rho^{-1} \Sigma \rho_i E(x, y)$, formed by the factor structure constants for the regression using scores in region 8, is shown in Fig. 4. The equivalent fields for the neighbouring regions 7 and 9 are also shown.

It is clear that in each region of the East Pacific/North American quadrant, the factor structure constants are describing a mode of variability that closely resembles the PNA mode. The pattern is well localized in regions 8 and 9, less so in region 7. The sign of the anomaly field is such that when a day-9 forecast field anomaly correlates positively with the appropriate pattern in Fig. 4, the r.m.s. error will be larger than average by one standard deviation in the appropriate region. Hence the results here are in agreement with the analysis of the extended range forecasts discussed above, though of course it should be recognized that only a relatively small fraction of the variance of individual forecast scores is explained by this mode. When the 5-day running mean is applied before the regression analysis is done (see above), the resulting patterns differ little from those shown in Fig. 4.

Figure 5 shows the forecast pattern corresponding to factor structure constants for skill defined firstly over region 12 and secondly over the whole hemisphere. A quite different pattern to the PNA emerges for region 12, with most of its amplitude near Europe. With hemispheric scores as predictand, the factor structure constant pattern shows both PNA and European modes, though with a relatively low correlation $\rho = .31$. The pattern correlating with hemisphere scores appears therefore as a composite of more regional modes of oscillation rather than representing a coherent hemispheric scale pattern.

As discussed in Palmer and Tibaldi (1986), forecast skill over the PNA region in short range forecasts does not correlate with any quasi-stationary PNA-like mode of variability of the forecast flow. For example, Fig. 5(e) shows the factor structure constant pattern for day 2 associated with r.m.s. error in region 8. The pattern is quite different from the day-9 counterpart in Fig. 4(b), and the correlation $\rho = .37$ is lower. The negative anomaly over region 8 is not a 'stationary' feature of the regression in the sense that when day-2 skill in region 7 or 9 is regressed against forecast flow, the negative anomaly centre moves upstream or downstream respectively. Palmer and Tibaldi (1986) speculate that this may indicate the overwhelming importance of transient baroclinic monopole cyclogenesis modes of instability influencing regional skill scores in the short range (see also discussion).
Figure 6 shows an example of the correspondence between the actual day-9 r.m.s. errors in region 8 (the dashed line), and the regressed skill scores (the solid line: effectively the amplitude of the mode in Fig. 4(b)) for the winter of 1980/81 (part of the dependent data set). A five-day running mean has been applied to both graphs in order to emphasize low frequency variability. During the first 20 or so days of January, the forecast model was particularly skilful over the east Pacific region, and correspondingly, the amplitude of the PNA mode was anomalously strong. As discussed above, both Mansfield and Miyakoda et al. have reported exceptionally skilful extended range forecasts for January 1981.
Figure 5. As Fig. 4 but (a) skill for region 12; (b) hemispheric skill; (c) day-2 skill for region 8.
Figure 6. Actual day-9 r.m.s. error in metres (dashed line) for the winter 1980/81 of the ECMWF operational model over the region 60°N–30°N, 150°W–120°W. Regressed day-9 r.m.s. error (solid line) for the same winter and same region based on statistics derived from 500 wintertime forecasts from 1980/81 to 1984/85. A 5-day running mean has been applied to both curves to emphasize low frequency variability.

3. A POSSIBLE PHYSICAL MECHANISM

In this section we shall discuss possible mechanisms that might explain the results above. Broadly speaking, candidate mechanisms could be classified as those related directly to systematic deficiencies in model formulation, and those related to intrinsic loss of predictability of the atmospheric flow due to its instability characteristics. Indeed the existence of model systematic errors can give rise to a 'trivial' relationship between forecast skill and variability in forecast flow. Consider, for example, a model with no predictive skill and a climate drift $\Delta C(t)$ where $t$ is forecast verification time. Clearly, over a large enough sample the r.m.s. error of those forecasts which are close (in an r.m.s. sense) to the observed climate will be smaller than those farther away. Hence forecasts with anomaly $-\Delta C(t)$ (relative to the model climate) will, over a large sample, have smaller r.m.s. error than those with anomaly $+\Delta C(t)$.

Figure 7 shows the ECMWF model mean (or systematic) day-9 wintertime error, averaged over all 500 forecasts that make up the regression sample. If the PNA-like mode of variability shown in Fig. 4 correlated strongly with this mean error pattern, there would be strong prima facie evidence that the impact of model systematic errors as discussed above was the dominant mechanism explaining the results in section 2. However, it can be seen that the principal anomaly centres of the PNA mode are not in phase with the centres of maximum systematic error.

In order to study the stability characteristics of flow patterns associated with either skilful or unskilful forecasts, it is worth recalling that in the medium and extended range, forecast errors are dominated by large-scale quasi-stationary patterns (Wallace et al. 1983). Simons et al. (1983) have shown that the climatological zonally varying flow is
barotropically unstable to just such large-scale quasi-stationary modes. (It should be noted that a similar instability had been identified in a barotropic model by Matsuno and Hirota (1966) using as basic state, observed 50 mb zonally varying flows.) Therefore, in order to test the importance of variability in the intrinsic predictability of forecast flows, the barotropic stability of basic state flows which differ from climatology by either the subtraction or addition of the anomaly field shown in Fig. 4(b) was studied. These two basic states are shown in Fig. 8. (More precisely these represent 300 mb streamfunction fields formed by adding to a climatological 300 mb streamfunction field, either minus (Fig. 8(a)) or plus (Fig. 8(b)), the geostrophic streamfunction anomaly formed from Fig. 4(b), scaled by the density factor 5/3.) One can see that with the mode subtracted, (basic state B-), the east Asian jet extends across the Pacific and is weakly diffuent over the east Pacific. The jet is then steered north by the enhanced Rockies ridge. With the mode added (basic state B+), the jet is more strongly diffuent over the central Pacific.

In order to present results in a way that corresponds most closely with the hypothesis that the variability in skill is related to the rate at which initial errors grow, we show in Fig. 9 results from two (nonlinear) initial value problems, where in both integrations an identical disturbance was introduced at 30°N 120°E. The model's drag coefficient was set to (10 days)^{-1}. (Similar types of initial value problems were reported by Simmons et al. 1983. As discussed by these authors, the quasi-stationary growth of an initial disturbance can be interpreted in terms of its projection onto the unstable normal modes of the basic state flow. Readers unfamiliar with this type of calculation, and its interpretation are therefore referred to Simmons et al. for details.)
Figure 8. 300 mb streamfunction basic states defined by subtracting ($B-$; (a)) or adding ($B+$; (b)) onto a 300 mb streamfunction climatology, the mode shown in Fig. 4(b) associated with variability in forecast skill. The contour interval is equivalent to a 120 m interval in geopotential height computed by applying the geostrophic approximation at 45°N.

Figure 9 shows the model streamfunction response at days 2, 4, 9 and 30 for the two basic states. At day 2, the initial disturbance is seen propagating downstream and the effects of the difference in basic states are small. At day 4 the disturbance has propagated further downstream, and, for example, the effect of the enhanced Rockies ridge in the basic state $B-$ can be seen steering the disturbance to the north. However, the major difference is the amplitude of the disturbance centre over the central Pacific. By day 9, this difference is more clearly marked; the downstream propagating wavetrain has largely dissipated and a quasi-stationary unstable mode is growing in the central Pacific. There is almost a factor of three difference in the amplitudes of the central Pacific response at day 9 between $B-$ and $B+$. This difference continues to grow throughout the thirty-day integration; by day 30 there is substantial difference in the response between $B-$ and $B+$. It should be mentioned that these integrations were repeated with initial disturbances $10^{-6}$ of the present values. Since the perturbation equations are nonlinear, the response did not scale precisely linearly. However, with contour interval also scaled by a factor of
Figure 9. Perturbation streamfunction from integrations of a barotropic model with the two basic states shown in Fig. 8, and identical perturbations positioned at 30°N 120°E. The top diagrams are for basic state B−, the bottom for B+. (a) Day 2; (b) day 4; (c) day 9; (d) day 30. For details see text and Simmons et al. (1983). The contour interval is equivalent to a 10 m interval in geopotential height computed by applying the geostrophic approximation at 45°N. The initial disturbance had an amplitude of 200 m. Integrations with reduced amplitude show that the dynamics of the disturbance is approximately linear (see text).

10^−6, the differences between the growths on basic states B− and B+ were essentially the same as shown here. Hence the perturbation dynamics are essentially linear.

The barotropic stability of various monthly mean height fields has also been cal-
culated by Simmons (1987) in terms of a normal mode analysis. Some of his results are reproduced here. Two 30-day-mean fields, for January 1981 and January 1983, were noted by Simmons as being quite stable in the sense that broadly similar patterns to the 30-day-mean anomaly are seen in the Pacific/North American sector for individual 10-day means within the period. Both months had a pronounced (positive index) PNA-like pattern (see Figs. 1(b), (c)).
Figure 10. Structure of the two most unstable northern hemisphere normal modes for: (a) and (b) January 1981 monthly mean 300 mb streamfunction field; and for (c) and (d) a 300 mb streamfunction field calculated by subtracting from climatology the 300 mb streamfunction anomaly for January 1981. The e-folding time and period of the mode are given. For nonstationary waves, the direction of the arrows denotes phase, with a clockwise rotation representing a passing of time. See Simmons et al. (1983) for further details.
Figure 10 (continued)
For each of these months Simmons calculated the normal mode structure both for the monthly mean field itself, and for a field constructed by subtracting the observed monthly mean anomaly from climatology. Figure 10 shows the most unstable modes in the northern hemisphere for January 1981 and the construction with reversed anomaly. The e-folding time of the most unstable mode for January 1981 is about twice that for the state with the anomaly reversed. Furthermore the most unstable mode is stationary with reversed anomaly, with greatest amplitude over the North Pacific.

![Graph](image)

Figure 11. e-folding times and periods of unstable northern hemisphere modes computed for basic states based on: (a) January 1981; and (b) January 1983. The open circles denote modes for the actual observed mean states, and the solid circles modes for basic states in which the signs of the observed anomalies were reversed. Only modes with e-folding times below 10 days are shown. From Simmons (1987).

For the January 1983 case, Simmons found that with reversed anomaly there were six normal modes with e-folding time smaller than with observed anomaly. This can be seen in Fig. 11 which shows a scatter plot of the period and e-folding time for normal modes of January 1981 and 1983 together with reversed anomaly states.

Normal mode analysis for the composite states in Fig. 8 did not reveal such rapidly growing modes as for the individual monthly mean basic states (typically the fastest growing modes had e-folding times around eight days for both basic states in Fig. 8). Nevertheless, as for the individual months, the fastest growing modes for the negative PNA index basic state (B+) had their amplitudes more geographically localized over the North Pacific than the modes for the positive PNA index basic state (B−). Hence it would appear that the faster growth on B+ arises because the perturbation projects more
strongly onto the most unstable normal modes. (More precisely, the local growth of perturbations in the initial value integration depends on their projection onto relevant adjoint eigenmodes (Branstator 1985; Z. J. Zhang, personal communication). Further details will appear elsewhere.)

Hence the precise normal mode structure does appear to vary both between months with similar PNA index, and between individual months and the composite shown in Fig. 4. A more profound theoretical understanding of the relationship between normal mode and basic state structure is therefore required. Nevertheless these results strongly suggest that the barotropic stability of the atmosphere is a function of the signed amplitude of the PNA mode and are sufficiently striking that it appears quite possible that the stability characteristics of the two 'composite' basic states could explain the forecast skill variability results above.

4. OBSERVATIONAL EVIDENCE

The results of section 3 suggest that there should be an asymmetry in the statistics of observational diagnostics associated with anomalously positive or negative phases of the PNA mode. In particular, if the circulation is more barotropically stable during spells of positive PNA index, then a time-series of the PNA index, suitably filtered to remove high frequency transience associated with individual baroclinic systems, should show larger variability when the PNA index is negative (positive height anomaly over the Aleutians) than when it is positive. A distribution of the index should therefore be skewed towards moderate positive values though with a relatively larger number of high amplitude negative cases.

<table>
<thead>
<tr>
<th>Amplitude of PNA mode (m)</th>
<th>Frequency of occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive PNA index</td>
</tr>
<tr>
<td>25–75</td>
<td>170</td>
</tr>
<tr>
<td>75–125</td>
<td>160</td>
</tr>
<tr>
<td>125–175</td>
<td>130</td>
</tr>
<tr>
<td>175–225</td>
<td>60</td>
</tr>
<tr>
<td>225–275</td>
<td>40</td>
</tr>
</tbody>
</table>

From 14 years of daily 500 mb height observations, low pass filtered. Adapted from Fig. 8 of Dole (1986).

Table 3 is adapted from Fig. 8(a) of Dole (1986). It shows (essentially) the frequency of occurrence of an anomalous PNA mode (defined by EOF analysis) of amplitude within a prescribed range, from a 14-year 90-day winter dataset of observed 500 mb heights. A low-pass filter has been applied to the time series to remove periods of less than 10 days. As above, it can be seen that for amplitudes within the ranges 25–75 m and 75–125 m, there are more occurrences of positive PNA index than negative, whilst for amplitude within the ranges 175–225 m and 225–275 m, the reverse is true. (Note that, what we term 'positive PNA index' Dole terms 'negative PAC anomaly'.) Similar results were reported in studies by Shukla and Mo (1983) and Dole and Gordon (1983). Hence the stability hypothesis is entirely consistent with earlier observational studies.
Figure 12. (a) Mean of all pentad 500 mb height fields from the 32-year dataset described in the text with PNA index greater than 0.5 standard deviations above zero. (Contour interval 100 m.)

(b) As (a) but with PNA index less than 0.5 standard deviations below zero.
(c) Difference field between (b) and (a). (Contour interval 50 m.)
(d) Standard deviation (m) of pentad-mean fields comprising the composite in (a). (Contour interval 25 m.)
(e) As (d) but for fields comprising the composite in (b).
A more direct verification of the stability hypothesis can be obtained by computing composite patterns of positive and negative PNA index. In Fig. 12(a) we show a composite of all pentad-mean fields from the 32-year dataset used to define the EOFs in the regression analysis, with PNA index greater than 0.5 standard deviations above zero (188 cases). Figure 12(b) shows the composite with index less than -0.5 standard deviations (184 cases). The difference (Fig. 12(c)) between (a) and (b) is a manifestation of the PNA mode.

Next the standard deviations of the pentad-mean fields within each of the two composites were calculated. These are shown in Figs. 12(d) and (e). Both (d) and (e) are similar over the Atlantic and Eurasia, but differ significantly over the east Pacific. The composite with positive PNA index shows a standard deviation of about 100 m whilst the composite with negative index shows a standard deviation in excess of 150 m. Hence there is about half as much variability again for pentad fields in the negative PNA index composite as for pentad fields in the positive PNA index composite—again consistent with the stability hypothesis.

A final diagnostic that is relevant to the discussion is shown in Fig. 13, a scatter diagram of the winter-mean 500 mb r.m.s. height error of the ECMWF day-9 forecast (ordinate) and a persistence forecast (abscissa) for the seven winter periods 1980/81 to 1986/87, for region 7 (containing one of the centres of action of the PNA pattern; see Figs. 2 and 3). It is clear that there is a marked correlation. Note that the two winters 1980/81 and 1982/83 whose Januaries we have noted to have large PNA index were particularly persistent (and well forecast), whilst 1981/82, for example, which had a mean negative PNA index for the whole winter was more variable and much less well forecast. Hence again we see evidence of a relationship between stability, the signed amplitude of the PNA mode, and forecast skill.

![Figure 13. Scatter plot of ECMWF winter-mean day-9 500 mb height r.m.s. error (m) in region 7 against the day-9 500 mb height r.m.s. error (m) of a persistence forecast in region 7. 1 = 1980/81, 2 = 1981/82, ..., 7 = 1986/87.](image-url)
5. Predictability on Seasonal Timescales

Predictive skill on timescales greater than a month will depend strongly on knowledge of anomalous boundary forcing for the atmosphere. In recent years there has been much research into the atmospheric response to the El Niño event, and it has been clearly established that (at least for models with realistic climatology) the response of a general circulation model (GCM) to a composite El Niño sea surface temperature (s.s.t.) anomaly includes a statistically significant excitation of the PNA mode (see, for example, Blackmon et al. 1983; Palmer and Mansfield 1986). Cubasch (1985) in addition to studying a number of 150-day integrations with a composite El Niño anomaly, also considered the response of the same (ECMWF) model to an s.s.t. anomaly of equal amplitude but opposite sign (an artificial 'anti-El-Niño' s.s.t. anomaly). With positive s.s.t. anomalies, Cubasch found a statistically significant response in the extratropical northern hemisphere that bore some resemblance to a PNA pattern. With negative s.s.t. anomalies, the response was broadly the opposite in the tropics, but the extratropical response to the negative s.s.t. anomalies was not statistically significant.

It is possible that Cubasch's results could be explained by the fact that the tropical diabatic heating anomaly was weaker for the anti-El-Niño anomaly, than for the El Niño anomaly. However, this effect cannot be used to explain the recent results of Pitcher et al. (1988) who studied a GCM response to s.s.t. anomalies in the extratropical North Pacific. With the observed cold s.s.t. anomalies across much of the North Pacific for the winter of 1976/77, Pitcher et al.'s results show that the model response is a statistically significant PNA pattern, with deeper than normal Aleutian low, in agreement with observations, and consistent with the Atlantic modelling studies of Palmer and Sun (1985). On the other hand, when the model was integrated with s.s.t. anomalies of equal amplitude but opposite sign, the model response was not statistically significant over the North Pacific. In this respect the results are similar to those of Cubasch, but, unlike his results, Pitcher et al. report that the diabatic heating anomalies over the North Pacific were approximately equal in magnitude but opposite in sign, in the two experiments.

A possible explanation of both Pitcher et al.'s and Cubasch's results could lie in the discussion above concerning the stability of the PNA mode. With positive El Niño anomalies, or negative North Pacific anomalies, it might be conjectured that the general circulation is being forced towards a more barotropically stable state. With anti-El-Niño anomalies, or with positive North Pacific anomalies, the general circulation is being forced towards a less stable state, and, as a consequence, the model internal variability may be too large to give a significant response.

One diagnostic from the GCM runs that would be relevant to the above discussion is the ratio of low frequency variance between integrations with positive and negative s.s.t. anomalies. If the above ideas are correct they suggest that there should be less low frequency variability in the El Niño/negative North Pacific s.s.t. anomaly experiments than in the anti-El-Niño/positive North Pacific s.s.t. anomaly experiments.

6. Discussion

We have found from a small sample of extended range forecasts, and a much larger set of medium range forecasts (but not for short range forecasts), that forecast skill over the Pacific/North American region is strongly correlated with the signed amplitude of the PNA mode. Using a barotropic model it was found that the growth of initial perturbations was stronger on basic states which had negative PNA index (positive height anomaly over the Aleutian Islands) than on basic states which had positive PNA index. Observational evidence supporting these results was given.
It was suggested therefore that the barotropic stability of the atmosphere is an important factor in determining medium and extended range forecast skill. On the other hand it might be argued that for individual forecasts, baroclinic instabilities might be overwhelmingly more important. Frederiksen (1983) has studied the three-dimensional instability properties of atmospheric flows, including instantaneous flows. He finds that monopole cyclogenesis modes are essentially baroclinic, small-scale blocking-dipole modes (characteristic of the southern hemisphere) are combined baroclinic–barotropic, whilst the final stages of development of large-scale blocking can be described by an essentially barotropic instability. The monopole cyclogenesis modes could certainly have substantial impact on forecast error growth for an individual forecast.

Now, whilst the regression analysis is based on individual forecasts, the modes of variability derived from the regression are, in a sense, relevant to a composite or average forecast, and, as can be seen by the correlation between regressed and actual scores, explain only a small fraction (~20%) of the variance of the actual variability of daily scores in the medium range. In section 2 it was suggested that these modes of variability may, for day-2 forecasts, be described in terms of transient baroclinic modes. However, as discussed in section 2, the day-9 regression analysis has been applied to 5-day-mean forecast fields (i.e. 5 consecutive day-9 forecasts) and the modes of variability derived from the regression analysis are essentially identical to those shown in Fig. 4. With this filter, the fraction of variance explained is up to 100% higher than the regression on daily fields. Whether this increase reflects primarily the suppression of degrees of freedom associated with the transient baroclinic cyclogenesis modes, or whether it reflects the filtering of day-to-day analysis error, remains to be determined. However, the similarity between the factor structure constants using either daily or pentad-mean forecast fields together with the ‘stationarity’ of the day-9 modes (i.e. the fact that the patterns are geographically fixed independent of precisely which region in the PNA area is chosen), suggest that if instability mechanisms are relevant, the relevant modes are likely to be either essentially barotropic or mixed barotropic/baroclinic. The uncertainty of the precise role of baroclinicity, however, indicates that quantitative estimates of error growth from the barotropic model are unreliable.

Another area of uncertainty is the physical ‘source’ of the perturbation whose projection onto the unstable normal modes causes growth in the barotropic model. Clearly analysis error is a candidate. (Perhaps more relevant in this sense would be the short range forecast error associated with the amplification downstream of analysis errors by baroclinic cyclogenesis modes of instability.) However, in addition, one cannot rule out the possibility that systematic model errors associated, for example, with mountain forcing over Asia, or convective forcing over Indonesia, would propagate into the PNA region and preferentially amplify according to the signed amplitude of the PNA mode. Research to try and isolate these effects is currently in progress.

Some discussion was given in section 5 to recent GCM experiments with s.s.t. anomalies of equal amplitude but opposite sign. It was conjectured that the dependence of the barotropic instability of the flow on the signed amplitude of the PNA index may be an important mechanism explaining the asymmetry of the model response to such anomalies. These experiments are relevant to extended range forecasting; s.s.t. anomalies are commonly used as predictors in statistical long range forecasting models. If these results are correct then one can infer that the reliability of such predictors will be a (possibly strong) function of their signed amplitude. For example, Namias (1986) has found that there is a statistically significant correlation between 700 mb wintertime height field anomalies with one-month lag during or just preceding El Niño years. No such effect is observed by Namias for anti-El-Niño years.
Finally it is planned that these modes of forecast variability identified by the regression analysis will form one of a series of predictor sets to be used to attempt a prediction of ECMWF forecast skill for forthcoming winters. A description of the forecast skill prediction scheme will appear elsewhere.

ACKNOWLEDGMENTS

I am grateful to S. Tibaldi for assistance in the calculation of the regression equations in section 2; to A. Simmons for help in the barotropic stability calculations in section 3; to F. Molteni for the help in calculating the composite PNA modes, and their internal variability, in section 4. Also I wish to thank A. Hollingsworth, B. Hoskins, K. Puri and P. Sardeshmukh for helpful discussions, and to G. Branstator, R. Dole, J. Frederiksen, R. Livezey, J. Namias and J. Shukla for valuable correspondence.

REFERENCES


Branstator, G 1985 Analysis of general circulation model sea-surface temperature anomaly simulations using a linear model. Part II: eigenanalysis. *ibid.*, 42, 2242–2254


Palmer, T. N. and Tibaldi, S.  
1986  ‘Forecast skill and predictability’. ECMWF Technical Memorandum No. 127

Pitcher, E. J., Blackmon, M. L., Bates, G. T. and Munoz, S.  

Quiroz, R. S.  

Shukla, J. and Mo, K. C.  
1983  Seasonal and geographical variation of blocking. *ibid.*, **111**, 388–402

Simmons, A. J.  
1987  ‘Barotropic instability, and anomalies of the extratropical northern winter circulation’. In Pontificae Academiae Scientiarum Scripta Varia: *Drought, El Niño and teleconnections*. Published by Pontificia Academia Scientiarum, Vatican City

Simmons, A. J., Wallace, J. M. and Branstator, G. W.  

Wallace, J. M. and Gutzler, D. S.  

Wallace, J. M., Tibaldi, S. and Simmons, A. J.  