RELATIONSHIPS BETWEEN INITIAL CIRCULATION ANOMALIES AND FORECAST ERRORS

by

Edward A. O'Lenic and Robert E. Livezey
Climate Analysis Center
NOAA/NWS/NMC
Washington, D. C. 20233

1. INTRODUCTION

The range of operational numerical forecasts now stands at about 5 days, although forecasts of from 6 to 10 days duration, or a forecast consisting of an average of these, are used as inputs to the medium range forecast and the 30-day Outlook produced at NMC. Efforts to push numerical forecast range even further have recently led to a resurgence of work on the problem of atmospheric predictability and the related problem of prediction of forecast skill.

Lorenz (1963) concluded that because the equations governing atmospheric flows exhibit a sensitive dependence upon initial conditions, it was unlikely that routine prediction beyond 2-3 weeks range would be feasible. Lorenz (1969a) further noted that in flows containing many scales of motion, each scale possesses its own intrinsic finite range of predictability, with smallest scale (high frequency) motions having the shortest range and largest scale (low frequency) motions having the longest range. Statistics describing the variation of real forecast skill clearly show that forecast accuracy is relatively high during winter, when the large scale quasi-stationary waves have their largest amplitude, and relatively low during summer, when smaller scale flows tend to prevail. However, notable episodes of low wintertime forecast skill and high summertime forecast skill do appear. Among recent studies designed to examine the variation of forecast skill is one by Palmer (1988). He found that the skill of medium range numerical forecasts during the Northern Hemisphere winter is relatively high when the forecasts contain a high-

amplitude version of the positive phase of the Pacific North American (PNA) upper air height anomaly pattern (Wallace and Gutzler [1981], Barnston and Livezey [1987] and Esbensen [1984]) and relatively low when the forecast flow is dominated by the pattern with opposite (negative) phase. Palmer (1988) and Palmer and Tibaldi (1987) respectively, also investigated the relationships between the low frequency component of the forecast model flow and the skill of medium range forecasts. They performed a linear regression of forecast skill against a set of either 9 or 23 EOF coefficients of the forecast 500 mb height field. They found that the year-to-year correlations between 9-day forecast error and 9-day persistence error were high, and that in general terms, forecast skill was also correlated with initial flow conditions. They also concluded that year-to-year differences in persistence error were much more closely related to variations in model error than were differences in model formulation. Barker and Horel (1988) examined the persistence of planetary scale circulation for 18 Northern Hemisphere winters and its relationship to SST anomalies in the equatorial Pacific. He compared the observed persistence of the 500 mb circulation with that simulated by several numerical models, and found that the more persistent winters tended to be those during which the positive phase of the PNA pattern was established.

The purpose of this paper is to demonstrate the feasibility of objectively specifying a priori the non-random or systematic error of daily medium range numerical forecasts based upon the existence of low-frequency upper air circulation anomalies in the initial conditions. We define the regimes for a given (2 month) season by performing a rotated principal component analysis (RPCA) upon a 37-year time series of low-pass filtered daily 700 mb height analyses. The RPCA-defined anomaly patterns correspond to quasi-stationary structures with relatively long lifetimes. A regime dependent climatology of forecast errors is created by compositing errors

in forecasts whose initial conditions contain high amplitude versions of each of the various RPCA-defined regimes. Finally, a number of techniques are used to determine which of the composite forecast error maps formed in this process have either unusually strong or unusually weak error patterns. Monte Carlo simulations are relied on most heavily in this regard. The results clearly show that for several modes during several seasons, including the PNA and the Northern Asian (NA) pattern, as described by Barnston and Livezey (1987), hereafter referred to as BL, during winter, and the subtropical-zonal (SZ) pattern discussed by BL during summer, the level of systematic forecast error can be predicted with a moderate degree of confidence. Forecast skill is found to be related not only to modes explaining relatively large fractions of the hemispheric variance, but also to modes whose influence is highly regionalized. The spatial signatures of the composite forecast errors in cases where the error amplitude is large tend to have a larger spatial scale than the parent pattern with error centers far removed from the original pattern in many cases.

The technique used here is different from that of either Palmer (1988) or Barker and Horel (1988) in that it permits us to examine the sensitivity of the model to those modes which the atmosphere preferentially produces during a given season. Branstator (1987), Richman (1986) and Horel (1981) and others have noted the particular advantage of RPCA in situations such as these.

This study must be considered to be primarily a feasibility study for two reasons. First, only a relatively small number of years of daily MRF data (1982-88) is available, consequently all modal circulation regimes will not be well sampled. Second, the MRF model underwent several major modifications during 1982-1988 which are likely to have some effect upon the results. These changes and possible effects upon the systematic error of the model will be discussed in greater detail later. The procedure we

describe here would undoubtedly prove more effective when applied to a more representative (but not necessarily larger) and more homogeneous (in terms of model climate) data set.

In Section 2 the data and the analysis techniques are discussed. Section 3 contains a discussion of the results, while in Section 4 conclusions and recommendations are presented.

2. THE DATA AND THE ANALYSIS TECHNIQUE

2.1 Data

The data used in this work come from two sources. The first set was derived from twice-daily analyses of Northern Hemisphere 700 mb height spanning 1950 to 1987 from the National Meteorological Center (NMC). The data were originally given in the form of 541 point gridded fields between $15^{o}N$ and $90^{o}N$. Missing maps were few, and were dealt with by temporal interpolation using data from adjacent days. Maps from 00Z and 12Z for a given day were then averaged to give a time resolution of one map per day. The data were then transformed to a 358 point grid containing approximately equal areas between grid points, regardless of latitude. This step avoids the problems which arise when RPCA is performed on irregularly spaced data (Karl, 1982). Data points at $10^{o}N$, $15^{o}N$, and at certain points at $20^{o}N$ between 20^{o} E and $90^{o}E$, and over the Himalayas are discarded on the equal area grid due to either missing data early in the record or fictitious data in the Himalayan region. BL give a detailed description of the construction of this grid. The data were then subjected to a low-pass filter (Blackmon, 1976) with an amplitude response of 0.9 and 0.2 for signals with periods of 20 and 11.5 days, respectively. This step effectively removes signals with periods of 10 days or less. Finally, the annual cycle is removed by computing the long term mean, day-by-day, at each grid point and subtracting it from the data. This step prevents the annual march of the daily mean, which is not of interest in this study,

from appearing as one of the leading modes of the RPCA analysis.

The second data set used in this work is called the MRF data set. It consists of a 6 to 7-year time series of daily analyses and forecasts in spectral form from NMC's medium range forecast (MRF) system from 1982 to 87 or 1988. Daily forecasts of from one to ten days' lead were available. All forecasts were made from ØØ GMT initial data. Although no explicit temporal or spatial averaging was performed on the MRF data, the fact that a rhomboidal truncation of only the first 12 spherical harmonics (R12) was stored results in some filtering of small-scale features from the data but it is difficult by eye to see any substantive differences between truncated and fully resolved maps.

It should be noted that the MRF data used in this study comes from an evolving forecast system. The current configuration of the MRF contains sophisticated physics and radiation software, increased spatial and temporal resolution and envelope topography, whereas the earliest version contained almost no physical parameterization, and was not even global for the last 5 days of the forecast. Such changes in model formulation are likely to cause changes in the systematic errors of the model with time. Palmer and Tibaldi (1987) however, noted that year-to-year differences in ECMWF model error were more closely related to variations in persistence error than they were to differences in model formulation. In fact, the model bias contains not only effects of changes to the model but also any tendency for a given low-frequency mode or modes to appear preferentially during the period shown. The contribution of the latter factor to the biases is very difficult to assess, but it is reasonable to assume that that effect would be more pronounced over shorter time periods than for longer ones. Thus, the mean error for 1982-1988 probably represents a reasonable approximation to the "true" model bias, since the modes, whose effect upon the model we wish to assess, will have more uniformly sized

samples over that period than over subperiods. The bias appropriate to each season and forecast duration was removed from each of the daily forecast error maps prior to compositing.

2.2. Rotated principal component analysis

The method of RPCA used in this study involves transforming a set of data time series, one for each grid point of the 358 point grid, into a set of principal components (PCs). Each PC is also a time series, and the entire set corresponds precisely in temporal and spatial distribution to the original set of data time series. Prior to transformation, the annual cycle is removed from the data time series by subtracting daily means and dividing by each day's standard deviation, resulting in standard variates. The transformation is obtained by finding the eigenvalues and eigenvectors of the cross-correlation matrix of the standardized data time series. As noted by BL, use of the correlation matrix ensures that the systematic, but low amplitude variance in lower latitudes is included in the analysis. Before rotation, the eigenvalue corresponding to each eigenvector reflects the fraction of the original variance it explains. Because each unrotated PC is constrained to explain the maximum amount of variance of the input data (excluding that of any PCs which explain more variance) over the entire domain, an eigenvector may not accurately correspond to a single cluster of points in hyperspace (Richman, 1986, North et al., 1982). The corresponding spatial pattern may then represent some mixed combination of two physically distinct patterns, each of which has a more regional character. Linearly transforming (rotating) (Richman, 1986) the eigenvectors so that each one corresponds to a single set of spatial points that vary coherently in time, reduces the possibility of mixing, producing a set of patterns which are more regionalized, more physically meaningful and robust (Horel, 1981; Richman, 1986; Richman and Lamb, 1985). Choosing the optimum number of PCs to rotate is an important consideration because

rotation of fewer than the total possible number of eigenmodes removes the variance explained by the modes which are discarded. Richman (1986), and O'Lenic and Livezey (1988), have shown that rotating either too few (discarding useful signal) or too many (retaining noise) PCs can result in serious distortion of the resulting spatial patterns. O'Lenic and Livezey (1988) found that a combination of objective criteria and trial and error can be used to choose a number of PCs to rotate which represents an acceptable compromise between over- and under-rotation. Using this technique, we chose to rotate 14 modes for January-February, 13 modes for July-August (JA), and October-November (ON) and 15 modes for April -May (AM).

In addition to partitioning the spatial structures in a data set which vary coherently in time into a set of spatial patterns (eigenvectors), RPCA also gives a detailed record of the time variation of each mode. The eigenvectors can be used as a set of basis functions onto which atmospheric data can be projected for diagnostic purposes. The time series of amplitudes of the various modes can be used as indices, or for compositing studies, as was done for example in Dole (1982), Branstator (1987), and Barker and Horel (1988).

2.3 Compositing techniques

A set of regime-dependent 5-, 7- and 10-day forecast errors was computed for 2 month seasons from 1982-88, (for January-February) or 1982-87 (other seasons) when the latter years of the RPCA time series for any particular mode overlapped the time when MRF data were available. Compositing was done objectively, as shown in Figure 1, by separately testing the amplitude time series for times when the amplitude of the regime in question (here, the PNA for JF) in the MRF initial conditions either exceeded (positive phase) or was less than (negative phase) the threshold criterion for compositing of the corresponding forecast errors. Compositing was not done

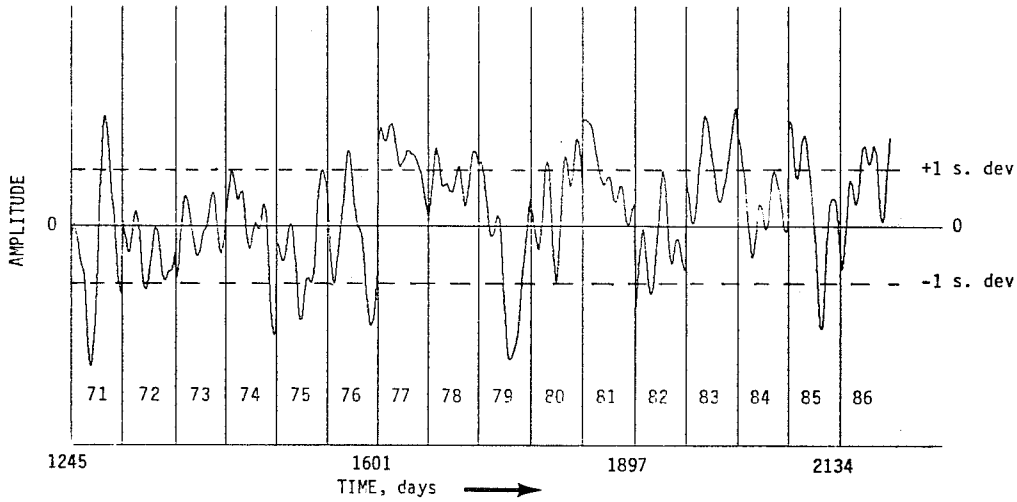


Fig. 1. Schematic illustrating compositing process using model amplitude time series. Vertical lines distinguish years, each spanning two adjacent months of days. Upper and lower dashed horizontal lines indicate the +/- 1 standard deviation of the time series.

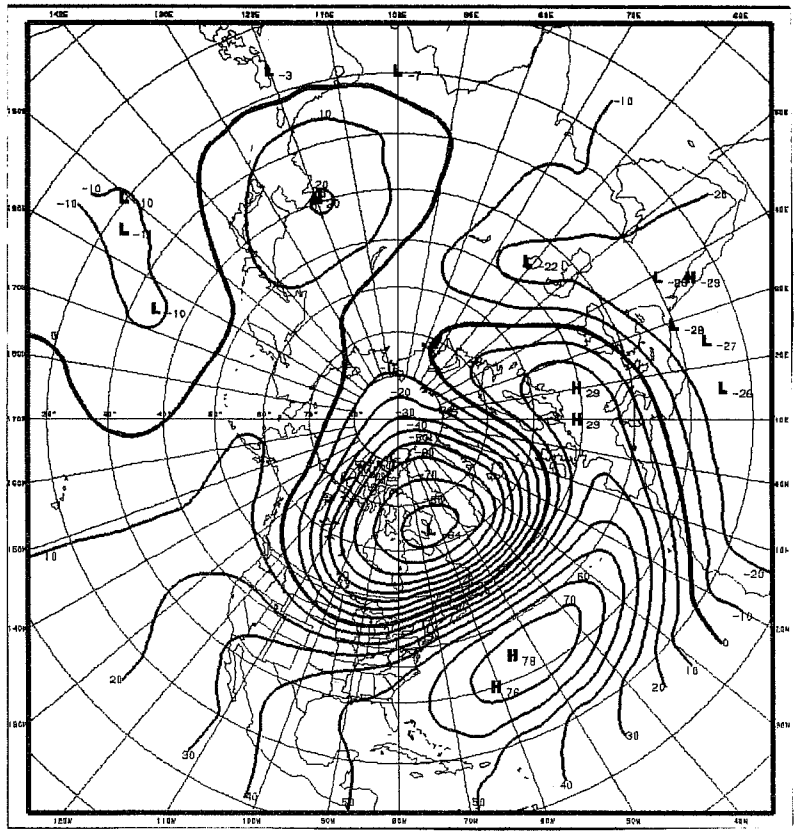


Fig. 2. Loading pattern of JF mode 1 (NAO). Contours are correlation coefficients scaled by 100. Contour interval is 15.

if a forecast error map was missing. Experimentation with various criteria led to the choice of +/- one standard deviation of the amplitude of the amplitude time series (ATS) for each mode as the compositing threshold for forecast errors related to that mode. This technique tends to sample (we hope) the tails of distributions made up of many forecast error maps. If it does, then it is likely that non-random signals in the error fields will be apparent in separate composites of the tails. Thus, some of the modes should lead to composite error maps that exhibit unusually large amplitude with respect to randomly constructed composites. In other situations where certain circulation patterns are not tied to systematic errors, composite error maps should exhibit unusually small amplitudes. This hypothesis will be tested by Monte Carlo simulations. Another characteristic of the response of the MRF model to high amplitude versions of a given mode that can be examined is its degree of linearity, that is, the extent to which the amplitude and phase of the composite errors for one sign (phase + or -) of a mode in the initial conditions are equal and opposite to those of the composite corresponding to the existence of the opposite sign. Linearity can be examined by simply subjectively comparing the composite maps for opposite signs of a given mode. Another method is to correlate these maps. A large negative correlation (at least -Ø.3 or less) would suggest a linear response. The occurrence of reasonably well-marked linearity would constitute an independent validation of the significance of the error signal, since such a correlation is unlikely to occur by chance. Both of these techniques were used in this study.

Finally, to get some idea of how much of the total error variance that unusually large regime-dependent error composites represent, for each composite we computed the ratio of the total squared error of the composite map to the average squared error of the maps in the composite. Because this statistic is for the entire hemisphere and includes day-to-day error

variability it will considerably understate both the local impact of a large error signature and the practical utility of the pattern for improving local and/or time averaged forecasts. Thus, we would expect maps with statistically strong error signatures and percents of total error variance on the order of ten or more to be of substantial value for correction of such forecasts.

In this context we also examined whether the average error variance (regime-dependent or otherwise) in each composite was particularly large or small by simply dividing it by the average error variance for the entire respective study period. These ratios were by and large well within sampling variability of unity, implying that cases of unusually small regime-dependent systematic error do not necessarily correspond to unusually small total error.

2.4. Monte Carlo simulations

The composite forecast error maps were tested to determine whether their amplitude patterns were either unusually strong or weak via Monte Carlo simulation. The map parameter tested for each chart is the total squared standardized error. Grid values on each composite map were divided by the corresponding grid values of the standard deviation of the forecast errors for the period of MRF data under study, squared, and then integrated to a single value. The standardization removes geographical biases (i.e., middle latitude centers of action and tropics versus middle latitudes) from the forecast error maps. The object of the Monte Carlo runs was to produce a large number of composites with the same number of maps randomly selected as the composite being tested. The random selection had to be done in a way that mimicked the actual selection process. In other words, the randomized series that are used to guide map selection must have the same properties, i.e., persistence and transition behavior, as the original ATS. First the original (observed) time series of eigenvector amplitudes (ATS)

for a given mode is converted to a time series with elements having only three possible values, a_1=1, a_2=Ø, or a_3=-1 based on whether the original time series element b had values $b \geq +1$, $-1 < b < +1$, or $b \leq -1$, respectively in units of the standard deviation of the original ATS. There are a total of nine possible transitions among the three a_i values corresponding to all permutations of these states. The fraction of the time each of the nine possible transitions occurs is determined by counting. This gives a set of transition statistics T_{ij} (i=1,3; j=1,3). Then for each of the three possible final states, a_j, the conditional probability of the occurrence of a transition given a particular initial state a_i is then given by Pr $(a_j|a_i)$.

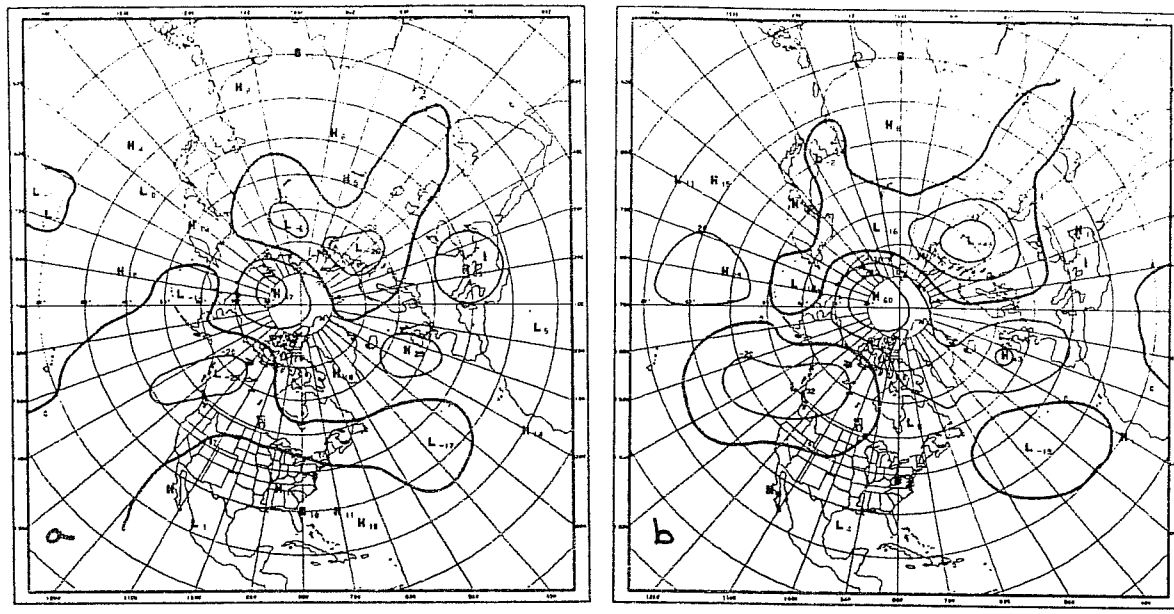
The quantities T_{ij} and Pr $(a_j|a_i)$ were determined for each mode during a given season for the particular subperiod being studied (1982-88 for January-February, and 1982-87 for other seasons). The Pr $(a_j|a_i)$ were then used to create random three-phase time series. Because of the way they were created, the randomized time series have the same statistical properties as the originals. A set of composite forecast error maps corresponding to each mode was then created by the technique described earlier but using the randomized ATS to govern the selection of days. This was done by starting the random selection of cases at randomly selected points over the study period, and then marching through the study period as often as necessary with as long a randomized an ATS as necessary to select the required number of cases for the composite. A random distribution of these integrated squared errors from 5ØØ composites was developed for each phase (+ or -) of each mode for each season. An appropriate measure then for whether or not an observed composite has an unusually strong or unusually weak error pattern will be the percent of the composites from the randomly generated distribution that have a larger or smaller integrated squared error, respectively.

3. RESULTS

Results of compositing and the significance tests are now presented for January-February (JF). Results for winter (JF) use RPCA for 1950-1988 and MRF data from 1982-88. The results for the other seasons use RPCA and MRF data ending with 1987. The use of an additional year of data for the JF analysis is justified since it increases the size of the forecast error map sample, and does not result in large changes between the 1982-87 and 1982-88 results. Results for seasons other than winter will not be explicitly presented here, but will be summarized in the conclusion.

3.1 Results for January-February

The leading mode (mode 1) in the JF RPCA, explaining 9.3 percent of the hemispheric variance of the input data set, was the North Atlantic Oscillation (NAO), whose loading pattern (positive phase, NAO+), is shown in Figure 3. This version of the NAO (and other modes) differs somewhat from that of BL due to differences in filtering, time resolution, and size of season between our analysis and theirs. The evolution of the corresponding composite forecast errors from day 5 to day 10 of the MRF model forecast are shown in Figs. 3 and 4, while map significance deduced from the Monte Carlo run are presented in Table 1 for all of the mode-linked JF composite maps. Also listed are correlations between error maps for both phases of each mode and percents of total composite error variance represented by each composite error map. Errors corresponding to NAO+ (Figures 3a-c) grow in place with little or no change in phase from day 5 to day 7. While the integrated squared error in each of these maps is relatively large, none of them exceeds the 20 percent significance level. The evolution of the NAO- composite error field is quite different from that of NAO+. It exhibits distinct phase changes and an increase in spatial scale from day 5 to day 10 (Figure 4a-c). According to Table 1 the amplitude of the error map goes from being unusually small, at the 0.6 percent level of significance at day 5 and the 6.2 percent level at day 7,



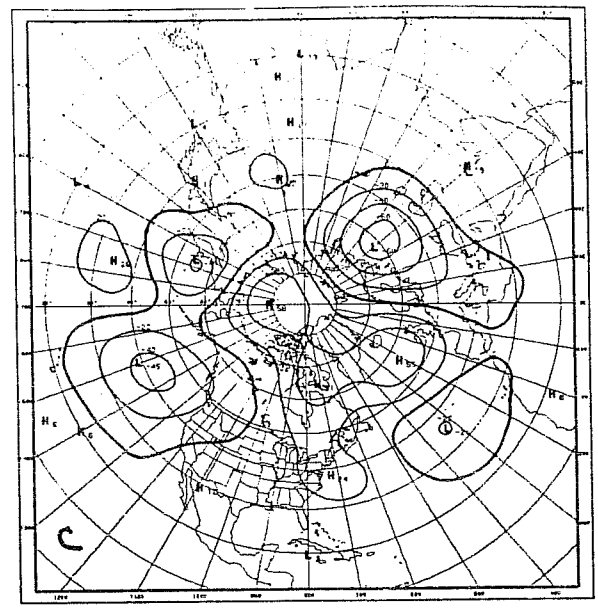
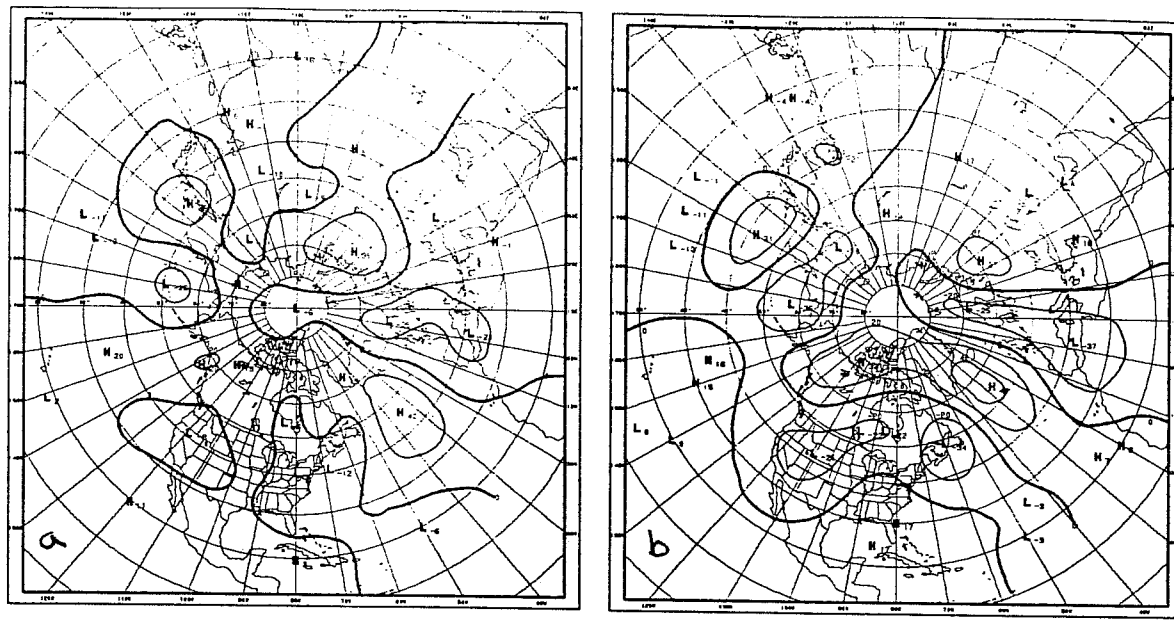


Fig. 3. a) Composite day-5 700 mb height forecast error for mode 1+ (NAO+), JF 1982-88. Contours are in meters, interval is 20.
b) Same as 3a but for day-7.
c) Same as 3a but for day-10.



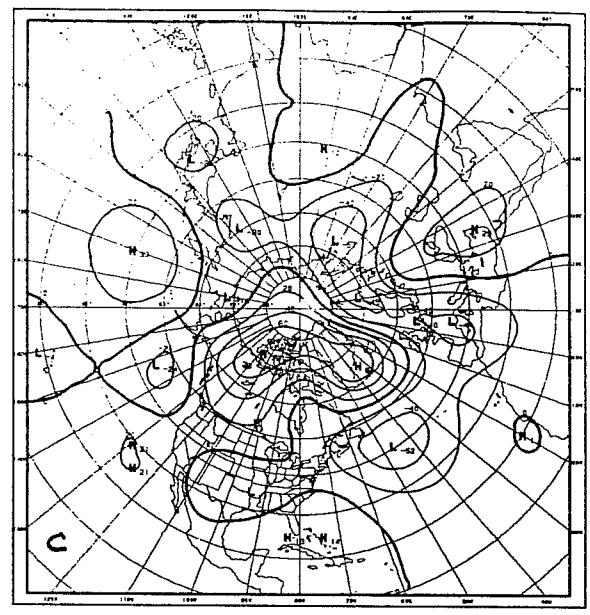


Fig. 4. a) Composite day-5 700 mb height forecast error for mode 1-NAO-), JF 1982-88. Contours are in meters, interval is 20.
b) Same as 4a but for day-7.
c) Same as 4a but for day-10.

to just average by day 10. This marked change during the last three days of the forecast is due mainly to growth of the error center over Canada and Greenland. The correlation coefficient between + and - NAO phase error maps is listed in the bottom row of Table 1. The +0.55 value for the NAO indicates a strong non-linear response, the only pronounced non-linear response for any of the 14 JF modes.

The second mode in order of explained variance for JF was the PNA, whose loading pattern appears in Figure 5. Forecast errors for PNA+ grow in amplitude very little from day 5 to day 10, (Fig. 6) so that by the latter time the error is relatively small. This result is in agreement with PT and Horel (1987). The fact that it confirms an a priori expectation means that we may loosen our definition of significance. Thus, our result, significant at the 13 percent level (Table 1) is at least suggestively strong, if not highly significant.

The error field for PNA indicates that the field also migrates toward a tail of its error map distribution from day 5 to day 10, Fig. 7) but in the opposite (large error) direction. When we compare Figure 6 with Figure 7 we observe a tendency for the error patterns corresponding to opposite phases of the mode to have opposite phase. From Table 2 it can be seen that the PNA + and - error maps are negatively correlated (-0.40) indicating some linearity in phase if not in amplitude. Results of similar calculations for the other JF modes (Table 1) are dominated by negative values, suggesting that regime dependent error response has an important linear component. In addition to that for the PNA, error maps for three other JF modes have correlations less than or equal to -0.4, four others are smaller than -0.3, and five have correlations in the noise level. The only mode with a decisively non-linear, non-random error response is the NAO.

The loading pattern for mode 6 (Fig. 8) consists of a north-south

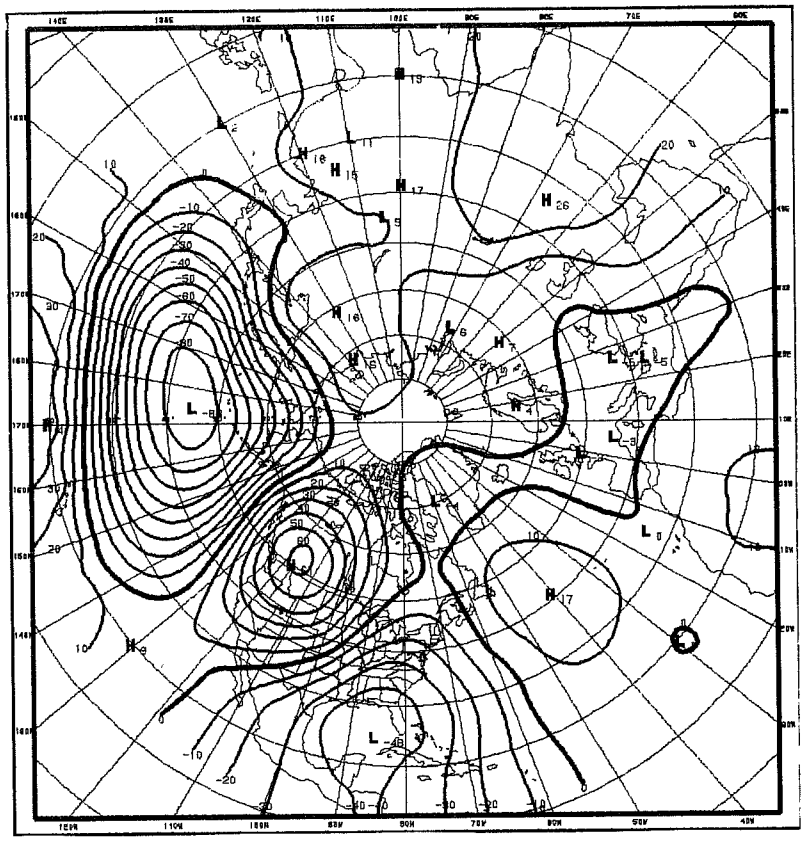


Fig. 5. Loading pattern of JF mode 2 (PNA). Contours are correlation coefficients scaled by 100. Contour interval is 15.

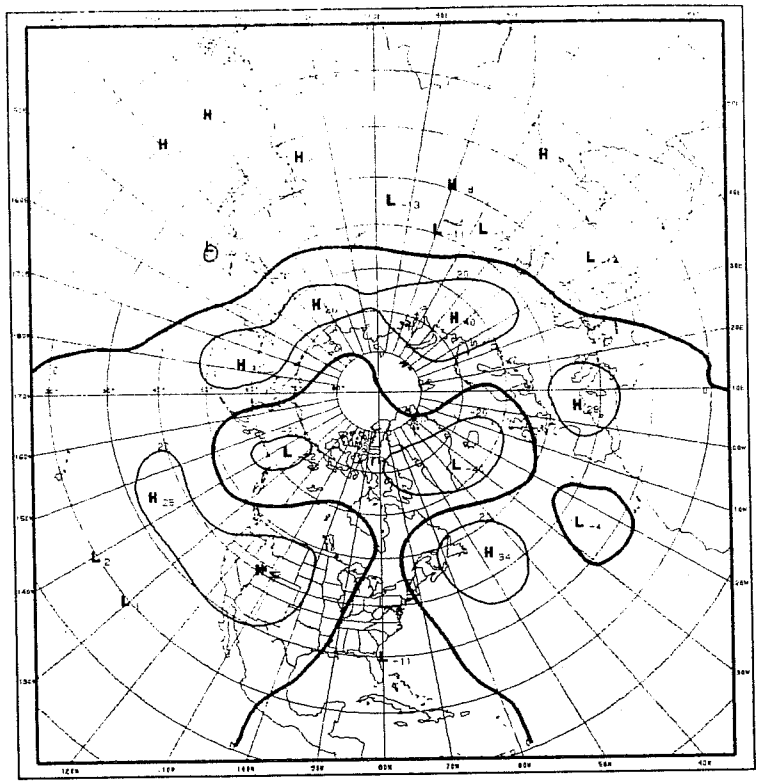


Fig. 6. Composite day-10 700 mb height forecast error for mode 2+ (PNA+), JF 1982-88. Contours are in meters, interval is 20.

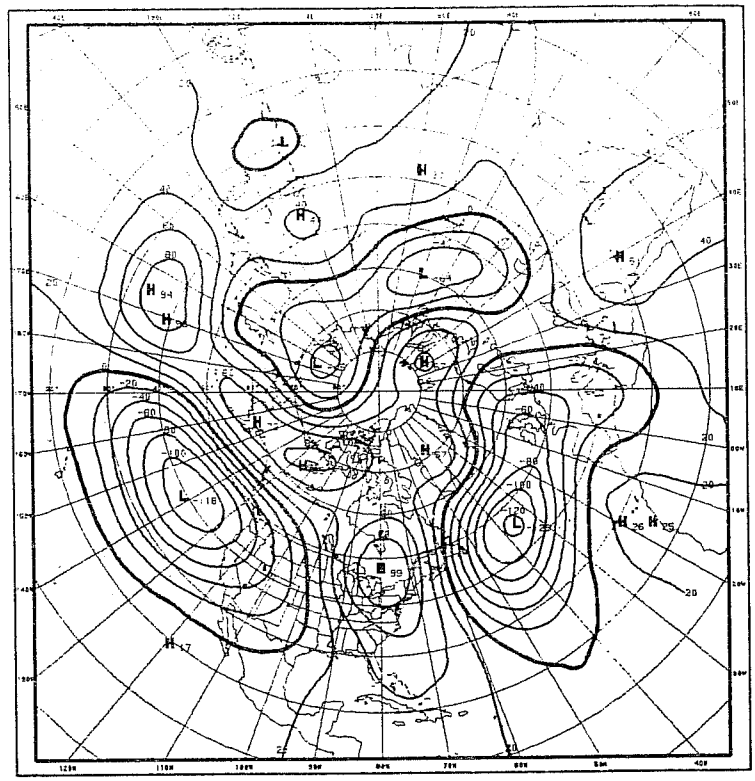


Fig. 7. Composite day-10 700 mb height forecast error for mode 2- (PNA-), JF 1982-88. Contours are in meters, interval is 20.

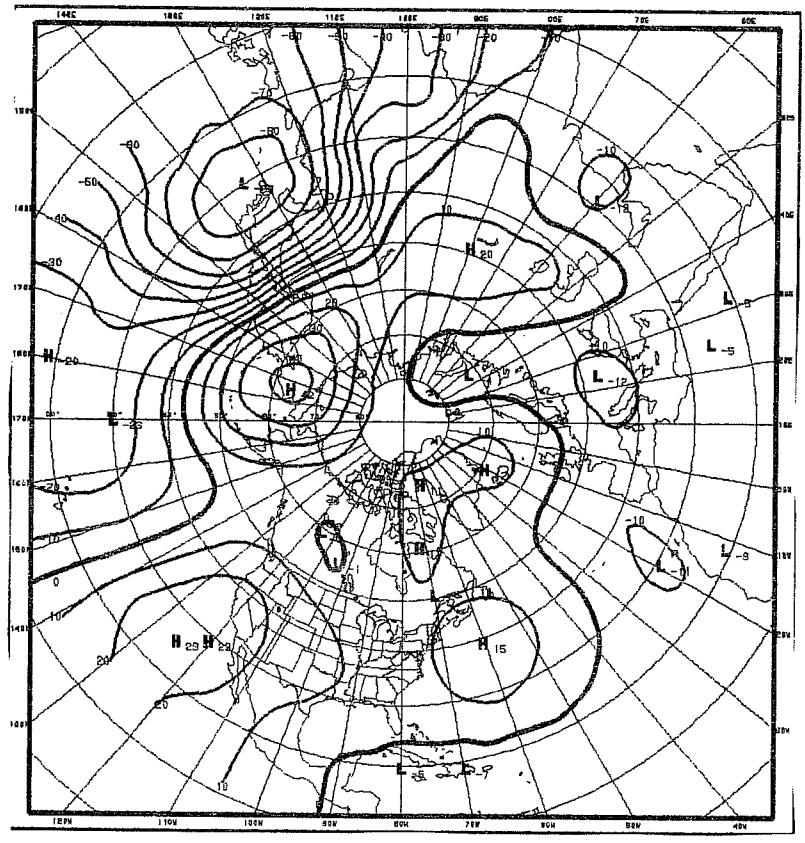


Fig. 8. Loading pattern for JF mode 6 (WP). Contours are correlation coefficients scaled by 100. Contour interval is 15.

TABLE 1: Significance level (%) of composite forecast error maps as function of mode, forecast length, and phase of the mode for January-February 1982-87. Value listed is the percentage of maps in each random distribution whose rms error score was greater than that of the map being tested. Levels at or below 10% and at or above 90% and correlations greater than or equal in absolute value than 0.30 are entered in boldface.

JANUARY-FEBRUARY							+ PHASE								
Mode	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
# maps per composite	72	101	70	67	48	50	39	28	68	26	61	63	77	97	
F 5	29.2	36.4	24.2	36.0	**90.6**	87.0	**93.4**	69.7	85.0	**90.0**	74.0	13.0	77.6	16.4	
F 7	27.4	56.8	42.4	52.8	74.4	54.6	**96.6**	41.6	71.8	46.8	89.6	**8.6**	79.2	29.8	
F 10	26.0	87.0	43.5	46.5	77.5	39.0	84.0	44.5	46.5	68.0	**97.0**	14.0	69.5	35.0	
% non-random variance *	6	2	2	3	5	7	5	15	3	7	2	9	3	4	
							- PHASE								
Mode	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
# maps per composite	41	23	56	61	56	73	66	54	94	57	99	38	96	63	
F 5	**99.4**	24.0	25.0	38.8	42.6	44.2	**92.0**	23.2	69.4	75.2	60.2	24.6	41.6	59.6	
F 7	**93.8**	20.8	32.4	49.2	26.2	58.2	**93.2**	18.6	89.8	62.8	53.8	26.6	50.2	34.0	
F 10	50.5	**8.5**	39.0	55.5	57.5	76.5	78.0	37.5	88.0	69.5	39.0	23.5	31.5	45.5	
% non-random variance *	8	18	3	5	4	3	3	4	2	4	3	12	3	6	
Correlation Coefficient + versus -	+.55	**-.40**	-.14	**-.30**	-.20	**-.63**	**-.48**	-.06	-.16	**-.42**	+0.08	**-.34**	**-.33**	-.35	

* = normalized percentage of the total composite variance that the composite map explains.

oriented dipole or seesaw pattern centered mainly over eastern portions of Asia and much of the western Pacific Ocean. This is the West Pacific (WP) pattern (Wallace and Gutzler, 1981). The systematic error patterns corresponding to the presence of either phase of this mode in the MRF model initial conditions tend to evolve with little change of phase from day 5 out to day 10 (Figs. 9a, b). These maps bear a strong inverse relationship to one another as indicated by the -0.61 correlation coefficient between them (Table 2). Comparison of Figs. 9a and b indicates that the model response indeed has a linear component in phase, but not in amplitude over much of the hemisphere. This is especially true over western Asia and Scandanavia, the eastern Pacific Ocean and Northern Canada, but is not true over the Kamchatka Peninsula, where a nonlinear response can be seen.

The loading pattern for mode 7, the Northern Asian pattern (NA), is shown in Figure 10. The amplitude of the composite error fields for both phases of this mode are relatively small at all forecast times, reach highest significance at day 7, and degrade somewhat from day 7 to day 10 (Table 1). This tendency for a large change in the amplitude of the integrated error of the composite error field from day 7 to day 10 was also observed for modes 1 (Figs. 3 and 4) and 2. According to Table 1, the model systematic error response to this mode is highly linear, particularly in the vicinity and upstream from the main modal center.

Mode 11 (Figure 11) explains about 4.3 percent of the hemispheric variance of the input map time series. The composite error field for the positive phase has relatively small amplitude at day 5 which apparently grows at a relatively slow rate, since its error falls well out in the (unusually small) tail of the random distribution by day 10. The PNA+ composite error field exhibited similar behavior (Figure 7). The negative phase composite error field progresses toward the opposite (large error)

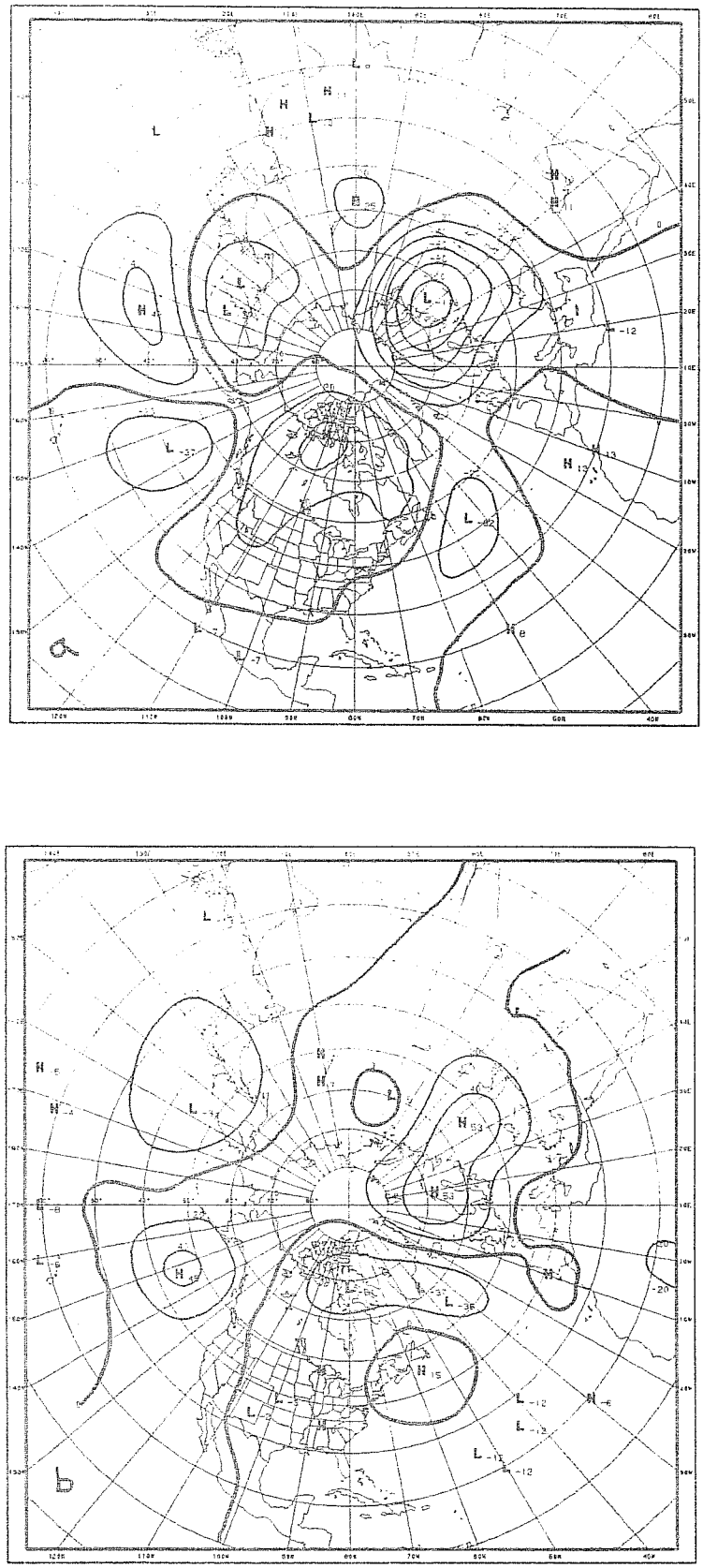


Fig. 9. a) Composite day-10 700 mb height forecast error for mode 6+, JF 1982-88. Contours are in meters, interval is 20.
b) Same as 9a but for mode 6-.

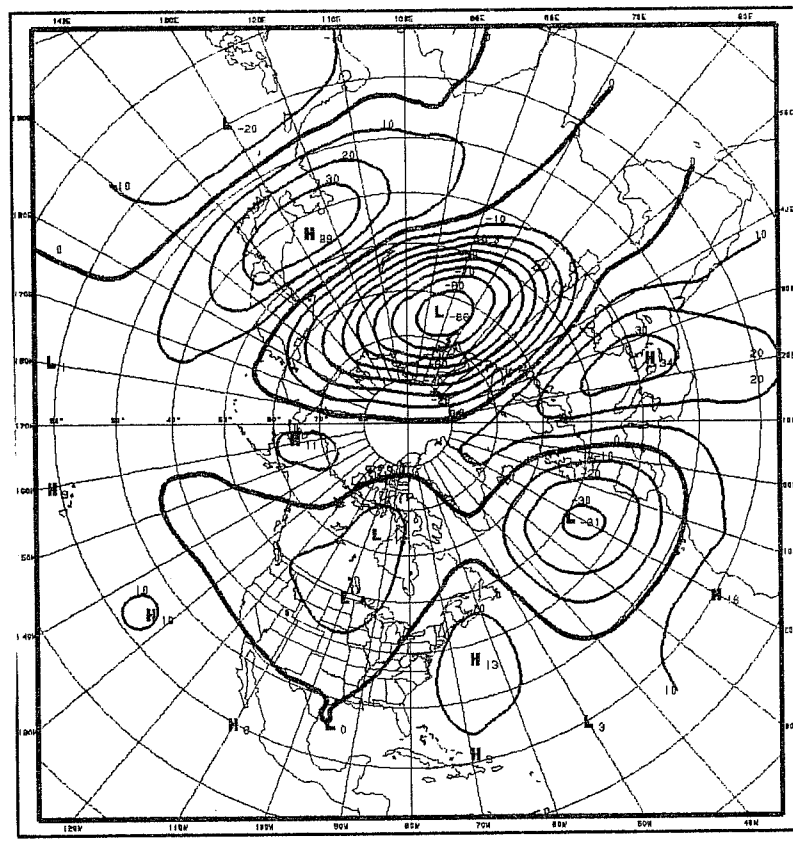


Fig. 10. Loading pattern for JF mode 7 (NA). Contours are correlation coefficients scaled by 100. Contour interval is 15.

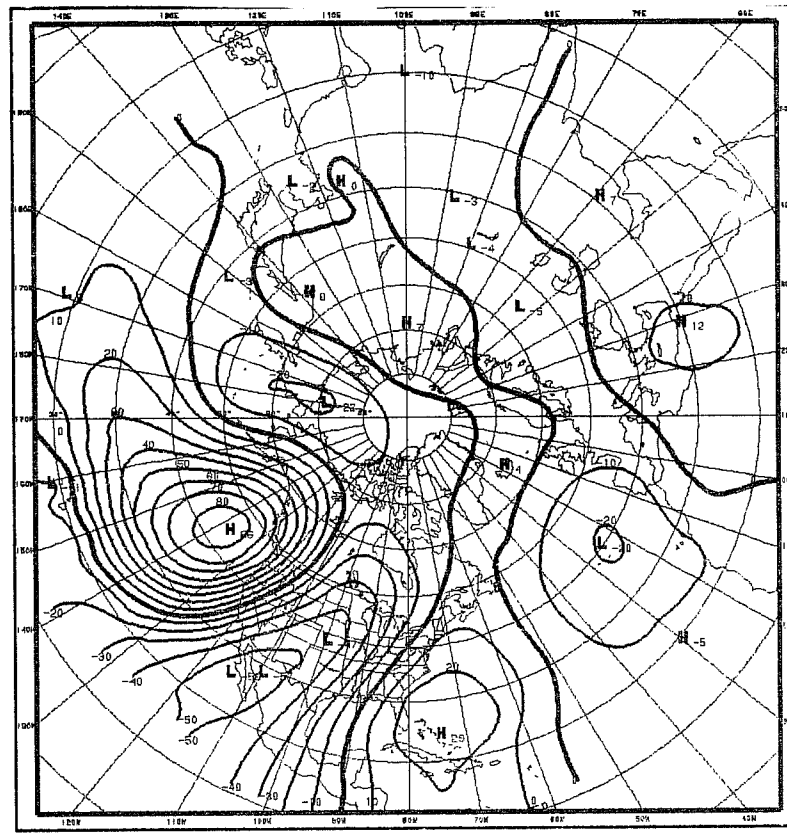


Fig. 11. Loading pattern for JF mode 11 (TNH). Contours are correlation coefficients scaled by 100. Contour interval is 15.

tail of its distribution. Again, this resembles the behavior of the PNA-error field (Figure 8). This was one of the few modes for which little can be said about linearity of the model response.

The + phase for Mode 12 for JF is shown in Fig. 12. According to Table 2, this phase of the mode tends to be related to unusually large forecast errors. This mode corresponds to a blocking pattern over the U.S. with anomalously low 700 mb heights over the West Coast and the Rockies, and anomalously high heights centered over the Great Lakes and New England. The composite forecast error pattern corresponding to this mode's existence in the initial conditions tends to be established by the fifth day of the forecast and to grow in place thereafter. Figure 13 shows that much of the F10 mode 12+ composite error is concentrated over the northern USSR and Scandanavia, but that broad scale centers of moderate amplitude also exist over much of the hemisphere. Table 1 indicates that the F10 error patterns for this mode are negatively correlated to a moderate degree, indicating some degree of linearity in the model response.

The values given in Table 1 reflect the relative position of each composite map in its randomly generated distribution. The data suggest that the positive phase of modes 2, 5, 7, 10, 11, and 12 and the negative phase of modes 1, 2, and 7 are significantly related (at the 10% level) to forecast errors. Furthermore, many modes related to significant composite maps at some time during the 10-day forecast period are also linked to F10 composites which bear a strong inverse relationship with their opposite phase counterpart. Those exhibiting the latter are modes 2, 6, 7, and 10, whose + and - composite are strongly negatively correlated with correlation coefficients of -0.40 or less, and modes 4, 12, 13, and 14 with correlation coefficients of -0.30 or less. It is unlikely that so marked a linear response could occur in so many cases by chance.

Table 1 also includes the percentage of the total error variance of

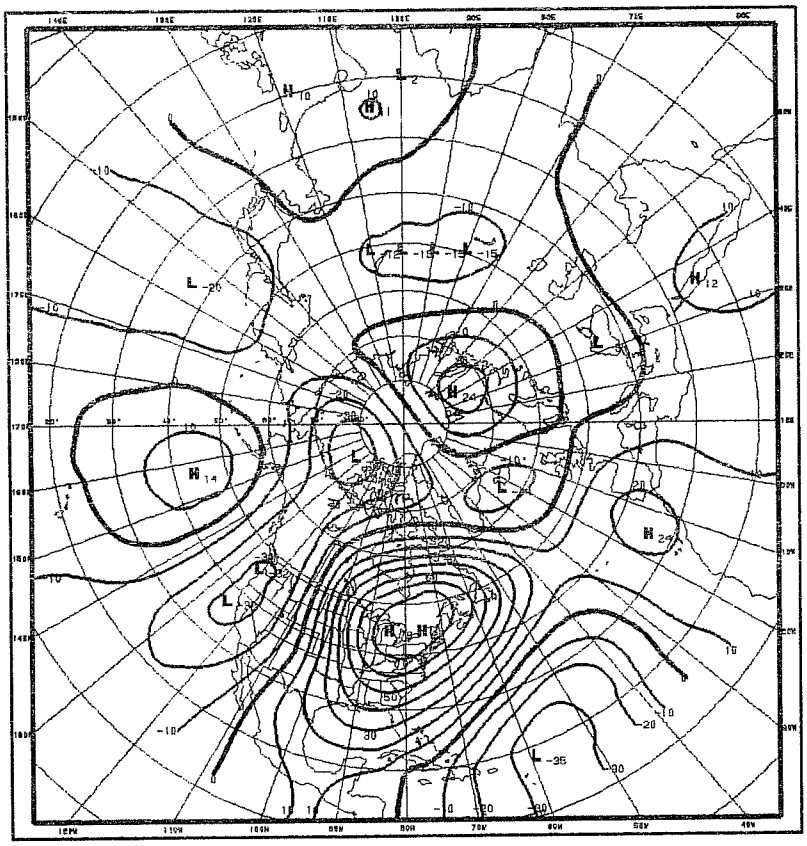


Fig. 12. Loading pattern for JF mode 12 (TNH). Contours are correlation coefficients scaled by 100. Contour interval is 15.

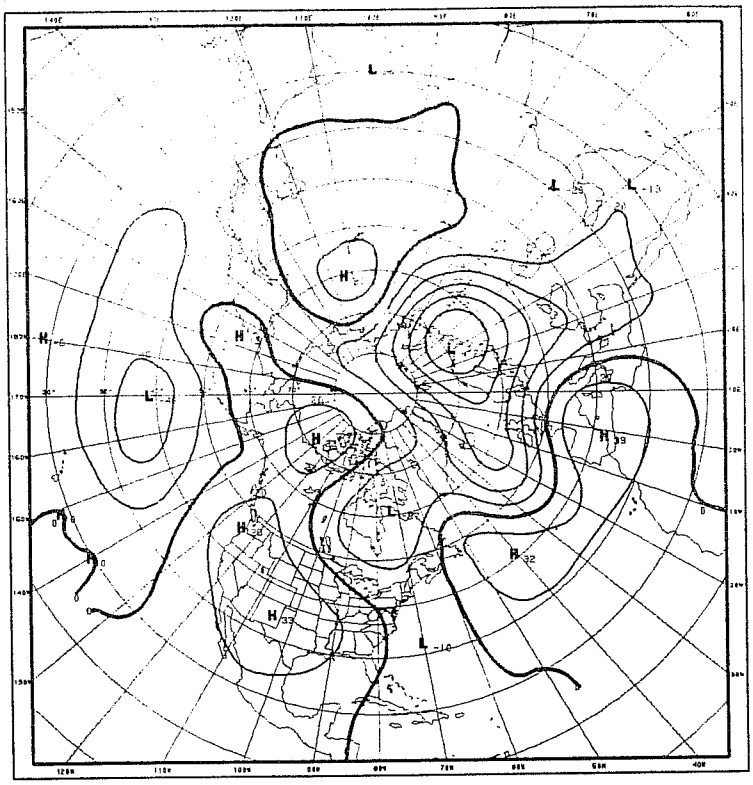


Fig. 13. Composite day-10 700 mb height forecast error for mode 12+, JF 1982-88. Contours are in meters, interval is 20.

a given composite that the variance of each F1Ø composite error map itself represents. This is a measure of the extent to which the variance contained in the composite map is representative of the overall error variance of the maps comprising the composite (i.e., the extent to which it is non-random). In this regard composite maps corresponding to modes 2-, 8+, possibly 12+, and 12- contain error variance that is a considerable part of the total error for their respective composite cases. More importantly, the error maps explain much more variance locally than the statistic in the table reflects.

Of 28 phases (2 phases per mode, 14 modes), only those for modes 2- and 12+ correspond to unusually large errors, whereas six fall in the lower (low error) 1Ø percent of their respective distribution. Apparently, during JF the largest part of the systematic error (which survived removal of the long-term bias) is concentrated in just a few modes, while it is practically absent for a greater number of modes, and neither exceptionally large nor small for most of them. In addition to problems resulting from small sample sizes, it is possible that some mode-error relationships are clouded by other sources of systematic error not related to quasi-stationary modes in the initial conditions, for example, the effects of model changes during 1982-88. If that is the case, systematic mode-error relationships will stand out much more clearly when a homogeneous set of model forecast error maps is used as input to this analysis technique.

4. CONCLUSION

The dependence of the systematic error of the NMC MRF model during 1982-87/88 upon the existence of certain seasonally dependent low frequency 7ØØ mb height anomalies in the initial conditions has been investigated. The analysis technique involves developing a set of basis functions using RPCA for four two-month seasons. The leading dozen or so RPCA modes correspond to the naturally occurring geographically regionalized modes of oscillation

which the atmosphere produces preferentially. Error maps of 5-, 7- and 10-day forecasts corresponding to the existence of large amplitude versions of each RPCA mode in the MRF initial conditions were produced by compositing, using the time series of the amplitudes of each mode as a guide. Each such map was then subjected to significance testing by Monte Carlo trials.

Despite large changes in the formulation of the MRF model during 1982-88, the technique described in this paper enabled us not only to reproduce the wintertime results of Palmer (1988) and Barker and Horel (1988), for the PNA, but also to identify many more modes, several in each season, whose presence in the MRF initial conditions are related, in many cases strongly, to errors in 5 to 10 day forecasts of 700 mb height. The MRF model response had a clear linear component for most of the modal patterns in all of the seasons. One notable exception in which a decisive non-linear response was found was for the NAO in winter.

In general, the phasing of the major centers of the composite forecast error patterns is established by the fifth day of the forecast and remains stationary thereafter. Like the RPCA patterns, the composite error patterns have a regional emphasis; however, the major centers on the error maps are often far removed from the locations of the anomaly centers of the parent RPCA patterns in the MRF initial conditions.

The numerous significant results appearing in this work are remarkable given the inhomogeneities in the MRF data we were forced to use. Several of the regime-dependent error signatures from each season found in this study can be used almost immediately to correct medium range forecasts out to ten days. These include those that have been shown to be unusually strong and those with moderate strength that reflect a pronounced linear component to the response. Examples of the former are PNA-, and mode 12- for winter, both phases of the Asian summer pattern, and the TNH+ pattern for spring, the subtropical zonal -, NAO-, and the EA+ for summer, and mode

11- for autumn. Examples of the latter are the PNA, WP, and mode 10 during winter, the WP, EA, EU1, EP, mode 12, TNH, and WA patterns during spring, the NAO, mode 7, and mode 11 for summer, and the NAO, modes 5, 6, 7, and 11 for fall.

Clearly, the technique we have outlined is feasible for discovering mode-forecast error relationships for subsequent use in error correction schemes. Use of a homogeneous, comprehensive forecast/analysis data set would almost certainly yield stronger signals and additional useful information. One possible way the potential usefulness of this approach could be estimated would be through the use of a very large sample of forecasts made with a simple model that contains much of the large-scale physics and dynamics of a more comprehensive model. If the results here or from the simple model experiments suggest useful gains in forecast skill are realizable, then it should be possible to construct homogeneous and representative samples of forecasts from the comprehensive model for analysis without freezing operational model configurations for unreasonable lengths of time (like several years), or running an inordinate number of additional cases. This can be accomplished by comparing the distribution of initial conditions for recent archived forecasts (since the last model change) to the long-term distribution of initial modes, and then running additional forecast cases for only those modes not well represented in the archived operational runs.

New computer systems are often used to run more sophisticated, higher resolution models in an effort to improve forecast skill. Why not invest some fraction of that computing power in a technique such as we have described which can, a priori, provide a useful means to reduce error for many different flow regimes? A number of other regimes have been identified that apparently lead to little in the way of systematic error but whose total error level is comparable to that for the important regime-

dependent modes. Error reduction for these cases will require ensemble or other techniques.

Barker, T. W., and J. D. Horel, 1988: Quasi-stationary regimes in the planetary circulation of the NCAR community climate model. J. Clim., 1, 406-417.

Barnston, A. G., R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. Mon. Wea. Rev., 115, 1083-1126.

Blackmon, M. L., 1976: A climatological spectral study of the 500 mb geopotential height of the Northern Hemisphere. J. Atmos. Sci., 33, 1607-1623.

_____, Y. H. Lee, J. M. Wallace, 1984: Horizontal structure of 500 mb height fluctuations with long, intermediate and short time scales. J. Atmos. Sci., 41, 961-979.

Branstator, G., 1987: A striking example of the atmosphere's leading travel pattern. J. Atmos. Sci., 44, 2310-2323.

Dole, R. M., 1982: Persistent anomalies of the extratropical Northern Hemisphere wintertime circulation. PhD. Thesis, Mass. Inst. of Technology, Cambridge, Mass.

Esbensen, S. K., 1984: A comparison of intermonthly and interannual teleconnections in the 700 mb geopotential height field during the Northern Hemisphere winter. Mon. Wea. Rev., 112, 2016-2032.

Guttman, L., 1954: Some necessary conditions for common-factor analysis. Psychometrika, 29, 347-362.

Horel, J. D., 1981: A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 mb height field. Mon. Wea. Rev., 109, 2080-2092.

_____, 1984: Complex principal component analysis: Theory and examples. J. Climate Appl. Meteor., 23, 1660-1673.

Kaiser, H. F., 1958: The Varimax criterion for analytic rotation in factor analysis. Psychometrika, 23, 187-200.

Karl, T. R., A. J. Koscielny and H. F. Diaz, 1982: Potential errors in the application of principal component (eigenvector) analysis to geophysical data. J. Appl. Meteor., 21, 1183-1186.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. J. Atmos. Sci., 20, 130-141.

_____, 1969a: The predictability of a flow which possesses many scales of motion. Tellus, 21, 289-307.

North, G. R., T. L. Bell, R. F. Cahalan and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. Mon. Wea. Rev., 110, 699-706.

O'Connor, J. F., 1969: Hemispheric teleconnections of mean circulation anomalies at 700 mb. ESSA Tech. Rep. WB10, 103 pp.

O'Lenic, E. and R. E., Livezey, 1988: Practical considerations in the use of rotated principal component analysis (RPCA) in diagnostic studies of upper-air height fields. Mon. Wea. Rev., 116, July (in press).

Palmer, T., 1988: Medium and extended range predictability and stability of the Pacific/North American mode. Quart. J. Roy. Met. Soc., 114, 691-713.

_____, and S. Tibaldi, 1987: Predictability studies in the medium and extended range. ECMWF Technical Memorandum No. 139.

Panofsky, H. A., and G. W. Brier, 1968: Some Applications of Statistics to Meteorology. The Pennsylvania State University, 224 pp.

Richman, M. B., 1981: Obliquely rotated principal components: an improved meteorological map typing technique. J. Appl. Meteor., 20, 1145-1159.

_____, 1986: Rotation of principal components. J. Climatol., 6, 293-335.

_____, and P. J. Lamb, 1985: Climatic pattern analysis of three- and seven-day rainfall in the central United States: Some methodological considerations and a regionalization. J. Climate Appl. Meteor., 24, 1325-1343.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. Mon. Wea. Rev., 109, 784-812.