Predictable Anomaly Patterns and the Forecast Skill of Northern Hemisphere Wintertime 500-mb Height Fields

JAMES A. RENWICK* AND JOHN M. WALLACE

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

(Manuscript received 21 March 1994, in final form 2 November 1994)

ABSTRACT

Techniques for identification of well-predicted spatial patterns in numerical weather prediction model output are outlined and applied to a 14-winter set of Northern Hemisphere 500-mb geopotential height analyses and 1–10-day forecasts produced by the ECMWF operational model. Three approaches are investigated: canonical correlation analysis (CCA), singular value decomposition analysis, and predictable component analysis, the products of which are related to the optimization of forecast–analysis correlation, covariance, and rms error, respectively. In confirmation of earlier results, the most predictable anomaly pattern identified by all three methods is found to be similar to the leading empirical orthogonal function of the analyzed 500-mb height anomaly field, which is dominated by the Pacific–North American pattern. The time series of forecast and verifying analysis projections onto the leading pattern have temporal correlations of at least 0.75 at all forecast intervals out to 10 days and greater than 0.85 for 5-day averages of 6–10-day forecasts and analyses. The leading pattern displays strong temporal persistence and is prominent on the interannual timescale. CCA is found to be the most desirable technique for identification of such patterns.

When CCA is applied to the first seven winters' data (as a dependent sample), the amplitude of the leading pattern is well predicted in either polarity and the skill of the full forecast field is shown to increase as the amplitude of the leading pattern increases, regardless of the polarity. However, when the analyzed and predicted fields from the second seven winters of the dataset (an independent sample) are projected onto the patterns derived from the first seven winters, the skill of the full forecast field does not appear to be well related to the amplitude of the leading predictable pattern. Slight decreases in rms errors were achieved by statistically correcting the independent data, but only at the expense of a considerable damping of forecast amplitude. It is concluded that continuing model improvements make such approaches to skill prediction and statistical correction of little value in an operational setting.

1. Introduction

The skill of medium- and extended-range numerical weather prediction varies substantially from day to day. Quantifying the uncertainties in numerical forecasts and predicting, in an operational setting, how they vary on a daily basis are among the principal objectives of current research in dynamical extended-range prediction (National Research Council 1991). Much of this research has tended to focus on the identification of either "predictable anomaly patterns" or "predictable flow regimes." The former may be viewed as *axes* in a multidimensional phase space defined, say, by the coefficients of the leading empirical orthogonal functions (EOFs) of the 500-mb height field, along which the forecast errors tend to be relatively small, whereas the latter correspond to *regions* of phase space characterized by relatively small errors and high predictability. Predictable anomaly patterns tend to be associated with above-average forecast skill in either polarity, whereas predictable flow regimes are associated with anomalies of one polarity. Relatively high forecast skill, in a domain-averaged sense, can be anticipated if one of the predictable anomaly patterns is expected to be present with large amplitude or, alternatively, if the circulation is expected to reside within one of its more predictable "flow regimes."

Early studies of atmospheric predictability, based on concepts of turbulence theory (Lilly 1969; Leith 1971; Leith and Kraichnan 1972) and on experiments with numerical models (Lorenz 1965; Charney et al. 1966; Smagorinsky 1969), placed an upper limit of approximately two weeks on the predictability of instantaneous fields and showed that errors in prediction of the smallest space scales and shortest timescales grow the fastest. More recent studies of model output (Shukla 1981; Lorenz 1982; Dalcher and Kalnay 1987) confirm that in terms of a spherical harmonic representation, the lowest wavenumbers are predicted most accurately out to the longest forecast intervals. The predictability of a

* Permanent affiliation: National Institute of Water and Atmospheric Research Ltd., Wellington, New Zealand.

Corresponding author address: Dr. James Renwick, JISAO, Box 354235, University of Washington, Seattle, WA 98195-4235.
E-mail: renwick@atmos.washington.edu

particular anomaly pattern is naturally associated with its persistence in time, the most slowly varying features being the most predictable (Palmer 1988; Chen 1989; Branstator et al., 1993).

A number of studies have linked the leading EOFs of the wintertime 500-mb height field with high forecast skill, either in terms of the prediction of the polarity and amplitude of the EOF pattern itself or in terms of the prediction of the full height anomaly field. Déqué (1988) isolated those flow patterns best forecast in an rms sense through "predictable component analysis" (PrCA). Based on several winters of Northern Hemisphere 500-mb geopotential height forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF), Déqué found that the mode in the analysis anomaly fields predicted with the smallest rms error projects strongly onto the leading EOF of the height anomalies, which is well correlated with the Pacific–North American pattern (PNA; Wallace and Gutzler 1981; Barnston and Livezey 1987). Such EOF-related results suggest more than simply a spatial-scale–predictability relationship as found by Shukla (1981) and others. The characteristic scale of the first few EOFs of the 500-mb height field, as defined by the distance between centers of action, the strength of gradients, etc., may be similar, yet the leading EOF is more accurately predicted than subsequent patterns. The strong persistence exhibited by the leading EOF suggests that predictability is not solely a function of spatial scale, but is closely related to temporal scale, with the amplitude of the leading PNA-like mode exhibiting substantial interannual variability (see also Branstator et al. 1993).

Using an approach based on forecast–analysis temporal correlation, Branstator et al. (1993) looked for well-predicted 500-mb height patterns in ECMWF model output and found a relationship between EOF number and forecast skill: the leading EOF, dominated by the PNA pattern in the Pacific–North American sector, corresponds closely with the most accurately forecast anomaly pattern. The polarity of this pattern is particularly well forecast, and its amplitude is predicted with approximately equal skill in either polarity. Hence, the PNA pattern may be viewed as a predictable anomaly pattern. In apparent contrast, the results of Palmer (1988), O'Lenic and Livezey (1989), and Chen (1990, 1992) show a polarity-dependent relationship between the PNA pattern and full forecast error, with negative-PNA states (ridge over the North Pacific) being relatively poorly forecast. In the context of their results, the positive-PNA state may be viewed as a predictable flow regime. Branstator et al. (1993) suggest that while the polarity of the PNA pattern may be an important indicator of forecast skill for the full anomaly field, the amplitude of the pattern itself is relatively predictable in either polarity. To investigate this point further, the present paper reports on relationships between the polarity and amplitude of the leading predictable patterns and the skill of the full forecast anomaly field.

Recent efforts to predict the skill of operational forecasts have met with some success. The temporal persistence of certain atmospheric flow regimes, and of skill scores themselves, combined with measures of the dispersion among "ensembles" of forecasts have been used as predictors of forecast skill (Hoffman and Kalnay 1983; Branstator 1986; Kalnay and Dalcher 1987; Chen 1989). The polarity and amplitude of specific elements of the forecast atmospheric state are also related directly to the skill of the forecast, as shown by Palmer (1988), Palmer and Tibaldi (1988), and Molteni and Palmer (1991). Multiple regression models relating rms errors to such predictors were shown to exhibit positive skill at forecast intervals between 3 and 9 days (Molteni and Palmer 1991). However, links between predicted and actual skill were found to be very weak at the longest forecast intervals, leading the authors to conclude that such linear statistical approaches are of limited value.

Two main objectives are pursued in the present paper: 1) to describe three techniques for identification of predictable anomaly patterns: canonical correlation analysis (CCA), singular value decomposition analysis (SVDA), and predictable component analysis (PrCA), with application to a set of Northern Hemisphere wintertime 500-mb height fields; and 2) to examine the robustness and utility of relationships between predictable pattern amplitude and the skill of prediction of the full forecast field, with a view toward operational prediction of forecast skill. A method for the statistical correction of forecasts is also presented. Investigations are carried out for instantaneous fields, and for 5-day averages of day 6 to day 10 forecasts. A description of data sources and a review of basic statistics is found in section 2. The main analysis techniques are outlined in section 3, including presentation of the leading 500-mb height predictable patterns and their relation to forecast skill. Section 4 discusses the effect of changes in model formulation and performance on the predictable patterns, and tests the prediction of forecast skill and the statistical correction of forecasts using independent datasets. A summary and discussion of results concludes the paper.

2. Data and basic statistics

The dataset investigated consists of 14 winters of operational ECMWF daily 500-mb height analyses and 1–10-day predictions (D1, D10, etc., hereafter), from the 1980/81 season through 1993/94. Each winter contains fields valid at 1200 UTC on each of 100 consecutive days starting with 1 December—a total of 1400 sets of forecasts and analyses. Subsets of these data have been used extensively in other studies, such as Lorenz (1982), Dalcher and Kalnay (1987), Déqué (1988), and Branstator et al. (1993). All fields were

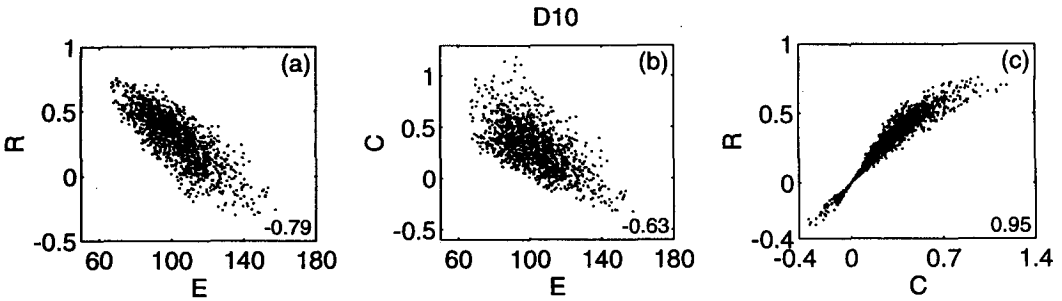

FIG. 1. Scatterplots of relationships among error/skill scores E, R, and C for D10 forecasts: (a) and (b) R and C (nondimensional) plotted against E (m); (c) R plotted as a function of C. Correlation coefficients are indicated in the lower right corners.

converted from spherical harmonic form to a latitude–longitude grid with a latitude resolution of 5° between 25° and 85°N, and a longitude resolution of 15° beginning at the Greenwich meridian (0°). Gridpoint values were weighted by the square root of cosine of latitude to take into account the convergence of the meridians when calculating covariances. There are substantial inhomogeneities in this dataset, due to model changes, as discussed in later sections. Most of the predictable anomaly patterns presented are based on the full dataset, in order to maximize the number of degrees of freedom in the statistics.

Separate "climatologies" were defined for the analyses and for each set of forecasts, as least-squares parabolic fits in time to the set of 100 14-yr mean heights at each grid point, for each forecast interval. When calculating predictable patterns, forecasts and analyses were taken as anomalies from their respective means, thereby removing systematic errors and the climatological mean annual march. When calculating integrated amplitude, error and skill, the analysis means were subtracted from both forecast and analysis fields, removing the annual march but retaining the systematic errors, in keeping with the definitions for rms error and anomaly correlation used in an operational setting. In practical terms, integrated error statistics proved to be relatively insensitive to systematic errors, since the systematic component of the spatially integrated ECMWF forecast error is comparatively small, even for D10 forecasts.

Spatially integrated model amplitude and error are presented in terms of analysis rms amplitude A, forecast rms amplitude F, and rms error E, defined as

$$A = \left[\frac{1}{n}\sum_i (a_i - c_i)^2\right]^{1/2} \tag{1a}$$

$$F = \left[\frac{1}{n}\sum_i (f_i - c_i)^2\right]^{1/2} \tag{1b}$$

$$E = \left[\frac{1}{n}\sum_i (f_i - a_i)^2\right]^{1/2}, \tag{1c}$$

where the summation is over space, f indicates forecast values, a indicates analyzed values, and c indicates the time-varying climatological mean values of the 500-mb height field described above. The number of points in the summation is n. Forecast skill is represented using two derived statistics, the anomaly correlation R, which may be expressed in terms of A, F, and E as

$$R = \frac{A^2 + F^2 - E^2}{2AF} \tag{2}$$

(Molteni and Palmer 1991), and the anomaly covariance C,

$$C = \frac{A^2 + F^2 - E^2}{2\bar{A}\bar{F}}, \tag{3}$$

where the overbars indicate temporal means. The denominator in (3) is a constant and is introduced solely as a scaling factor intended to restrict the range of C approximately to $[-1,1]$. Statistics E and R are widely used measures of forecast error and skill. The less familiar anomaly covariance score C is introduced here to facilitate a more complete discussion and the comparison of the techniques used for defining predictable patterns. It is not put forward as a useful forecast skill score.[1]

Relationships among the scores E, R, and C are similar at all forecast intervals, as illustrated for D10 forecasts in Fig. 1. Here R and C are both negatively correlated with E, and positively correlated with one another. The more subtle features in the scatterplots are readily explainable in terms of (2) and (3). Scatterplots of E, R, and C against the analysis and forecast

[1] The value of anomaly covariance as a measure of forecast skill is compromised by its dependence on forecast amplitude. Here C can be increased without limit simply by inflating the amplitude F of the forecast anomaly field. Also, rms error suffers from a less serious dependence: unless the anomaly correlation is strongly positive, E can usually be decreased (improved) by damping the anomaly patterns in the forecast fields.

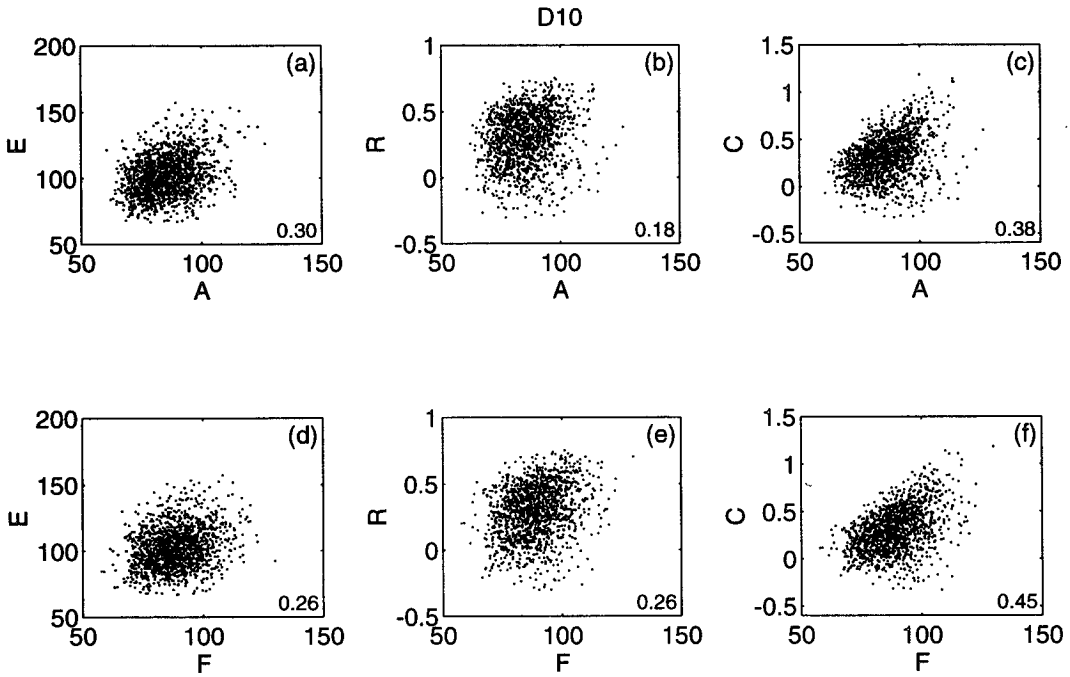

FIG. 2. Scatterplots showing relationships between error/skill scores E, R, and C and analyzed/forecast rms amplitudes A and F, for D10 forecasts: (a), (b), and (c) E (m), R, and C (nondimensional), respectively, plotted against A (m); (d), (e), and (f) E, R, and C, respectively, plotted against F (m). Correlation coefficients are indicated in the lower right corners.

rms amplitudes A and F are shown for D10 in Fig. 2, which is indicative of only a weak correlation between full field amplitude and the scores E and R. Results in sections 3 and 4 will document relationships between these statistics and the amplitude of the leading predictable anomaly patterns.

The procedures discussed in this paper are related to EOF analysis, a technique described in Déqué (1988). The two leading analysis anomaly EOFs are shown in Fig. 3 in the form of covariance maps (contours of the covariances between gridpoint height anomalies and the normalized EOF expansion coefficients). The first EOF is related to the PNA pattern and to the North Atlantic oscillation (NAO; van Loon and Rogers 1978; Barnston and Livezey 1987). An NAO signature is also apparent in EOF 2. The two leading EOFs together

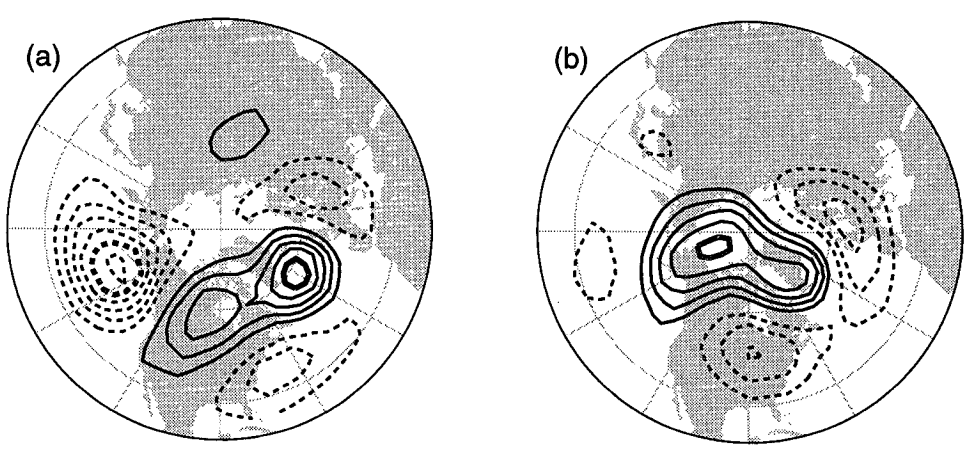

FIG. 3. Leading two EOFs of ECMWF analyzed 500-mb height anomalies, plotted as contours of covariances between original gridpoint data and normalized EOF expansion coefficients. Contours are in meters with a 20-m contour interval, negative contours dashed, ±100-m contours thickened, and zero contour omitted.

account for 18% of the hemispherically averaged 500-mb height variance.

3. Analysis techniques and predictable anomaly patterns

a. Singular-value decomposition analysis

The term "singular-value decomposition analysis" (SVDA; Bretherton et al. 1992; Wallace et al. 1992) is used here to denote the application of the matrix operation of singular-value decomposition to cross-covariance matrices derived from paired sets of fields. Using forecast and analyzed 500-mb height anomalies as input, SVDA yields pairs of forecast and analyzed spatial patterns whose projection coefficient time series have maximal covariance. If the leading pair of patterns are highly correlated in space, the large temporal covariance between their coefficient time series implies large anomaly covariance scores (on average) for fields projecting strongly onto those patterns. If the leading pair of patterns are not highly correlated in space, such a relationship cannot be inferred, but differences between forecast and analysis patterns indicate systematic biases in the forecasts, which may be removed by statistical correction procedures.

If the matrix of forecast 500-mb height anomalies is denoted $\mathbf{X}$ (n rows, one for each forecast time, by p columns for the p grid points in the forecast vector) and the matrix of verifying analysis anomalies is denoted $\mathbf{Y}$ ($n \times q$), with $p \geq q$, SVDA is a decomposition of the sample forecast–analysis covariance matrix $\mathbf{C}_{XY}$:

$$\mathbf{C}_{XY} = \frac{1}{n}\mathbf{X}^{\mathrm{T}}\mathbf{Y} = \mathbf{U}\begin{bmatrix}\mathbf{S}\\ \mathbf{O}\end{bmatrix}\mathbf{V}^{\mathrm{T}} = \sum_{i=1}^{q} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathrm{T}}, \quad (4)$$

where the columns of the orthogonal matrix $\mathbf{U}$ are the left singular vectors $\mathbf{u}_i$, the columns of the orthogonal matrix $\mathbf{V}$ are the right singular vectors $\mathbf{v}_i$, $\mathbf{O}$ is a $(p - q) \times q$ matrix of zeros, and $\mathbf{S}$ is a real diagonal matrix of singular values $\{\sigma_i\}$, $i = 1, 2, \cdots q$, with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_q \geq 0$. The left and right vectors $\mathbf{u}_1$ and $\mathbf{v}_1$ associated with the largest singular value σ_1 identify the pair of patterns that account for the maximal covariance between $\mathbf{X}$ and $\mathbf{Y}$. Time series $\mathbf{a}_i$ and $\mathbf{b}_i$ of the expansion coefficients of the singular vectors are calculated by projecting $\mathbf{X}$ onto the $\mathbf{u}_i$ and $\mathbf{Y}$ onto the $\mathbf{v}_i$. Orthogonality relationships are

$$\mathrm{cov}(\mathbf{a}_k, \mathbf{b}_l) = \sigma_k \delta_{kl}$$

$$\mathrm{cov}(\mathbf{a}_k, \mathbf{a}_l), \mathrm{cov}(\mathbf{b}_k, \mathbf{b}_l) \neq \delta_{kl}$$

$$\mathbf{u}_k \cdot \mathbf{u}_l = \mathbf{v}_k \cdot \mathbf{v}_l = \delta_{kl},$$

where cov denotes covariance and δ_{kl} is the Kronecker delta (Bretherton et al. 1992).

b. Canonical correlation analysis

Canonical correlation analysis (CCA; Glahn 1968; Barnett and Preisendorfer 1987) is used here to yield pairs of forecast and analyzed spatial patterns whose expansion coefficient time series have maximal correlation. As with SVDA, if the leading pair of spatial patterns have a similar form, forecasts projecting strongly onto them will on average have large anomaly correlations with verifying analyses. CCA is similar in many respects to SVDA in that it involves a singular value decomposition of a modified form of the forecast–analysis covariance matrix (see Bretherton et al. 1992). The analogs of the singular values σ_i in SVDA are the canonical correlations ρ_i, the correlations between pairs of expansion coefficient time series.

The number of spatial degrees of freedom in the 500-mb height field is much less than the number of grid points used here to represent it, possibly leading to "overfitting" if the full gridpoint representation is used in the CCA. The datasets are, therefore, truncated in spatial dimension by projection onto a number of their respective leading EOFs, as suggested by Barnett and Preisendorfer (1987). CCA is then a singular value decomposition of the matrix of covariances between the normalized time series of the leading EOFs of each set of fields. The resulting singular vector pairs are used as weights to rotate the EOFs to the directions of maximal correlation in physical space. The CCA pattern pairs (in physical space) are not constrained to be spatially orthogonal, but their associated expansion coefficient time series $\mathbf{a}_i$ and $\mathbf{b}_i$ obey the relations

$$\mathrm{cov}(\mathbf{a}_k, \mathbf{b}_l) = \rho_k \delta_{kl}$$

$$\mathrm{cov}(\mathbf{a}_k, \mathbf{a}_l) = \mathrm{cov}(\mathbf{b}_k, \mathbf{b}_l) = \delta_{kl},$$

where ρ_k is the "canonical" correlation between pairs of time series $\mathbf{a}_k$ and $\mathbf{b}_k$, which are assumed to have unit variance.

c. Predictable component analysis

The forecast error decomposition technique known as predictable component analysis (PrCA, Déqué 1988) yields spatial patterns forecast with minimal mean-square error and is hence associated with the minimization of the rms error E. PrCA finds a single set of spatial patterns in the analyses whose associated time series are predicted with minimal rms error, through a diagonalization of a form of the forecast error variance matrix. As with CCA, the data are reduced toward their "true" spatial dimension by projection onto a number of the leading EOFs of the analysis fields, and forecast errors are taken as the differences between the normalized forecast and analysis EOF expansion coefficients. The eigenvector associated with the smallest eigenvalue of the error variance matrix defines the weights used to rotate the analysis EOFs to the direction of minimal error in physical space.

d. Interpretation

Each of the techniques described in the previous subsections, and summarized in Table 1, focuses on a dif-

TABLE 1. Summary of the main points of the three analysis techniques used in this study.

Name	Acronym	Method	Quantity optimized	Comments
Singular-value decomposition analysis	SVDA	Perform SVD of forecast–analysis covariance matrix.	Forecast–analysis covariance	Large covariance may be ambiguous in terms of forecast skill. Bias toward leading forecast–analysis EOFs.
Canonical correlation analysis	CCA	Project forecast and analysis maps onto respective leading EOFs. Perform SVD of covariance matrix of normalized expansion coefficients.	Forecast–analysis correlation	Need to choose optimal EOF-space dimension.
Predictable component analysis	PrCA	Project forecast and analysis maps onto leading analysis EOFs. Diagonalize the normalized expansion coefficient error variance matrix.	rms error	Need to choose optimal EOF-space dimension. Single set of patterns produced.

ferent aspect of forecast error or skill. If predictability is evaluated in terms of forecast–analysis correlation, the "predictable patterns" derived from SVDA are inherently biased toward the leading EOFs of the analysis and forecast fields. Consider the situation in which each of the EOFs of the forecast fields is identical to its counterpart in the analysis fields and all EOFs are equally predictable. In this case, it is readily apparent that the CCA modes are not unique, but the leading EOF (if one exists) will emerge as the "most predictable mode" in the SVDA expansion, when in reality it is no more predictable than any other mode. In view of this bias, and the dependence of C and E on forecast (and analysis) amplitude as mentioned in footnote 1, CCA appears to be the method of choice for identifying predictable anomaly patterns. However, all three techniques are used in this study, to give an indication of the sensitivity of the results.

For display, spatial patterns are shown as covariance maps, which are contours of the covariances between the height anomalies at each grid point and the normalized time series of the expansion coefficient of the pattern in question. For EOF and PrCA patterns, such maps indicate average anomaly amplitudes in the analysis or forecast fields associated with a unit projection onto each pattern. For SVDA or CCA one can envision homogeneous covariance maps relating forecast (analysis) fields and forecast (analysis) expansion coefficients, and heterogeneous covariance maps relating forecast (analysis) fields and analysis (forecast) expansion coefficients. Homogeneous maps are the analogs of the EOF covariance maps in Fig. 3, indicating height anomalies in one set of fields associated with unit projections onto patterns in the same set of fields. Heterogeneous maps indicate anomalies in one set of fields associated with unit projections onto patterns in the other set of fields and are a measure of the forecast–analysis covariance associated with those patterns. Forecast–analysis mode pairs are displayed as analysis homogeneous covariance maps and forecast heterogeneous covariance maps, based on the analysis time series in both cases. The former portray the patterns in the verifying analyses, represented with typical amplitude, while the latter show the typical amplitude of the "response" in the forecasts.

The importance of individual patterns is assessed in terms of the fraction of variance of the analysis fields accounted for. For PrCA modes, the rms error is calculated as the rms difference between forecast and analysis expansion coefficient time series, normalized by the standard deviation of the analysis time series. The strength of the coupling of SVDA and CCA modes is measured using the squared covariance fraction (SCF_k), the normalized rms squared covariance (c_k), the correlation coefficient between the time series of the paired expansion coefficients (r_k), and the spatial correlation coefficient between forecast and analysis patterns (s_k), where k indicates the mode number. All range between 0 and 1 in magnitude, with +1 indicating perfect agreement between forecast and analysis modes. All four statistics are defined and discussed briefly in the appendix. Further discussion is found in Bretherton et al. (1992) and Wallace et al. (1993).

e. Application to 500-mb height anomalies

All three techniques were applied to forecast and analysis height anomalies for each forecast interval from D1 through D10 and to D6–D10 averages. Most results displayed are based on CCA, for reasons discussed in the previous subsection. Both CCA and PrCA were carried out for a variety of EOF truncations, to assess their sensitivity to the number of EOFs retained (Barnett and Preisendorfer 1987; Déqué 1988). The results obtained here are relatively robust, retaining essentially the same spatial features (for at least the leading two modes) for EOF truncations between 5 and 20. Table 2 shows summary statistics for the leading D10

TABLE 2. Summary statistics for the first D10 forecast–verifying CCA and PrCA modes, at various EOF truncations (column 1). Column 2 shows the total amount of analysis variance retained. CCA summary statistics are analysis variance accounted for by first analysis mode (Ex. V), squared covariance fraction (SCF_1), normalized rms squared covariance (c_1), forecast–verifying expansion coefficient temporal correlation (r_1), and forecast–verifying weight vector spatial correlation (s_1). PrCA statistics are fraction of analysis variance accounted for and normalized rms error. All statistics are nondimensional. The D10 SVD explained variance is that accounted for by the first D10 analysis SVD mode in the analysis dataset, and the rms error is between the leading forecast and analysis mode time series.

		CCA					PrCA	
EOF dimension	Tot. var (%)	Ex. V. (%)	SCF_1 (%)	c_1	r_1	s_1	Ex. V. (%)	rmse
5	36	10	36	0.062	0.64	0.93	9	0.83
7	45	9	37	0.063	0.67	0.93	9	0.79
10	55	9	38	0.063	0.70	0.94	9	0.77
15	67	8	39	0.064	0.73	0.95	9	0.75
20	75	8	39	0.064	0.74	0.94	9	0.74
30	84	8	38	0.064	0.77	0.84	8	0.72
40	89	7	38	0.063	0.79	0.79	8	0.70
50	92	7	38	0.063	0.79	0.72	8	0.70
SVDA	100	10	40	0.065	0.68	0.95	10	0.81

CCA and PrCA patterns at dimensional truncations between 5 and 50. There is a noticeable decline in the CCA pattern spatial correlation for the leading mode (s_1) beyond the 20 EOF truncation. With the retention of up to 20 EOFs, the leading CCA pattern retains a PNA-like appearance, but beyond that point it begins to drift toward a more zonally symmetric form that appears to be partly related to interannual variations in the position and intensity of the Asian jet. For display, a truncation of 20 EOFs was selected, which retains at least 75% of the total variance in both the analyses and forecasts. At this level of truncation, the PrCA rms error of 0.74 for the leading mode agrees well with that reported by Déqué (1988), based on a somewhat shorter record of D10 forecasts.

Covariance maps for the two leading D5 and D10 CCA modes are shown in Figs. 4 and 5, and summary statistics for the three leading CCA modes are presented in Table 3. The two leading D10 SVDA covariance maps are shown in Fig. 6 and those for the two leading D10 PrCA modes in Fig. 7. Summary statistics for the leading SVDA modes are given in Table 4. The leading D10 pattern defined by all three techniques is qualitatively similar to EOF 1. The agreement between forecast and analysis patterns is very strong for the two leading modes, at all forecast intervals (all spatial correlations between CCA pattern pairs are at least 0.90), suggesting little systematic bias in prediction of these two patterns of variation. Forecasts projecting strongly onto the leading pattern should therefore achieve relatively high anomaly correlations, on average. Mode 1 accounts for an increasing fraction of the squared covariance as the forecast interval increases, although the total covariance accounted for decreases monotonically with forecast interval. The loss of some of the features on the Atlantic side of both modes one and two at later forecast intervals suggests that the predictable anomaly patterns retain their integrity the longest for planetary-scale features over the Pacific sector. The D10 patterns shown in Fig. 5 agree well with the D10 "easily forecast patterns" found by Branstator et al. (1993, their Figs. 4e and 4f) and with the leading "predictable component" shown in Déqué (1988, Fig. 2).

The above results were compared with their counterparts for persistence forecasts, by replacing forecasts with their initial analyses in the calculations. The leading persistence CCA mode (not shown) is again related to analysis EOF 1 but initial–verifying analysis relationships are weaker than those for the forecast model. (CCA summary statistics are listed in Table 5.) Persistence mode two is rather weakly represented in the initial analyses and changes shape as the "forecast interval" increases, suggesting that only the first mode (related to the PNA pattern) is strongly persistent out to 10 days.

The polarity dependence of the predictability of the leading pattern was tested by separating the expansion coefficient time series of the leading D10 CCA mode and of the leading analysis EOF into positive and negative projections, based on the sign of the verifying analysis expansion coefficient. Table 6 shows the correlation coefficient and rms difference between forecast and analysis time series for the two polarities, and for all cases together. The correlation coefficients based on the whole dataset are higher than those for either subset, indicating that the polarity of the leading pattern is well forecast, in agreement with the results of Branstator et al. (1993). In terms of forecast–analysis correlation, there appears to be no sensitivity to the polarity of the leading pattern. Higher rms errors in the negative polarity are partly associated with generally higher amplitudes for negative projections onto the leading pattern.

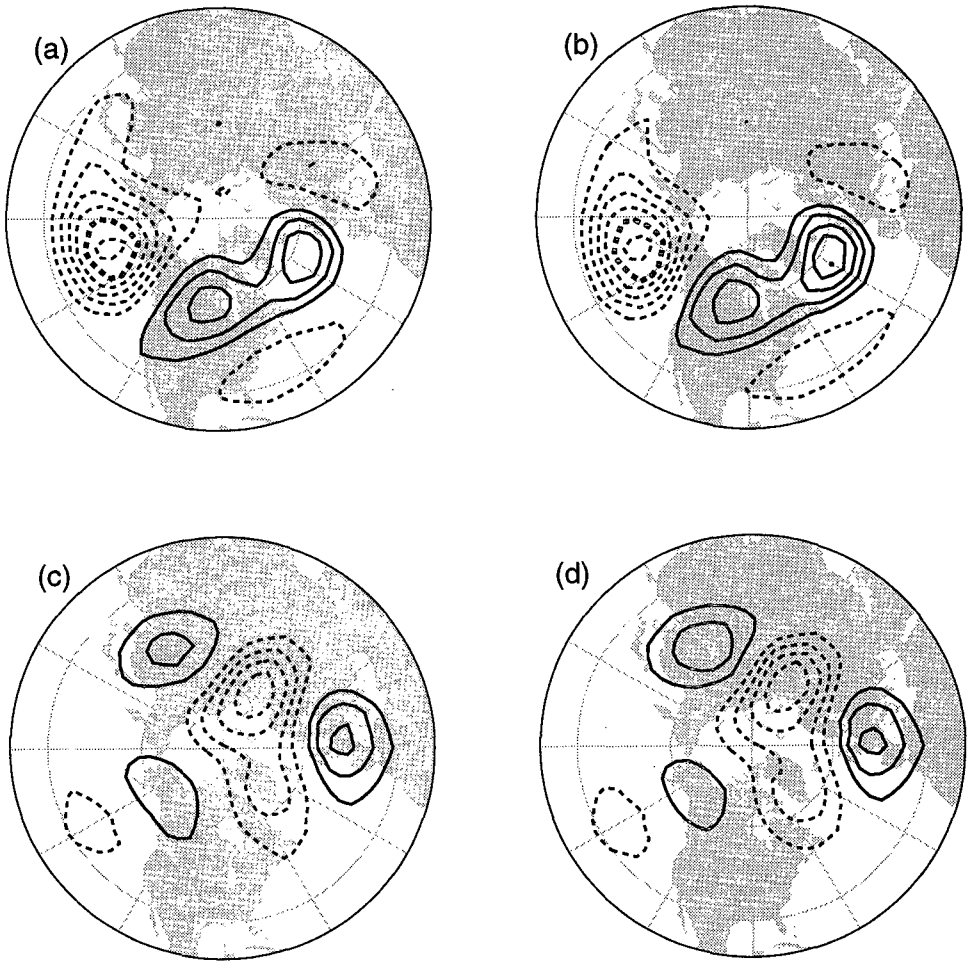

FIG. 4. Leading two modes from D5 forecast–verifying analysis CCA, displayed as covariance maps based on the normalized analysis pattern time series in each case. (a) The leading forecast heterogeneous covariance map, that is, contours of covariance between D5 forecast height anomalies and the normalized expansion coefficients of analysis mode 1. (b) The leading analysis homogeneous covariance map, that is, contours of covariance between analysis height anomalies and the normalized expansion coefficients of analysis mode 1. (c) and (d) The analogous maps for mode 2. Contour conventions as in Fig. 3.

Predictable anomaly patterns were also calculated for 5-day means of D6–D10 forecasts (from the same initial time) and averaged verifying analyses. A simple running mean was used, in preference to a more sophisticated filtering technique, since it is likely to be the best available option in an operational setting. The results of all three techniques were again in close agreement, with the first two CCA modes (not shown) being similar in form to the unaveraged D10 results of Fig. 5, but with stronger forecast–analysis links, as listed in Table 7.

As a measure of the contribution of the leading predictable patterns to a standard forecast skill statistic, the anomaly correlation R was calculated for fields degraded by the removal of the variance associated with the 10 leading CCA predictable patterns and with the 10 leading analysis EOFs. The contribution of the leading patterns to the anomaly correlation increases with forecast interval: Table 8 shows results for D10 forecasts and for D6–D10 averages. Removal of the leading CCA modes has a slightly larger cumulative effect on R, although the differences are not statistically significant. For D6–D10 averages, the fractional amount of total skill accounted for by the leading mode is roughly half that of the leading D10 mode. The averaged forecasts are comparatively more skillful, and the trailing modes carry relatively more information than they do for instantaneous fields.

f. Pattern amplitude and full forecast skill

The results discussed above suggest that EOF 1 is a particularly well-forecast structure, in both its positive and negative polarities. To investigate the form of the

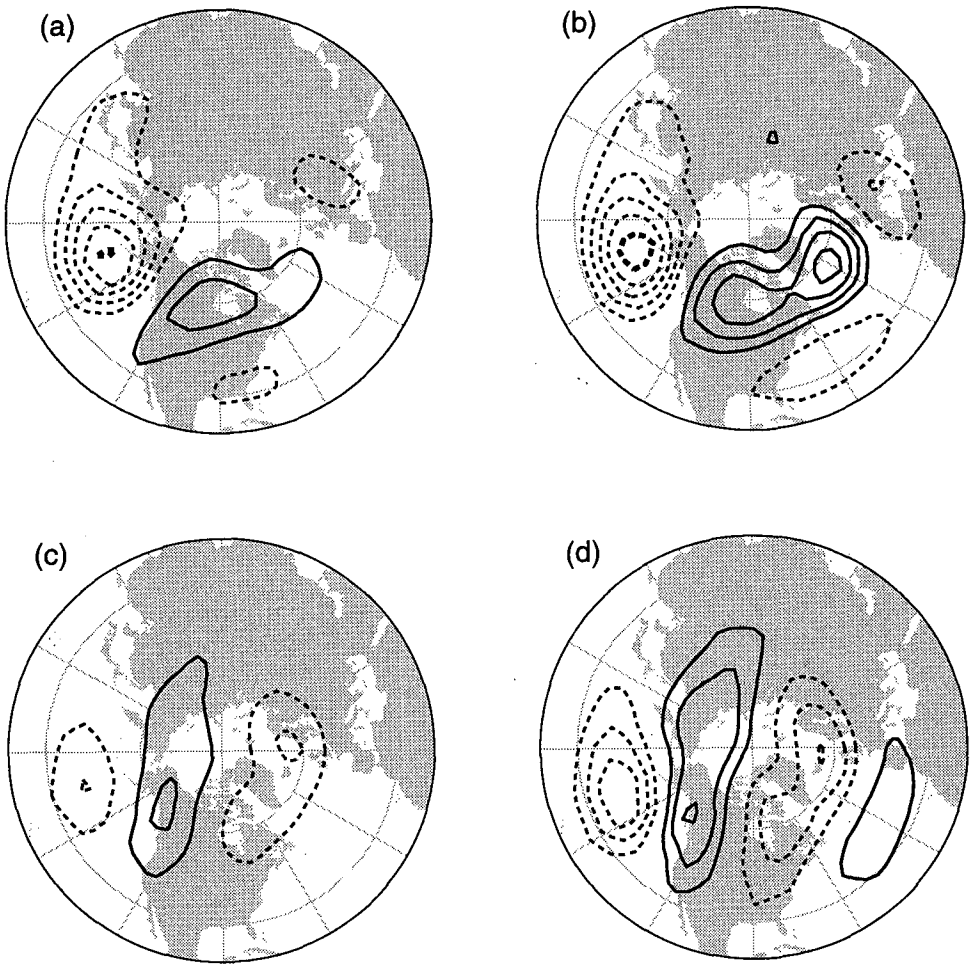

FIG. 5. As in Fig. 4 but for D10 forecasts and verifying analyses.

relationships between predictable pattern amplitude and full forecast amplitude and skill, daily measures of forecast error and height anomaly amplitude have been considered as functions of the expansion coefficients of the leading CCA predictable anomaly patterns. Figure 8 shows scatter diagrams of amplitude, error, and skill quantities A, F, E, R, and C, plotted against the amplitude of the leading forecast D10 CCA pattern. The straight lines in Fig. 8b represent the forecast rms amplitude due solely to forecast CCA mode 1 for given values of its expansion coefficient. There is a weak absolute-value relationship between expansion coefficient 1 and all the quantities, except for E. The slight asymmetries evident in Figs. 8d and 8e are related to asym-

TABLE 3. Summary statistics for the first three modes, forecast–verifying analysis CCA, based on a 20-EOF truncation, for forecast D5 and D10. Columns are analysis variance accounted for by analysis pattern (Ex. V), squared covariance fraction (SCF_k), normalized rms squared covariance (c_k), forecast–verifying expansion coefficient temporal correlation (r_k), and forecast–verifying singular vector spatial correlation (s_k). All statistics are nondimensional.

Mode number k	D5					D10				
	Ex. V. (%)	SCF_k (%)	c_k	r_k	s_k	Ex. V. (%)	SCF_k (%)	c_k	r_k	s_k
1	9	26	0.09	0.94	0.98	8	39	0.06	0.75	0.94
2	6	10	0.06	0.92	0.98	6	14	0.04	0.62	0.92
3	7	16	0.07	0.91	0.98	7	15	0.04	0.57	0.77

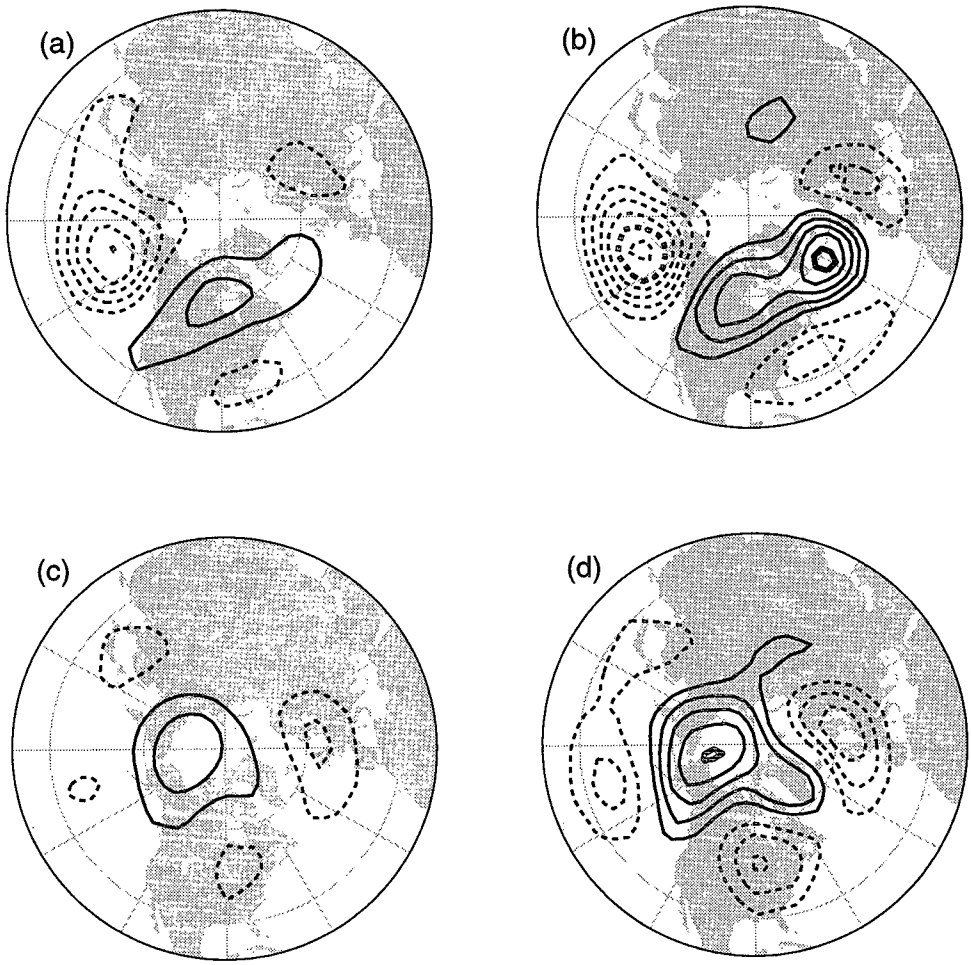

FIG. 6. As in Fig. 4 but for D10 forecast–verifying analysis SVDA.

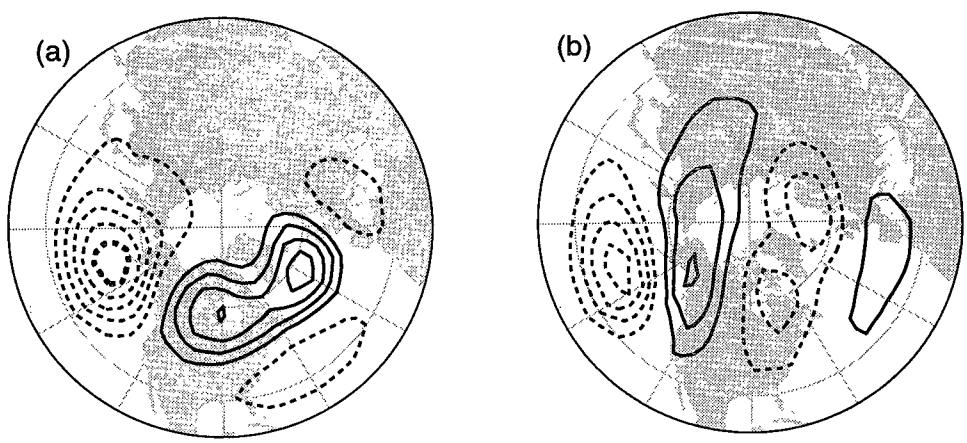

FIG. 7. Leading two PrCA modes, based on a 20-EOF truncation of D10 forecast and verifying analyses. Patterns are shown as contours of covariances between analysis gridpoint data and normalized predictable vector expansion coefficients. Contours are in meters with a 20-m contour interval, negative contours dashed, ±100-m contours thickened, and zero contour omitted.

TABLE 4. As in Table 3 but for D5 and D10 SVDA. All statistics are nondimensional.

Mode number k	D5					D10				
	Ex. V. (%)	SCF_k (%)	c_k	r_k	s_k	Ex. V. (%)	SCF_k (%)	c_k	r_k	s_k
1	10	28	0.10	0.93	0.99	10	40	0.07	0.68	0.95
2	9	19	0.08	0.90	0.99	8	18	0.04	0.56	0.95
3	6	11	0.06	0.89	0.99	7	12	0.04	0.57	0.93

metries in forecast and analysis amplitudes: extreme negative projections onto the leading predictable pattern, associated with the negative polarity of the PNA pattern, have larger amplitude than extreme positive projections.

The anomaly correlation R is positively correlated with both the full forecast amplitude F and the absolute value of the amplitude of CCA D10 forecast mode 1. An interesting question is whether the correlation between F and R leads to the correlation between predictable pattern amplitude and R (since the amplitude of the leading predictable pattern is a relatively large contributor to F) or whether the converse is true, that the enhanced predictability of the leading CCA pattern leads to the F–R correlation. A study of data subsamples stratified by field amplitude, and of the partial correlations between all three variables, did not yield a clear answer to this question. In support of the former view, it seems that as the analysis anomalies move farther from "climatology," R may be expected to increase since the model is likely to correctly forecast at least the sign of large anomalies, given that it is relatively unbiased over all. Equally, one might argue that since the major contributors to anomaly amplitude (the leading EOFs) happen to be relatively persistent in time, and hence relatively predictable, a relation between total anomaly amplitude and forecast skill must follow.

Figure 8d suggests that increasing forecast pattern amplitude is associated with an increasing probability of high anomaly correlations. For example, for D6–D10 averaged forecasts, 43% of all forecasts achieved values of $R > 0.6$, compared to 73% of those forecasts where the projection onto the leading predictable pattern was greater than 1.5 in absolute value. The use of such relationships for prediction of forecast skill is tested and discussed in the following section.

PrCA-based scatterplots analogous to those in Fig. 8 (not shown) indicate that the expansion coefficient of the leading PrCA mode is only very weakly related to E, with a slight tendency for decreased rms errors for large absolute values of the expansion coefficient. PrCA actually minimizes the mean-square error, and plots such as Fig. 8c using E^2 rather than E do show somewhat clearer relationships with the amplitude of the leading predictable pattern. In terms of the full fields, E generally increases with A and F (Figs. 2a and 2d). It is notable that the leading PrCA pattern matches so closely the leading EOF of analyzed 500-mb height anomalies, since this is the pattern associated with the largest height variance (anomaly amplitude). The leading PNA-like pattern is simultaneously associated with large variance and relatively small rms errors, in opposition to the weak positive correlation between E and A noted in Fig. 2.

Smoothed curves relating D6–D10 forecast predictable pattern amplitude and forecast skill are shown in Fig. 9, which were generated by ranking the individual values in accordance with the expansion coefficient of the leading predictable pattern and then smoothing them with a 51-point running mean. The R and C curves exhibit noticeable U shapes, but there is little suggestion of an "inverted-U" relationship between E and the forecast time series. The leading CCA mode of the averaged fields clearly defines a predictable "axis" in terms of anomaly correlation, with generally more accurate forecasts for large amplitudes in either polarity and little suggestion of a region of enhanced predictability on one side of the origin, apart from the asymmetries associated with differences in amplitude be-

TABLE 5. Summary statistics, as in Table 3 but for initial–verifying analysis CCA, based on a 20-EOF truncation, for lags of 5 and 10 days.

Mode number k	5-day lag					10-day lag				
	Ex. V. (%)	SCF_k (%)	c_k	r_k	s_k	Ex. V. (%)	SCF_k (%)	c_k	r_k	s_k
1	8	32	0.07	0.80	0.84	7	39	0.05	0.68	0.79
2	7	21	0.05	0.73	0.58	5	17	0.03	0.55	0.66
3	5	10	0.04	0.70	0.67	4	8	0.02	0.53	0.56

TABLE 6. Normalized rms error and correlation coefficient between D10 forecast and verifying analysis time series. The first two columns are based on the leading CCA predictable pattern and columns 3 and 4 are based on forecast and analysis projections onto the leading analysis EOF. The first row is for all days. Subsequent rows are stratified by the sign of the analysis projection.

	CCA		EOF	
	rmse	corr	rmse	corr
All	0.71	0.75	0.86	0.61
Proj < 0	0.76	0.56	0.88	0.41
Proj > 0	0.66	0.55	0.84	0.41

TABLE 7. Summary statistics, as in Table 3, for the first three modes of D6–D10 average forecast-verifying analysis CCA, based on a 20-EOF truncation. The last line shows the average values of the mode 1 CCA results for individual forecast days 6–10.

Mode number k	Ex. V. (%)	SCF_k (%)	c_k	r_k	s_k
1	11	36	0.10	0.87	0.97
2	7	13	0.06	0.79	0.93
3	8	15	0.07	0.77	0.84
D6–D10 mode 1 avg.	9	33	0.08	0.83	0.97

tween large positive and negative projections onto mode 1.

4. Impact of model changes

a. *Predictable anomaly patterns*

The ECMWF forecast model has been evolving steadily since it was put into operational use in 1979. Many of the significant changes are described in Déqué (1988) and Bengtsson (1991). A crude appreciation of the impacts of model improvement can be obtained by splitting the data into two halves, the 1980/81–1986/87 winters (the seven "early" winters), and the 1987/88–1993/94 winters (the seven "late" winters). This division is interesting since the frequency of "blocking" states has increased noticeably in medium-range forecasts, relative to the number of analyzed blocks, since 1986/87 (Tibaldi et al. 1993), and there is a large difference in the form of seasonal mean errors between the two periods (Fig. 10). For comparison, the data were also split into halves by choosing the seven "even" winters (those beginning in December 1980, 1982, etc.) and the seven "odd" winters (beginning in December 1981, 1983, etc.) separately. Comparison of results based on the even–odd split and based on the early–late split give some indication of the relative sensitivity of the results to sampling variability and to model changes.

The main features of the leading predictable pattern are relatively reproducible, as indicated in Fig. 11, which shows the leading D10 CCA-based analysis pattern for early, late, even, and odd subsets. (The leading D10 forecast patterns closely resemble their analysis counterparts in all cases.) There is a shift in emphasis between the Pacific and Atlantic from the early to the late winters. The Pacific center of action moves a little northward, making it less clearly related to the PNA pattern. A similar change is apparent over the North Pacific between even and odd subsets, suggesting that this variation may be attributable to sampling. The spatial correlation between the leading analysis patterns in the early and late winters is 0.58, versus 0.70 between even and odd winters.

Relationships between the amplitude of the leading forecast predictable pattern and forecast skill are shown in Table 9 as correlation coefficients, between the absolute amplitude of the leading forecast pattern and full-field anomaly correlations, for D10 and D6–D10 average forecasts. In the "dependent" samples, the correlations are based on the same period of record as was used for the CCA calculations, whereas in the "independent" samples, the correlations are based on the part of the record excluded from the CCA calculations. In terms of the dependent samples, amplitude–skill relationships appear to be significantly stronger during the later years, since the increase in correlation from the early to the late winters is roughly twice as large as the correlation difference between the even and odd winters. When the withheld winters' forecasts are projected onto the leading CCA pattern, amplitude–skill correlations generally decrease, as expected. Figures in the first row of Table 9 are of the most interest since, in operational use, patterns derived from "early" years would be used with "later" years forecasts. Assuming one independent data point every 5 days (approximately 140 independent days in each subsample), correlations coefficients above approximately 0.18 may be considered significantly greater than zero with 95%

TABLE 8. Anomaly correlation coefficient R for D10 and D6–D10 average forecasts and analyses. The first line shows the mean value of R for unmodified fields. Subsequent lines show the mean R after removal of the variance associated with the first few predictable patterns, defined as the leading CCA modes and as the leading analysis EOFs. Column 1 shows the number of modes removed. Estimated standard deviations are approximately 0.20 in all cases.

Modes removed	D10		D6–D10	
	CCA	EOF	CCA	EOF
0	0.30	0.30	0.55	0.55
1	0.27	0.27	0.52	0.52
1–2	0.24	0.25	0.49	0.50
1–3	0.22	0.23	0.47	0.47
1–4	0.21	0.22	0.44	0.45
1–5	0.19	0.21	0.43	0.44
1–10	0.13	0.15	0.33	0.35

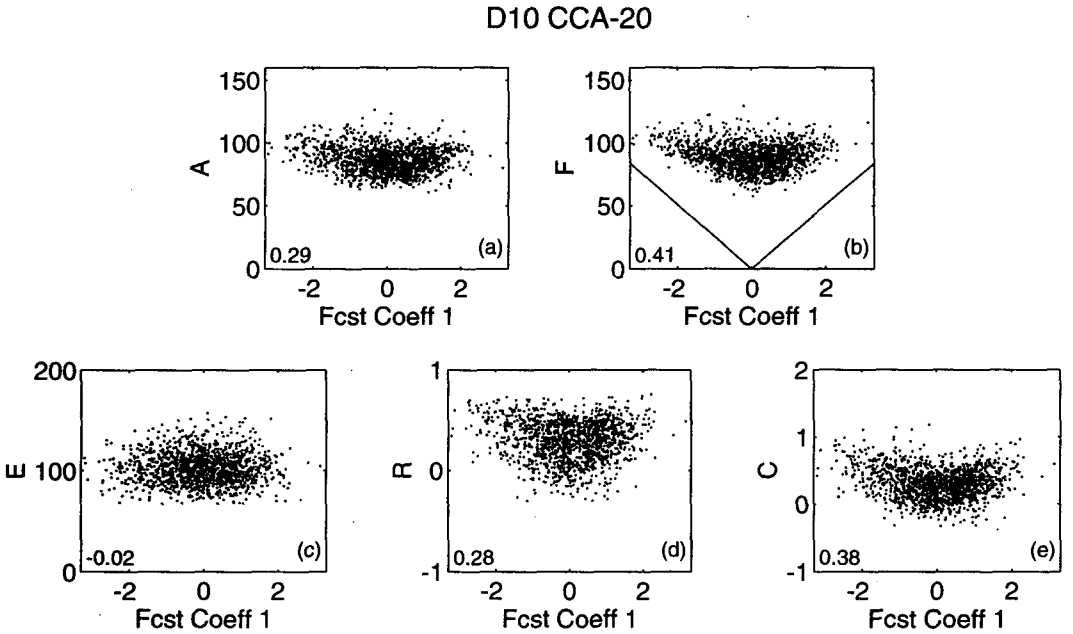

FIG. 8. Daily values of (a) analysis height anomaly rms amplitude A (m), (b) D10 forecast height anomaly rms amplitude F (m), (c) D10 rms error E (m), (d) D10 anomaly correlation R, and (e) D10 anomaly covariance C, plotted as a function of the normalized expansion coefficient of the leading D10 forecast CCA mode. The straight lines in (b) show the contribution of mode 1 to the forecast rms amplitude. Correlation coefficients between each statistic and the absolute value of the forecast mode expansion coefficient are shown in the lower left corner of each panel.

confidence. Based on this measure, the amplitude of the patterns derived from the early winters are only marginally related to predictability during the later winters. Apparent relationships between forecast pattern amplitude and the probability of high forecast skill do not apply well to independent datasets.

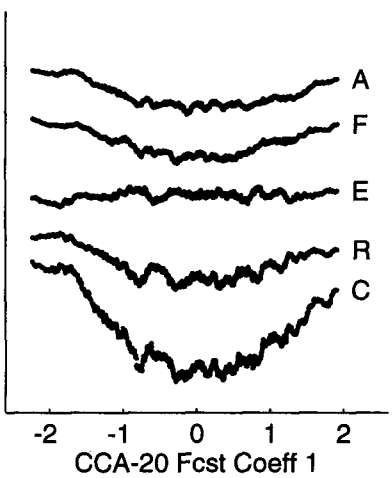

FIG. 9. Scaled daily values of A, F, E, R, and C for D6–D10 averaged fields, plotted against the normalized expansion coefficient of the leading D6–D10 CCA forecast pattern, after smoothing with a 51-point running mean.

b. Statistical correction of forecasts

Paired CCA patterns may be used for statistical correction of model output, since a given amplitude of each forecast pattern implies (on average) a certain amplitude of the corresponding pattern in the verifying analysis. As mentioned in section 3, this formalism may be useful for correcting obvious distortions in the prediction of the leading patterns. Although this is not the case here, the approach is described and is tested on the data subsamples described above.

The fields are written approximately as

$$\mathbf{f}(x, t) \approx \sum_{i=1}^{r} a_i(t)\mathbf{p}_i(x) \tag{5}$$

and

$$\mathbf{v}(x, t) \approx \sum_{i=1}^{r} b_i(t)\mathbf{q}_i(x)\,, \tag{6}$$

where $\mathbf{f}(x, t)$ is the forecast anomaly field at time t, approximated as the sum of the first r forecast modes $\mathbf{p}_i(x)$ weighted by amplitudes $a_i(t)$, and $\mathbf{v}(x, t)$ is the verifying analysis anomaly field at time t, approximated as the sum of the first r analysis modes $\mathbf{q}_i(x)$ weighted by amplitudes $b_i(t)$. The corrected forecast $\mathbf{f}'(x, t)$ is

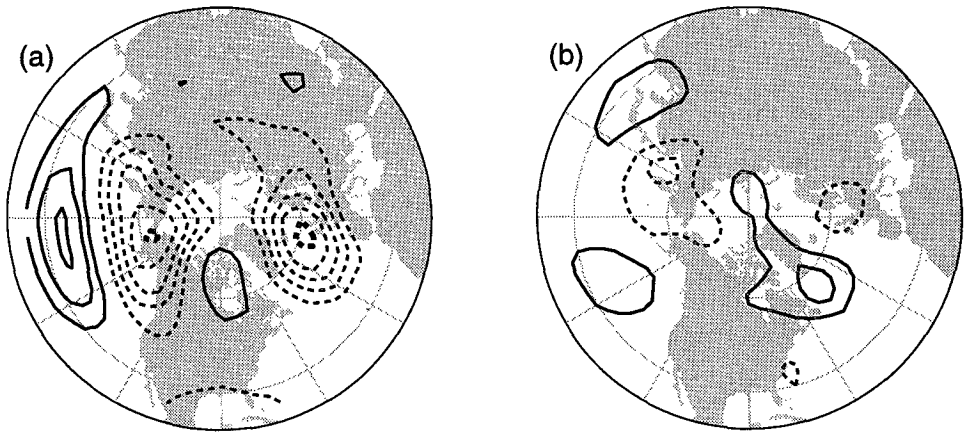

FIG. 10. The D10 mean forecast error for the winters of (a) 1980/81–1986/87 and (b) 1987/88–1993/94. Contours are in meters with a 20-m contour interval, negative contours dashed, ±100-m contour thickened, and zero contour omitted.

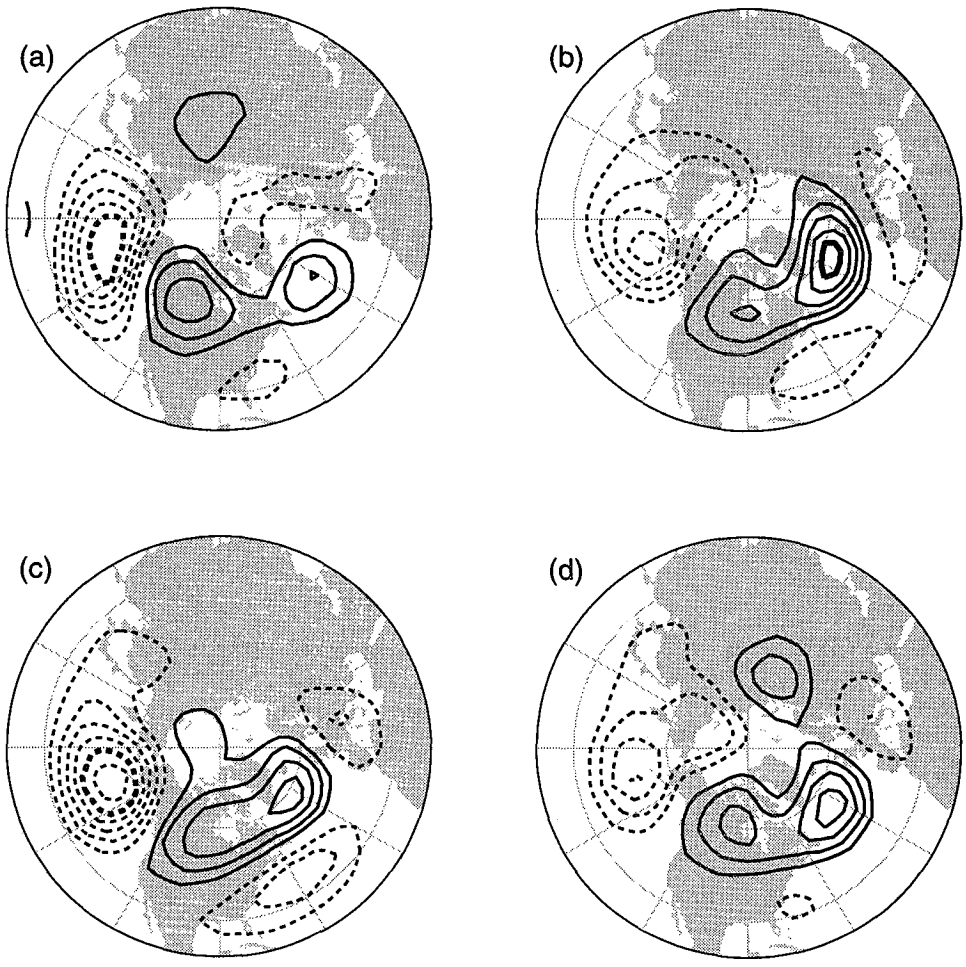

FIG. 11. Homogeneous covariance maps of the leading analysis mode from D10 forecast–verifying analysis CCA, for different subsamples: (a) the seven early winters, (b) the seven late winters, (c) the seven even winters, and (d) the seven odd winters. Contour conventions as in Fig. 10.

TABLE 9. Correlation coefficients between the absolute value of the amplitude of the leading CCA-based forecast predictable pattern and the anomaly correlation score. The "dependent" columns show figures for the dependent data samples listed. The "independent" column shows the correlations between pattern amplitude and R for the winters excluded from the dependent sets. The last line shows correlations based on the whole 14-winter set as a dependent sample.

	D10		D6–D10	
	Dept.	Indept.	Dept.	Indept.
Early	0.19	0.15	0.23	0.18
Late	0.33	0.23	0.37	0.26
Even	0.32	0.21	0.34	0.31
Odd	0.24	0.30	0.27	0.28
All data	0.28	—	0.30	—

$$\mathbf{f}'(x, t) = \sum_{i=1}^{r} c_i(t)\mathbf{q}_i(x). \quad (7)$$

The amplitudes $c_i(t)$ are found by regression of the analysis pattern amplitudes on those of the forecast. For CCA-based patterns, the best linear correction that can be made to the forecast pattern amplitudes is to multiply them by the appropriate canonical correlations; that is, $c_i(t) = \rho_i a_i(t)$, where ρ_i is the ith canonical correlation coefficient. Since the ρ_i must be less than one, this procedure damps the amplitude of the forecast anomalies, especially of the most poorly forecast elements. Beyond D6, the canonical correlations fall off quickly with increasing predictable mode number, so the corrected forecast is dominated by fluctuations in only the leading few modes.

Corrections were applied to each of the seven-winter subsamples, using the leading 20 CCA patterns calculated from the remaining seven seasons. Forecasts were compared to the full analyses, and to analyses filtered to retain only the leading 20 analysis CCA patterns, matching those modes of variation that are retained in the corrected forecasts. Table 10 shows results for each of the independent samples, for D10 and D6–D10 average forecasts, and results for all winters as dependent data, for comparison. Mean rms errors showed statistically significant reductions in all cases (significant at the 99% level, one-tailed t test, assuming one independent observation every 5 days), but at the expense of reducing forecast field amplitude by up to 60%, perhaps rendering the procedure unattractive in an operational setting. Although dependent data statistics suggested that significant improvements in anomaly correlations are possible, the correction procedure either had no effect on, or degraded, mean anomaly correlations in the independent samples. The largest reductions in R occurred for the late winters, suggesting that model improvements invalidate some of the assumptions behind this procedure. In the comparisons against filtered analyses, apparent skill improvements are attributable mostly to the filtering procedure and not to the statistical correction. Of course, given the similarity between the leading pairs of patterns for these data, the full "correction" is little more than a filtering procedure.

5. Summary and discussion

Predictable anomaly patterns have been identified through the application of CCA, SVDA, and PrCA,

TABLE 10. Independent sample values of forecast amplitude F (m), rms error E (m), and anomaly correlation R for the four seven-winter sets listed, and for the whole dataset as a dependent sample. Lines labeled "raw" indicate values before statistical correction. Those labeled "full analysis" show values for corrected forecasts compared to unfiltered analyses, and those labeled "filtered analysis" are for corrected forecasts compared to analyses filtered to retain only the leading 20 analysis predictable patterns.

		D10			D6–D10		
		F	E	R	F	E	R
Early	raw	87	106	0.25	75	73	0.49
	full analysis	32	84	0.26	43	63	0.48
	filtered analysis	32	69	0.31	43	54	0.54
Late	raw	89	99	0.34	74	63	0.60
	full analysis	35	82	0.27	42	59	0.52
	filtered analysis	35	67	0.32	42	50	0.58
Even	raw	89	104	0.29	76	69	0.54
	full analysis	35	83	0.29	44	61	0.52
	filtered analysis	35	69	0.34	44	52	0.58
Odd	raw	88	101	0.31	73	66	0.55
	full analysis	33	81	0.27	41	59	0.51
	filtered analysis	33	68	0.33	41	50	0.57
Dependent	raw	89	102	0.30	74	68	0.55
	full analysis	33	79	0.38	43	56	0.60
	filtered analysis	33	65	0.44	43	48	0.65

each of which optimizes a different measure of forecast error or skill. All techniques agree that a pattern related to the leading EOF of analyzed 500-mb height anomalies is the most accurately forecast, confirming the results of Déqué (1988) and Branstator et al. (1993). Consideration of the scores and biases associated with the different techniques leads to the conclusion that CCA, which maximizes the temporal correlation between the expansion coefficients of the forecast and analyzed patterns, is the most useful of the three techniques for identification of predictable anomaly patterns in model output. Model changes and sampling variability do influence the form of the leading pattern, but a pattern close to the leading EOF emerges as the most predictable in all samples used.

As noted in section 3f, it seems remarkable that the most prominent spatial pattern in the wintertime 500-mb height field (the leading EOF) also turns out to be the most predictable in terms of both anomaly correlation and mean-square error. One reason for the relatively high predictability of this pattern is its persistence, as evidenced by the autocorrelation statistics plotted in Fig. 12. Statistics were averaged for EOFs 3, 5, and 6 and for EOFs 7–10, since the individual curves for these grouped modes are similar. The strong persistence of EOF 1 is consistent with the fact that the PNA pattern is more prominent in the interannual variability than in the intraseasonal variability (Kushnir and Wallace 1989; Wallace et al. 1993). In other words, it relaxes back toward "climatology" the slowest, rendering it the most predictable, in agreement with Chen (1989) and Branstator et al. (1993). Out to about a 10-day lag, EOF 4 (concentrated over northern Russia) is the next most persistent of the leading patterns, and is identified as the most predictable in the D1 CCA results.

Removal of interannual variability by subtraction of individual winter means reduces the importance of the leading pattern somewhat. The leading CCA pattern (not shown) retains a shape similar to those shown in Figs. 4 and 5, but with lower amplitude, and the associated squared covariance fractions are reduced by approximately 20%. This suggests that although variations on timescales much longer than the forecast interval do not have a strong influence on day-to-day forecasts, they do contribute to the prominence of the leading predictable pattern.[2]

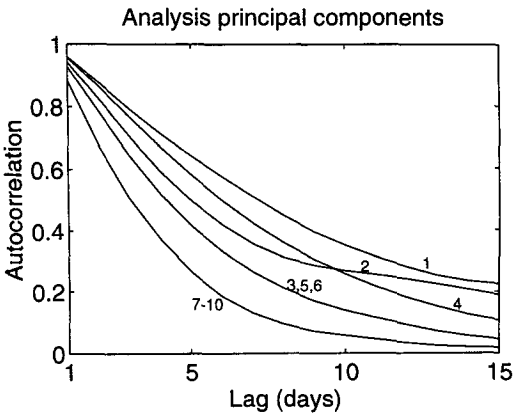

FIG. 12. Autocorrelation coefficients for the principal components (expansion coefficient time series) of the leading analysis EOFs, calculated at lags between 1 and 15 days. Values for EOFs 3, 5, and 6 have been averaged, as have those for EOFs 7–10.

The amplitude of the leading predictable pattern is relatively well forecast in both polarities, in agreement with the finding of Branstator et al. (1993). Over the whole dataset, the skill of the full forecast field increases with increasing forecast amplitude of the leading predictable pattern, approximately equally in either polarity, in apparent contrast with the results of Palmer (1988), O'Lenic and Livezey (1989), and Chen (1990, 1992). In fact, Fig. 10 suggests slightly below average rms errors for large negative amplitudes of the leading predictable pattern, the reverse of the situation described by Palmer.[3]

When the relationship between the "PNA index" and forecast error was investigated using the dependent/independent data subsets described in section 4, it was found to be sensitive to model changes. The same was true of the predictable pattern amplitude–forecast skill relationships described in the previous paragraph. The regression procedure described in section 2b of Palmer (1988) and section 3b of Palmer and Tibaldi (1988) was applied to D9 rms errors over Palmer's "region 8" (between 150° and 120°W), using the seven early winters' data. The resulting pattern was found to be very similar to that shown in Palmer and Tibaldi's Fig. 8c: its amplitude times series (the "regressed rms error") was found to be correlated at a level of 0.43 with the actual rms errors, consistent with

[2] The influence of variability on seasonal to interannual timescales was also apparent in the results of sensitivity tests involving the treatment of the annual march. Test calculations were performed both before and after removal of the climatological mean variability within the winter season. SVD and PrCA results were found to be insensitive to the inclusion or removal of the annual signal, with both procedures yielding a PNA-like leading mode in either case. However, when the anomalies are defined as departures from seasonal mean analysis and forecast fields, rather than from time-varying mean fields, the CCA-based patterns become dominated by the annual march in the position and intensity of the Asian jet, which is indeed the most predictable mode of variation in the analysis fields, but is of only minor interest in an operational setting, since it has little influence on day-to-day variations in forecast skill.

[3] Branstator (1994, personal communication) confirms that in the discussion section of Branstator et al. (1993), the reported dependence of forecast skill on the polarity of their EOF1 is also the reverse of Palmer's result, though it was mistakenly interpreted as being in agreement with it. The misunderstanding comes from a sign reversal, between the two papers, of what is deemed to be the "positive" PNA polarity.

Palmer and Tibaldi's result (their Table 1). When the data for the seven late winters were projected onto the same pattern, the correlation between the pattern amplitude and rms errors dropped to 0.10. These results do not refute the finding that the forecast polarity of the PNA pattern was an indicator of forecast skill over the North Pacific during the period examined by Palmer and Tibaldi, but they do raise questions as to whether this relationship would have been any use for predicting forecast skill in later seasons.

Scope for prediction of the probability of high forecast skill and for statistical correction of forecasts is limited due to the effect of changes in model formulation. It is possible that results for the independent data samples would be improved if the calculations could be repeated on a long record of forecasts and analyses derived from a fixed, state-of-the-art data assimilation/forecast system. Inclusion of a forecasting component into future data reanalysis efforts would provide valuable datasets for studies of predictability and model output statistics.

The D6–D10 average forecasts considered here showed levels of skill well above those of instantaneous D10 fields. To extend studies of this type out to the limits of existing predictability, one would have to repeat these calculations on longer range forecasts such as 5-day averages centered on D10. To provide datasets for such studies, it would be necessary to extend routine operational forecasts out to day 12, or perhaps even day 15. It may be worth considering whether more can be learned about predictability by creating a large (15+ years), homogeneous database consisting of medium-range (12 day) forecasts than by creating a much more limited database consisting of extended-range (e.g., 60 day) forecasts.

Acknowledgments. The ECMWF data used here were kindly provided by Dr. Tim Palmer at ECMWF, and an earlier version was provided through the National Center for Atmospheric Research by Drs. Grant Branstator and Andy Mai. This work has benefited from useful discussions with Drs. Tim Palmer and Gerhard Schmitz. Useful comments and suggestions from two anonymous reviewers helped to focus the discussion considerably. We are particularly appreciative of the two careful readings by Dr. Grant Branstator, who raised some very interesting and instructive points. Support was provided by the National Science Foundation through the Climate Dynamics Program Office under Grant ATM-9215512. J. Renwick was also supported by the New Zealand National Institute of Water and Atmospheric Research Ltd.

APPENDIX

Definition of CCA and SVDA Summary Statistics

The squared covariance fraction (SCF_k) is defined as

$$SCF_k = 1 - \frac{\|\mathbf{C}_{XY} - \tilde{\mathbf{C}}_k\|_F^2}{\|\mathbf{C}_{XY}\|_F^2},$$

where $\tilde{\mathbf{C}}_k$ is the contribution to the covariance matrix $\mathbf{C}_{XY}$ constructed using a given singular mode k [using the right-hand side of Eq. (4)] and $\| \ \|_F$ denotes the Frobenius matrix norm (Bretherton et al. 1992). It is easily shown that the square of the Frobenius norm of $\mathbf{C}_{XY}$ is equal to the sum of the squares of its singular values and that the SCF of any mode is the ratio of the square of its singular value to the sum of squares of all the singular values. Although this statistic is analogous to the fraction of variance explained, ranging in value from 0 to 1, the use of *squares* of singular values tends to exaggerate the importance of the leading modes.

For the purposes of more general comparisons from one analysis to another, Wallace et al. (1993) defined the normalized rms squared covariance

$$c_k = \left[\frac{\sigma_k^2}{\sum\sum s_i^2 s_j^2}\right]^{1/2},$$

where σ_k^2 is the square of the singular value of the kth mode and the denominator represents the product of the total variance in the left field (s_i^2) with that in the right field (s_j^2). The possible range of c_k is 0 to f, where f is the minimum of the respective fractions of variance accounted for by the leading EOF of either dataset. Hence, while the maximum possible value of c_k is 1, its value rarely exceeds 0.2.

The two other measures used are both correlation coefficients, spatial and temporal. From the definitions in (4), for SVDA or CCA mode k, the temporal correlation r_k between the time series of mode k is

$$r_k = \frac{1}{n}\mathbf{a}_k \cdot \mathbf{b}_k,$$

assuming that the time series $\mathbf{a}_k$ and $\mathbf{b}_k$ are normalized to zero mean and unit variance. The spatial correlation s_k between the patterns of mode k is

$$s_k = \frac{\frac{1}{n}\sum_i (u_{ki} - \bar{u}_k)(v_{ki} - \bar{v}_k)}{\sigma_{u_k}\sigma_{v_k}},$$

where u_{ki} is the ith element of left pattern $\mathbf{u}_k$, and $\bar{u}_k$ and σ_{u_k} are the mean and standard deviation of the elements of $\mathbf{u}_k$, respectively (similarly for right pattern $\mathbf{v}_k$). Both r_k and s_k range in value from -1 to $+1$.

REFERENCES

Barnett, T. P., and R. W. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.*, **115**, 1825–1850.

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

Bengtsson, L., 1991: Advances and prospects in numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **117,** 855–902.

Branstator, G., 1986: The variability in skill of 72-hour global-scale NMC forecasts. *Mon. Wea. Rev.,* **114,** 2628–2639.

——, A. Mai, and D. P. Baumhefner, 1993: Identification of highly predictable flow elements for spatial filtering of medium- and extended-range numerical forecasts. *Mon. Wea. Rev.,* **121,** 1786–1802.

Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate,* **5,** 541–560.

Charney, J. G., R. G. Fleagle, V. E. Lally, H. Riehl, and D. Q. Wark, 1966: The feasibility of a global observation and analysis experiment. *Bull. Amer. Meteor. Soc.,* **47,** 200–220.

Chen, W. Y., 1989: Another approach to forecasting forecast skill. *Mon. Wea. Rev.,* **117,** 427–435.

——, 1990: Interannual variability of skill of NMC medium-range forecasts over the Pacific/North America sector. *Mon. Wea. Rev.,* **118,** 179–188.

——, 1992: Dynamical prediction at medium-extended ranges employing low-frequency PNA mode. *Mon. Wea. Rev.,* **120,** 2641–2652.

Dalcher, A., and E. Kalnay, 1987: Error growth and predictability in operational ECMWF forecasts. *Tellus,* **39A,** 474–491.

Déqué, M., 1988: 10-day predictability of the Northern Hemisphere winter 500-mb height by the ECMWF operational model. *Tellus,* **40A,** 26–36.

Glahn, H. R., 1968: Canonical correlation and its relationship to discriminant analysis and multiple regression. *J. Atmos. Sci.,* **25,** 23–31.

Hoffman, R. N., and E. Kalnay, 1983: Lagged-averaged forecasting, an alternative to Monte Carlo forecasting. *Tellus,* **35A,** 100–118.

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

Kushnir, Y., and J. M. Wallace, 1989: Low-frequency variability in the Northern Hemisphere winter: Geographical distribution, structure and time-scale dependence. *J. Atmos. Sci.,* **46,** 3122–3142.

Leith, C. E., 1971: Atmospheric predictability and two-dimensional turbulence. *J. Atmos. Sci.,* **28,** 145–161.

——, and R. H. Kraichnan, 1972: Predictability of turbulent flow. *J. Atmos. Sci.,* **29,** 1041–1058.

Lilly, D. K., 1969: Numerical simulation of two-dimensional turbulence. *Phys. Fluids,* **12,** (Suppl. 2), 240–249.

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

——, 1982: Atmospheric predictability experiments with a large numerical model. *Tellus,* **34,** 505–513.

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

National Research Council, Panel on Dynamic Extended-Range Forecasting, 1991: Prospects for extending the range of prediction of the global atmosphere. Nat. Acad. Press, Washington, DC, 33 pp.

O'Lenic, E. A., and R. E. Livezey, 1989: Relationships between systematic errors in medium range numerical forecasts and some of the principal modes of low-frequency variability of the Northern Hemisphere 700 mb circulation. *Mon. Wea. Rev.,* **117,** 1262–1280.

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

——, and S. Tibaldi, 1988: On the prediction of forecast skill. *Mon. Wea. Rev.,* **116,** 2453–2480.

Shukla, J., 1981: Dynamical predictability of monthly means. *J. Atmos. Sci.,* **38,** 2547–2572.

Smagorinsky, J., 1969: Problems and promises of deterministic extended range forecasting. *Mon. Wea. Rev.,* **91,** 99–164.

Tibaldi, S., P. Ruti, E. Tosi, and M. Maruca, 1993: Operational predictability of winter blocking: An ECMWF update. *Proc. ECMWF Seminars on Validation of Forecasts and Large-Scale Simulations over Europe.* Reading, Berkshire, United Kingdom, ECMWF, 91–105.

van Loon, H., and J. C. Rogers, 1978: The seesaw in winter temperatures between Greenland and northern Europe. Part I: General description. *Mon. Wea. Rev.,* **106,** 296–310.

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.

——, C. Smith, and C. S. Bretherton, 1992: Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies. *J. Climate,* **5,** 561–576.

——, Y. Zhang, and K.-H. Lau, 1993: Structure and seasonality of interannual and interdecadal variability of the geopotential height and temperature fields in the Northern Hemisphere troposphere. *J. Climate,* **6,** 2063–2082.