PERSISTENT OSCILLATION PATTERNS IN CLIMATOLOGICAL FIELDS

Charles Kooperberg, Finbarr O'Sullivan University of Washington*
Charles Kooperberg, Dept of Stat GN-22, Univ of Washington, Seattle WA 98195

KEY WORDS: principal oscillation patterns, principal component analysis, forecast errors.

1. Introduction

In the fields of atmospheric science and oceanography, the study of the fluid motion is of considerable interest. For the most part the physical principals and the equations of motion describing the small scale or *microscopic* evolution of such systems is quite well established (e.g. Smagorinsky 1969 and Holton 1992). However summary characterization of the *macroscopic* behavior of the system in terms of spatial and temporal variation is typically rather difficult to predict, except for simple problems such as the wave equation in a homogeneous medium. Thus there has been substantial interest in the development of empirical methods which are capable of describing the phenomenological aspects of fluid motion in more macroscopic terms.

The empirical identification of spatial patterns of oscillations has been realized using a collection of statistical of techniques including principal components analysis, canonical correlations analysis and first order Markov modeling (Barnett and Preisendorfer 1987, Hasselmann 1988, Von Storch *et al.* 1988, Bretherton *et al.* 1992). Although these methods are well established, for a given data set it is often unclear which technique and in what form is most appropriate. As a result there is often a substantial degree of subjectivity involved in the application of these methods. In this paper we propose a statistical methodology based on a forecast error criterion. The approach involves constructing 'optimal' one-step ahead forecast rules on half the data and then computing the performance of those rules on the remaining half of the data.

In this paper, we apply various methods to a 47 year record (January 1946 through June 1993) of Northern Hemisphere extra-tropical geopotential (500 hPa) height based on daily operational analyses from the National Meteorological Center (NMC). The full resolution NMC grid is a 1977 point octagonal superimposed on a polar stereographic projection of the Northern Hemisphere, extending to 20 degrees north (Wallace *et al.* 1992). The data are projected onto a 445-point half-resolution grid. The series was made approximately stationary by removing the mean field and the first three harmonics of the annual cycle, separately for every grid point. The data were then time-averaged to produce five-day averages. In this paper we examine the complete series of 3467 five-day averages.

We begin by presenting a brief description of two methods to determine the spatial patterns, empirical orthogonal functions (EOF, also known as principal components analysis, PCA), and principal oscillation patterns (POPs). In Section 3 we describe the idea of the one-step ahead prediction error, as well as its implementation. In Section 4 we describe a method that is intended to minimize the one step ahead prediction error. The ideas presented in this paper are discussed in more detail in two related papers (Kooperberg and O'Sullivan 1994a, 1994b).

2. Spatial-Temporal Decompositions

In the following we assume that the climatological field being studied has been preprocessed, and that it can be assumed to be stationary. In particular, let $x(t)$, $t = 1, 2, \ldots, T$, be the vector of length p, representing the value of the (preprocessed) field at the (discrete) time t.

2.1. Empirical Orthogonal Functions

Principal components analysis is a well established multivariate statistical method widely used in a variety of applications (e.g. Mardia *et al.* 1979). Its development for multivariate time series analysis owes

*Charles Kooperberg was supported in part by NSF grant DMS-94.03371. Finbarr O'Sullivan was supported in part by NIH grant CA-57903. The authors want to thank Professor John M. Wallace, Gregor Nitsche and James Renwick of the University of Washington, Department of Atmospheric Sciences for encouragement and many helpful discussions.

much to the work of of Brillinger and Priestley and co-authors (Brillinger 1981, Priestley, 1987). In the atmospheric science literature principal components are more frequently known as empirical orthogonal functions (Barnett and Preisendorfer 1987, Jolliffe 1986, Lorenz 1956). Generalizations of EOF for the examination of coupled in fields, include canonical correlation analysis and singular value decomposition analysis (Barnett and Preisendorfer 1987, Bretherton *et al.* 1991, Wallace *et al.* 1991).

Empirical orthogonal function analysis (EOF, also known as principal component analysis) is based on the spectral or eigenvalue decomposition of the (marginal) covariance of the measured field, *i.e.* $\Sigma_0 = \mathrm{Var}(x(t))$. The covariance matrix is Hermetian, its spectral decomposition is

$$\Sigma_0 = UDU^* = \sum_{j=1}^{p} d_j U_j U_j^*.$$

The symbol $*$ denotes the transpose of the complex conjugate. The matrix U is orthonormal, with columns U_j for $j = 1, 2 \ldots p$, and D is a diagonal matrix with as elements the eigenvalues of Σ_0, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$. The columns of U represent orthogonal linear combinations of the process with maximal marginal variance. Projecting onto the first K columns gives

$$x(t) = \sum_{k=1}^{K} U_k z_k(t) + \eta_K(t), \qquad (1)$$

where $z_k(t) = U_k^* x(t)$ and $\eta_K(t) = \sum_{j>K} U_j z_j(t)$. Note that $z_k(t)$ is a scalar, while $\eta_K(t)$ is a vector of length p. For K sufficiently large, the total marginal variance of η_K can be made small. With this we have a representation of the process in terms of a sum of a fixed number of K spatial patterns (U_k's) forced by statistically uncorrelated (marginally) temporal oscillation patterns $z_k(t)$.

In Figure 1 we show the two EOFs, U_1 and U_2 as well as the corresponding temporal oscillation patterns $z_1(t)$ and $z_2(t)$ for the 500 hPa geopotential height field for the first 1700 five day intervals (1946-1969). Note that although the EOF analysis ignores the temporal correlations in the data, the spectrum of the temporal patterns, $z_k(t)$ is red, as can be seen from the spectral estimates and autocorrelations.

The first two EOFs are recognized as mixtures of the North Atlantic Oscillation (NAO) and the Pacific North American (PNA) pattern. These patterns are well established in the atmospheric sciences (Wallace and Gutzler 1981).

2.2. Principal Oscillating Patterns

This technique, also known as POPs, was introduced by Hasselmann (1988) and by Von Storch *et al.* (1988) and has become widely used in the atmospheric science literature. The method is based on an AR(1) representation for the process

$$x(t) = Bx(t-1) + \epsilon(t)$$

where $\epsilon(t)$ is assumed to be a mean zero uncorrelated (in time) process. The coefficient matrix B is expanded in terms of its eigenvectors, U_j, and corresponding eigenvalues λ_j, both of which may be complex. The eigenvalues are ordered by modulus, $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_p|$.

Typically, the coefficient matrix B is estimated by $\hat{B} = \Sigma_0^{-1}\Sigma_1$. In particular, the eigenvalue decomposition of $\hat{B}$ is now

$$\hat{B} = \Sigma_1 \Sigma_0^{-1} = UDU^* = \sum_{j=1}^{P} d_j U_j U_j^*. \qquad (2)$$

Typically, some of the eigenvectors of $\hat{B}$ are real, while some pairs may be complex. Van Storch *et al.* (1988) interpret the real eigenvectors as standing oscillatory patterns while they interpret the complex eigenvectors as propagating patterns.

Note that in contrary to the spatial patterns for EOF, the eigenvectors for $\hat{B}$ are not orthonormal. However, it is still straightforward to decompose $x(t)$ in a similar fashion as (1).

The patterns that are obtained from (2) are often not very satisfactory. This is because the estimate of Σ_0 is often quite noisy, in fact, it is often almost singular. Thus when inverted, the smallest, and least interesting, eigenvalues and eigenvectors of Σ_0 dominate the decomposition. To circumvent this problem, it is customary to first project the data on the first k empirical orthogonal functions of $x(t)$, before applying the canonical correlations algorithm. This is known as the Barnett-Preisendorfer (BP) approach (Barnett and Preisendorfer 1987).

The interpretations of the spatial patterns, U_j's, in EOF and POPs are quite different. The EOF patterns are defined without regard to the temporal structure of their forcing terms z_k's. Indeed there is the tacit assumption that these terms are serially uncorrelated. By contrast, the POPs analysis, are attempting to extract patterns whose forcing terms have a richer and potentially predictable stochastic evolution characteristic. In the case that the process is rescaled so that the marginal covariance is the identity, the results of the POPs and EOF tend to be more complementary.

In Figure 2 we show the first two principal oscillation patterns for the first half of the complete series, both of which turned out to be real. As can be seed from this figure, the POPs patterns are considerably more noisy than the EOF patterns.

3. The one-step ahead forecast error

Which method to use? Suppose that one is analyzing a climatological field, should one use EOF, CCA, POPs, BP-CCA or BP-POPs to compute the spatial patterns that determine a spatial-temporal decomposition? Another question is how many patterns one should consider? Although one would typically only interpret a few patterns, for a larger climatological modeling effort one would want to include any patterns that have more signal than noise. Indeed, the number of spatial oscillation patterns, K, acts as a *regularization* parameter. If K is too small there will be substantial bias in the model; on the other hand if K is too large the model will tend to overfit the data. As one approach in answering these two questions we consider the one-step ahead forecast error.

Let $x(t)$ be a climatological field, and let U be a $p \times p$ matrix, containing the spatial patterns as columns in decreasing order of importance. Let W_K be the $p \times K$ matrix consisting of the first K columns of U. Let $z^K(t)$ be the least squares projection of $x(t)$ on W_K.

Suppose that we know the temporal oscillation patterns z^K up to time t, and we wish to predict the complete field at time $t+1$, $x(t+1)$. A measure of how good of a summary of the field z^K and W_K are would be the difference between the prediction $\tilde{x}(t+1)$ of $x(t+1)$ on the basis of the information contained in the projection of $x(t)$ onto the patterns W_K, *i.e.* $z^K(s)$ for $s \leq t$. To formalize this, let $FE(K,U)$ be the expected value of the squared norm of the prediction error, that is, $\mathrm{E}||\tilde{x}(t+1) - x(t+1)||^2/\mathrm{E}||x(t+1)||^2$. The interpretation of $FE(K,U)$ is, that it is the fraction of the variance that cannot be explained if we forecast one-step ahead based only upon the information contained in the projection of $x(t)$ on W_K for $s \leq t$.

This definition of the forecast error is motivated by the desire to identify patterns with some degree of temporal persistence, since those patterns seem best suited for tasks like numerical weather forecasting. We could thus answer the questions posed at the beginning of this section using the one-step ahead forecast error: when we have to choose between two decompositions U and U', both with K patterns based on the forecast error, we choose U over U' if $FE(K,U) < FE(K,U')$. Similarly, if we have to determine how many spatial patterns K of a decomposition U we are going to use, we might choose that value of K that minimizes $FE(K,U)$.

To apply the idea of minimizing the one-step ahead forecast error, the patterns are estimated on the first half of the data and a corresponding optimized linear one-step ahead forecast rule is also estimated there. Subsequently, this forecast rule is evaluated on the second half of the data providing an estimate of the predictive performance of the model. The approach is described in detail in Kooperberg and O'Sullivan (1994a).

4. Predictive Oscillation Paterns

As an alternatice to EOFs or POPs we can define the patterns such that the one-step ahead forecast error is minimized. We refer to Kooperberg and O'Sullivan (1994b) for the details. The main ingredient in the derivation is Kolmogorov's formula (see for example Chapter 10 of Priestley, 1987), which states that the minimal forecast error for $z(t)$ is

$$\mathrm{E}[(z(t+1)-\tilde{z}(t+1))^2] = 2\pi \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log h(\omega)d\omega\right\},$$

where $h(\omega)$ is the spectral density of $z(t)$. The first predictive oscillation pattern A is now found as the maximizer of

$$\frac{A'\widehat{\Sigma}_0 A}{A'A}\left[I - 2\pi\exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left(\frac{A' f_{xx}^{(h)}(\omega)A}{A'\bar{f}_{xx}^{(h)}A}\right)d\omega\right\}\right]$$

where $\widehat{\Sigma}_0$ is the variance covariance matrix of the field and $f^{(h)}$ is a consistent estimate of the spectral density matrix of the field. Higher PrOPs, orthogonal to the earlier PrOPs, are derived in a similar fashion.

In Figure 3 we give the spatial patterns for the first two PrOPs. As can be seen, the first PrOP is very similar to the first EOF. The second PrOP is actually a combination of the second and the third EOF. Here we should keep in mind that while the EOF decomposition ignores the temporal dependence of the field, the PrOP decomposition takes this into account. The similaroty between the PrOP and EOF patterns suggests that the forecast errors based on these two decompositions may be very similar. And indeed, in Figure 4 we notice that the one-step ahead forecast errors for EOF and PrOP are indistinguishable. Both of these methods produce much smaller errors than POP does.

5. Discussion

In the atmospheric sciences there has been considerable interest in the development of empirical methods which are capable of decomposing spatial and temporal variation into a small number of fixed patterns. Techniques that are used include empirical orthogonal functions, canonical correlation analysis and principal oscillation patterns. All of these methods have their advantages. Patterns that are obtained using empirical orthogonal functions explain a large fraction of the variance of the field, however, the spatial patterns need not reflect anything to do with the evolution of the field, which is often of prime interest for modeling purposes. On the other hand, canonical correlations analysis and principal oscillation patterns focus on components with a strong temporal auto-correlation pattern, but there is no constraint that the patterns are necessarily highly correlated with the original field.

In this paper we discuss a criterion that can be a tool when comparing different methods. The one-step ahead forecast error measures how well one is able to predict the next instance of the field, based on the summary patterns up to now. Clearly a good prediction requires both high correlation between the patterns and the field and a strong temporal auto-correlations of the patterns. The decomposition based on predictive oscillation patterns, described in detail in Kooperberg and O'Sullivan (1994b) minimizes the one-step ahead prediction error.

When applied to a 47 year record of Northern Hemisphere extra-tropical geopotential height, we found that EOF and PrOP performed equally well, with respect to the one-step ahead forecast error, while both performed better than POPs.

References

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

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.

Brillinger, D. R., 1981: *Time Series: Data Analysis and Theory.* Holden Day, 540pp.

Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93 D9**, 11,015–11,021.

Holton, J. R., 1992: *An introduction to dynamic meteorology, 3rd ed.* Academic Press, 511pp.

Jolliffe, I.T., 1986: *Principal Component Analysis*, Springer, 271pp.

Kooperberg, C. and F. O'Sullivan, 1994a: The use of a statistical forecast criterion to evaluate alternative empirical spatial oscillation pattern decomposition methods in climatological fields. *Technical Report No. 276, Department of Statistics, University of Washington, Seattle.*

Kooperberg, C. and F. O'Sullivan, 1994b: Predictive Oscillation Patterns: a synthesis of methods for spatial-temporal decomposition of random fields. *Technical Report No. 277, Department of Statistics, University of Washington, Seattle.*

Lorenz, E. N., 1956: *Empirical orthogonal functions and statistical weather prediction.* MIT, Department of Meteorology, Sci. Rep. 1, 49pp.

Mardia, K. V., J. T. Kent, and J. M. Bibby, 1979: *Multivariate Analysis.* Academic Press, 521pp.

Priestley, M. B., 1987: *Spectral Analysis and Time Series.* Academic Press, 890pp.

Smagorinsky, J., 1969: Problems and promises of deterministic extended range forecasting. *Bul. Amer. Meteor. Soc.*, **50**, 286–310.

Von Storch, H., T. Bruns, I. Fischer-Bruns, and K. Hasselmann, 1988: Principal oscillation pattern analysis of the 30- to 60-day oscillation in general circulation model equatorial troposphere. *J. Geophys. Res.*, **93 D9**, 11,022–11,036.

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

Wallace, J. M., 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.

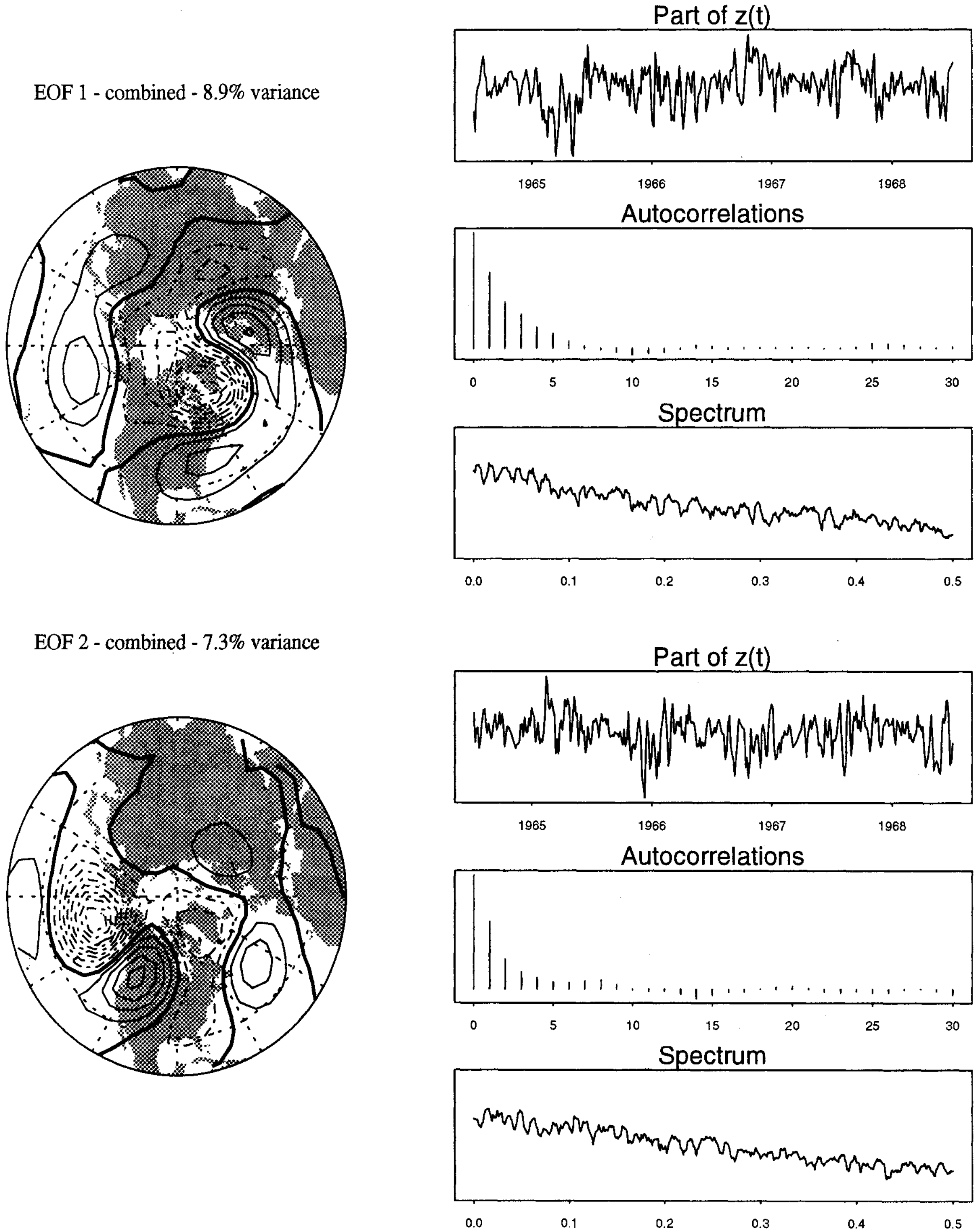

Figure 1. The first two empirical orthogonal functions and their corresponding temporal oscillation patterns for the geopotential height. On the right side of the figure we show for a part of $z_k(t)$, as well as the first 30 autocorrelations and an estimate of the log-spectral density of $z_k(t)$.

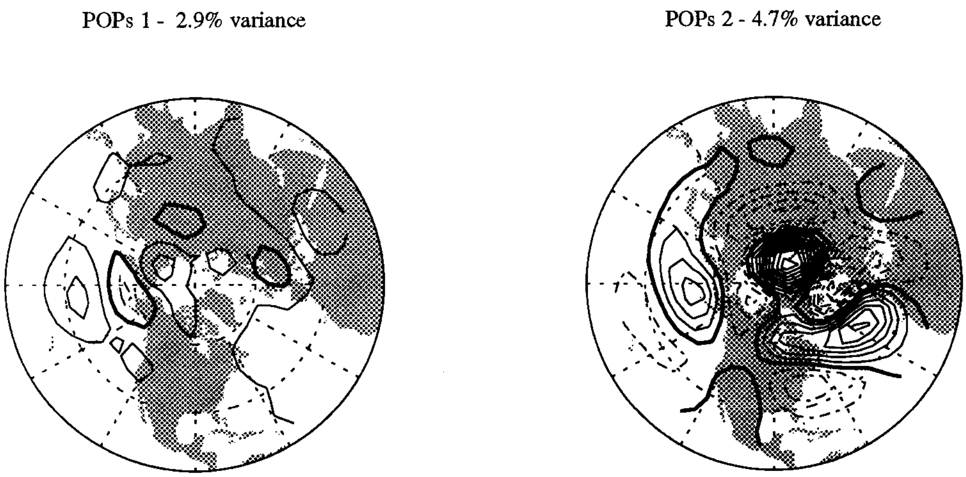

Figure 2. The first two principal oscillation patterns for the geopotential height.

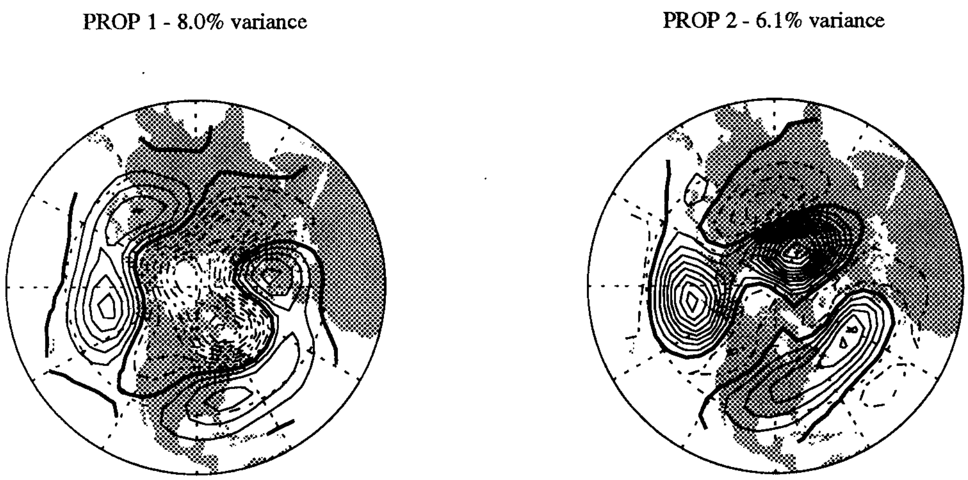

Figure 3. The first two predictive oscillation patterns for the geopotential height.

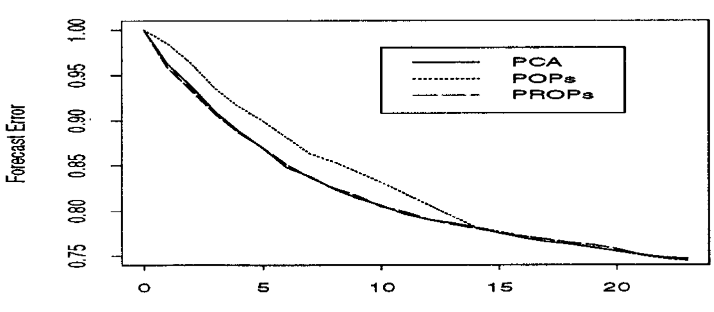

Figure 4. Forecast error performance of the alternative decomposition methods as a function of the number of patterns included. There is only a slight improvement in the forecast error after 25 components - hence these are not included.