

Review

Empirical orthogonal functions and related techniques in atmospheric science: A review

A. Hannachi,* I. T. Jolliffe and D. B. Stephenson

Department of Meteorology, University of Reading, Reading RG6 6BB, UK

Abstract:

Climate and weather constitute a typical example where high dimensional and complex phenomena meet. The atmospheric system is the result of highly complex interactions between many degrees of freedom or modes. In order to gain insight in understanding the dynamical/physical behaviour involved it is useful to attempt to understand their interactions in terms of a much smaller number of prominent modes of variability. This has led to the development by atmospheric researchers of methods that give a space display and a time display of large space-time atmospheric data.

Empirical orthogonal functions (EOFs) were first used in meteorology in the late 1940s. The method, which decomposes a space-time field into spatial patterns and associated time indices, contributed much in advancing our knowledge of the atmosphere. However, since the atmosphere contains all sorts of features, e.g. stationary and propagating, EOFs are unable to provide a full picture. For example, EOFs tend, in general, to be difficult to interpret because of their geometric properties, such as their global feature, and their orthogonality in space and time. To obtain more localised features, modifications, e.g. rotated EOFs (REOFs), have been introduced. At the same time, because these methods cannot deal with propagating features, since they only use spatial correlation of the field, it was necessary to use both spatial and time information in order to identify such features. Extended and complex EOFs were introduced to serve that purpose.

Because of the importance of EOFs and closely related methods in atmospheric science, and because the existing reviews of the subject are slightly out of date, there seems to be a need to update our knowledge by including new developments that could not be presented in previous reviews. This review proposes to achieve precisely this goal. The basic theory of the main types of EOFs is reviewed, and a wide range of applications using various data sets are also provided. Copyright © 2007 Royal Meteorological Society

KEY WORDS empirical orthogonal functions; simplified EOFs; extended EOFs; complex EOFs; North Atlantic Oscillation; Madden Julian oscillation; Quasi-biennial oscillation

Received 4 May 2006; Revised 22 December 2006; Accepted 27 December 2006

INTRODUCTION

Climate is regarded as the aggregation of (random) daily weather, and as pointed out by Lorenz (1970), climate may be defined, in mathematical terms, as the collection of all long-term statistical properties of the atmospheric state. It is therefore the long-term statistics of weather. Heinlein (1973) says ‘climate is what we expect but weather is what we get’ (this quotation is in the section ‘More from the Notebooks of Lazarus Long’ of Robert A. Heinlein’s novel, but some sources, however, attribute it to the American writer/lecturer Samuel L. Clemens known by the pen name Mark Twain (1835–1910)). Climate variations are also the result of exceedingly complex non-linear interactions between very many degrees of freedom or modes. Both weather and climate are

characterised by non-linearity and high dimensionality. Consequently, a challenging task is to find ways to reduce the dimensionality of the system to a few modes if possible. A further, yet challenging, task is to link these modes to the dynamics/physics of the system.

Empirical orthogonal function (EOF) analysis (Fukuoka, 1951; Lorenz, 1956) is among the most widely and extensively used methods in atmospheric science. The method is in essence an exploratory (i.e. non-model orientated) tool, which allows a time display and a space display of the space-time field that may be useful to the atmospheric scientist. The existence of fast and efficient algorithms also helped its widespread use. EOFs are multipurpose and have been used for example in dimensionality reduction and patterns extraction.

Since the early review of Kutzbach (1967) on EOFs, and apart from a few textbooks (Preisendorfer, 1988; von Storch and Zwiers, 1999; Jolliffe, 2002; Wilks, 2006), the subject has not been systematically reviewed in

*Correspondence to: A. Hannachi, Department of Meteorology, King Abdulaziz University, PO Box 80208, Jeddah 21589, Saudi Arabia.
 E-mail: ahannachi@kau.edu.sa

the atmospheric science literature to address the various recent developments in the field. Given the importance of EOFs and related methods in climate research, we believe that the limited literature reviews on the subject do not meet the need of climate researchers. This introductory review on EOFs is a contribution to fill in this gap, but is by no means exhaustive. Further references and more detailed material on the subject will be provided as we go through the text. The manuscript is intended for research students and also researchers starting in the field of weather/climate analysis, and can be used for educational purpose.

Given any space-time meteorological field, EOF analysis finds a set of orthogonal spatial patterns along with a set of associated uncorrelated time series or principal components (PCs). The geometrical constraints characterising EOFs and PCs can be very useful in practice since the covariance matrix of any subset of retained PCs is always diagonal. These same constraints, however, can also be restrictive in other contexts. Take spatial orthogonality, for instance. Because it is a global property, the orthogonality constraint can cause the EOFs to have structures over most of the domain and with significant amplitude, when in fact one expects the patterns to be more localised. Horel (1981), for example, points out that if the first EOF has a constant sign over its domain then the second one will generally have both signs with the zero line going through the maxima of the first EOF. This also yields the problem of domain-dependence and non-locality (Horel, 1981; Richman, 1986, 1987). These problems can cause difficulties in interpreting the obtained patterns (Amabaum *et al.*, 2001, 2002; Dommegård and Latif, 2002; Jolliffe *et al.*, 2003) because physical modes are not necessarily orthogonal. Normal modes derived, for example, from linearised dynamical/physical models, such as barotropic models (Simmons *et al.*, 1983) are not orthogonal since physical processes are not uncorrelated.

The previous prevailing difficulties associated with interpreting EOFs have led researchers to develop tools to overcome these difficulties. Linear transformations of EOFs, based on rotation, have been introduced and yield the concept of rotated empirical orthogonal functions (REOFs) (Horel, 1981; Richman, 1981, 1986; Cheng *et al.*, 1995, Xinhua and Dunkerton 1995). The REOF method yields in general localised structures by compromising some of the EOFs' geometric properties such as orthogonality. Rotation attempts to yield simpler patterns than EOFs. In rotation various decisions have to be made. For instance, the rotation, or simplicity, criterion is not unique. In addition, the number of EOFs used for rotation cannot be fixed a priori but remains arbitrary. Other alternatives to rotation have been proposed. A particularly interesting one, simplified EOFs, is a method based on the least absolute shrinkage and selection operator (LASSO) approach introduced in the context of regression by Tibshirani (1996) and adapted to EOFs by Jolliffe *et al.* (2003). The method attempts to achieve simultaneously the desirable property of large variance and simplicity. The method can yield loadings,

i.e. EOF elements, that are identically zero, resulting in localised structures without any rotation.

EOFs and REOFs are mainly based on using the spatial correlation of the field, an important feature of climate data. Auto- and cross-correlation in time between grid points, however, are ignored in those techniques. Extended EOF analysis (Weare and Nasstrom, 1982) is a technique that attempts to incorporate both the spatial and the temporal correlation. The method has, since its introduction, become a useful tool to extract dynamical structure, e.g. trends, oscillations, propagating structures, and to filter data (Broomhead and King, 1986a,b; Fraedrich, 1986a,b; Kimoto *et al.*, 1991; Plaut and Vautard, 1994).

Another related method that attempts to find propagating patterns is complex Hilbert empirical orthogonal function (HEOF) analysis. The frequency domain empirical orthogonal function (FDEOF) method finds EOFs based on the cross-spectrum matrix averaged over a specific small frequency band (Wallace and Dickinson, 1972; Wallace, 1972; Brillinger, 1981). FDEOFs generalise conventional EOFs in the sense that the covariance matrix used for EOFs is only related to the real part of the cross-spectrum matrix and hence does not use the whole information from the complex cross-spectrum matrix. The HEOF method (Rasmusson *et al.*, 1981; Barnett, 1983; Horel, 1984; von Storch and Zwiers, 1999) is an alternative to FDEOFs and uses the Hilbert, or quadrature transform of the field. This transform allows HEOFs to deal with propagating structures/waves in the time domain using complexified fields.

The manuscript reviews the *exploratory* methods mentioned above. Various other extensions to EOF/PC analysis are briefly discussed toward the end, and references are provided for further details. Other methods such as principal oscillation patterns (POPs) and principal interaction patterns, which are model-orientated, i.e. non-exploratory, methods are not discussed here. Also, methods involving covariability between two or more fields such as in coupled patterns, e.g. canonical correlation analysis (Bretherton *et al.*, 1992), are not presented here. The manuscript is organised as follows. Section 2 reviews the concept of EOFs with application to winter monthly sea level pressure (SLP) reanalyses. Section 3 presents ways of simplification of EOFs, focussing mainly on rotated and simplified EOFs, with application to winter SLP. Extended EOFs with application to outgoing long wave radiation (OLR) are presented in section 4 while section 5 deals with complex EOFs with application to the quasi-biennial oscillation (QBO). Section 6 briefly discusses various other extensions to EOF analysis including very recent ones. A summary and conclusions are presented in the final section.

EOFs

Historical background

EOFs have been used in atmospheric science since the late 1940's by Obukhov (1947, 1960), Fukuoka (1951),

Lorenz (1956), and Kutzbach (1967). See, for example, Craddock (1973) for a discussion of eigenanalysis in meteorology. Since then EOFs have become popular analysis tools and are widely used in climate research. EOF techniques have their roots in social science, and go back to Pearson (1902), and later to Hotelling (1933, 1935) who introduced principal component analysis (PCA), the more common name for EOF analysis. EOFs, however, are not restricted to multivariate statistics or atmospheric sciences. They extend to feature extraction (Fukunaga and Koontz, 1970) and the analysis of stochastic fields in the mathematical literature where they are known under the name Karhunen-Loëve basis functions (Loëve, 1978). The original aim of EOFs (Obukhov, 1947; Fukuoka, 1951; Lorenz, 1956) was to achieve a decomposition of a continuous space-time field $X(t, \mathbf{s})$, where t and \mathbf{s} denote respectively time and spatial position, as

$$X(t, \mathbf{s}) = \sum_{k=1}^M c_k(t) u_k(\mathbf{s}), \quad (1)$$

where M is the number of modes contained in the field, using an optimal set of basis functions of space $u_k(\mathbf{s})$ and expansion functions of time $c_k(t)$. In practice the EOF/PCA technique aims at finding a new set of variables that capture most of the observed variance from the data through linear combinations of the original variables.

The EOF terminology is due to Lorenz (1956) who applied it in a forecasting project at the Massachusetts Institute of Technology. The method, however, had been applied in meteorology a decade earlier by Obukhov (1947) for smoothing purposes, and was mentioned by Fukuoka (1951) in a forecasting context, see Craddock (1973) for a little further historical account. In addition to smoothing and prediction, EOFs have also been used to reduce the large number of variables of the original data to a few variables, but without compromising much of the variability of the data (e.g. Hannachi and O'Neill, 2001.) Recently EOF analysis has been used to extract individual modes of variability that can be physically relevant such as the Arctic Oscillation (AO), (Pavan *et al.*, 2000), known as teleconnections (Angström, 1935; Bjerknes, 1969; Wallace and Gutzler, 1981; Wallace and Thompson 2002, etc.) Today, EOF methods are commonly used in most meteorological centres to compare observations and reanalyses to climate model simulations.

EOFs have been extensively studied in the literature, and for a detailed analysis the reader is referred to the following textbooks, mostly orientated toward atmospheric science applications: Preisendorfer (1988), von Storch and Zwiers (1999), and Wilks (2006). For more general application of PCA analysis, the reader is referred, e.g. to the textbooks by Seal (1967), Morrison (1976), Anderson (1984), Chatfield and Collins (1989), Mardia *et al.* (1979), Krzanowski (2000), Jackson (1991), and Jolliffe (2002) and more references therein.

Data formatting

We suppose that we have a gridded data set composed of a space-time field $X(t, \mathbf{s})$ representing the value of the field X , such as SLP, at time t and spatial position \mathbf{s} . The value of the field at discrete time t_i and grid point \mathbf{s}_j is denoted x_{ij} for $i = 1, \dots, n$ and $j = 1, \dots, p$. The observed field is then represented by the data matrix:

$$X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \quad (2)$$

where $\mathbf{x}_t = (x_{t1}, x_{t2}, \dots, x_{tp})^T$, $t = 1, \dots, n$, represents the map, or the value of the field at time t . Let us denote by $\bar{x}_{\cdot i}$ the time average of the field at the i 'th spatial grid point. This time average is given by:

$$\bar{x}_{\cdot i} = \frac{1}{n} \sum_{k=1}^n x_{ki}. \quad (3)$$

The climatology of the field is defined by

$$\bar{\mathbf{x}} = (\bar{x}_{\cdot 1}, \dots, \bar{x}_{\cdot p}) = \frac{1}{n} \mathbf{1}_n^T X \quad (4)$$

where $\mathbf{1}_n = (1, \dots, 1)^T$ is the (column) vector of length n containing only ones. The anomaly field, or departure from the climatology is defined at (t, \mathbf{s}_k) , $t = 1, \dots, n$, and $k = 1, \dots, p$, by:

$$x'_{tk} = x_{tk} - \bar{x}_{\cdot k} \quad (5)$$

or in matrix form:

$$X' = X - \mathbf{1}_n \bar{\mathbf{x}} = \left(I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) X = H X \quad (6)$$

where I_n is the $n \times n$ identity matrix, and H is the centring matrix of order n (Mardia *et al.*, 1979). To keep the notation simple, from now on and unless otherwise stated, the dash in (6) will be dropped and X will simply denote the anomaly data matrix.

Formulation and computation of EOFs

We present below a description of how to obtain EOFs, and for more details the reader is referred, for example, to von Storch (1995), von Storch and Zwiers (1999), Jolliffe (2002), and Wilks (2006). Once the anomaly data matrix (6) is determined, the sample covariance matrix is then defined by:

$$S = \frac{1}{n} X^T X, \quad (7)$$

which contains the covariances s_{ij} , $i, j = 1, \dots, p$, between the time series of the field at any pair of grid points $(\mathbf{s}_i, \mathbf{s}_j)$, i.e.

$$s_{ij} = [S]_{ij} = \frac{1}{n} \sum_{t=1}^n x_{ti} x_{tj}. \quad (8)$$

The aim of EOF analysis/PCA is to find uncorrelated linear combinations of the different variables that explain maximum variance, that is to find a unit-length direction $\mathbf{u} = (u_1, \dots, u_p)^T$ such that $X\mathbf{u}$ has maximum variability. This readily yields:

$$\max(\mathbf{u}^T S \mathbf{u}), \text{ s.t. } \mathbf{u}^T \mathbf{u} = 1 \quad (9)$$

The EOFs are therefore obtained as the solution to the eigenvalue problem:

$$S\mathbf{u} = \lambda^2 \mathbf{u} \quad (10)$$

The k 'th EOF is simply the k 'th eigenvector \mathbf{u}_k of S . The corresponding eigenvalue λ_k^2 , $k = 1, \dots, p$ is then

$$\lambda_k^2 = \mathbf{u}_k^T S \mathbf{u}_k = \frac{1}{n} \|X\mathbf{u}_k\|^2 \quad (11)$$

and hence gives a measure of the variance of the data accounted for in the direction \mathbf{u}_k . After finding the eigen elements of the sample covariance matrix S in (9), the eigenvalues are normally sorted in decreasing order as $\lambda_1^2 \geq \lambda_2^2 \dots \geq \lambda_p^2$. It is usual to write the variance accounted for in percentage as:

$$\frac{100\lambda_k^2}{p} \%. \quad (12)$$

$$\sum_{k=1}^p \lambda_k^2$$

The projection of the anomaly field X onto the k 'th EOF $\mathbf{u}_k = (u_{k1}, u_{k2}, \dots, u_{kp})^T$, i.e. $\mathbf{a}_k = X\mathbf{u}_k$ is the k 'th PC whose elements a_{tk} , $t = 1, \dots, n$, are given by:

$$a_{tk} = \sum_{j=1}^p x_{tj} u_{kj}. \quad (13)$$

So the k 'th eigenvalue λ_k^2 represents the variance of the k 'th PC $\mathbf{a}_k = (a_{1k}, a_{2k}, \dots, a_{nk})^T$. The relationship between Eq (13) and Eq (1) can now be noted. The time function $c_k(t)$ and the space function $u_k(\mathbf{s})$ in (1) are represented by x_{tj} and u_{kj} in (13) respectively. In various literatures the EOFs are also known as the PC loadings, and sometimes simply PCs. The PCs on the other hand are also known as EOF expansion coefficients, EOF amplitudes, PC time series, and PC scores. In this manuscript we use the terminology EOFs and PCs for the spatial and temporal patterns respectively.

In practice we do not need to compute the covariance matrix (7) and solve the eigenvalue problem (10). We use a powerful tool from linear algebra namely the singular value decomposition (SVD), (Golub and van Loan, 1996). Any $n \times p$ data matrix X can be decomposed as:

$$X = A \Lambda U^T. \quad (14)$$

In (14) A and U are respectively $n \times r$ and $r \times p$ unitary matrices, i.e. $U^T U = A^T A = I_r$ where $r \leq \min(n, p)$ is the rank of X and I_r is the identity matrix of order r .

The matrix Λ is diagonal, i.e. $\Lambda = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_r)$. The diagonal elements $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \geq 0$ of Λ are the singular values of X . The columns $\mathbf{a}_1, \dots, \mathbf{a}_r$ of A , and $\mathbf{u}_1, \dots, \mathbf{u}_r$ of U are respectively the left and right singular vectors of the data matrix X . There are other ways to express the SVD but (14) provides a compact representation because it drops unnecessary zero singular values.

The application of the SVD to the data matrix $\sqrt{n}X$ yields for the covariance matrix (7) the decomposition:

$$S = U \Lambda^2 U^T, \quad (15)$$

where $\Lambda^2 = \text{Diag}(\lambda_1^2, \lambda_2^2, \dots, \lambda_r^2)$ and where the singular values have been sorted in decreasing order. Note that the constant n has been absorbed by the diagonal matrix Λ^2 . The EOFs $\mathbf{u}_1, \dots, \mathbf{u}_r$ and the PCs $\mathbf{a}_1, \dots, \mathbf{a}_r$ are therefore the right and left singular vectors of the data matrix anomaly X . Note that for the decomposition (14) to be efficient in computation the data matrix X has to be transposed to yield $\min(n, p)$ as its first dimension. The EOFs are therefore orthogonal and the PCs uncorrelated, and this is a major characteristic of conventional EOFs. The orthogonality is a useful property since it provides a complete basis for the data matrix. Equation (14) yields in fact the decomposition:

$$X = \sum_{k=1}^r \lambda_k \mathbf{a}_k \mathbf{u}_k^T. \quad (16)$$

Component-wise, the previous decomposition expresses the map $\mathbf{x}_t = (x_{t1}, x_{t2}, \dots, x_{tp})^T$ of the field X at time t by

$$\mathbf{x}_t = \sum_{k=1}^r \lambda_k a_{tk} \mathbf{u}_k \quad (17)$$

where a_{tk} is the element of the k 'th PC \mathbf{a}_k at time t . Note again the link between Eq (17), which is simply the vector form of Eq (13), to Eq (1).

Equation (17) is particularly useful when EOFs are used to reduce the dimensionality of the data. This can be achieved simply by truncating the above sum by keeping, say, the first M terms where M is generally much smaller than the rank r of X . There is no universal rule, however, for truncation, and the choice of M is in general arbitrary. In practice the truncation order is often obtained by fixing the amount of represented variance, e.g. 80%, and choosing the set of the M leading EOFs that explain altogether at least this amount of variance.

The spectrum of the covariance matrix S composed of the eigenvalues $\lambda_1^2, \dots, \lambda_r^2$ provides information on the distribution of power (energy) as a function of scale, and on the separation/degeneracy of the EOF patterns. For example, high/low power are associated respectively with low/high frequency variability. Hence low frequency and large scale patterns tend to capture most of the variance observed in the system. The non-degeneracy of the eigen-spectrum is particularly an important property and can be

very useful when interpreting EOFs. For example, if two or more eigenvalues are degenerate, i.e. indistinguishable within their uncertainties, then the corresponding patterns do mix the population counterparts arbitrarily. Moreover, their actual structures may not be particularly interesting since any linear combination of these patterns is as significant as each one of them. There are always exceptions as will become clear in later sections, but the message should be clear.

The investigation of the degeneracy of the covariance matrix spectrum requires a measure of uncertainty of each eigenvalue that reflects sampling and this is quite difficult to get. This uncertainty is normally based on asymptotic results of the previous eigenvalue problem (10) in the limit of large samples (Anderson, 1963). In practice there are mainly two ways to compute the uncertainty of the eigenvalues and/or the eigenvectors of S . The first one is based on asymptotic results (Girshick, 1939; Lawley, 1956) summarised by a rule of thumb (North *et al.*, 1982):

$$\begin{aligned}\Delta\lambda_k^2 &\sim \lambda_k^2 \sqrt{\frac{2}{n^*}} \\ \Delta\mathbf{u}_k &\sim \frac{\Delta\lambda_k^2}{\lambda_j^2 - \lambda_k^2} \mathbf{u}_j\end{aligned}\quad (18)$$

where λ_j^2 is the closest eigenvalue to λ_k^2 , and n^* is the number of independent observations in the sample, also known as the effective sample size, or the number of degrees of freedom (Trenberth, 1984; Thiébaux and Zwiers, 1984). For example, the 95% confidence interval of λ_k^2 is given by $\lambda_k^2 \left(1 \pm \sqrt{\frac{2}{n^*}}\right)$. The effective sample size of a time series of length n involves in general the autocorrelation structure of the series. For example, the sum of the autocorrelation function, $1 + 2 \sum_{k \geq 1} \rho(k)$, provides a measure of the decorrelation time, and an estimate of n^* is given by (Thiébaux and Zwiers, 1984); $n^* = n \left(1 + 2 \sum_{k=1}^{n-1} (1 - k/n) \rho(k)\right)^{-1}$.

Another alternative is to use Monte Carlo simulations, (see for example Björnsson and Venegas 1997). This can be achieved by forming surrogate data by resampling a part of the data using randomisation. An example would be to randomly select a subsample and apply EOFs, then select another subsample etc. This operation, which can be repeated many times, yields various realisations of the eigenelements from which one can estimate the uncertainties. Another example would be to fix a subset of variables then scramble them by breaking the chronological order then apply EOFs, and so on. Further Monte Carlo alternatives exist to assess uncertainty on the spectrum of the covariance matrix. One could for example scramble blocks of the data, for example two- or three-year blocks of monthly data keeping thus some parts of the autocorrelation structure, (e.g. Peng and Fife 1996). To keep the whole autocorrelation structure of the data the phase randomisation method (Kaplan and Glass, 1995) can be used. The method generates data

with the same Fourier spectrum as the original data. Uncertainties on the data can even be incorporated into the EOF machinery (e.g. Thacker, 1996.) The uncertainty in the eigenvalues is useful when attempting to physically interpret a pattern or in dimension reduction or when one is looking for a break in the spectrum (Overland and Preisendorfer, 1982). For example to keep the leading three EOFs to reduce the dimensionality it is recommended that the third eigenvalue should not be too close to the fourth or higher eigenvalues.

Application

We have applied EOFs to winter monthly SLP over the Northern Hemisphere (NH). The data come from the National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalyses (Kalnay *et al.*, 1996; Kistler *et al.*, 2001). They are available on a $2.5^\circ \times 2.5^\circ$ regular grid, and span the period January 1948 to December 2000. The mean annual cycle is first calculated by averaging the monthly data over the years, then subtracted from the data to yield SLP anomalies. We are only interested in analysing the winter season defined by December to February (DJF). The data are therefore obtained by concatenating the winter monthly means for all years. Finally a weighting by the square root of the cosine of the corresponding latitude is applied to each grid point to account for the converging longitudes poleward. The data over the NH north of 20°N are used to compute EOFs. Note that the examples presented here have also been used in Hannachi *et al.* (2006).

Figure 1 shows the spectrum of the covariance matrix along with their standard errors as given by the first equation of (18) with sample size $n = 3 \times 52 = 156$. The leading two eigenvalues seem nondegenerate and separated from the rest, but overall the spectrum looks in general smooth, which makes truncation difficult. Figure 2 shows the first two EOFs. These EOFs explain

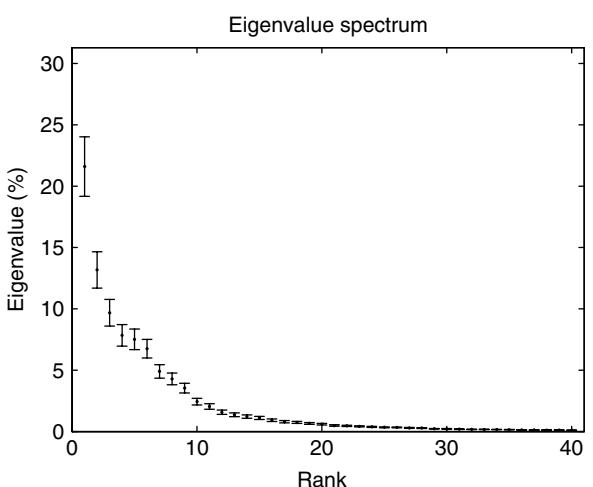


Figure 1. Spectrum, in percentage, of the covariance matrix of winter monthly (DJF) SLP. Vertical bars show approximate 95% confidence limits given by the rule of thumb (18). Only the leading 40 eigenvalues are shown.

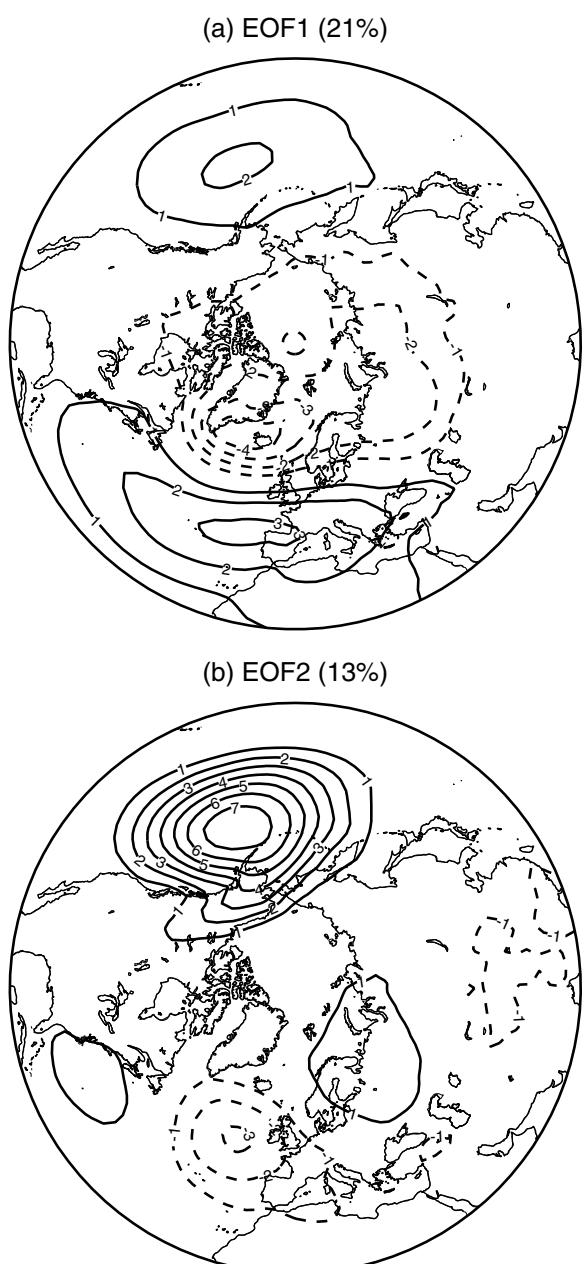


Figure 2. The first (a) and the second (b) EOFs of DJF monthly mean SLP. Positive contours solid, negative contours dashed. EOFs have been multiplied by 100.

respectively 21% and 13% of the total (monthly winter) variance. EOF 1 (2a) shows a high over the North Pole and two low centres over the Mediterranean-North East Atlantic and over the North Pacific. This is the familiar AO mode (Thompson and Wallace, 1998, 2000; Wallace and Thompson 2002). EOF2 (2b) shows two separated centres of opposite signs over the North Pacific and North East Atlantic respectively. Figure 3 shows the first two PCs associated with the leading two EOFs. These PCs are uncorrelated at zero-lag but not at other lags. A trend signature can be noted in both the PCs. This is again due to the way EOFs process the data. In fact, EOFs do not look for trends, and if there is one then it is likely that it will be spread over more than one PC.

Figure 4 shows the autocorrelations of PC1 (4a) and PC2 (4b) of DJF SLP. Note that the data are monthly, and not seasonal means. The autocorrelations indicate a short memory behaviour with one or two months lag. There seems to be a small autocorrelation around 24 months lag in PC1 (Figure 4(a)) and around 17 months lag in PC2 (Figure 4(b)). From this limited sample it is difficult to know the exact origin of these autocorrelations. However, it is possible that the previous autocorrelation observed in PC1 could be due to the effect of El Nino Southern Oscillation (ENSO) cycle, and that observed in PC2 could be the effect of the QBO (Trenberth and Shin, 1984).

There is an ongoing debate within the climate community on whether the Arctic Oscillation (Figure 2(a)) or the North Atlantic Oscillation is the most physically-relevant mode of variability of the NH SLP. Because of the nature of the method, this debate cannot be resolved using EOFs alone. EOFs have a serious difficulty when it comes to interpretation. For example, because of the spatial and/or temporal autocorrelation, the coherent-like large scale EOF patterns obtained only reflect the effect of the correlations of neighbouring grid-points. Various methods have been proposed to ease this difficulty and aid interpretation. In this review we present two alternatives, the first one is familiar to climate researchers and is based on rotation, and the second one, simplified EOFs approach, is relatively new and is based on the LASSO approach.

SIMPLIFICATION METHODS

Rotation of EOFs is perhaps the most used method in atmospheric science due in part to its simplicity. Simplified EOFs method provides also another useful and new way of simplification in addition to its nice formulation and natural link to EOFs. Both of these methods are discussed below. See also Hannachi *et al.* (2006), which is principally devoted to simplification. Other less known methods of simplification have been proposed. These methods will be discussed briefly toward the end of this section with references provided for the interested readers.

Rotated EOFs

What is it and why?. Spatial orthogonality and temporal uncorrelation of EOFs and PCs respectively impose limits on physical interpretability of EOF patterns. This is because physical processes are not independent, and therefore physical modes are expected in general to be non-orthogonal. As an example, normal modes derived from linearised physical models, such as the barotropic vorticity equation (Simmons *et al.*, 1983) are non-orthogonal. Furthermore, EOFs tend to be dependent on the size and shape of the data domain (Richman, 1986). For instance, the first EOF pattern tends to have wavenumber one sitting on the whole domain. The second EOF, on the other hand, tends to have wavenumber two and be orthogonal to EOF1 regardless of the nature

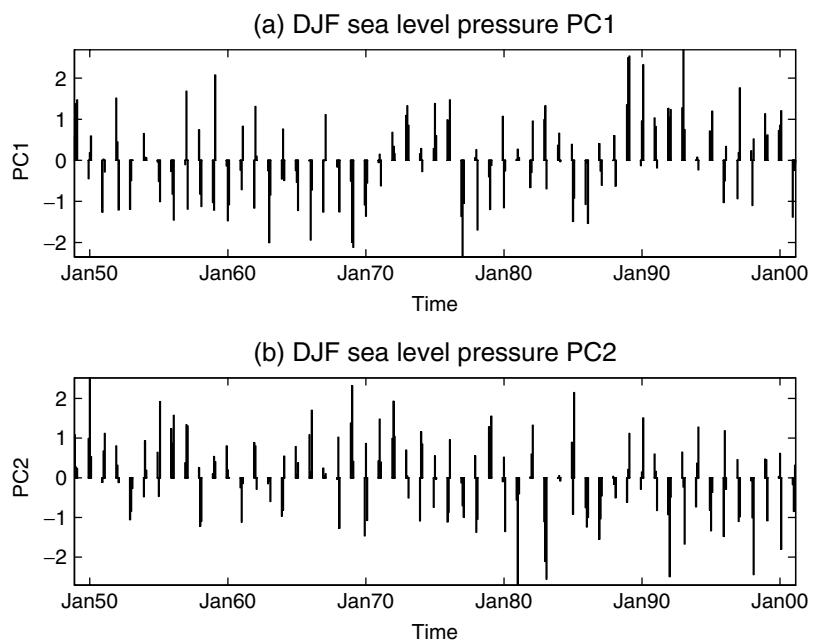


Figure 3. The leading two scaled PCs corresponding to the leading two EOFs of Figure 2.

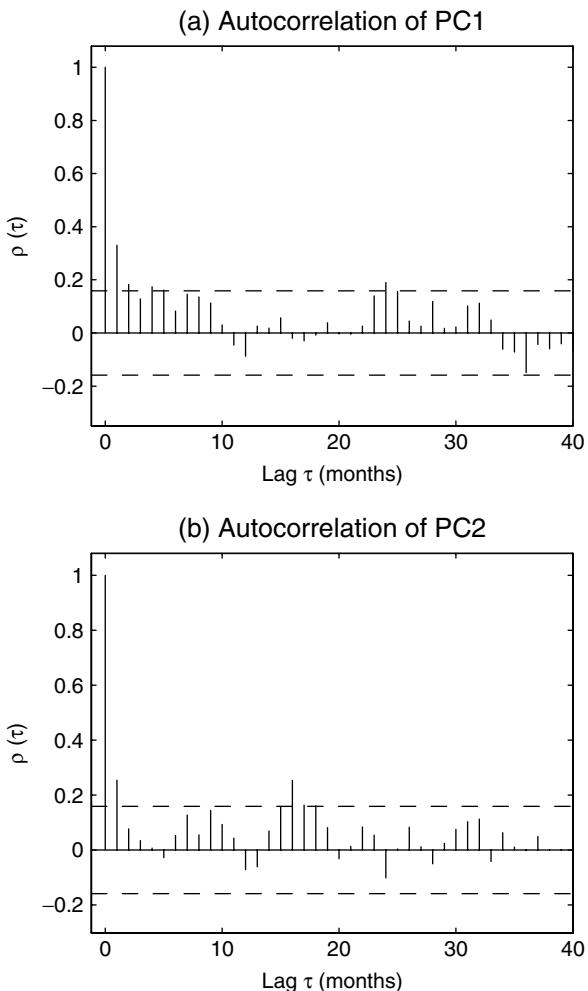


Figure 4. Autocorrelation functions of DJF SLP PC1 and PC 2. Horizontal lines show approximate 95% confidence limits.

of the physical process involved in producing the data, and this applies in general to subsequent EOFs.

To help overcome these difficulties and gain easy interpretation, a number of methods have been proposed. Among these methods, REOFs, based simply on rotating the EOF patterns, seems to be the most widely used method in atmospheric science mainly because of its relative simplicity. REOF techniques have been adopted by atmospheric scientists since the early 1980s (Horel, 1981; Richman, 1981, 1986; Jolliffe, 1987). The technique, however, is much older and was known in factor analysis as factor rotation, and has been applied extensively in social science (Carroll, 1953; Kaiser, 1958). The main objectives of REOFs are to:

- alleviate the strong constraints of EOFs, namely orthogonality/uncorrelation of EOFs/PCs, and domain dependence of EOF patterns (see e.g. Dommegård and Latif, 2002),
- obtain simple structures,
- ease the interpretation of obtained patterns.

Formulation and computation of REOFs

Rotation of the EOF patterns can systematically alter the structures of EOFs. By constraining the rotation to maximise a simplicity criterion the REOF patterns can be made simple. Given a $p \times m$ matrix $U_m = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m)$ of the leading m EOFs (or loadings), the rotation is formally achieved by seeking an $m \times m$ rotation matrix R to construct the REOFs B according to:

$$B = U_m R, \quad (19)$$

where R is either R or $(R^T)^{-1}$ depending on the type of rotation as detailed below. The criterion for choosing

the rotation matrix R is what constitutes the rotation algorithm or the simplicity criterion, and is expressed by the maximisation problem:

$$\max f(U_m \mathcal{R}) \quad (20)$$

over a specified subset or class of $m \times m$ square rotation matrices \mathcal{R} . The functional $f()$ represents the rotation criterion. Note that besides rotating the EOFs U_m as in (19), one could equally rotate the EOFs scaled by the square root of the corresponding eigenvalues, i.e. using $U_m \Lambda_m$, where $\Lambda_m = (\lambda_1, \dots, \lambda_m)$ is the diagonal matrix containing the leading singular values. Alternatively, one can also rotate PCs instead. Various rotation criteria exist in the literature (Richman, 1986; Harman, 1976; Reymant and Jöreskog, 1996). Richman (1986), for example, lists more than ten simplicity criteria. Broadly speaking there are two large families of rotation: orthogonal and oblique rotations.

a. Orthogonal rotation

In orthogonal rotation (Kaiser, 1958; Jennrich, 2001) the rotation matrix R in (19) is chosen to be orthogonal, and $\mathcal{R} = R$. The problem is to solve (20) subject to the condition:

$$RR^T = R^T R = I_m \quad (21)$$

where I_m is the $m \times m$ identity matrix.

The most well-known and used rotation algorithm is the VARIMAX criterion (Kaiser, 1958). Let us designate by b_{ij} , $i = 1, \dots, p$, and $j = 1, \dots, m$, the elements of the REOFs matrix B in (19), i.e. $b_{ij} = [B]_{ij}$, then the VARIMAX orthogonal rotation maximises a simplicity criterion according to:

$$\max \left(f(B) = \sum_{k=1}^m \left[p \sum_{j=1}^p b_{jk}^4 - \left(\sum_{j=1}^p b_{jk}^2 \right)^2 \right] \right) \quad (22)$$

where m is the number of EOFs chosen for rotation. The quantity inside the square brackets in (22) is proportional to the (spatial) variance of the square of the rotated vector $\mathbf{b}_k = (b_{1k}, \dots, b_{pk})^T$. Therefore VARIMAX attempts to simplify the structure of the patterns by pushing the loadings coefficients towards zero, or ± 1 . In some cases, the loadings of the REOFs B are weighted by the communalities of the different variables (Walsh and Richman, 1981). The communalities h_j^2 , $j = 1, \dots, p$, are directly proportional to the sum of squares, $\sum_{k=1}^m u_{jk}^2$, of the loadings for a particular variable. Hence if $C = \text{Diag}(U_m U_m^T)^{-1/2}$, then in the weighted or normalised VARIMAX, the matrix B as used in (22) is simply replaced by BC . This normalisation is generally used to reduce the bias toward the first EOF with the largest eigenvalue.

Another familiar orthogonal rotation method is based on the QUARTIMAX criterion. It seeks to maximise the variance of the patterns:

$$f(B) = \frac{1}{mp} \sum_{k=1}^m \sum_{j=1}^p \left[b_{jk}^2 - \frac{1}{mp} \sum_{k=1}^m \sum_{j=1}^p b_{jk}^2 \right]^2. \quad (23)$$

Because of the orthogonality property (21) required by R , the REOFs matrix also satisfies $B^T B = I_m$ when the (unscaled) EOFs are rotated, and the sum of the squared elements of B is constant. Therefore the QUARTIMAX simply boils down to maximising the fourth order moment of the loadings, hence the term QUARTIMAX:

$$\max \left[f(B) = \frac{1}{mp} \sum_{k=1}^m \sum_{j=1}^p b_{jk}^4 \right] \quad (24)$$

Eq (22) or (24) are then to be optimised subject to the orthogonality constraint (21). VARIMAX is in general preferred to the QUARTIMAX because it is slightly less sensitive to changes in the number of variables (Richman, 1986), although the difference in practice is not significant.

b. Oblique rotation

Here the rotation matrix R is chosen to be non-orthogonal (Harman, 1976; Kiers, 1994; Jennrich, 2002) normalised to have unit-length columns, and where $\mathcal{R} = (R^T)^{-1}$. The oblique rotation matrix is obtained by solving (20) subject to the previous constraints. Among the familiar examples of oblique criteria one finds the QUARTIMIN (Carroll, 1953; Harman, 1976), which corresponds to the following criterion:

$$f(B) = \frac{1}{4} \sum_{r \neq s} \sum_k b_{kr}^2 b_{ks}^2. \quad (25)$$

Illustration. We have applied both orthogonal and oblique rotations to the (unscaled) EOFs and the EOFs scaled by the corresponding singular values. We have therefore four cases to be discussed: (i) orthogonal rotation of EOFs, (ii) orthogonal rotation of scaled EOFs, (iii) oblique rotation of EOFs, and (iv) oblique rotation of scaled EOFs. Various rotation criteria have been applied, but we focus our discussion on the results obtained from three criteria, namely VARIMAX, QUARTIMAX, and QUARTIMIN. The discussion also includes the effect of changing the number m of EOFs to be rotated, (see also Hannachi *et al.* 2006).

The first observation is that orthogonal rotation is more efficient, computationally, than oblique rotation, due to matrix inversion in the latter. Using various rotation criteria and various values of the parameter m , we have found that case (i) and (iii) give virtually the same result. Figure 5 shows a scatter plot of rotated loadings using VARIMAX versus QUARTIMIN for $m = 30$. A similar feature has also been obtained with other criteria (not shown). This seems to indicate that orthogonal/oblique rotation of (unscaled) EOFs is a robust feature. Unfortunately, this is not true. The rotated patterns change as m changes. For example, when we rotate 3 EOFs, the NAO and the North Pacific patterns emerge as the most prominent patterns with associated time series having leading variances. As m increases, however, these features disappear progressively in favour

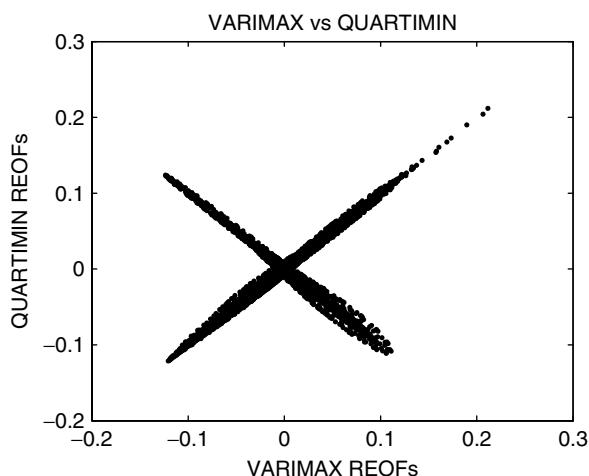


Figure 5. Scatter plot of VARIMAX REOFs versus QUARTIMIN REOFs using $m = 30$ EOFs. Scatter with negative slopes corresponds to similar REOFs but with opposite sign.

of other structures with smaller scales. Figures 6 and 7 show examples of VARIMAX REOFs using respectively $m = 6$ and $m = 20$. These patterns have been selected visually so to be as close as possible to familiar large scale modes of variability. Note that as m increases the patterns become more and more localised. This non-invariance of the leading rotated patterns to changes in m can be explained by the fact that in the rotation process, there is no preferential order or varying weights attached to the EOFs. All the EOFs are equivalent, since no variance is included, and the final solution is only dictated by condition (20).

To overcome the previous difficulty, the alternative is to weigh the EOFs by the square root of the associated eigenvalues. In this case the norm-squared of each (scaled) EOF is precisely the variance of the corresponding time series. This automatically yields case (ii) when the rotation is orthogonal, which we discuss now. As would be expected, this case produces leading REOFs that are invariant to changes in m . This is because low ranked EOFs contribute less to the leading REOFs. Note also that because of orthogonality the order of rotated patterns is provided by their squared norms, which play the role of associated variances. These ‘variances’ are not, however, additive because the associated time series XB are not uncorrelated. There is no non-trivial rotation that conserves spatial orthogonality and temporal uncorrelatedness. Figure 8 shows the leading three VARIMAX REOFs using $m = 20$ identified respectively as the NAO, the North Pacific pattern, and the Scandinavian pattern (Barnston and Livezey, 1987). The same result is obtained for the leading QUARTIMAX REOFs. This invariance property breaks down for low-ranked rotated patterns. For example when $m = 30$, the leading 15 REOFs are similar between the two orthogonal rotations used here.

For the last case (iv) we have found that the algorithm runs into convergence problems due to bad conditioning in the matrix inversion process. This does not happen

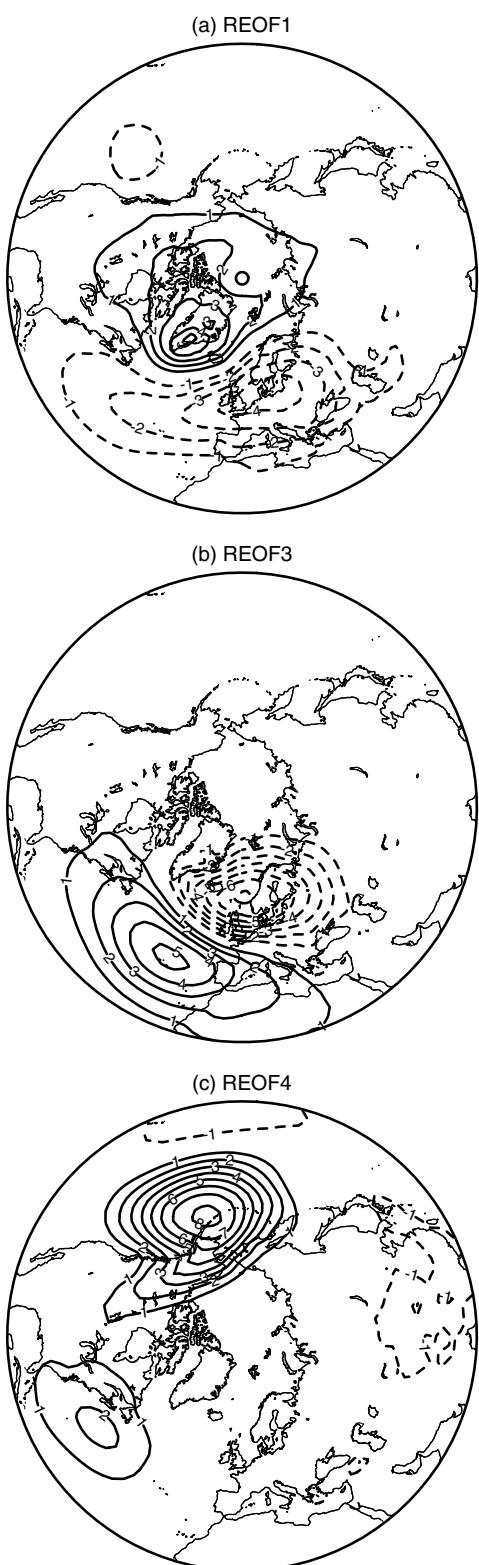


Figure 6. Three VARIMAX rotated EOFs using the leading 6 DJF SLP EOFs showing the first (a), the third, NAO, (b), and the fourth, North Pacific pattern, (c). The order is fixed using the variances of the corresponding time series. Positive contours solid, negative contours dashed. Loadings have been multiplied by 100 as in Figure 2.

in the orthogonal rotation since orthogonal matrices are easily obtained using SVD.

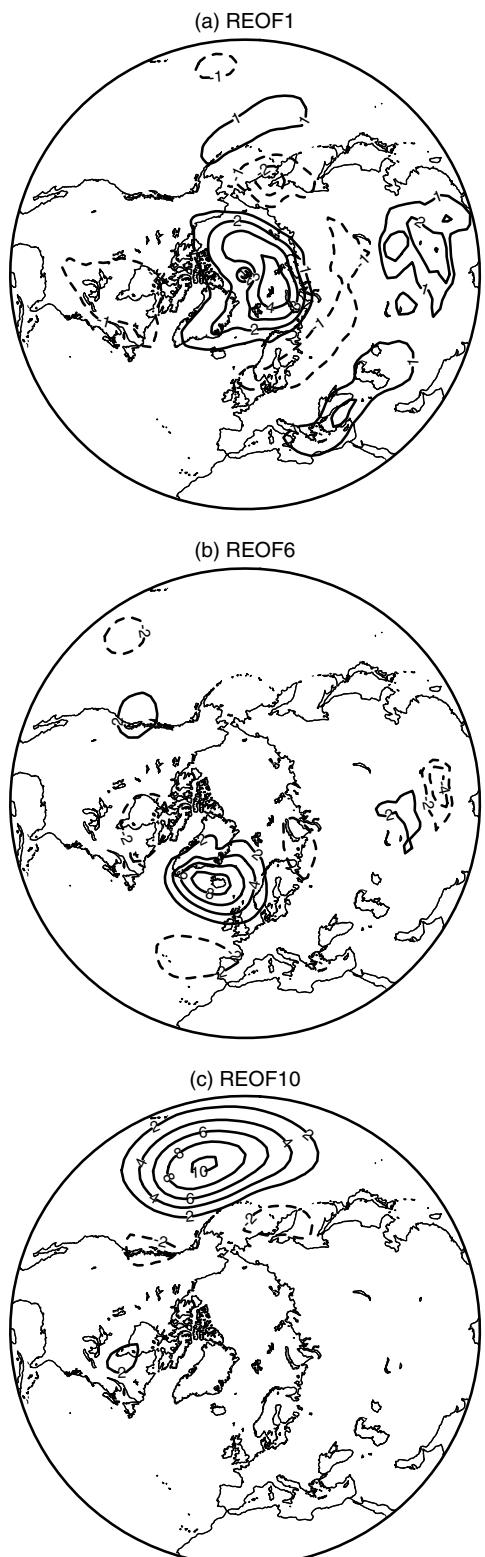


Figure 7. Same as in Figure 6 but with $m = 20$. The orders of the patterns are 1 (a), 6 (b), and 10 (c).

Simplified EOFs

Background. REOFs have been introduced mainly to improve interpretation through obtaining simpler patterns than EOFs. Building objective simplicity criteria, however, turns out to be a difficult problem. Jolliffe *et al.*

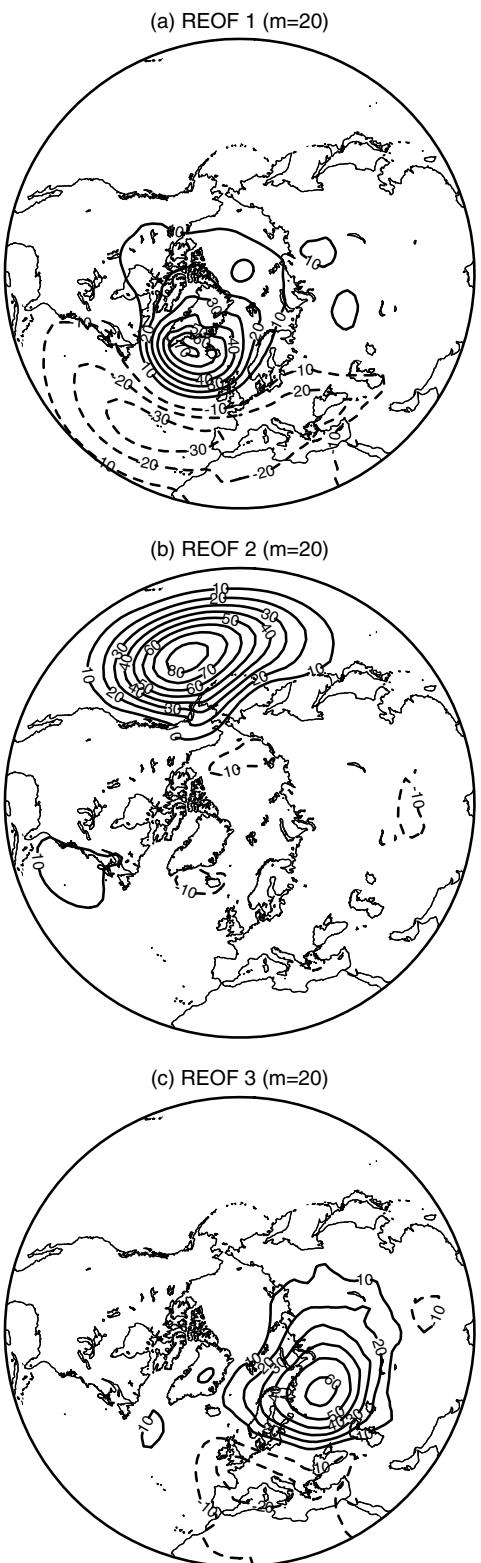


Figure 8. The leading three VARIMAX REOFs obtained using the leading $m = 20$ EOFs scaled by the square root of the corresponding eigenvalues. Loadings have been multiplied by 100 as in Figure 2.

(2002) point out that concentrating the EOF coefficients close to 0 or ± 1 is not the only possible definition of simplicity. For example a pattern with only ones is simple though it could rarely be of much interest in atmospheric science. Although REOFs attempt to achieve

this using a simple and practical criterion they have a number of difficulties which make the method quite controversial (Richman, 1986, 1987; Jolliffe, 1987, 1995; Mestas-Nuñez, 2000).

When we apply the rotation procedure we are usually faced with the following questions:

- how to fix the number of EOFs or PCs to be rotated?
- what type of rotation, e.g. orthogonal or oblique, should be used?
- which of the large number of simplicity criteria should be used? and
- how to chose the normalisation constraint (Jolliffe, 1995)?

Although the results of the previous section shed some light on which rotation criterion to choose, namely orthogonal rotation of scaled EOFs, there is still the issue of non-invariance of the low-ranked rotated patterns for large m . A simplification technique that retains some of the useful properties of EOFs has been proposed by Jolliffe *et al.* (2003) as an alternative to rotation. This technique is simplified EOFs and is described next.

LASSO-based simplified EOFs. Various simplification techniques have been suggested to obtain simple structures (e.g. Chapter 11 of Jolliffe 2002). Most of these techniques attempt to reduce the two stages of rotated PCA into just one step. Here we discuss a particularly interesting method of simplicity that is rooted in regression analysis. A common problem that arises in multiple linear regression is the instability of regression coefficients because of collinearity or high dimensionality. Tibshirani (1996) has investigated this problem and proposed a technique known as the LASSO. The LASSO approach attempts to shrink some regression coefficients exactly to zero, hence implicitly selecting variables. The same idea was adapted in the PCA context by Jolliffe *et al.* (2003) who labelled it ‘Simplified Component Technique-LASSO’ (SCoT-LASSO). For brevity we refer to the SCoT-LASSO EOF method as simplified EOFs (SEOFs), but it should be borne in mind that this is not the only form of simplicity (see Jolliffe 2002). Jolliffe *et al.* (2003) applied the method to a toy example, and Hannachi *et al.* (2006) applied it to a moderately large climate dataset, the same one as here.

The SEOF method attempts to use the main properties of EOFs and REOFs simultaneously by successively maximising variance and constraining the patterns to be orthogonal and simple. Simplicity here means that the loadings of each pattern have either small, i.e. close to zero, or large, i.e. close to one, magnitude and no intermediate values. The objective of SEOFs is to seek directions $\mathbf{u}_k = (u_{k1}, u_{k2}, \dots, u_{kp})^T$, $k = 1, \dots, p$ maximising:

$$F(\mathbf{u}_k) = \mathbf{u}_k^T \mathbf{S} \mathbf{u}_k \quad (26)$$

subject to

$$\mathbf{u}_k^T \mathbf{u}_l = \delta_{kl}. \quad (27)$$

In addition, to achieve simplicity the lasso technique requires the following extra constraint to be satisfied (Jolliffe *et al.*, 2003):

$$\|\mathbf{u}_k\|_1 = \sum_{j=1}^d |u_{kj}| = \mathbf{u}_k^T \text{sign}(\mathbf{u}_k) \leq \tau \quad (28)$$

for some tunable threshold parameter τ . In (28) $\text{sign}(\mathbf{u}_k) = (\text{sign}(u_{k1}), \dots, \text{sign}(u_{kp}))^T$ is the sign of \mathbf{u}_k . The following properties can be easily verified:

- No solution exists for the optimisation problem (26–28) when $\tau < 1$.
- any simplified \mathbf{u}_k satisfies $\|\mathbf{u}_k\| \leq \sqrt{p}$
- for $\tau \geq \sqrt{p}$, \mathbf{u}_k , $k = 1, \dots, p$ are simply the EOFs.

The last property indicates that EOFs are a particular case of simplified EOFs.

The numerical solution to Eqs (26)–(28) is presented in Trendafilov and Jolliffe (2005) who applied it to a small problem, and Hannachi *et al.* (2006) who applied it to the NH SLP as is done here. The approach of finding the k^{th} SEOF \mathbf{u}_k is based on integrating the following system of ordinary differential equations (ODEs):

$$\frac{d}{dt} \mathbf{u}_k = (I_d - \mathbf{u}_k \mathbf{u}_k^T) \nabla F_\mu^{(k)}(\mathbf{u}_k) \quad (29)$$

forward in time for “sufficiently” long time interval using suitably chosen initial conditions (e.g. Hir. and Smale 1974). In Eq (29) the function $F_\mu^{(k)}$ is defined by:

$$F_\mu(\mathbf{u}_k) = \frac{1}{2} \mathbf{u}_k^T \mathbf{S} \mathbf{u}_k - \mu H(\mathbf{u}_k^T \tanh(\gamma \mathbf{u}_k) - \tau) \quad (30)$$

with $H(x) = \frac{1}{2}x(1 + \tanh \gamma x)$, μ and γ are fixed large positive numbers, and the matrix S_k given by:

$$S_k = \left(I_d - \sum_{l=0}^{k-1} \mathbf{u}_l \mathbf{u}_l^T \right) S \left(I_d - \sum_{l=0}^{k-1} \mathbf{u}_l \mathbf{u}_l^T \right). \quad (31)$$

Hence the k^{th} SEOF \mathbf{u}_k is the limit, when $t \rightarrow \infty$, of the solution to eq (29), i.e. the stationary solution to the same equation. In the application section below we follow Hannachi *et al.* (2006) to which the reader is referred for more details and further references.

Application. We have computed the SEOFs of the DJF monthly NH SLP field for various values of the threshold parameter τ from 8 to 30. For a given value of the threshold parameter τ the SEOFs are obtained by integrating eq (29) forward in time using MATLAB function ODES15, which can solve stiff ODEs. The constants γ and μ are fixed as in Trendafilov and Jolliffe (2005) and Hannachi *et al.* (2006) to 1000 and 800 respectively. The solution is found to be virtually invariant to changes in those parameters. In fact, these parameters are not part of the problem, and their role is pretty universal, (see Hannachi *et al.* 2006 for computational details). We also

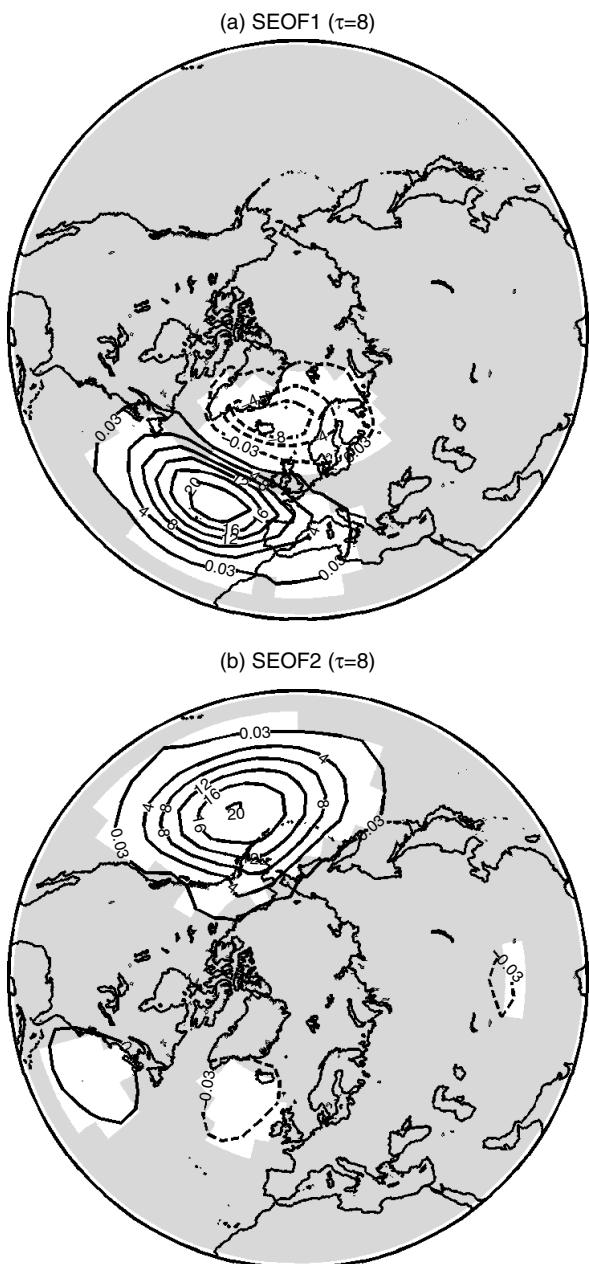


Figure 9. The leading SEOF1 (a) and SEOF2 (b) obtained for a threshold parameter $\tau = 8$. Regions of zero loadings are shaded. The patterns have been multiplied by 100 as in Figure 2.

follow Hannachi *et al.* (2006) by computing the few leading SEOFs, using only a coarse grid of $5^\circ \times 5^\circ$, and by using the same example.

Figure 9 shows the leading SEOF1 (9a) and SEOF2 (9b) for $\tau = 8$. Regions where the loadings are zero are shaded. Figure 9(a) clearly shows the NAO pattern with its distinctive dipolar structure (Hurrell, 1996; Thompson *et al.*, 2000; Hurrell *et al.*, 2003) whereas Figure 9(b) shows the North Pacific pattern, a monopolar structure centered over the North mid-Pacific. As τ increases the shaded regions in Figure 9 shrink and the patterns become more and more non-local. Figure 10 shows the leading two SEOFs for $\tau = 18$, where one can still see the NAO and North Pacific patterns but with structure

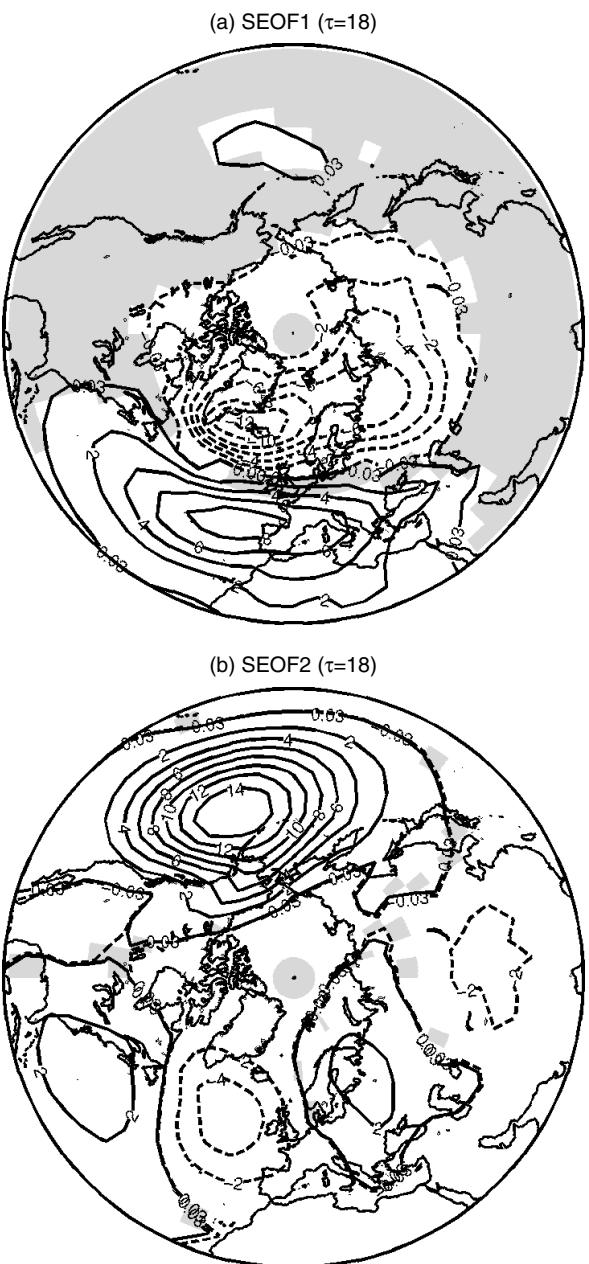


Figure 10. As in Figure 9 but for $\tau = 18$.

enlargement compared to Figure 9. When τ reaches 26 the leading SEOFs (not shown) start to converge to the corresponding EOFs patterns. Hannachi *et al.* (2006) find that for $\tau \leq \frac{1}{2}\sqrt{p}$, the variances of the time series corresponding to the leading two SEOFs are not too different. This may explain why the simple patterns found by SEOFs method appear combined in a single pattern when using EOF analysis (Figure 2(a)).

The third SEOF pattern is found to represent the Scandinavian pattern for τ smaller than approximately $\frac{1}{2}\sqrt{p}$. Figure 11 shows SEOF3 for $\tau = 12$ and $\tau = 16$. For the latter value of the threshold parameter the pattern becomes nearly hemispheric with the emergence of a third centre over the North Atlantic basin, and is close to EOF3 (not shown). Note in particular the resemblance

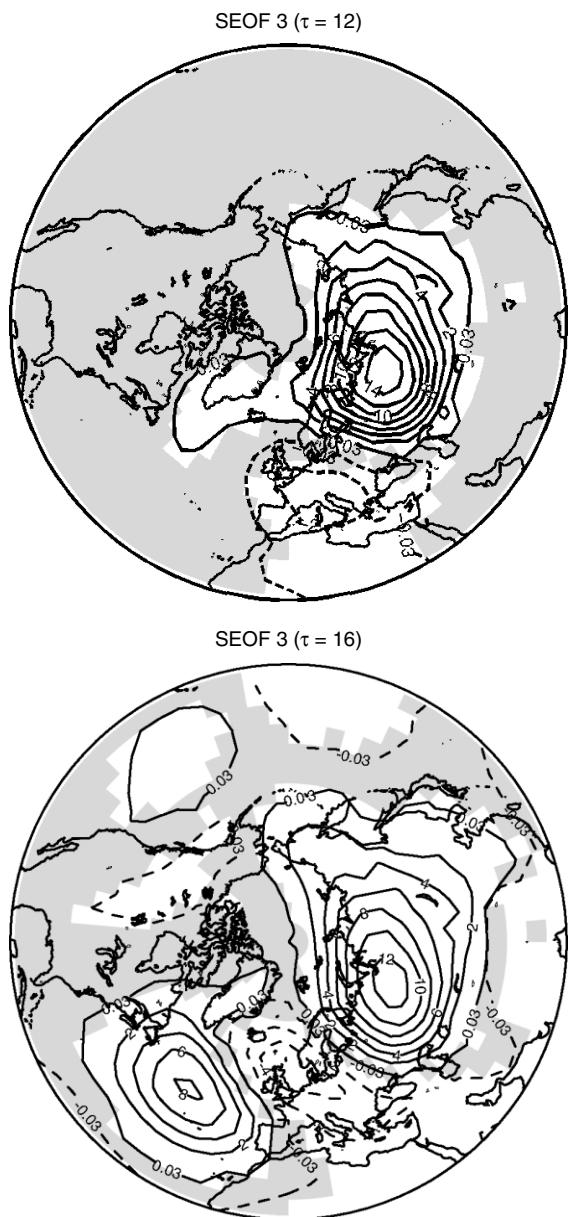


Figure 11. SEOF3 for $\tau = 12$ (top) and $\tau = 16$ (bottom). The patterns have been multiplied by 100 as in Figure 9.

between SEOF3 corresponding to $\tau = 12$ (Figure 11) and REOF3 shown in Figure 6(c). As τ decreases further, the patterns seem to keep their structure except that they become smaller in spatial extent, and of course lose variance. The loss in variance, however, is justified by the increase in simplicity. Figure 12 shows the ratio between the variances of the times series corresponding to SEOF1 and EOF1 respectively *versus* the parameter τ . Figure 12 indicates that convergence to EOFs starts around $\tau = \frac{2}{3}\sqrt{p}$.

SEOF patterns seem to produce invariant features vis-a-vis changes in the threshold parameter for $\tau < \frac{1}{2}\sqrt{p}$. Hannachi *et al.* (2006) propose $\tau = \frac{1}{3}\sqrt{p}$ to be a reasonably good choice, regarding balance between variance maximisation and locality or simplicity. The method, however, is more expensive, computationally,

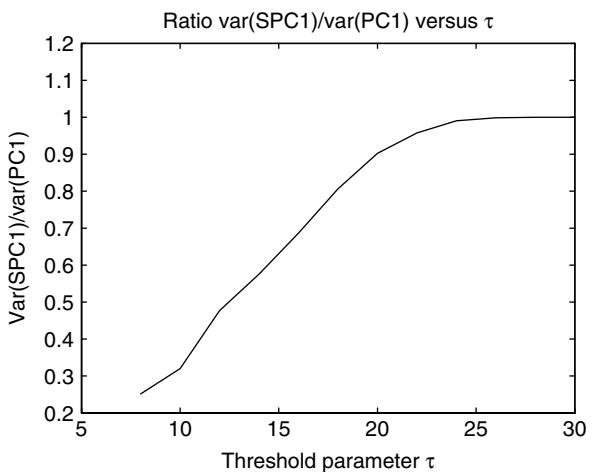


Figure 12. Variance ratio of SPC1 to that of PC1 *versus* the simplicity parameter τ .

compared to EOFs or REOFs when the number of EOFs selected for rotation is not very large.

Other forms of simplification

Various other methods of simplifications have been proposed in the PCA literature. Most of these methods impose extra constraints on the variables in a similar manner to the SEOF method. In a number of those techniques the loadings are restricted to be integers taking the values 0, and ± 1 (Hausmann, 1982.) Vines (2000) also uses a similar procedure that she labels simple components. The method starts from the natural basis of the variables-space and proceeds by orthogonally rotating them pairwise such that the variance of the pair is increased and simplicity preserved. The latter is achieved by choosing angles that yield vectors whose components are proportional to integers. See Sun (2005) for a detailed analysis of simple component analysis with an account of the various algorithms used. The SEOF method presented above is rather an extension of these discrete values methods in that the coefficients can vary smoothly and not be restricted to integers. Green (1977), Bibby (1980), and Jackson (1991) present another way of simplification based on considering the usual PCs, then proceed to their ‘simplification’ by rounding them to the first digit (Green, 1977; Bibby, 1980) and also to the closest integer (Bibby, 1980; Jackson, 1991).

A different alternative to rotation and simplification was presented by Van den Dool *et al.* (2000), which they label empirical orthogonal teleconnection (EOT). The method proceeds as follows. The grid point s_j is first obtained that maximises $\sum_{k=1}^p \text{corr}^2(s_j, s_k) \text{var}(s_k)$. The first EOT is then obtained as the regression coefficient between the grid s_j and all other grid points. The next EOT is obtained in a similar way, using the residuals from the separate regression used for EOT1, and are orthogonal to the previous EOT. The EOT method is not as simple as the other methods of simplification, but Van den Dool *et al.* (2000) argue that it helps the physical interpretation of the patterns. Jolliffe (2002) points out that the first

EOT is a compromise between the first principal variables obtained using the covariance and correlation matrices.

Although simplification methods have been presented as a way to overcome the drawbacks of EOFs to yield perhaps physically relevant patterns, such as teleconnections, the problem remains, however, whether teleconnections exist and if they do, how they can be identified (Jolliffe, 2002; Dommenget and Latif, 2002)? Rotation and other simplification procedures remain after all simple mathematical tools that may or may not speak the language of nature. Section 6 discusses possible alternatives to evaluate EOFs and to identify possible teleconnection patterns from observed climate data.

EXTENDED EOFs

Background

The previous sections have dealt with patterns that maximise variance using spatial correlation structures provided by the covariance matrix. The matrix uses only simultaneous information (in time) between different grid points, but forgets about any lagged information. Since lagged correlations constitute an important characteristic feature of climate data, it is important to incorporate this information into the analysis. An important method using such information is based on extended empirical orthogonal function (EEOF). EEOFs constitute an extension of the traditional EOF technique to deal not only with spatial- but also with temporal correlations observed in weather/climate data. The method was first introduced by Weare and Nasstrom (1982) who applied it to the 300-mb relative vorticity to identify propagating structures.

A similar approach was developed later to deal with dynamical reconstruction of low order chaotic systems by Broomhead and King (1986a,b) who called it *singular system analysis* (SSA). At the same time Fraedrich (1986) also used the same approach to compute dimensions of chaotic attractors from climate data. SSA was also used to find oscillations from climate records (Vautard *et al.*, 1992). It was extended to deal with multivariate, or multichannel, (MSSA) time series (Broomhead and King, 1986a,b) in a way similar to EEOF analysis. MSSA (or EEOF) was applied later by Kimoto *et al.* (1991) and Plaut and Vautard (1994) to find propagating structures from 500-mb heights reanalyses.

The use of the lagged information from time series, e.g. lagged auto-covariance matrix, to find propagating or periodic signals, goes back to the middle of the century with Whittle (1951) and a few others. The method has been applied to observed time series first by Basilevsky and Hum (1979) to find an embedded periodic signal in the data. In the one-dimensional case the procedure is as follows. Given a single-channel time series w_t , $t = 1, 2, \dots, n$, the method consists first in constructing an M -dimensional time series \mathbf{w}_t , $t = 1, 2, \dots, n - M + 1$, using the *delay* coordinate as

$$\mathbf{w}_t = (w_t, w_{t+1}, \dots, w_{t+M-1})^T. \quad (32)$$

The parameter M in Eq (32) is known as *window length* or *delay parameter*. This parameter is also known as *embedding dimension*, a concept that is rooted in the theory of dynamical systems (e.g. Takens 1981), and has to be chosen beforehand. The lagged covariance matrix \mathcal{C} of this newly formed multi-channel time series is given by:

$$\mathcal{C} = \frac{1}{n - m + 1} \sum_{t=1}^{n-M+1} \mathbf{w}_t \mathbf{w}_t^T. \quad (33)$$

Periodic signals are then identified through the existence of pairs of degenerate eigenvalues of \mathcal{C} , that are separated from the rest of the spectrum. When the single-channel time series is stationary the auto-covariance matrix \mathcal{C} in Eq (33) has a Toeplitz structure, that is constant over the diagonals, and is known to have useful properties (see e.g. Graybill 1969). The sample estimate, however, will not have an exactly Toeplitz structure unless it is imposed. The method can also be applied to nonstationary time series, (see e.g. Elsner and Tsonis 1996, and Golyandina *et al.* 2001) for further details. The multivariate extension of this method yields the EEOFs or MSSA, and is detailed next. The next two sections provide a technical background of EEOFs. In section 4.4 we have chosen to apply the method to identify the Madden-Julian oscillation (MJO). This is because the MJO is well studied and well documented in the literature since it was first identified by Madden and Julian (1972) using spectral techniques.

Definition and computation of EEOFs

In EEOF analysis the atmospheric state vector at time t , i.e. $\mathbf{x}_t = (x_{t1}, \dots, x_{tp})$, $t = 1, \dots, n$, used in traditional EOF, is extended to include temporal information as

$$\mathbf{x}_t = (x_{t1}, \dots, x_{t+M-1,1}, x_{t2}, \dots, x_{t+M-1,2}, \dots, x_{t,p}, \dots, x_{t+M-1,p}) \quad (34)$$

with $t = 1, \dots, n - M + 1$. The new data matrix now takes the form

$$\mathcal{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_{n-M+1} \end{bmatrix} \quad (35)$$

It is now clear from (34) that time is incorporated in the state vector side by side with the spatial dimension. If we denote by

$$\mathbf{x}_t^s = (x_{ts}, x_{t+1,s}, \dots, x_{t+M-1,s}) \quad (36)$$

then the extended state vector (34) is written in a similar form to the conventional state vector, i.e.

$$\mathbf{x}_t = (\mathbf{x}_t^1, \mathbf{x}_t^2, \dots, \mathbf{x}_t^p) \quad (37)$$

except that now the elements \mathbf{x}_t^k , $k = 1, \dots, p$, of this grand state vector (Equation 37) are themselves temporal lagged values. The data matrix \mathcal{X} in (36) now takes the form

$$\mathcal{X} = \begin{bmatrix} \mathbf{x}_1^1 & \mathbf{x}_1^2 & \dots & \mathbf{x}_1^p \\ \vdots & \vdots & & \vdots \\ \mathbf{x}_{n-M+1}^1 & \mathbf{x}_{n-M+1}^2 & \dots & \mathbf{x}_{n-M+1}^p \end{bmatrix} \quad (38)$$

which is again similar to the traditional data matrix X in (2) except that now its elements are (temporal) vectors.

The vector \mathbf{x}_t^s in (36) is normally referred to as the *delayed vector* obtained from the time series (x_t^s) , $t = 1, \dots, n$ of the field value at grid point s . The new data matrix (38) is now of order $(n - M + 1) \times pM$ which is significantly larger than the original matrix dimension.

We suppose that \mathcal{X} in (38) has been centered and weighted. The covariance matrix of (37) is:

$$\Sigma = \frac{1}{n - M + 1} \mathcal{X}^T \mathcal{X} = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1p} \\ C_{21} & C_{22} & \dots & C_{2p} \\ \vdots & \vdots & & \vdots \\ C_{p1} & C_{p2} & \dots & C_{pp} \end{bmatrix} \quad (39)$$

where each C_{ij} , $1 \leq i, j \leq p$ is a lagged covariance matrix between gridpoint i and gridpoint j , given by:

$$C_{ij} = \frac{1}{n - M + 1} \sum_{t=1}^{n-M+1} \mathbf{x}_t^{i^T} \mathbf{x}_t^j. \quad (40)$$

Other alternatives to compute C_{ij} also exist and they are related to the way the lagged covariance between two time series is computed (see e.g. Priestley 1981 and Jenkins and Watts 1968). If the multivariate time series is stationary, then the population version of each block matrix of (39) is symmetric Toeplitz. The sample version C_{ij} from (39) is nearly symmetric Toeplitz for large sample size. The symmetric grand covariance matrix Σ is not in general Toeplitz because the different blocks represent covariances between different pairs of grid points. An alternative form of the data matrix is provided by writing the state vector (34) in the form

$$\mathbf{x}_t = (x_{t1}, \dots, x_{tp}, x_{t+1,1}, \dots, x_{t+1,p}, \dots x_{t+M-1,1}, \dots, x_{t+M-1,p}) \quad (41)$$

that is

$$\mathbf{x}_t = (\mathbf{x}_t, \mathbf{x}_{t+1}, \dots, \mathbf{x}_{t+M-1}) \quad (42)$$

where $\mathbf{x}_t = (x_{t1}, \dots, x_{tp})$ is the state vector at time t , $t = 1, \dots, n - M + 1$. Hence the matrix (38) now takes the following alternative form, used by Weare and Nasstrom (1982):

$$\mathcal{X}_1 = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_M \\ \vdots & \vdots & & \vdots \\ \mathbf{x}_{n-M+1} & \mathbf{x}_{n-M+2} & \dots & \mathbf{x}_n \end{bmatrix} \quad (43)$$

This form is exactly equivalent to (38) since it is obtained from (38) by a permutation of the columns as

$$\mathcal{X}_1 = \mathcal{X}P \quad (44)$$

where $P = (p_{ij})$, $i, j = 1, \dots, Mp$, is a permutation matrix (which is orthogonal, i.e. $PP^T = P^TP = I$, and contains exactly 1 in every line and every column and zeros elsewhere) given by

$$p_{ij} = \delta_{i,\alpha} \quad (45)$$

where α is a function of j given by $\alpha = rM + \left[\frac{j}{p} \right] + 1$ where $j - 1 \equiv r(p)$, and $[x]$ is the integer part of x .

The covariance matrix (39) represents a conventional version based on the grand data matrix (38). This is the trajectory matrix method (Broomhead and King, 1986a,b; Ghil *et al.*, 2002). One could also compute an alternative grand block ‘covariance’ matrix $\mathcal{T} = (\mathbf{T}_{ij})$ whose blocks \mathbf{T}_{ij} , $i, j = 1, \dots, p$ represent lagged covariances between grid points i and j . This version is used by Plaut and Vautard (1994) who considered the longest possible segment of each channel (grid point) to compute the elements of each block \mathbf{T}_{ij} . The use of this (Toeplitz) version to compute EEOFs can result in a big matrix, which can be computationally expensive to diagonalise. The conventional (trajectory) version, however, can use SVD efficiently particularly when the sample size n is much smaller than the original number of variables p .

EEOFs are the EOFs of the extended data matrix (35) or (38), i.e. the eigenvectors of the grand covariance matrix Σ given in (39). They can be obtained directly by computing the eigenvalues/eigenvectors of (39). Alternatively, one can use SVD of the grand data matrix \mathcal{X} in (38) in a similar way to (14). Note that now we have $d = Mp$ new variables, i.e. the number of columns of the grand data matrix. The SVD of (38) yields:

$$\mathcal{X} = V\Theta U^T \quad (46)$$

where the $d \times d$ matrix $U = (u_{ij}) = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)$ represents the matrix of the Mp extended EOFs or right singular vectors of \mathcal{X} . The diagonal matrix Θ contains the singular values $\theta_1, \dots, \theta_d$ of \mathcal{X} , and $V = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_d)$ is the matrix of the left singular vectors or extended PC’s where the k ’th extended PC is $\mathbf{v}_k = (v_k(1), \dots, v_k(n - M + 1))^T$. These extended EOFs and PCs can be used to filter the data by removing the contribution from nonsignificant components and also for reconstruction purposes as detailed below

Data filtering and oscillation reconstruction

The extended EOFs U can be used as a filter exactly like EOFs. For instance the SVD decomposition (46) yields the expansion of each row \mathbf{x}_t of \mathcal{X} in (38)

$$\mathbf{x}_t^T = \sum_{k=1}^d \theta_k v_k(t) \mathbf{u}_k \quad (47)$$

for $t = 1, \dots, n - M + 1$, or in terms of the original variables \mathbf{x}_t , see Eq (42), as

$$\mathbf{x}_{t+j-1}^T = \sum_{k=1}^d \theta_k v_k(t) \mathbf{u}_k^j \quad (48)$$

for $j = 1, \dots, M$, and where

$$\mathbf{u}_k^j = (u_{j,k}, u_{j+M,k}, \dots, u_{(p-1)M,k})^T. \quad (49)$$

Note that the expression of the vector \mathbf{u}_k^j depends on the form of the data matrix. The one given above corresponds to (38), whereas when the data matrix \mathcal{X}_1 is used in (46) one gets

$$\mathbf{u}_k^j = (u_{(j-1)p+1,k}, u_{(j-1)p+2,k}, \dots, u_{jp,k})^T \quad (50)$$

Note also that when we filter out higher EEOFs, expression (48) is to be truncated to the required order $d_1 < d$.

The expansion (48) is exact by construction. However, very often one wants to truncate it by keeping a much smaller number of EEOFs for filtering purposes. This happens for instance when one reconstructs the field components from a single EEOF, or a pair of EEOFs corresponding for example to an oscillation. When this happens, the obtained expansion does not give a complete picture. This is because when (48) is truncated to a smaller subset K of EEOFs giving:

$$\mathbf{y}_{t+j-1}^T = \sum_{k \in K} \theta_k v_k(t) \mathbf{u}_k^j, \quad (51)$$

where $\mathbf{y}_t = (y_{t,1}, \dots, y_{t,p})$ is the filtered or reconstructed state space vector, then one obtains a multivalue function. For example, for $t = 1$ and $j = 2$, one gets one value of $y_{t,1}$ and for $t = 2$ and $j = 1$ one gets another value of $y_{t,1}$. This also occurs with single channel SSA. This occurs because EEOFs have time lagged components. To get a single reconstructed value we can simply take the average of those multiple values, but one could equally construct a ‘better’ weighted average. The number of multiple values depends on the value of time $t = 1, \dots, n$ (these numbers can be obtained by constructing an $M \times n$ array $A = (a_{jt})$ with entries $a_{jt} = t - j + 1$, then all entries that are nonpositive or greater than $n - M + 1$ are to be equated to zero, and finally for each time t take all the indices j with positive entries). The reconstructed variables using a subset K of EEOFs are then easily obtained from (51) by

$$\mathbf{y}_t^T = \begin{cases} \frac{1}{t} \sum_{j=1}^t \sum_K \theta_k v_k(t - j + 1) \mathbf{u}_k^j \\ \text{for } 1 \leq t \leq M - 1 \\ \frac{1}{M} \sum_{j=1}^M \sum_K \theta_k v_k(t - j + 1) \mathbf{u}_k^j \\ \text{for } M \leq t \leq n - M + 1 \\ \frac{1}{n - t + 1} \sum_{j=t-n+M}^M \sum_K \theta_k v_k(t - j + 1) \mathbf{u}_k^j \\ \text{for } n - M + 2 \leq t \leq n \end{cases} \quad (52)$$

Note that these reconstructions can also be obtained in a least square sense (see, e.g. Vautard *et al.*, 1992, and

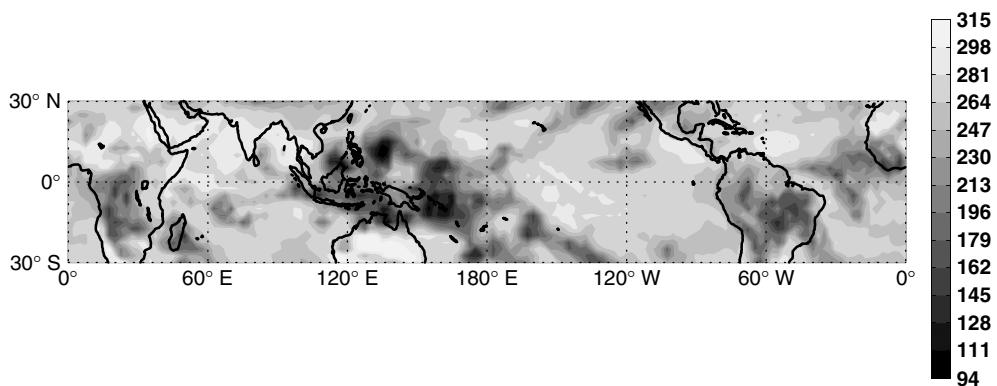
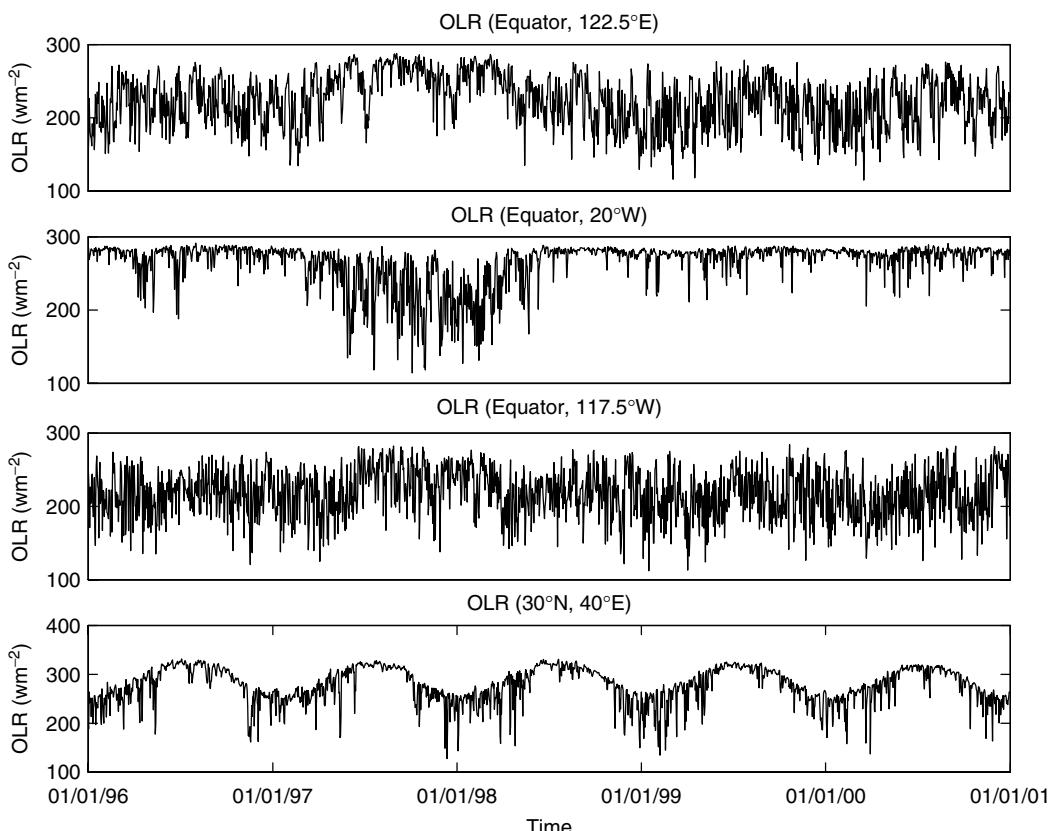
Ghil *et al.*, 2002). The reconstructed components can also be restricted to any subset of the eigen elements of the grand data matrix (38). For example to reconstruct the time series associated with oscillatory eigen elements, i.e. a pair of degenerate eigenvalues, the subset K in the sum (51) is limited to that pair.

The reconstructed multivariate time series \mathbf{y}_t , $t = 1, \dots, n$, can represent the reconstructed (or filtered) values of the original field at the original p grid points. In general, however, the number of grid points is too large to warrant an eigen-decomposition of the grand data, or covariance, matrix. In this case a dimension reduction of the data is first applied by using say the leading p_0 PCs and then applying a MSSA to these retained PCs. In this case the dimension of \mathcal{X} becomes $(n - M + 1) \times Mp_0$, which may be made considerably smaller than the original dimension. To get the reconstructed space-time field one can then use the reconstructed PCs (RPCs) in conjunction with the p_0 leading EOFs.

EEOFs can be efficient in detecting propagating structures. However, there are cases where the interpretation of individual EEOFs can be difficult and should be done with care (Chen and Harr, 1993). This happens particularly when the data contain a strong standing wave. Monahan *et al.* (1999) show that if the lag chosen in EEOF analysis is too close to the first zero of the sample autocorrelation function of the standing wave time series, then the wave signal obtained from EEOF can be substantially degraded and the interpretation can be difficult and misleading.

Application to outgoing long wave radiation

The EEOF method is applied here to identify the MJO. The MJO, an eastward propagating planetary-scale wave of tropical convective anomalies, is a well-established dominant mode of intra-seasonal tropical variability, and hence constitutes a convenient test-bed. The oscillation has a quite broad band with a period between about 40 and 60 days (Knutson and Weickmann, 1987; Hendon and Salby, 1994; Madden and Julian, 1994). It has been identified from various fields such as zonal and divergent wind, SLP, and OLR in the tropics (Madden and Julian, 1972; Kiladis and Weickmann, 1992), and here we choose to use OLR. For details on the mechanisms involved in MJO the reader is referred, for example, to Hendon and Salby (1994), Matthews (2000), and Krishnamurthi *et al.* (2003). The OLR data used here come from NCEP/NCAR reanalyses over the tropical region from 30°S to 30°N. This analysis focuses on a 5-year period of daily data from 1 January, 1996 to 31 December, 2000. Figure 13 shows the OLR field on 25 December 1996. Note the low-value regions particularly over the warm pool, an area of large and intensive convection, and to a lesser extent over the Amazonian and tropical African regions. The OLR data are not very homogeneous, and have a quite complicated variability as well as seasonality. To illustrate this complication, Figure 14 shows the OLR time series at four different

Figure 13. OLR distribution (w/m^2) over the tropics on the 25-Dec-1996.Figure 14. Time series of OLR (wm^{-2}) at four different locations.

locations. A clear indication of nonstationarity of the times series is obvious. Note for example the high values at the equator and $122.5^\circ E$ in northern winter 1997/98. This period corresponds to a strong El-Niño event where the convection shifts eastward to the mid-Pacific allowing more long wave radiation to be lost to space over the maritime continent. A decrease of OLR accompanied by a strong variability can also be seen during the same period at $20^\circ W$ on the Equator (Figure 14). This is due to the lower surface pressure over the Tropical Atlantic ocean during El-Niño, compared to normal conditions, see e.g. Webster and Chang (1988) or Holton (1992, Figure 11.10.) Note also the strong seasonal component at $30^\circ N$, with a particular stronger variability in the

winter time compared to other seasons. Before applying EEOF analysis, the data were first subjected to an EOF analysis to reduce the dimension of the data. The leading 10 EOFs/PCs of the anomaly field with respect to the long term average (climatology) were retained for the analysis. Figure 15 shows the leading EOF mode. This pattern explains about 15% of the total variability and is associated with the seasonal cycle. Figure 15 clearly indicates that the seasonal cycle is mostly explained by the Inter Tropical Convergence Zone (ITCZ) and some monsoonal activities.

The first 10 EOFs/PCs used for the EEOF/MSSA analysis together explain about 32% of the total variability. A window length corresponding to $M = 80$ days is used

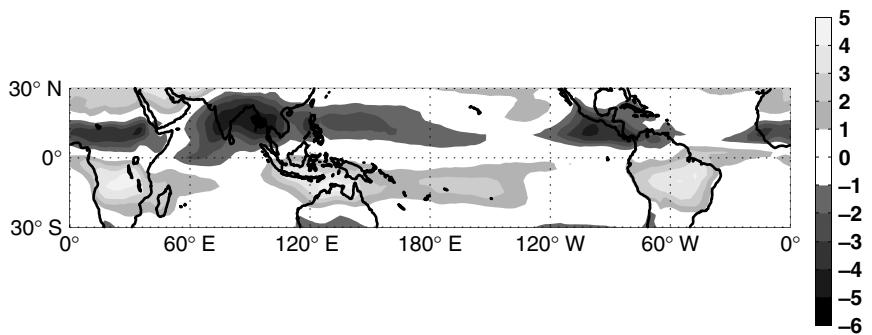


Figure 15. The leading EOF of OLR anomalies. Units are arbitrary.

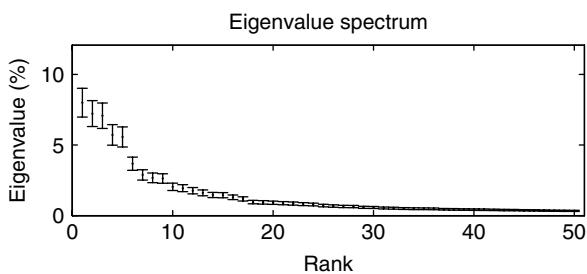


Figure 16. Spectrum of the grand covariance matrix (39). Approximate standard errors are derived from (18) using an effective sample size of 116.

to construct the grand data matrix (38). This choice was motivated by the desire to capture the MJO identified first by Madden and Julian (1971). This choice also avoids the problem of interpretation in the presence of a standing wave (Monahan *et al.*, 1999). The extended (grand) data matrix is then factorised using SVD. Figure 16 shows the spectrum of the grand covariance matrix Σ (39). The standard errors are obtained using (18) with a heuristic effective sample size of 116, corresponding to a decorrelation lag of 15 days.

The first two eigenvalues of the spectrum (Figure 16) correspond to the oscillation of the seasonal cycle. They do not look nearly equal and well separated from the rest of the spectrum, however, and this is due to the choice of the window length, which is much smaller than the

length of the seasonal cycle. Despite this, the first two extended PCs (EPCs) show a pair of sine waves perfectly in quadrature, and represent the seasonal cycle. Figure 17 shows in fact the raw PC1 along with the reconstructed or smoothed PC1. The reconstruction is based on (52) using the leading 5 EPCs.

Beside the annual cycle, the (degenerate) fourth and fifth eigenvalues constitute also another oscillatory pair corresponding to the semi-annual cycle. This degeneracy can be seen by using the rule of thumb (18) but without serial correlation. The left panel of Figure 18 shows a time plot of EPC4 and EPC5 whereas the phase diagram of EPC4 versus EPC5 is shown in the right panel of Figure 18. The figure clearly shows the semi-annual oscillation (SAO) in OLR. The phase diagram (right panel) also shows the SAO with slight irregularities due to interannual variability. The next oscillatory pair corresponds to the 8th and 9th eigenvalues (Figure 16). This pair of (nearly) equal eigenvalues corresponds to the familiar MJO, and corresponds to a period of about 50 days.

Figure 19 shows extended EOF 8 along 10°N as a function of time lag. This sort of diagram where the space and time axes are shown is known as Hövmoller diagram. Figure 19 shows clearly the eastward moving oscillation with an average phase speed around 7°/day, making the wave travel from west to east in roughly 50 days, i.e. approximately 9 m s⁻¹ as observed in Knutson and Weickmann (1987). This is also comparable to the speed

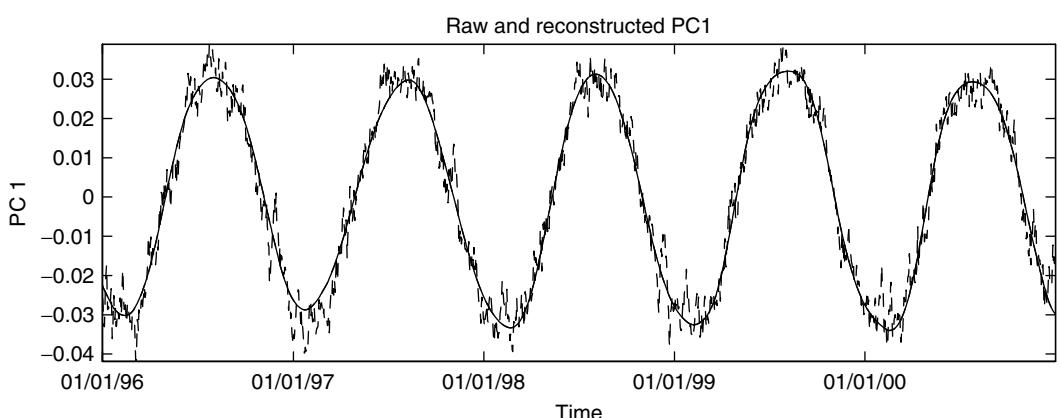


Figure 17. Time series of raw and reconstructed PC1.

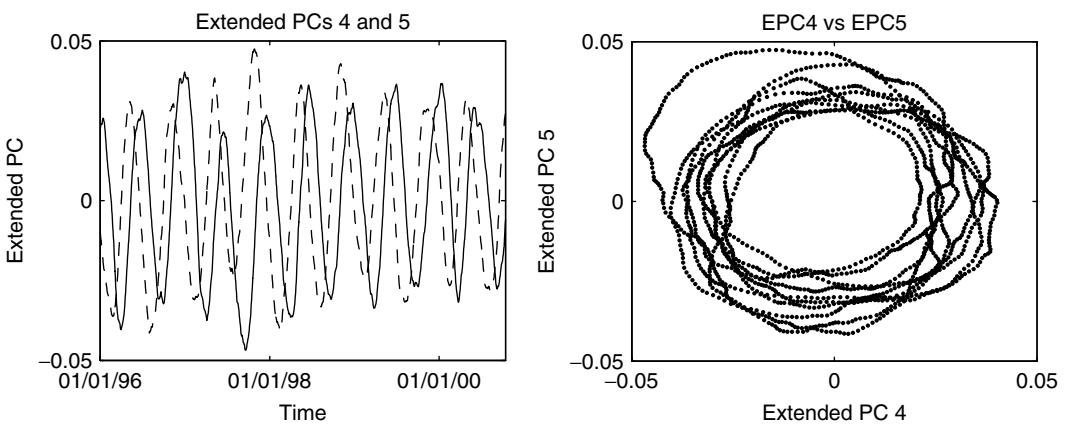


Figure 18. Time plot of EPC4 (—) and EPC5 (---) (left panel) and phase diagram of EPC4 versus EPC5 (right panel).

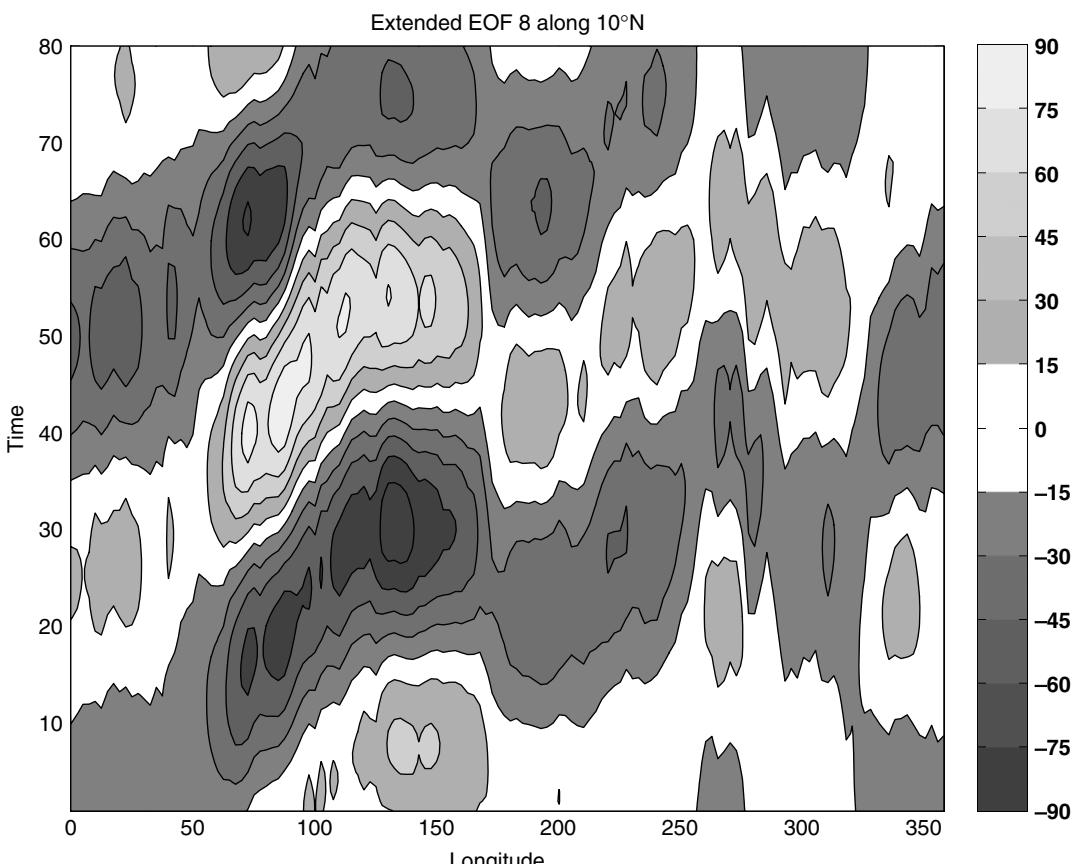


Figure 19. Extended EOF 8 along 10°N as a function of time lag. Units arbitrary.

of the Kelvin wave, obtained as a radiating response and propagating away from the convective anomaly, by Hendon and Salby (1994). Figure 20 shows the time plot of the extended PCs 8 and 9 along with the phase diagram of EPC 8 versus EPC 9. The EPCs are oscillating in quadrature, with a period of about 50 days and with varying intensity. The power spectrum of EPC8 (not shown) peaks around 50 days with a slight broad band structure.

The MJO can be found in the reconstructed PCs using the extended EOFs/PCs. Figure 21, for example, shows a time plot of the reconstructed PC2 using the

8th EEOF/EPC. The figure shows clearly the oscillation with varying amplitude. Note in particular the strong oscillation during winter 1997 *versus* the weak oscillation from winter to early autumn 1998. The same behaviour is also observed in the phase diagram in the RPC 6 versus RPC 7 (Figure 22).

Since EOF analysis does not focus on a particular frequency band, one would expect MJO to project onto more than one PC. In fact, MJO is found to project onto various PCs but with differing energies (amplitudes). For instance SSA analysis of single PCs reveals that MJO is mostly present in higher PCs, e.g.

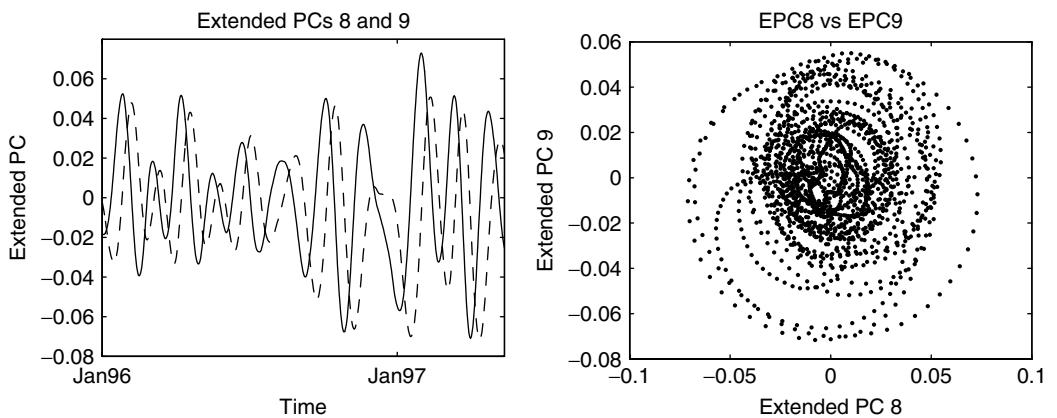


Figure 20. Extended PCs 8 and 9 (left panel) and phase diagram of EPC8 *versus* EPC9 (right panel).

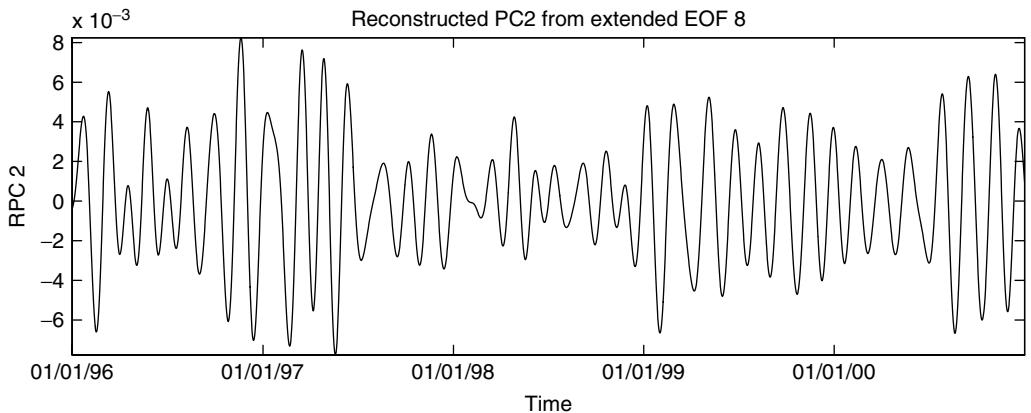


Figure 21. Reconstructed PC2 using extended EOF/PC 8.

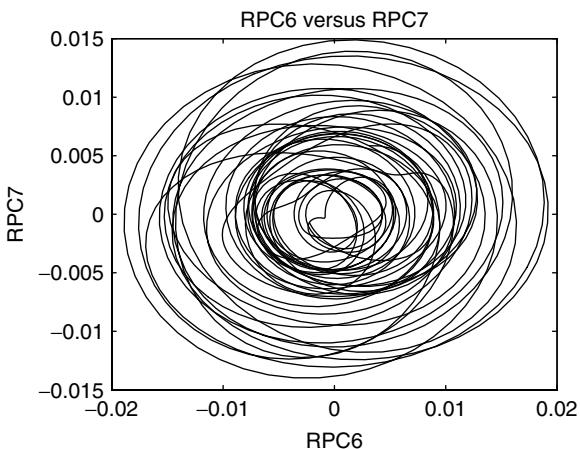


Figure 22. Phase diagram of reconstructed PC 6 *versus* reconstructed PC 7.

PCs 6 and 7. This observation is also revealed from analysing the reconstructed PC's using the extended EOFs/PCs 8 and 9 associated with the oscillation. This is well illustrated in Figure 23, which shows the first 8 reconstructed PCs. Reconstructed PCs 5–8 are the most energetic components representing MJO. Note in particular the weak projection of MJO onto the first PC, which represents only the seasonal cycle.

The reconstructed or filtered PCs allow a reconstruction of the original space-time OLR field. Figure 24 shows a Hövmoller diagram of the reconstructed OLR field from 3 March, 1997 to 14 May, 1997, at 5°N using the MJO-related extended EOFs/PCs 8 and 9. The reconstructed PCs 1–8 (Figure 23) are used in conjunction with the EOFs as in (17), but with the sum truncated to the leading 8 EOFs, to get the MJO reconstructed field. It is very important to note here that a similar diagram, but using the raw data, does not reveal any feature, and the picture (not shown) looks noisy and featureless. It is clear from Figure 24 that MJO is triggered around 25–30°E over the African jet region. It reaches its mature stage over the Indian ocean in the Bay of Bengal and starts to decay thereafter. The MJO becomes particularly damp near 150°E over the convective region in the warm pool. Figure 24 shows also the dispersive nature of the MJO with a stronger phase speed during the growth phase compared to that observed during the decay phase.

COMPLEX/HILBERT EOFS

Background

Conventional EOF analysis can be applied to a single space-time field or a combination of fields. EOF analysis finds “stationary” patterns in the sense that they are not

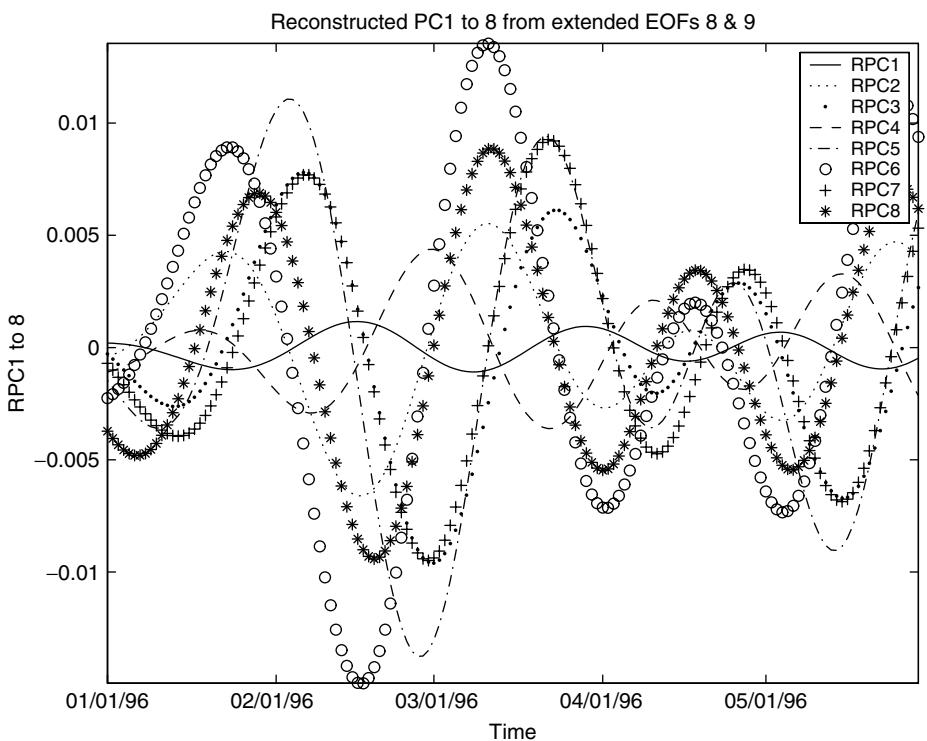


Figure 23. Reconstructed PCs 1 to 8 using extended EOFs/PCs 8 and 9.

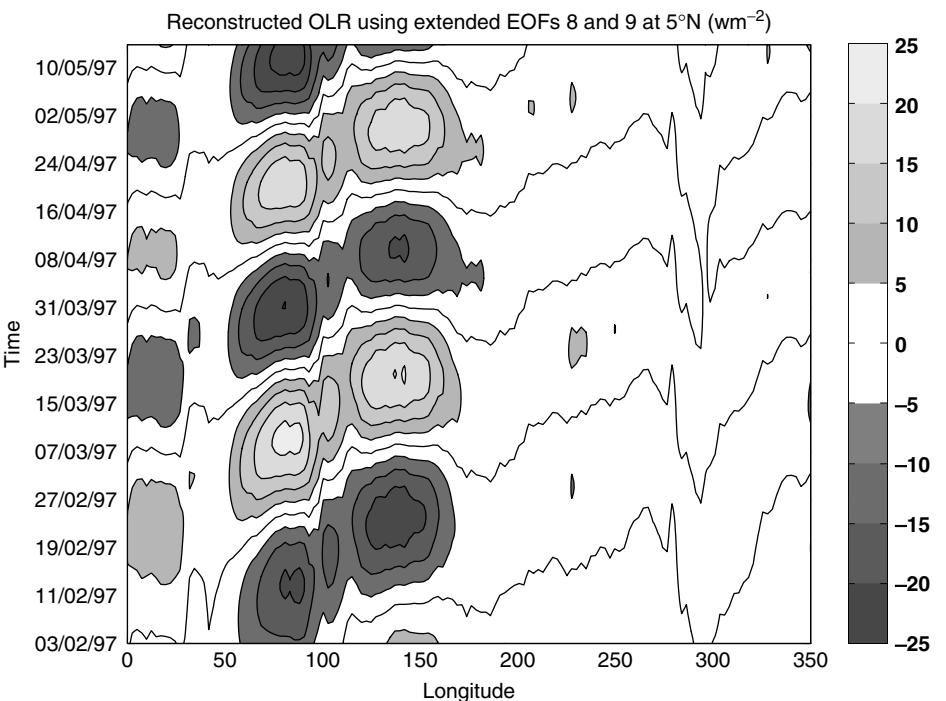


Figure 24. Reconstructed OLR field using reconstructed EOFs/PCs 1 to 8 shown in Figure 23.

evolving. It yields a varying time series for any obtained EOF pattern, which means that the spatial EOF pattern will only decrease or increase in magnitude whereas the spatial structure remains the same. Because EOFs are based on (simultaneous) covariances, the way time is arranged is irrelevant. In fact, if x_t and y_t , $t = 1, \dots, n$, are two univariate time series then any permutation of x_t

and y_t will yield the same covariance, i.e.

$$\rho_{xy} = \text{cov}(x_t, y_t) = \text{cov}(x_{\pi(t)}, y_{\pi(t)}) \quad (53)$$

where π is any permutation of the set of time indices $\{1, 2, \dots, n\}$.

This automatically means that any propagating structure in the field will not be captured by EOFs. We have

seen in the previous section that EEOFs can extract such information in the time domain by analysing a relatively large data matrix. In addition, one has to fix a priori the lag-window. There are also other ways, such as POP analysis (e.g. von Storch and Zwiers 1999), that can find these structures. Here we review another method similar to EOF analysis but based on the complexified field. The method does not explicitly involve the lagged information, hence avoiding the use of a similar grand data matrix to that of EEOFs, and also avoiding the problem of choosing the lag-window.

It is known that any wave can be expressed using complex representation as:

$$x(t) = ae^{i\omega t + \phi} \quad (54)$$

where a is the wave amplitude and ω and ϕ are respectively its frequency and phase shift (at the origin). Complex empirical orthogonal functions (CEOFs) are based on this representation. There are, in principle, two ways to perform complex EOFs, namely ‘conventional’ complex EOFs and ‘Hilbert’ EOFs. When we deal with a pair of associated climate fields then conventional complex EOFs are obtained. HEOFs correspond to the case when we deal with a single variable, i.e. single field, and where we are interested in finding propagating patterns. In this case the field has to be complexified by introducing an imaginary part, which is a transform of the actual field.

Conventional complex EOFs

Pairs of scalar fields. The method is similar to conventional EOFs except that it is applied to the complex field obtained from a pair of associated variables such as the zonal and meridional components u and v of the wind field $\mathcal{U} = (u, v)$ (Kundu and Allen, 1976; Hardy and Walton, 1978; Salstein *et al.*, 1983; Brink and Muench, 1986; von Storch and Zwiers, 1999; Preisendorfer, 1988). The wind field $\mathcal{U}_{tl} = \mathcal{U}(t, \mathbf{s}_l)$, defined at each location \mathbf{s}_l , $l = 1, \dots, p$, and time t , $t = 1, \dots, n$, can be written using a compact complex form as:

$$\mathcal{U}_{tl} = u(t, \mathbf{s}_l) + i v(t, \mathbf{s}_l) = u_{tl} + i v_{tl}. \quad (55)$$

The complex covariance matrix is obtained using the data matrix $\mathcal{U} = (\mathcal{U}_{tl})$, and is given by:

$$S = \frac{1}{n-1} \mathcal{U}^* \mathcal{U}, \quad (56)$$

The elements s_{kl} , $k, l = 1, \dots, p$, of S in (56) are given by:

$$s_{kl} = \frac{1}{n} \sum_{t=1}^n \mathcal{U}_{tk}^* \mathcal{U}_{tl} \quad (57)$$

where $(*)$ is the complex conjugate operator. The (complex) covariance matrix (56) is Hermitian, i.e. $S^{*T} = S$, and is therefore diagonalisable. The matrix has therefore a set of orthonormal complex eigenvectors $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_p)$, and a real non-negative eigenspectrum

$\lambda_1^2, \dots, \lambda_p^2$. The complex amplitude of the k th EOF is the k th complex principal component (CPC) \mathbf{e}_k , and is given by:

$$\mathbf{e}_k = \mathcal{U} \mathbf{u}_k^* \quad (58)$$

This immediately yields the non-correlation between the CPCs:

$$\mathbf{e}_k^* \mathbf{e}_l = \lambda_k^2 \delta_{kl}. \quad (59)$$

Complex EOFs and associated complex PCs can also be obtained using the SVD of \mathcal{U} . Any CEOF \mathbf{u}_k has a pattern amplitude and phase. The phase and amplitude at each grid are obtained using the real and imaginary parts of the loadings.

This method of doing CEOFs seems to have been originally applied by Kundu and Allen (1976) to the velocity field of the Oregon coastal current. The conventional CEOFs are similar to conventional EOFs in the sense that time ordering is irrelevant, and hence the method is mostly useful to capture covarying spatial patterns between the two fields.

Single field. If one is dealing with a single field $\mathbf{x}_t = (x_{t1}, \dots, x_{tp})^T$, $t = 1, 2, \dots, n$, such as sea surface temperature, and one is interested in propagating patterns one can still use the conventional complex EOFs applied to the complexified field obtained from a pair of lagged variables $(\mathbf{x}_t, \mathbf{x}_{t+\tau})$ for some chosen lag τ . The complex field is defined by:

$$\mathbf{y}_t = \mathbf{x}_t + i \mathbf{x}_{t+\tau}. \quad (60)$$

This is a natural way to define a homogeneous complexified field using lagged information. The corresponding complex data matrix defined from (60) is then given at each grid point \mathbf{s}_l and each time t by $Y_{tl} = (x_{tl} + i x_{t+\tau,l})$. The obtained complex data matrix $\mathbf{Y} = (Y_{tl})$ can then be submitted to the same complex EOF analysis as in the previous section.

The obtained CEOFs provide the propagating structures and the corresponding CPCs provide the phase information. This procedure is based on the choice of the lag time τ , which reflects the characteristic time of the propagating feature. In general, however, this parameter is not precisely known, and requires some experience. The choice of this parameter remains, in practice, subject to some arbitrariness. To avoid this difficulty in choosing the lag in the time domain, the Hilbert transform provides an alternative, based on phase shift in the frequency domain. This is presented next.

Complex Hilbert EOFs

Frequency domain EOFs. It appears that the earliest introduction of complex EOFs in an atmospheric context dates back to the early seventies with Wallace and Dickinson on frequency domain EOFs (FDEOFs). Their work stimulated the introduction later of HEOFs, and we start by reviewing FDEOFs first. The spectrum gives a

measure of the contribution to the variance across the whole frequency range. EOF analysis in the frequency domain (Wallace and Dickinson, 1972; Wallace, 1972; Johnson and McPhaden, 1993; see also Brillinger 1981 for details) attempts to analyse propagating disturbances by concentrating on a specific frequency-band allowing thus the decomposition of variance in this band while retaining phase relationships between locations.

FDEOFs are based on performing an eigenanalysis of the cross-spectrum matrix calculated in a small frequency band. Let $\mathbf{u}(\omega)$ be the Fourier transform (FT) of the (centered) field \mathbf{x}_t , $t = 1, \dots, n$ at frequency ω , i.e.

$$\mathbf{u}(\omega) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{x}_t e^{-i\omega t} \quad (61)$$

The cross-spectral matrix at ω is $\Gamma_{xx}(\omega) = \mathbf{u}(\omega)\mathbf{u}(\omega)^T$, and can be written in terms of the lagged covariance matrix

$$S_{xx}(\tau) = \frac{1}{n} \sum_{t=1}^{n-\tau} \mathbf{x}_t \mathbf{x}_{t+\tau}^T \quad (62)$$

as:

$$\Gamma_{xx}(\omega) = \sum_{\tau} S_{xx}(\tau) e^{-i\omega\tau} = \Gamma_{xx}^R(\omega) + \Gamma_{xx}^I(\omega). \quad (63)$$

The real part of the cross-spectrum, $\Gamma_{xx}^R(\omega)$, is the co-spectrum, and the imaginary part, $\Gamma_{xx}^I(\omega)$, is the quadrature spectrum. Note that the covariance matrix of the field satisfies

$$S = \int_{-\omega_N}^{\omega_N} \Gamma_{xx}(\omega) d\omega = 2 \int_0^{\omega_N} \Gamma_{xx}^R(\omega) d\omega, \quad (64)$$

where $\omega_N = \frac{1}{2\Delta t}$ is the Nyquist frequency and Δt is the time interval between observations. Therefore the spectrum gives a measure of the contribution to the variance across the whole frequency range. The average of the cross-spectral matrix over the frequency band $[\omega_0, \omega_1]$ is given by

$$C = \int_{\omega_0}^{\omega_1} \Gamma_{xx}(\omega) d\omega \quad (65)$$

and provides a measure of the contribution to the covariance matrix in that frequency band.

Now since waves are coherent structures with consistent phase relationship at various lags, and given that FDEOFs represent patterns that are uniform across a frequency band, the leading FDEOF provides coherent structures with most wave variance. The FDEOFs are then obtained as the EOFs of C , (see Brillinger 1981) for more details. Johnson and McPhaden (1993) have applied FDEOFs to study the spatial structure of intraseasonal Kelvin wave structure in the Equatorial Pacific ocean. They identified coherent wave structures with periods of 59–125 days. Because most spectra of climate data look

reddish, FDEOF analysis may be cumbersome in practice (Horel, 1984). This is particularly the case if the power spectrum of an EOF is spread, for example, over a wide frequency band, requiring an averaging of the cross-spectrum over this wide frequency range, where the theory behind FDEOFs is no longer applicable (Wallace and Dickinson, 1972). This difficulty has resulted in the method being abandoned in climate research in favour of HEOFs described next.

Complex Hilbert EOFs. An elegant alternative to FDEOFs is the complex EOFs in the time domain introduced into atmospheric science by Rasmusson *et al.* (1981) using Hilbert singular decomposition. The method was refined later by Barnett (1983) and applied to the monsoon (Barnett, 1984a,b), atmospheric angular momentum (Anderson and Rosen, 1983), the QBO in northern hemispheric SLP (Trenberth and Shin, 1984), and coastal ocean currents (Merrifield and Winant, 1989). The method is based on the Hilbert transform and is therefore referred to as HEOF analysis.

Let $\mathbf{x}_t = (x_{t1}, \dots, x_{tp})^T$, $t = 1, \dots, n$, be a scalar field, with Fourier representation:

$$\mathbf{x}_t = \sum_{\omega} \mathbf{a}(\omega) \cos \omega t + \mathbf{b}(\omega) \sin \omega t \quad (66)$$

where $\mathbf{a}(\omega)$ and $\mathbf{b}(\omega)$ are vector Fourier coefficients. Since propagating disturbances require complex representation as in (54), Eq (66) can be transformed to yield the general (complex) Fourier decomposition:

$$\mathbf{y}_t = \sum_{\omega} \mathbf{c}(\omega) e^{-i\omega t} \quad (67)$$

where precisely $Re(\mathbf{y}_t) = \mathbf{x}_t$, and $\mathbf{c}(\omega) = \mathbf{a}(\omega) + i\mathbf{b}(\omega)$. The new complex field $\mathbf{y}_t = (y_{t1}, \dots, y_{tp})^T$ can therefore be written as:

$$\mathbf{y}_t = \mathbf{x}_t + i\mathcal{H}(\mathbf{x}_t). \quad (68)$$

The imaginary part of \mathbf{y}_t is given by:

$$\mathcal{H}(\mathbf{x}_t) = \sum_{\omega} \mathbf{b}(\omega) \cos \omega t - \mathbf{a}(\omega) \sin \omega t, \quad (69)$$

and is precisely the Hilbert transform, or quadrature function of the scalar field \mathbf{x}_t and is seen to represent a simple phase shift by $\frac{\pi}{2}$ in time. In fact, it can be seen that the Hilbert transform, considered as a filter, removes the zero frequency without affecting the modulus of all the others, and is as such a unit gain filter. Note that if the time series (66) contains only one frequency then the Hilbert transform is simply proportional to the time derivative of the time series. Therefore, locally in the frequency domain $\mathcal{H}(\mathbf{x}_t)$ provides information about the rate of change of \mathbf{x}_t with respect to time t . In formal terms the Hilbert transform of a continuous time series

$x(t)$ (Thomas, 1969; Brillinger, 1981) is defined by the convolution

$$\mathcal{H}(x(t)) = \frac{1}{\pi} \int \frac{x(u)}{t-u} du \quad (70)$$

where the integral is taken to mean the Cauchy principal value. In the discrete case the Hilbert transform can be derived, in time domain, by applying a rectangular rule to (70) to yield (Kress and Martensen, 1970; see also Weideman, 1995)

$$\mathcal{H}(x(t)) \approx \frac{2}{\pi} \sum_{k=-\infty}^{\infty} \frac{x(t + (2k+1)h))}{2k+1} \quad (71)$$

where h is the step size. When Eq (71) is applied to a discrete time series x_t , $t = 0, \pm 1, \dots$, one gets (see also Bloomfield and Davis, 1994; and von Storch and Zwiers 1999):

$$\begin{aligned} \mathcal{H}(x_t) &= \frac{2}{\pi} \sum_{k=-\infty}^{\infty} \frac{x_{t+2k+1}}{2k+1} \\ &= \frac{2}{\pi} \sum_{k \geq 0} \frac{1}{2k+1} (x_{t+2k+1} - x_{t-2k-1}). \end{aligned} \quad (72)$$

In practice, various methods exist to compute the finite Hilbert transform. For a scalar field \mathbf{x}_t of finite length n , the Hilbert transform $\mathcal{H}(\mathbf{x}_t)$ can be estimated using the discrete FT (69) in which ω becomes $\omega_k = \frac{2\pi k}{n}$, $k = 1, \dots, \left[\frac{n}{2}\right]$. Alternatively, $\mathcal{H}(x_t)$ can be obtained by truncating the infinite sum in Eq (72). This truncation can also be written using a convolution or a linear filter as (see e.g. Hannan, 1970):

$$\mathcal{H}(x_t) = \sum_{k=-L}^L \alpha_k \mathbf{x}_{t-k} \quad (73)$$

with the filter weights

$$\alpha_k = \frac{2}{k\pi} \sin^2 \frac{\pi k}{2}. \quad (74)$$

Barnett (1983) found that $7 \leq L \leq 25$ provides adequate values for L . For example for $L = 23$ the frequency response function is a band pass filter with periods between 6 and 190 time units with a particular excellent response obtained between 19 and 42 time units (Trenberth and Shin, 1984). The Hilbert transform has also been extended to vector fields, i.e. two or more fields, through concatenation of the respective complexified fields (Barnett, 1983). Another interesting method to compute Hilbert transform of a time series is presented by Weideman (1995), using a series expansion in rational eigenfunctions of the Hilbert transform operator (70).

The HEOFs \mathbf{u}_k , $k = 1, \dots, p$, are then obtained as the eigenvectors of the Hermitian covariance matrix

$$\mathcal{S}_{yy} = \frac{1}{n} \sum_{k=1}^n \mathbf{y}_t \mathbf{y}_t^{*T} = 2(S_{xx} + i S_{\mathcal{H}(x)x}), \quad (75)$$

where $S_{\mathcal{H}(x)x}$ is the cross-covariance matrix between $\mathcal{H}(\mathbf{x}_t)$ and \mathbf{x}_t . They can be obtained also as the right complex singular vectors of the data matrix $Y = (y_{tk})$. The uncorrelated complex principal components (CPCs) \mathbf{z}_k , for $k = 1, \dots, p$, are then obtained similarly to (58). From this decomposition we also get the spatial amplitude and phase functions respectively:

$$\begin{aligned} \mathbf{a}_k &= \mathbf{u}_k \bullet \mathbf{u}_k^* = \text{Diag}(\mathbf{u}_k \mathbf{u}_k^{*T}) \\ \theta_k &= \text{atan} \left[\frac{\text{Im}(\mathbf{u}_k)}{\text{Re}(\mathbf{u}_k)} \right] \end{aligned} \quad (76)$$

where the vector product and division are performed component-wise, and where $\text{Re}()$ and $\text{Im}()$ represent respectively the real and imaginary parts. Similarly, one also gets the temporal amplitude and phase functions as:

$$\begin{aligned} \mathbf{b}_k &= \mathbf{z}_k \bullet \mathbf{z}_k^* = \text{Diag}(\mathbf{z}_k \mathbf{z}_k^{*T}) \\ \phi_k &= \text{atan} \left[\frac{\text{Im}(\mathbf{z}_k)}{\text{Re}(\mathbf{z}_k)} \right], \end{aligned} \quad (77)$$

The function θ_k gives information on the relative phase. For “simple” fields, its spatial derivative provides a measure of the local wavenumber. Its interpretation for moderately complex fields/waves can be difficult (Wallace, 1972), and can be made easier by applying a prior filtering (Barnett, 1983). Also for simple waves, the time derivative of the temporal phase gives a measure of the instantaneous frequency. Note that the phase speed of the wave at time t and position x can be measured by $\frac{\theta_{k(x)}}{\phi_{k(t)}}$.

The covariance matrix \mathcal{S}_{yy} in (75) can be shown to be related to the cross spectrum matrix

$$\Gamma_{yy}(\omega) = \frac{1}{n} \sum_{\tau} \sum_{t=1}^{n-k} \mathbf{y}_{t+\tau} \mathbf{y}_t^{*T} e^{i\omega\tau} \quad (78)$$

according to:

$$\mathcal{S}_{yy} = \int_{-\omega_N}^{\omega_N} \Gamma_{yy}(\omega) d\omega = 4 \int_0^{\omega_N} \Gamma_{xx}(\omega) d\omega \quad (79)$$

Because the covariance matrix \mathcal{S}_{xx} is only related to the co-spectrum Γ_{xx}^R , it is clear that conventional EOFs do not take into consideration the quadrature part of the cross-spectrum matrix. Using therefore the cross-spectrum matrix, as given by (79), it is seen that HEOFs generalise conventional EOFs.

It is also clear from (79) that HEOFs are equivalent to FDEOFs with the cross-spectrum integrated over all frequencies. Note that the frequency band of interest can be controlled by prior smoothing. Horel (1984)

points out that HEOFs can fail to detect irregularly occurring progressive waves (see also Merrifield and Winant 1989). Merrifield and Guza (1990) have showed that complex EOF analysis in the time domain HEOFs is not appropriate for non-dispersive and broad-banded waves in wavenumber $\Delta\kappa$ relative to the largest separation measured (array size Δx). In fact Merrifield and Guza (1990) (see also Johnson and McPhaden 1993), identified $\Delta\kappa \Delta x$ as the main parameter causing spread of propagating variability into more than one HEOF mode, and the larger the parameter, the smaller the data variance is captured.

Application

In this section we apply HEOFs to the QBO using stratospheric zonal winds. The section is divided into two subsections. In the first we describe the data and the zonal wind, and in the second subsection we present the application.

Structure of the zonal wind. The data used here are taken from the European Reanalyses (ERA40), which come from the European Centre for Medium Range Weather Forecasting (ECMWF), and consist of tropospheric and stratospheric monthly zonal wind. The data are provided on a regular $2.5^\circ \times 2.5^\circ$ horizontal grid on 23 vertical pressure levels from 1000 mb to 1 mb. The data span the period January 1958 to December 2001. Stratospheric and upper tropospheric zonal winds tend to be nearly zonally symmetric. We have therefore focussed only on the zonal-mean zonal wind, which we also refer to (for simplicity) as zonal wind hereafter.

Zonal wind in the stratosphere has a different pattern of variability to that in the troposphere for various reasons, not least the fact that the stratosphere is a free atmosphere, i.e. far from land/sea influence, e.g. friction; add to this the smaller density of the stratosphere, and the forcing effect of vertically propagating Rossby and gravity waves. Figure 25 shows the climatology of the zonal wind for January (25a) and July (25b). Various features can be noted from Figure 25. First, the tropospheric upper westerly jets can be seen in both hemispheres around 250 mb. In January, the Northern Hemispheric (NH) jet is only slightly stronger than its Southern Hemispheric (SH) counterpart. In July, however, the NH jet has weakened and the SH jet becomes much stronger. The SH jet is more active and stronger than the NH one due in part to the absence of boundary layer friction caused by land/mountains. In the stratosphere, however, the flow is entirely different. Here one can see both the easterly and the westerly flows. During NH winter, stratospheric westerly flow is over most extratropical NH and shows basically the polar vortex whereas SH stratospheric flow is mostly easterly. During NH summer the picture is reversed. Again, over the SH, stratospheric westerly reaching 100 m/s is much stronger than its NH analogue (around 40 m/s). Stratospheric easterlies, on the other hand, have nearly the same magnitude, around 50 m/s on both hemispheres.

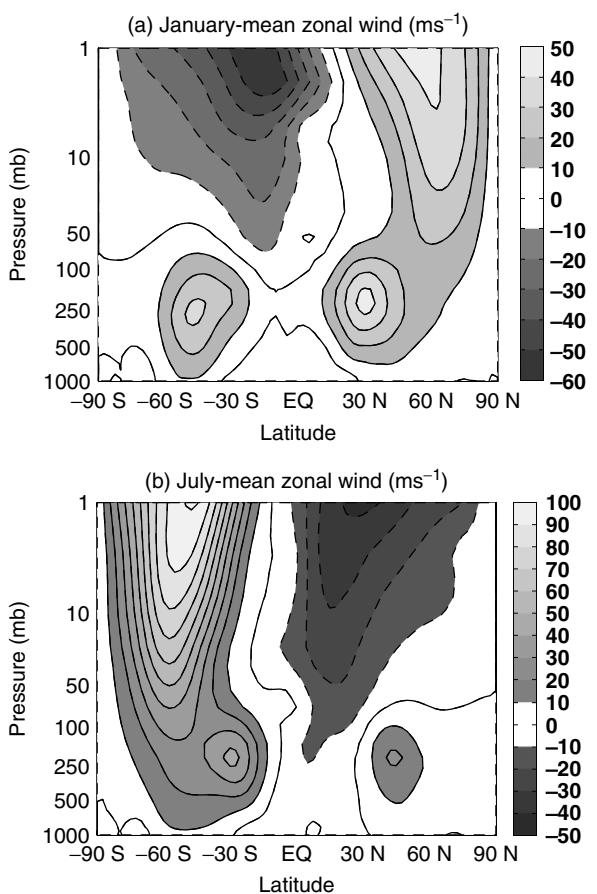


Figure 25. Climatology of ECMWF zonal mean zonal wind for January (a), and July (b). The period of observation spans the years 1958 through to 2001.

These stratospheric winds have been known since the 19th century when the German meteorologist A. Berson found them in 1908 through balloon measurements, and were known as Berson westerlies (Baldwin *et al.*, 2001). It was only in the 1950s that alternating winds were known to exist (Palmer, 1954; Graystone, 1959). These changes in stratospheric winds are merely due to seasonal variations.

To remove the effect of seasonality, we first calculate the mean annual cycle, which is then subtracted from the data to give the anomalies. When seasonal variations are removed, the picture becomes different. Figure 26 shows the variance of the zonal wind anomalies over the observed period. Most of the variance is concentrated around a narrow latitudinal band from approximately 15°S to 15°N , and extending from around 70 mb up to the 1 mb level. This region is of great importance to climate researchers working on the stratosphere. Figure 27 shows a time-height plot of the zonal wind anomalies at the Equator from January 1994 to December 2001.

A clear downward propagating signal can be seen from Figure 27. The propagation region is between around 3 mb and 70 mb, which is the approximate height of the tropopause at the equator. The descent speed is variable and is roughly between 1 and 1.4 km/month. The alternating wind region at any given vertical level

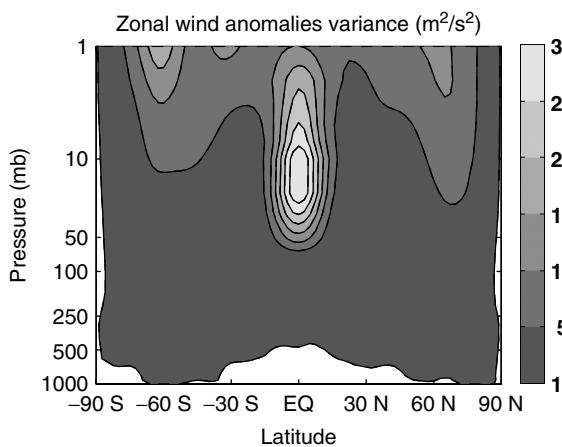


Figure 26. Variance of monthly zonal mean zonal wind anomalies for the observed period.

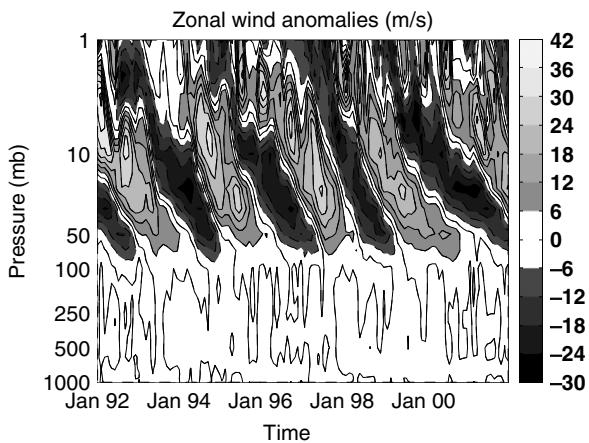


Figure 27. Hövmoller plot of equatorial zonal mean zonal wind anomalies at various vertical levels for the period January 1992–December 2001.

between 3 mb and 70 mb also varies, with a period between about 24 and 34 months. This quasi-oscillation has become known as the QBO, a term that was coined by Angell and Korshover (1964), but the quasi-biennial periodicity and the downward propagation were only discovered a little earlier by Ebdon (1960) and Reed *et al.* (1961). For reviews of the QBO, see Maruyama (1997) and Labitzke and van Loon (1999), and Baldwin *et al.* (2001) and references therein.

Hilbert EOF analysis of zonal wind. Since most variability of the QBO is concentrated around the Equator we have focussed on the mean zonal wind anomalies over the latitudinal band between 15S and 15N at all pressure levels. The Hilbert transformed data are obtained by computing the Hilbert transform of the field at each grid point, then forming the complexified field according to Eq (68). The Hilbert transform is computed using Eq (69) via a fast FT, which we found more efficient, in terms of CPU time, than the time domain transform (Eq 73). The eigenvalues spectra and corresponding eigenvectors are computed using the SVD (Eq 14) of the complex data

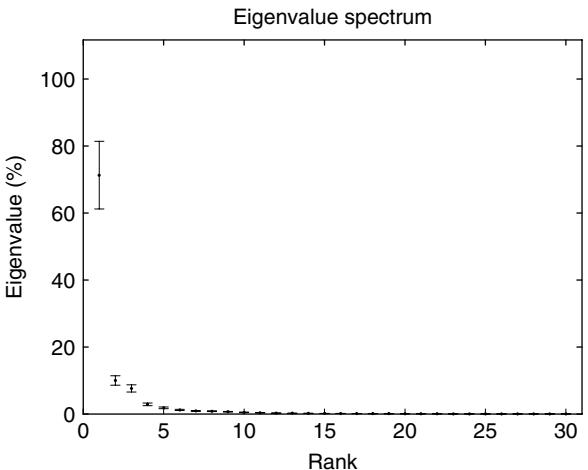


Figure 28. Spectrum of the covariance matrix S_{yy} , Equation (75), of zonal mean zonal wind anomalies. Vertical bars represent approximate 95% limits using the rule of thumb (18) with a heuristic sample size of 100.

matrix Y obtained from the complexified field (68). The CEOFs and corresponding CPCs are respectively given by the left and right complex singular vectors of the complex data matrix Y .

Figure 28 shows the spectrum of the covariance matrix $S_{yy} = \frac{1}{n} YY^*T$, Eq (75), where only the leading 30 eigenvalues, expressed in percentage represented variance, are shown. The uncertainties on the eigenvalues are obtained using the rule of thumb (18) with a heuristic sample size of 100. It is difficult to talk about the number of degrees of freedom here because of the strong autocorrelation due to the presence of the QBO cycle, but the point here is simply to stress the leading role of the first eigenvalue. It is clear that the first eigenvalue, to which we restrict our discussion here, is well separated from the rest of the spectrum and explains a substantial amount of variance. Figure 29 shows the real (29a) and the imaginary (29b) parts of CEOF1 respectively. The leading CEOF (Figure 29) is composed of a pair of patterns in quadrature, an indication of the propagating nature of the patterns as is shown later using the corresponding complex principal component CPC1. The direction of propagation is also given by the axis linking the low and high centers of actions.

Figure 30 shows the real and imaginary parts of the leading CPC (30a), the phase portrait (30b) of CPC1 and the power spectrum of its real part (30c). The real and imaginary parts are in quadrature and show a near periodic signal (Figure 30(a),(b)) reflecting the propagation of the corresponding pattern. The period is also shown in the power spectrum (Figure 30(c)) to be around 30–32 months. Note that, for this particular case, if a conventional EOF analysis were performed instead, a degenerate leading pair of eigenvalues would have been obtained, and whose EOFs would be similar to those shown in Figure 29, and similarly for the PCs. In general, however, this might not be the case particularly when the

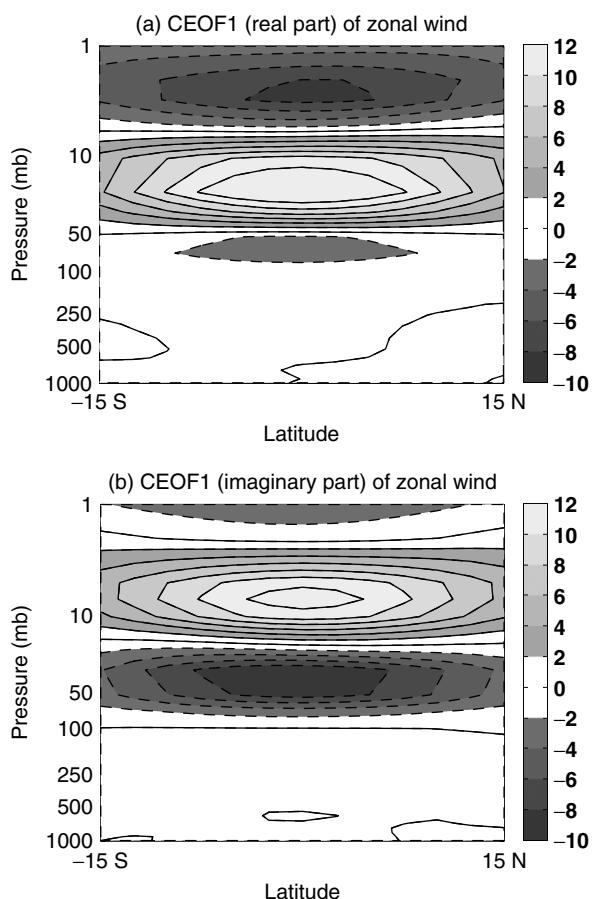


Figure 29. Real (a) and imaginary (b) parts of complex EOF1 of zonal mean zonal wind.

scale of the propagating feature is not well separated from the rest.

Table I shows the percentage of the cumulative variance accounted for by the leading one to ten eigenvalues of the EOFs and CEOFs, respectively. This Table can be used to compare the efficiency of the CEOF method at reducing the dimension of the data to that of the EOF method, particularly for the few leading patterns because of the existence of a propagating structure. One recalls that the time coefficients of EOFs are uncorrelated but not at non-zero lags. For example, the first two conventional PCs will be lagged correlated, hence EOFs would fail at efficiently reducing the data dimension compared to CEOFs when using the same number of EOFs and CEOFs, when there is a propagating feature. Note, however, that the CEOFs are twice as big. So in general, e.g. with no propagating disturbance, CEOFs would be less efficient than EOFs unless they carry at least twice the information on average.

Spatial amplitude and phase of the leading CEOF are computed following Eq (76). Figure 31 shows the spatial amplitude (31a) and the spatial phase (31b) of CEOF1. Figure 31(a) shows a clear indication of the maximum wave activity around 25 mb on the Equator. It also shows the asymmetry in the amplitude gradient, which is stronger in the lower part of the region where the propagation is inhibited. The spatial phase (Figure 31(b)),

on the other hand, tends to be not very easy to interpret (Wallace, 1972; Barnett, 1983). Here, however, because of the unambiguous wave propagation the interpretation is made easy. Figure 31(b) shows banded structures from 1 mb down to around 50 mb, below which there is no propagation.

The banded structure between 1 mb and around 50 mb indicates the direction of propagation of the disturbance where the phase of the wave changes between -180° and $+180^\circ$ in the course of a complete cycle. There are two main regions; the first one is between 1 mb and around 5 mb and corresponds to a first cycle, and the second one is between 5 mb and 25 mb and also corresponds to another cycle. These apparent cycles are associated with small disturbances that can be seen in the raw zonal wind anomalies (Figure 27) and give the impression that the wave grows and decays in the first region then grows and decays again in the second region.

Similarly, the temporal amplitude and phase have also been computed according to Eq (77). Figure 32 shows the temporal amplitude (32a) and temporal phase (32b) of CPC1 for the period January 1992 to December 2001. The temporal amplitude (Figure 32(a)) gives information on how the wave amplitude varies with time, see also the phase portrait (Figure 30(b)). It can be seen, for example from Figure 32(a) that the amplitude is largest in the middle of the wave life cycle. The temporal phase (Figure 32(b)), on the other hand, gives information on the phase of the wave of the zonal wind anomalies as a function of time. For nearly every (quasi-biennial) life cycle the phase is nearly quasi-linear, whose (constant) slope, or time derivative provides a measure of the instantaneous frequency.

The first complex EOF/PC can be used to filter the propagating QBO signal. Figure 33 shows the filtered anomalies using the leading CEOF/CPC for the period January 1992 to December 2001. The downward propagating signal is now clear, with a speed of roughly 1 km/month. The wave amplitude varies slightly with time. One can also note sometimes that the propagating bands can have more than one maximum, leaving the impression that the downward propagating disturbance can grow and decay more than once, a point that has been noted earlier, and which yields the actual structure of the spatial phase (Figure 31(b)).

OTHER EXTENSIONS OF EOFS

We have only discussed a relatively small number of the many techniques related to EOF analysis, but we have concentrated mainly on those that are most useful in atmospheric science. The EOF method has also been extended to deal with cyclostationarity. This is the case when the data contain for example a seasonal cycle. A cyclostationary EOF method was presented by Kim and Huang (1996) and Kim and North (1999). The method is based on EOFs of concatenated vector amplitudes obtained from a FT of the data. Jolliffe (2002) points out

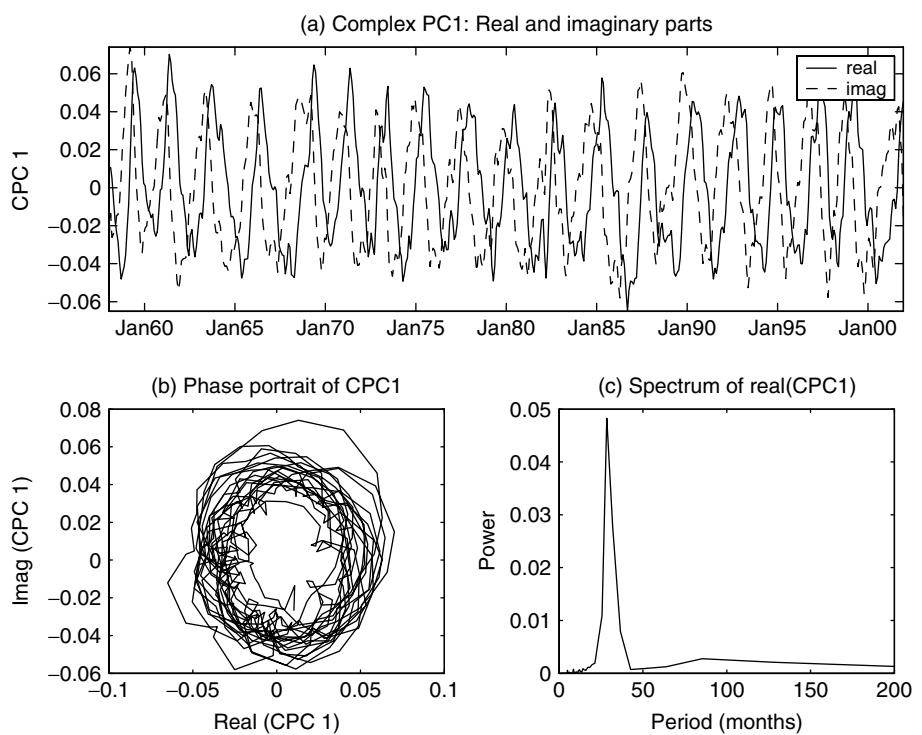


Figure 30. Real and imaginary parts of complex PC1 (a), phase portrait of CPC1 (b), and power spectrum of the real part of CPC1 (c).

Table I. The cumulative explained variance (explained variance of the EOFs and HEOFs of the zonal mean zonal wind anomalies.

Leading eigenvalues	1	2	3	4	5	6	7	8	9	10
EOFs	39.4	76.8	84.5	90.0	92.3	93.9	95.0	96.0	96.6	97.0
HEOFs	71.32	81.3	89.0	91.8	93.7	94.9	95.8	96.6	97.3	97.8

that the method is less transparent in its justification than cyclostationary POP analysis (Blumenthal, 1991; von Storch *et al.*, 1995.) Kim and Wu (1999) describe another EOF method dealing with periodicity in general, which they label periodically extended empirical orthogonal functions (PXEOFs). PXEOFs are the eigenvectors of a large covariance matrix, which is derived by dividing the data into a number of periodic segments and treating them as different variables. The method is akin to extended EOF analysis, but differs from it in that averages in PXEOFs are performed over times at the same point within each block across blocks.

Ordinary EOF analysis can be extended to the case where the data consist, for example, of two or more groups, and more layers such as different time periods. This type of analysis has led to the concept of three-mode (see, e.g. Magnus and Neudecker, 1995) or multiple group PCA. In the atmospheric science context, the ‘three’ in three-mode refers to the indices spanning the data, namely spatial locations, times, and set of atmospheric variables respectively. Note also that in the general three-mode PCA, time need not be ordered nor be equally spaced. Spatial locations and atmospheric variables can be grouped together and ordinary EOF analysis can be applied. A common approach in weather/climate

analysis is to fix the atmospheric variables and perform EOF analysis on the other pair. This yields the so-called S-mode analysis when locations are identified as variables and times as observations, or the T-mode analysis for the reverse, (see Jolliffe 2002 for further discussions and references). Further extension of EOF/PC analysis has been proposed for various other types of data. For example when the data are curves one obtains functional PCA (Ramsey and Silverman, 1997), and involves solutions to an eigenvalue problem of integro-differential type.

The EOF method has also been extended to deal with trends. Recently, Hannachi (2007) has presented an EOF-based method to identify trends in gridded climate data. The method is based on an eigenanalysis of the correlation/covariance matrix of time positions from the sorted data. Trend EOFs (TEOFs) are then identified as the EOFs associated with the leading non-degenerate eigenvalues of the matrix obtained using correlations between time positions of the sorted data. Because the TEOFs are not in the data state space, the data are first projected onto the TEOFs and the obtained time series are then regressed back onto the data to yield the trend patterns. The method, which has been applied to various low-order systems and to reanalyses data, provides a systematic decomposition of the data into

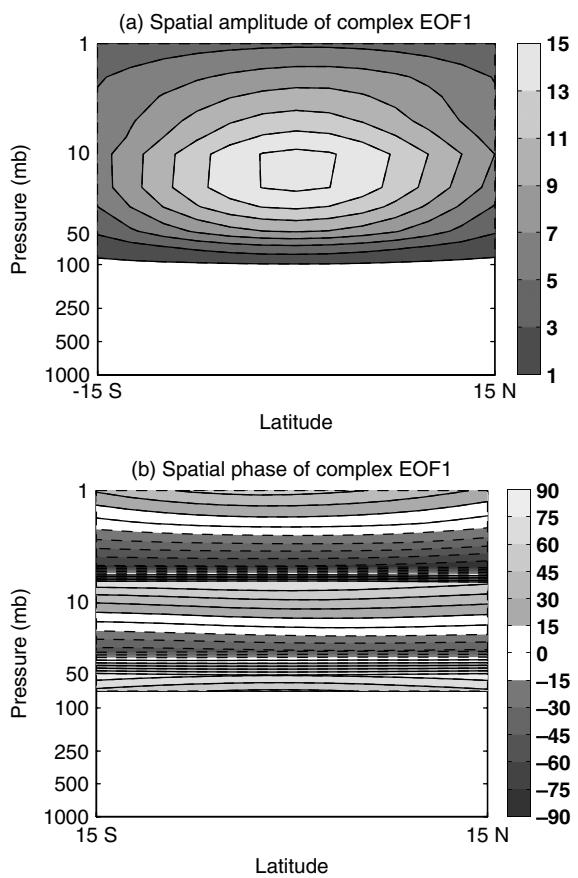


Figure 31. Spatial amplitude (a) and spatial phase (b) of complex EOF1.

a small number of trend patterns and a remaining set with no trend when at least two grid points contain a trend. The application to NCEP/NCAR SLP, for example, clearly yields two trend patterns; the NAO (Hurrell *et al.*, 2003) and the Siberian High (Panagiotopoulos *et al.*, 2005.) The Siberian High is particularly interesting; it was not captured in previous studies using various forms of conventional EOFs because of its local trend character, and also because of its confinement to the planetary

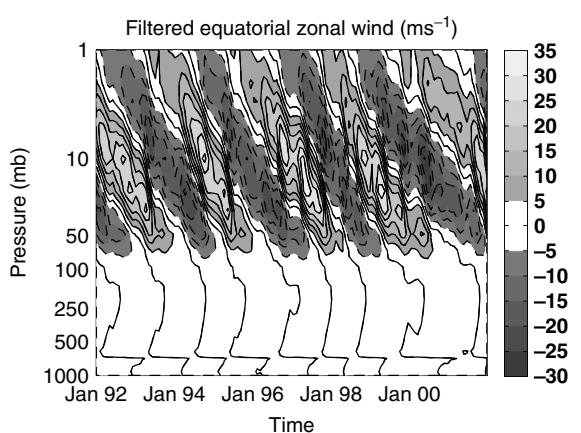


Figure 33. Filtered zonal mean zonal wind anomalies for the period January 1992–December 2001 using the leading complex EOF and the corresponding complex PC.

boundary layer (see Hannachi 2007 for more details).

It is mentioned in section 3 that simplification methods, e.g. rotation, aim to find physically relevant patterns of which teleconnections are good examples. These teleconnections are in general assumed to be the coherent part of the data. This assumption is implicit in the general interpretation of EOFs based on a factor analysis approach (Jolliffe, 2002; Dommenech and Latif, 2002.) Furthermore, it has become clear that in many recent investigations the leading EOF patterns are often interpreted as the leading teleconnections (Thompson and Wallace, 1998, 2000; Saji *et al.*, 1999; Thompson *et al.*, 2000; Wallace and Thompson, 2002.) As pointed out in Jolliffe (2002) and Dommenech and Latif (2002), however, it is unclear how to identify genuine teleconnections. In addition, simplification procedures like rotation may or may not lead to teleconnections. In this context stochastic null hypotheses are perhaps the correct tools for assessing and evaluating climate modes of variability. Cahalan *et al.* (1996) presented a spatial first order autoregressive model, AR(1),

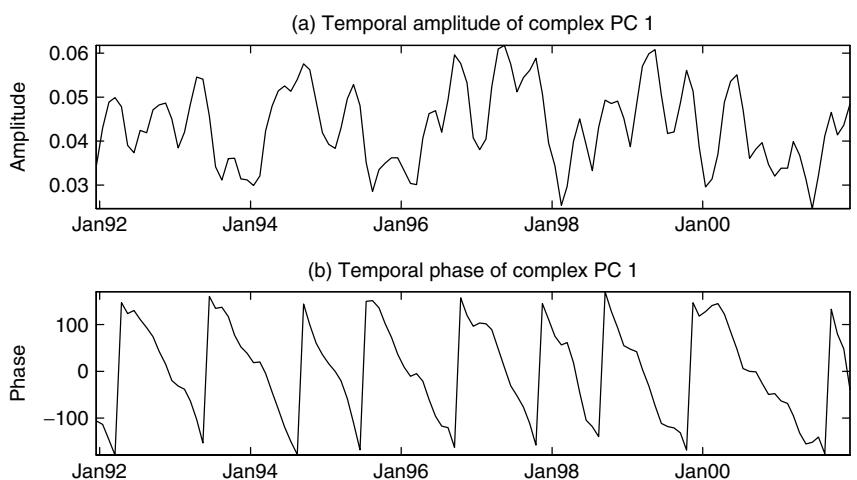


Figure 32. Temporal amplitude (a) and temporal phase (b) of complex CPC1 between January 1992 through to December 2001.

as a null hypothesis. The following diffusion model

$$\frac{d}{dt}\mathbf{u} = -\lambda\mathbf{u} + \nu\nabla^2\mathbf{u} + \mathbf{f}_t \quad (80)$$

has been considered by various authors such as Leung and North (1991) to study the variability of a simplified atmospheric model, and North (1984) to compare EOFs and normal modes. In Eq. (80) λ and ν are damping and diffusion parameters respectively, and \mathbf{f}_t is a spatial and temporal white noise forcing. The isotropic diffusion process (80), which is an extension of the simple AR(1) model, has recently been used by Dommeneget (2007) as a null hypothesis for climate modes of variability. To separate possible teleconnections from a homogeneous diffusive noise background the model is first fitted to the observed data then the leading EOFs of the obtained model compared to the EOFs of the data. Dommeneget (2007) applied the test to various observed fields and showed that the leading Tropical Pacific sea surface temperature (SST) and most NH large scale SLP structure are significantly different from an isotropic diffusion process, whereas the Tropical Indian Ocean SST variability seems to be well described by such a process. Other models could also be envisaged to deal with non-diffusive fields such as SLP, (see e.g. Gerber and Vallis 2005, and Dommeneget (2007) for more details and further references).

The previous EOF/PC methods are linear in the sense that they involve solving an eigenvalue problem. Non-linear extensions of EOFs have also been proposed in the literature. The most well known examples are non-linear PCA, e.g. Monahan (2001), and Hsieh (2001), and independent component analysis, e.g. Hyvärinen and Oja (2000). These methods are computing-power hungry, and the obtained results depend in general on the descent algorithm used to minimise the chosen (non-quadratic) costfunction.

SUMMARY AND CONCLUSION

We presented in this manuscript an introductory review of the state of the art of using EOFs and closely related methods as a means to find prominent patterns of variability, smoothing and reducing the high dimensionality of large scale climate variables, and reconstructing particularly revealing patterns. The review has focussed on five different methods based on EOFs to analyse various climate data. The methods considered here are conventional EOFs, REOFs, simplified EOFs, extended EOFs, and complex/Hilbert EOFs.

We began by reviewing the conventional EOFs method. In particular we have highlighted its benefits, such as easy computation, efficient data reduction, and useful geometric properties. We have also presented its major drawbacks, such as predictable relations between EOFs, and physical interpretability. The method has been illustrated with winter monthly SLP. The leading EOF pattern

shows the familiar and debatable AO pattern. The spectrum of the covariance matrix indicates that this EOF is non-degenerate, but the remaining EOF patterns look degenerate and yields therefore difficulties in interpretation.

Next we have reviewed REOFs, which have been introduced to overcome some of the previous drawbacks related to orthogonality/uncorrelatedness of EOFs/PCs respectively and also interpretation. The method attempts to rotate a fixed number of EOF patterns using either an orthogonal or oblique rotation matrix subject to maximising a simplicity criterion. The EOFs can be either unscaled or scaled by the square root of the associated eigenvalues. Various criteria exist in the literature, but the overall result is that all rotations can be classified into four classes: (i) orthogonal rotation of EOFs, (ii) orthogonal rotation of EOFs scaled by the square root of the associated eigenvalues, (iii) oblique rotation of EOFs, and (iv) oblique rotation of scaled EOFs. In this study we have applied various rotation types/criteria, but the discussion was mainly focussed on three criteria, namely orthogonal VARIMAX, QUARTIMAX, and oblique QUARTIMIN. The results of rotation applied to SLP EOFs indicate that (i) and (iii) give similar results, but the rotated patterns are sensitively dependent on the number m of EOFs selected for rotation. Case (ii) is found to give similar results across a range of criteria, but for large m the similarity between the low ranked REOFs across the various criteria cannot be guaranteed. Finally, case (iv) is found to be the most unstable. Therefore orthogonal rotation of scaled EOFs seems to offer the most robust rotation.

Simplified EOFs, a competitor to the REOF approach, has also been reviewed. The method makes use of some useful properties of EOFs and REOFs simultaneously. It attempts to achieve simultaneously successive variance maximisation, spatial orthogonality of EOFs, and simplicity of REOFs. This is achieved by solving the same eigenvalue problem of EOFs but with an extra constraint of simplicity that depends on a threshold parameter. The obtained optimisation is non-quadratic and involves using advanced numerical methods based on numerical solutions of ODEs. A threshold parameter of the order of $\sqrt{\frac{p}{3}}$, where p is the number of variables, is found to provide a reasonable balance between variance maximisation and simplicity of the patterns. The leading three SEOF patterns are identified respectively as the NAO, the North Pacific, and the Scandinavian patterns. The obtained patterns have invariant structures vis-a-vis changes in the threshold parameter. For example when this parameter gets smaller the patterns keep their structures but their spatial extension gets reduced. The method, however, is computationally intensive but can still be very useful to gain insight if we are particularly interested in the leading few patterns for interpretation.

Extended EOFs (or MSSA) are presented as a way to overcome some of the shortcomings of EOFs, namely the use of only spatial correlation. MSSA makes use of

spatial as well as temporal correlation by extending the familiar state vector by including explicitly the time information after choosing the delay parameter. The method can be used as a tool to filter the data, isolate a trend, or even separate an oscillatory component buried in the noisy data. We have illustrated the approach with 5 years of daily OLR data. We have in particular identified the seasonal cycle and the semi-annual oscillation. The MJO was also identified from the spectrum of the extended data matrix. Various characteristics of MJO have also been identified such as the most active region of growth and decay phase, approximate period, and phase speed.

The complex EOFs method constitutes another approach used to identify propagating disturbances. Unlike EEOFs, complex EOFs use complex formulation of propagating waves and involve the Hilbert transform of the field to form the complexified field with no parameter to fix. Complex HEOFs of ECMWF monthly zonal-mean zonal wind from January 1958 to December 2001 have been calculated. Complex EOFs/PCs have been used to filter the data and reduce their dimension. The QBO propagating signal has been filtered out using the leading complex EOF/PC. Spatial as well as temporal amplitude and phase associated with CEOF1/CPC1 have also been computed and help in interpreting the downward propagating QBO signal.

We have discussed briefly some other extensions of EOFs including cyclostationary, PXEOFs, the S-mode EOF analysis, trend EOFs, and nonlinear extensions of PCA. We have also discussed the methods used to separate possible teleconnections from the mere homogeneous diffusion process background. These extensions are not treated in detail, but we have provided references for interested readers.

Acknowledgements

This work was supported by the NERC Centre for Global Atmospheric Modeling (CGAM) at the Department of Meteorology, the University of Reading. We thank Dr I. Oliveira for her comments on a previous version of the paper, Dr C. Ferro for pointing to us Heinlein's reference, and Dr D. Dommengen for bringing to our attention his work on EOF evaluation. We also thank one reviewer for his/her comments that helped improve the manuscript.

References

- Ambaum MHP, Hoskins BJ, Stephenson DB. 2001. Arctic oscillation or North Atlantic Oscillation? *Journal of Climate* **14**: 3495–3507.
- Ambaum MHP, Hoskins BJ, Stephenson DB. 2002. Corrigendum: arctic oscillation or North Atlantic Oscillation? *Journal of Climate* **15**: 553.
- Anderson TW. 1963. Asymptotic theory for principle component analysis. *Annals of Mathematical Statistics* **34**: 122–148.
- Anderson TW. 1984. *An Introduction to Multivariate Statistical Analysis*, 2nd edn. John Wiley: New York.
- Anderson JR, Rosen RD. 1983. The latitude-height structure of 40–50 day variations in atmospheric angular momentum. *Journal of the Atmospheric Sciences* **40**: 1584–1591.
- Angell JK, Korshover J. 1964. Quasi-biennial variations in temperature, total ozone, and tropopause height. *Journal of the Atmospheric Sciences* **21**: 479–492.
- Angström A. 1935. Teleconnections of climatic changes in present time. *Geografiska Annaler* **17**: 243–258.
- Baldwin MP, Gray LG, Dunkerton TJ, Hamilton K, Haynes PH, Randel WJ, Holton JR, Alexander MJ, Hirota I, Horinouchi T, Jones DBA, Kinnersley JS, Marquardt C, Sao K, Takahashi M. 2001. The quasi-biennial oscillation. *Review in Geophysics* **39**: 179–229.
- Barnett TP. 1983. Interaction of the monsoon and pacific trade wind system at interannual time scales. Part I: The equatorial case. *Monthly Weather Review* **111**: 756–773.
- Barnett TP. 1984a. Interaction of the monsoon and the Pacific trade wind systems at interannual time scales. Part II: The tropical band. *Monthly Weather Review* **112**: 2380–2387.
- Barnett TP. 1984b. Interaction of the monsoon and the Pacific trade wind systems at interannual time scales. Part III: A partial anatomy of the Southern Oscillation. *Monthly Weather Review* **112**: 2388–2400.
- Barnston AG, Livezey BE. 1987. Classification, seasonality, and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review* **115**: 1083–1126.
- Basilevsky A, Hum PJ. 1979. Karhunen-Loève analysis of historical time series with application to Plantation birth in Jamaica. *Journal of the American Statistical Association* **74**: 284–290.
- Bjerknes J. 1969. Atmospheric teleconnections from the equatorial Pacific. *Monthly Weather Review* **97**: 163–172.
- Björnsson H, Venegas SA. 1997. A Manual for EOF and SVD Analyses of Climate Data. Report No 97-1, Department of Atmospheric and Oceanic Sciences and Centre for Climate and Global Change Research, McGill University. 52.
- Bibby J. 1980. Some effects of rounding optimal estimates. *Sankhya B* **42**: 165–178.
- Bloomfield P, Davis JM. 1994. Orthogonal rotation of complex principal components. *International Journal of Climatology* **14**: 759–775.
- Blumenthal MB. 1991. Predictability of a coupled ocean-atmosphere model. *Journal of Climate* **4**: 766–784.
- Bretherton CS, Smith C, Wallace JM. 1992. An intercomparison of methods for finding coupled patterns in climate data. *Journal of Climate* **5**: 541–560.
- Brillinger DR. 1981. *Time Series-Data: Analysis and Theory*. Holden-Day: San Francisco, CA.
- Brink KH, Muench RD. 1986. Circulation in the point conception-Santa Barbara channel region. *Journal of Geophysical Research C* **91**: 877–895.
- Broomhead DS, King GP. 1986a. Extracting qualitative dynamics from experimental data. *Physica D* **20**: 217–236.
- Broomhead DS, King GP. 1986b. On the qualitative analysis of experimental dynamical systems. In *Nonlinear Phenomena and Chaos*, Sarkar S (ed). Adam Hilger: Bristol; 113–144.
- Cahalan RF, Wharton LE, Wu W-L. 1996. Empirical orthogonal functions of monthly precipitation and temperature over the United States and homogeneous stochastic models. *Journal of Geophysical Research* **101**: 26309–26318.
- Carroll JB. 1953. An analytical solution for approximating simple structure in factor analysis. *Psychometrika* **18**: 23–38.
- Craddock JM. 1973. Problems and prospects for eigenvector analysis in meteorology. *The Statistician* **22**: 133–145.
- Chatfield C, Collins AJ. 1989. *Introduction to Multivariate Analysis*. Chapman and Hall: London.
- Chen J-M, Harr PA. 1993. Interpretation of Extended Empirical Orthogonal Function (EEOF) analysis. *Monthly Weather Review* **121**: 2631–2636.
- Cheng X, Nitsche G, Wallace JM. 1995. Robustness of low-frequency circulation patterns derived from EOF and rotated EOF analysis. *Journal of Climate* **8**: 1709–1720.
- Dommengen D. 2007. Evaluating EOF modes against a stochastic null hypothesis. *Climate Dynamics* **28**: 517–531.
- Dommengen D, Latif M. 2002. A cautionary note on the interpretation of EOFs. *Journal of Climate* **15**: 216–225.
- Ebdon RA. 1960. Notes on the wind flow at 50 mb in tropical and subtropical regions in January 1957 and in 1958. *Quarterly Journal of the Royal Meteorological Society* **86**: 540–542.
- Elsner JB, Tsonis AA. 1996. *Singular Spectrum Analysis: A New Tool in Time series Analysis*. Plenum Press: New York.
- Fraedrich K. 1986. Estimating the dimensions of weather and climate attractors. *Journal of the Atmospheric Sciences* **43**: 419–432.
- Fukunaga K, Koontz WLG. 1970. Application of the Karhunen-Loève expansion to feature selection and ordering. *IEEE Transactions on Computers* **C-19**: 311–318.
- Fukuoka A. 1951. *A Study of 10-day Forecast (A Synthetic Report)*, Vol. XXII. The Geophysical Magazine: Tokyo; 177–218.

- Gerber EP, Vallis GK. 2005. A stochastic model for the spatial structure of annular patterns of variability and the North Atlantic oscillation. *Journal of Climate* **18**: 2102–2118.
- Ghil M, Allen MR, Dettinger MD, Ide K, Kondrashov D, Mann ME, Robertson AW, Saunders A, Tian Y, Varadi F, Yiou P. 2002. Advanced spectral methods for climatic time series. *Reviews of Geophysics* **40**: 1.1–1.41.
- Girshick MA. 1939. On the sampling theory of roots of determinantal equations. *Annals of Mathematical Statistics* **43**: 128–136.
- Golub GH, van Loan CF. 1996. *Matrix Computation*. John Hopkins University Press: Baltimore, MD.
- Golyandina NE, Nekrutin VV, Zhitljavsky AA. 2001. *Analysis of Time Series Structure. SSA and Related Techniques*. Chapman and Hall: Boca Raton, FL.
- Graybill FA. 1969. *Introduction to Matrices with Application in Statistics*. Wadsworth: Belmont, CA.
- Graystone P. 1959. Meteorological office discussion-tropical meteorology. *Meteorological Magazine* **88**: 113–119.
- Green BF. 1977. Parameter sensitivity in multivariate methods. *Journal of Multivariate Behavioral Research* **12**: 263–287.
- Hannachi A. 2007. Pattern hunting in climate: a new method for finding trends in gridded climate data. *International Journal of Climatology* **27**: 1–15.
- Hannachi A, Jolliffe IT, Stephenson DB, Trendafilov N. 2006. In search of simple structures in climate: simplifying EOFs. *International Journal of Climatology* **26**: 7–28.
- Hannachi A, O'Neill A. 2001. Atmospheric multiple equilibria and non-Gaussian behaviour in model simulations. *Quarterly Journal of the Royal Meteorological Society* **127**: 939–958.
- Hannan EJ. 1970. *Multiple Time Series*. John Wiley: New York.
- Hardy DM, Walton JJ. 1978. Principal components analysis of vector wind measurements. *Journal of Applied Meteorology* **17**: 1153–1162.
- Harman HH. 1976. *Modern Factor Analysis*, 3rd edn. The University of Chicago Press: Chicago, IL.
- Hausmann R. 1982. Constrained multivariate analysis. In *Optimisation in Statistics*, Zanakis SH, Rustagi JS (eds). North-Holland: Amsterdam; 137–151.
- Heinlein RA. 1973. *Time Enough for Love*. New English Library: London.
- Hendon HH, Salby ML. 1994. The life cycle of the madden-Julian oscillation. *Journal of the Atmospheric Sciences* **51**: 2225–2237.
- Hirsch MW, Smale S. 1974. *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press: London.
- Holton JR. 1992. *An Introduction to Dynamic Meteorology*, 3rd edn. Academic Press: London.
- Horel JD. 1981. A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 mb height field. *Monthly Weather Review* **109**: 2080–2092.
- Horel JD. 1984. Complex principal component analysis: theory and examples. *Journal of Climate and Applied Meteorology* **23**: 1660–1673.
- Hotelling H. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* **24**: 417–520.
- Hotelling H. 1935. The most predictable criterion. *Journal of Educational Psychology* **26**: 139–142.
- Hsieh WW. 2001. Nonlinear principal component analysis by neural networks. *Tellus* **53A**: 599–615.
- Hurrell JW. 1996. Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. *Geophysical Research Letter* **23**: 665–668.
- Hurrell JW, Kushnir Y, Ottersen G, Visbeck M. 2003. An overview of the North Atlantic Oscillation. In *The North Atlantic Oscillation, Climate Significance and Environmental Impact*, Geophysical Monograph **134**, Hurrell JW, Kushnir Y, Ottersen G, Visbeck M (eds). American Geophysical Union: Washington, DC; 1–35.
- Hyyvärinen A, Oja E. 2000. Independent component analysis. Algorithms and applications. *Neural Networks* **13**: 4–5.
- Jackson JE. 1991. *A User's Guide to Principal Components*. Wiley: New York.
- Jenkins JM, Watts DG. 1968. *Spectral Analysis and its Applications*. Holden-Day: San Francisco, CA.
- Jennrich RI. 2001. A simple general procedure for orthogonal rotation. *Psychometrika* **66**: 289–306.
- Jennrich RI. 2002. A simple general procedure for oblique rotation. *Psychometrika* **67**: 7–19.
- Johnson ES, McPhaden MJ. 1993. Structure of intraseasonal Kelvin waves in the equatorial Pacific Ocean. *Journal of Physical Oceanography* **23**: 608–625.
- Jolliffe IT. 1987. Rotation of principal components: some comments. *Journal of Climatology* **7**: 507–510.
- Jolliffe IT. 1995. Rotation of principal components: choice of normalization constraints. *Journal of Applied Statistics* **22**: 29–35.
- Jolliffe IT. 2002. *Principal Component Analysis*, 2nd edn. Springer: New York.
- Jolliffe IT, Uddin M, Vines SK. 2002. Simplified EOFs-three alternatives to retain. *Climate Research* **20**: 271–279.
- Jolliffe IT, Trendafilov NT, Uddin M. 2003. A Modified principal component technique based on the LASSO. *Journal of Computational and Graphical Statistics* **12**: 531–547.
- Kaiser HF. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* **23**: 187–200.
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetma A, Reynolds B, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437–471.
- Kaplan D, Glass L. 1995. *Understanding Nonlinear Dynamics*. Springer-Verlag: New York.
- Kiers HAL. 1994. Simplimax: oblique rotation to an optimal target with simple structure. *Psychometrika* **59**: 567–579.
- Kiladis GN, Weickmann KM. 1992. Circulation anomalies associated with tropical convection during northern winter. *Monthly Weather Review* **120**: 1900–1923.
- Kim K-W, Huang J. 1996. EOFs of one-dimensional cyclostationary time series: computations, examples, and stochastic modelling. *Journal of the Atmospheric Sciences* **53**: 1007–1017.
- Kim K-W, Wu Q. 1999. A comparison study of EOF techniques: analysis of nonstationary data with periodic statistics. *Journal of Climate* **12**: 185–199.
- Kim K-W, North GR. 1999. EOF-based linear prediction algorithm: examples. *Journal of Climate* **12**: 2076–2092.
- Kimoto M, Ghil M, Mo KC. 1991. Spatial structure of the extratropical 40-day oscillation. *Proceedings of the 8th Conference on Atmospheric and Oceanic Waves and Stability*, American Meteorological Society: Boston, MA; 115–116.
- Kistler R, Kalnay E, Collins W, Saha S, White G, Woollen J, Chelliah J, Ebisuzaki W, Kanamitsu M, Koukya V, van den Dool H, Jenne R, Fiorino M. 2001. The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society* **82**: 247–267.
- Knutson TR, Weickmann KM. 1987. 30–60 day atmospheric oscillation: composite life cycles of convection and circulation anomalies. *Monthly Weather Review* **115**: 1407–1436.
- Kress R, Martensen E. 1970. Anwendung der rechteckregel auf die reelle Hilbertransformation mit unendlichem intervall. *Zeitschrift für Angewandte Mathematik und Mechanik* **50**: 61–64.
- Krishnamurthi TN, Chakraborty DR, Cubukcu N, Stefanova L, Vijaya Kumar TSV. 2003. A mechanism of the madden-Julian oscillation based on interactions in the frequency domain. *Quarterly Journal of the Royal Meteorological Society* **129**: 2559–2590.
- Krzanowski WJ. 2000. *Principles of Multivariate Analysis: A User's Perspective*, 2nd edn. Oxford University Press: Oxford.
- Kundu PK, Allen JS. 1976. Some three-dimensional characteristics of low-frequency current fluctuations near the Oregon coast. *Journal of Physical Oceanography* **6**: 181–199.
- Kutzbach JE. 1967. Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. *Journal of Applied Meteorology* **6**: 791–802.
- Labitzke K, van Loon H. 1999. *The Stratosphere*. Springer-Verlag: New York.
- Lawley DN. 1956. Tests of significance for the latent roots of covariance and correlation matrices. *Biometrika* **43**: 128–136.
- Leung L-Y, North GR. 1991. Atmospheric variability on a zonally symmetric land planet. *Journal of Climate* **4**: 753–765.
- Loève M. 1978. *Probability Theory*, Vol. 2, 4th edn. Springer Verlag: New York.
- Lorenz EN. 1956. Empirical Orthogonal Functions and Statistical Weather Prediction. Technical report, Statistical Forecast Project Report 1, Dep of Meteor, MIT: 49.
- Lorenz EN. 1970. Climate change as a mathematical problem. *Journal of Applied Meteorology* **9**: 325–329.

- Madden RA, Julian PR. 1971. Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *Journal of the Atmospheric Sciences* **28**: 702–708.
- Madden RA, Julian PR. 1972. Description of global-scale circulation cells in the tropics with a 40–50 day period. *Journal of the Atmospheric Sciences* **29**: 1109–1123.
- Madden RA, Julian PR. 1994. Observations of the 40–50-day tropical oscillation—a review. *Monthly Weather Review* **122**: 814–837.
- Magnus JR, Neudecker H. 1995. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley and Sons: Chichester.
- Mardia KV, Kent JT, Bibby JM. 1979. *Multivariate Analysis*. Academic Press: London.
- Maruyama T. 1997. The quasi-biennial oscillation (QBO) and equatorial waves—a historical review. *Papers in Meteorology and Geophysics* **47**: 1–17.
- Matthews AJ. 2000. Propagation mechanisms for the Madden-Julian oscillation. *Quarterly Journal of the Royal Meteorological Society* **126**: 2637–2651.
- Merrifield MA, Winant CD. 1989. Shelf circulation in the gulf of California: a description of the variability. *Journal of Geophysical Research* **94**: 18133–18160.
- Merrifield MA, Guza RT. 1990. Detecting propagating signals with complex empirical orthogonal functions: a cautionary note. *Journal of Physical Oceanography* **20**: 1628–1633.
- Mestas-Núñez AM. 2000. Orthogonality properties of rotated empirical modes. *International Journal of Climatology* **20**: 1509–1516.
- Monahan AH. 2001. Nonlinear principal component analysis: tropical Indo-Pacific sea surface temperature and sea level pressure. *Journal of Climate* **14**: 219–233.
- Monahan AH, Tangang FT, Hsieh WW. 1999. A potential problem with extended EOF analysis of standing wave fields. *Atmosphere-Ocean* **3**: 241–254.
- Morrison DF. 1976. *Multivariate Statistical Methods*, 2nd edn. McGraw-Hill: New York.
- North GR. 1984. Empirical orthogonal functions and normal modes. *Journal of the Atmospheric Sciences* **41**: 879–887.
- North GR, Bell TL, Cahalan RF, Moeng FJ. 1982. Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review* **110**: 699–706.
- Obukhov AM. 1947. Statistically homogeneous fields on a sphere. *Uspeshi Matematicheskikh Nauk* **2**: 196–198.
- Obukhov AM. 1960. The statistically orthogonal expansion of empirical functions. *Bulletin of the Academy of Sciences of the USSR. Geophysics Series (English Transl.)* **1**: 288–291.
- Overland JE, Preisendorfer RW. 1982. A significance test for principal components applied to a cyclone climatology. *Monthly Weather Review* **110**: 1–4.
- Palmer CE. 1954. The general circulation between 200 mb and 10 mb over the equatorial Pacific. *Weather* **9**: 3541–3549.
- Panagiotopoulos F, Shahgedanova M, Hannachi A, Stephenson DB. 2005. Observed trends and teleconnections of the Siberian high: a recently declining center of action. *Journal of Climate* **18**: 1411–1422.
- Pavan V, Tibaldi S, Brankovich C. 2000. Seasonal prediction of blocking frequency: results from winter ensemble experiments. *Quarterly Journal of the Royal Meteorological Society* **126**: 2125–2142.
- Pearson K. 1902. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* **2**: 559–572.
- Peng S, Fife G. 1996. The coupled patterns between sea level pressure and sea surface temperature in the mid-latitude North Atlantic. *Journal of Climate* **9**: 1824–1839.
- Plaut G, Vautard R. 1994. Spells of low-frequency oscillations and weather regimes in the northern hemisphere. *Journal of the Atmospheric Sciences* **51**: 210–236.
- Preisendorfer RW. 1988. *Principal Component Analysis in Meteorology and Oceanography*. Elsevier: Amsterdam.
- Priestley MB. 1981. *Spectral Analysis of Time Series*. Academic Press: London.
- Ramsey JD, Silverman BW. 1997. *Functional Data Analysis*. Springer Verlag: New York.
- Rasmusson EM, Arkin PA, Chen W-Y, Jalickee JB. 1981. Biennial variations in surface temperature over the United States as revealed by singular decomposition. *Monthly Weather Review* **109**: 587–598.
- Reed RJ, Campbell WJ, Rasmusson LA, Rogers RG. 1961. Evidence of a downward propagating annual wind reversal in the equatorial stratosphere. *Journal of Geophysical Research* **66**: 813–818.
- Reymert RA, Jöreskog KG. 1996. *Applied Factor Analysis in the Natural Sciences*. Cambridge University Press: Cambridge.
- Richman MB. 1981. Obliquely rotated principal components: an improved meteorological map typing technique. *Journal of Applied Meteorology* **20**: 1145–1159.
- Richman MB. 1986. Rotation of principal components. *Journal of Climatology* **6**: 293–335.
- Richman MB. 1987. Rotation of principal components: a reply. *Journal of Climatology* **7**: 511–520.
- Saji NH, Goswami BN, Vinayachandran PN, Yamagata T. 1999. A dipole mode in the tropical Indian Ocean. *Nature* **401**: 360–363.
- Salstein DA, Rosen RD, Peixoto JP. 1983. Modes of variability in annual hemispheric water vapor and transport fields. *Journal of the Atmospheric Sciences* **40**: 788–803.
- Seal HL. 1967. *Multivariate Statistical Analysis for Biologists*. Methuen: London.
- Simmons AJ, Wallace JM, Branstator GW. 1983. Barotropic wave propagation and instability, and atmospheric teleconnection patterns. *Journal of the Atmospheric Sciences* **40**: 1363–1392.
- Sun L. 2005. Simple Principal Components. PhD thesis. Department of Statistics, Faculty of Mathematics and Computing, The Open University, Milton Keynes.
- Takens F. 1981. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence. Lecture Notes in Mathematics* 898, Rand D, Young LS (eds). Springer-Verlag: New York; 366–381.
- Thacker WC. 1996. Metric-based principal components. *Tellus* **46A**: 584–592.
- Thiébaux HJ, Zwiers FW. 1984. The interpretation and estimation of effective sample sizes. *Journal of Climate and Applied Meteorology* **23**: 800–811.
- Thomas JB. 1969. *An Introduction to Statistical Communication Theory*. Wiley: New York.
- Thompson DWJ, Wallace JM. 1998. The Arctic oscillation signature in wintertime geopotential height and temperature fields. *Geophysical Research Letters* **25**: 1297–1300.
- Thompson DWJ, Wallace JM. 2000. Annular modes in the extratropical circulation. Part I: Month-to-month variability. *Journal of Climate* **13**: 1000–1016.
- Thompson DWJ, Wallace JM, Hegerl GC. 2000. Annular modes in the extratropical circulation, Part II: Trends. *Journal of Climate* **13**: 1018–1036.
- Tibshirani R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society B* **58**: 267–288.
- Trenberth KE. 1984. Some effects of finite sample size and persistence in meteorological statistics. Part I: Autocorrelations. *Monthly Weather Review* **112**: 2359–2368.
- Trenberth KE, Shin W-TK. 1984. Quasi-biennial fluctuations in sea level pressures over the Northern Hemisphere. *Monthly Weather Review* **111**: 761–777.
- Trendafilov NT, Jolliffe IT. 2005. Numerical solution of the SCoTCLASS. *Computational Statistics and Data Analysis* **50**: 242–253.
- Van den Dool HM, Saha S, Johansson A. 2000. Empirical orthogonal teleconnections. *Journal of Climate* **13**: 1421–1435.
- Vautard R, Yiou P, Ghil M. 1992. Singular spectrum analysis: a toolkit for short, noisy chaotic signals. *Physica D* **58**: 95–126.
- Vines SK. 2000. Simple principal components. *Applied Statistics* **49**: 441–451.
- von Storch H. 1995. Spatial Patterns: EOFs and CCA. In *Analysis of Climate Variability: Application of Statistical Techniques*, von Storch H, Navarra A (eds). Springer Verlag: Berlin; 227–257.
- von Storch H, Zwiers FW. 1999. *Statistical Analysis in Climate Research*. Cambridge University Press: Cambridge.
- von Storch H, Bürger G, Schnur R, Storch J-S. 1995. Principal oscillation patterns. A review. *Journal of Climate* **8**: 377–400.
- Wallace JM. 1972. Empirical orthogonal representation of time series in the frequency domain. Part II: Application to the study of tropical wave disturbances. *Journal of Applied Meteorology* **11**: 893–900.
- Wallace JM, Dickinson RE. 1972. Empirical orthogonal representation of time series in the frequency domain. Part I: Theoretical consideration. *Journal of Applied Meteorology* **11**: 887–892.
- Wallace JM, Gutzler DS. 1981. Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Monthly Weather Review* **109**: 784–812.
- Wallace JM, Thompson DWJ. 2002. The Pacific Center of Action of the Northern Hemisphere annular mode: real or artifact? *Journal of Climate* **15**: 1987–1991.

- Walsh JE, Richman MB. 1981. Seasonality in the associations between surface temperatures over the United States and the North Pacific Ocean. *Monthly Weather Review* **109**: 767–783.
- Weare BC, Nassstrom JS. 1982. Examples of extended empirical orthogonal function analysis. *Monthly Weather Review* **110**: 481–485.
- Webster PJ, Chang H-R. 1988. Equatorial energy accumulation and emanation regions: Impacts of a zonally varying basic state. *Journal of the Atmospheric Sciences* **45**: 803–829.
- Weideman JAC. 1995. Computing the Hilbert transform on the real line. *Mathematics of Computation* **64**: 745–762.
- Whittle P. 1951. *Hypothesis Testing in Time Series*. Almqvist and Wiksell: Upsala.
- Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences*, 2nd edn. Academic Press: Amsterdam.
- Xinhua C, Dunkerton TJ. 1995. Orthogonal rotation of spatial patterns derived from singular value decomposition analysis. *Journal of Climate* **8**: 2631–2643.