

In search of simple structures in climate: Simplifying EOFs

by

A. Hannachi^{1*}, I. T. Jolliffe¹, D. B. Stephenson¹, and N. Trendafilov²

¹Department of Meteorology, University of Reading
Reading RG6 6BB, U.K.

²Department of Probability and Statistics,
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences,
8, Acad. G. Bontchev Str., Sofia 1113, BULGARIA.

Revised version submitted to the Int. J. of Climatol.

June 15, 2005

*Corresponding author: A. Hannachi, Department of Meteorology, the University of Reading, Earley Gate PO Box 243, Reading RG6 6BB, U.K.
E-mail: A.Hannachi@reading.ac.uk
Tel: +44 (0) 118 378 7218
Fax: +44 (0) 118 378 8316

Abstract

Empirical orthogonal functions (EOFs) are widely used in climate research to identify dominant patterns of variability and to reduce the dimensionality of climate data. EOFs, however, can be difficult to interpret. Rotated EOFs have been proposed as more physical entities with simpler patterns than EOFs. This study presents a new approach for finding climate patterns with simple structures that overcomes the problems encountered with rotation. The method achieves simplicity of the patterns by using the main properties of EOFs and Rotated EOFs simultaneously. Orthogonal patterns are found that maximise variance subject to a constraint that induces a form of simplicity. The simplified EOF patterns, being more "local", are constrained to have zero loadings outside the main centre of action. The method is applied to winter Northern Hemisphere monthly mean sea level pressure reanalyses over the period 1948-2000. The "simplified" leading patterns of variability are identified and compared to the leading patterns obtained from EOFs and REOFs.

KEYWORDS: Principal components Simplified empirical orthogonal functions Ordinary differential equations Gradient methods North Atlantic Oscillation Arctic Oscillation

1 Introduction

Atmospheric scientists have introduced and developed various ways to find patterns of variability in high dimensional weather/climate system. The main objectives of these methods are twofold: to reduce the dimensionality of the system by retaining a much smaller set of dominant patterns (e.g. Hannachi and O'Neill 2001), and to obtain a few, and in some cases just one, leading pattern of variability that is physically relevant (e.g. Wallace and Thompson 2002). Since their introduction in atmospheric research (Obukhov 1947, 1960; Fukuoka 1951; Lorenz 1956), empirical orthogonal functions (EOFs) have been widely used to analyse atmospheric data (Kutzbach 1967; Preisendorfer 1988). A major reason for their widespread use in atmospheric science is that they allow a space display and a time display, which would seem to be relevant to climate researchers. The existence of fast and efficient algorithms that can handle large space-time data sets also helped their widespread use.

EOFs coefficients, or principal components (PCs), of a gridded spatio-temporal

field are linear combinations of the different variables of the field that maximise variance. The EOFs are the loadings of the PCs, but can also be defined without defining PCs first. They can also be defined as linear combinations of the different maps of the field that maximise norm. They have been applied to various atmospheric fields, e.g. surface temperature, sea level pressure, sea surface temperature, for various purposes such as analysing the leading modes/patterns of variability (e.g. Kutzbach 1967), predicting the weather (Lorenz 1956; Ward and Folland 1991), and comparing model simulations to observations and reanalyses as routinely practised in most climate research centres. For a definition and discussion on the differences between regimes, modes, patterns, etc. see Stephenson et al. (2004).

By construction, EOFs of climate data yield sets of orthogonal spatial patterns and sets of uncorrelated time series, see e.g. Preisendorfer (1988); Wilks (1995); von Storch and Zwiers (1999); and Jolliffe (2002) and references therein. These are geometric properties that can be very useful in modelling studies using PCs. For example, the covariance matrix of any subset of retained PCs is always diagonal. These constraints, however, yield partially predictable relationships between an EOF and the previous ones. For instance, as pointed out by Horel (1981), 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. The orthogonality constraint also makes the EOFs domain-dependent and can be too non-local (Horel 1981; Richman 1986). Consequently, these constraints can cause limitations to any possible physical interpretation of the obtained patterns (Ambaum et al. 2001, 2002; Dommenech and Latif 2002; Jolliffe 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.

To overcome some of the drawbacks caused by the geometrical constraints, researchers have looked for an alternative through linear transformation of the EOFs. The concept of rotation emerged in factor analysis and has been proposed since the late 1940s in social sciences. In atmospheric sciences, rotated EOFs (REOFs) have been applied nearly three decades later and continue to be widely used (Horel 1981; Richman 1981, 1986; Preisendorfer 1988; Cheng et al. 1995). The review of Richman (1986) provides a particularly detailed discussion of the characteristics of unrotated EOFs. REOFs yield simpler structures, compared to EOFs, by rotating the vector of loadings or EOFs hence losing some of the useful geometric properties of EOFs, in favour of yielding better interpretation. REOFs, however, have their own shortcomings summarised in how to (1) decide the rotation criteria that specify the simplicity,

(2) choose the normalisation, (3) the type of projection, and also (4) the number of EOFs to be rotated, though this latter can be regarded by some as a more general issue. Of course choosing the number of EOFs, for example, can be difficult, but adding an extra one does not change those found already, unlike REOFs.

Jolliffe et al. (2002) developed an alternative way to construct simple structure patterns without compromising the useful properties of EOFs, namely variance maximisation and EOF orthogonality. Although giving a best low-dimensional representation in a least square sense, variance maximisation does not guarantee physical interpretability. However, there is no way of knowing this in advance of an analysis, unless the form of the physical modes is already known, in which case *any* analysis is pointless.

Our main objective in this manuscript is to demonstrate the usefulness of this method for a moderately large climate application. Another motivation of using this approach is the fact that propagating planetary waves (Hoskins and Karoly 1981) tend to follow wave guides (Hoskins and Ambrizzi 1993; Ambrizzi and Hoskins 1995) because of the presence of critical lines (Held 1983; Nigam and Held 1983; Killworth and McIntyre 1985). One would therefore expect physically-relevant patterns to be more local or simple, i.e. with zeros outside the main centres of action.

The method, which is computationally intensive, helps interpret the nature of the leading modes of variability of sea level pressure and also contributes to the current AO/NAO debate. The manuscript is organised as follows. In section 2 we briefly review EOFs and rotated EOFs and then introduce the simplified EOFs in section 3. The application to winter monthly means of sea level pressure is presented in section 4. Conclusions are presented in the final section.

2 EOFs and REOFs: Background

2.1 EOFs: Formulation and algorithms

Empirical orthogonal function analysis, also known as principal component analysis (PCA), see, e.g. Preisendorfer (1988) and Jolliffe (2002), has been widely used since its introduction in atmospheric science by Obukhov (1947, 1960), Fukuoka (1951), and Lorenz (1956). Here we consider a finite set of a p -dimensional time series $\mathbf{x}_t = (x_{t1}, x_{t2}, \dots, x_{tp})^T$, $t = 1, \dots, n$, where the superscript T denotes the transpose operator. Each vector \mathbf{x}_t , $t = 1, \dots, n$, is a spatial field consisting of p grid point values. This multivariate time series may represent for example the time evolution of the geopotential height or temperature at p different stations on the Earth surface

recorded at regular time interval over a finite period.

A sequence of spatial fields can be represented by the following $n \times p$ data matrix, see e.g. Mardia et al. (1979):

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

where the rows constitute the time samples and the columns represent the grid point variables. For example, the k 'th column constitutes the time series at the k 'th variable (or grid point), and the j 'th row constitutes a map, i.e. the sample values at time j . EOFs normally describe the variability of the spatio-temporal field with respect to a base state taken in general to be the climate mean (or climatology). The $n \times p$ anomaly data matrix X' is defined as the departure of X from the climatology or long-term mean

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{t=1}^n \mathbf{x}_t = \frac{1}{n} X^T \mathbf{1}_n$$

and is given by

$$X' = \left(I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) X \quad (2)$$

where $\mathbf{1}_n = (1, 1, \dots, 1)^T$ is the column vector of length n containing only ones and I_n is the $n \times n$ identity matrix. To keep the notation simple the dash in the matrix (2) will be removed and the notation X will simply refer to the anomaly data matrix for the rest of the manuscript. The sample covariance matrix is given by:

$$S = \frac{1}{n} X^T X. \quad (3)$$

The EOFs $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p\}$ are defined as directions, (unit vectors \mathbf{a}) in the p -dimensional state space. The principal components, i.e. the time series associated with these EOFs, are uncorrelated linear combinations $X\mathbf{a}$, of the variables that successively have maximum variance. The EOFs are therefore obtained as the solution to the quadratic optimisation problem:

$$\begin{aligned} \max F(\mathbf{a}) &= \mathbf{a}^T S \mathbf{a} \\ \text{subject to } \mathbf{a}^T \mathbf{a} &= 1, \end{aligned} \quad (4)$$

and subject to each EOF being orthogonal to previous EOFs. This is equivalent to maximising $(\mathbf{a}^T S \mathbf{a}) / (\mathbf{a}^T \mathbf{a})$. The straightforward solution to this quadratic optimisation problem is given by solving the eigenvalue problem

$$S\mathbf{a} = \lambda^2 \mathbf{a}.$$

The EOFs are the eigenvectors of the sample covariance matrix S arranged in decreasing order of the eigenvalues. The first eigenvector \mathbf{a}_1 gives the first principal component, i.e. the linear function $X\mathbf{a}_1$, with the largest variance; the second EOF \mathbf{a}_2 gives the second principal component with the next largest variance subject to being orthogonal to \mathbf{a}_1 , etc.

In a similar manner to the formulation (4), the EOFs can also be defined as linear combinations $\mathbf{v}X$, where \mathbf{v} is a vector of weights, of the different maps of the field that maximise the norm squared. Applying this definition one obtains a similar equation to (4), namely:

$$\max (\mathbf{v}^T P \mathbf{v}) / (\mathbf{v}^T \mathbf{v}) \quad (5)$$

where $P = XX^T$ is the matrix of scalar product between the different maps. Eq. (5) yields automatically the (standardised) principal components (instead of the EOFs as in Eq. (4)). Note that Eq. (5) is formulated using a duality argument to Eq. (4), and can be useful for numerical purposes when for example the sample size is smaller than the number of variables. The EOFs $A = (\mathbf{a}_1, \dots, \mathbf{a}_r)$, where r is the rank of X , are the eigenvectors of (3), or equally the right singular vectors of X derived from the singular value decomposition (SVD) (see e.g. Golub and van Loan (1996)) of X , i.e.

$$X = V\Lambda A^T \quad (6)$$

with $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_r)$ containing the singular values $\lambda_1 \geq \lambda_2 \dots \geq \lambda_r \geq 0$ of X .

Because it is quadratic the optimisation problem (4) can be satisfactorily addressed using efficient algorithms such as SVD. However, as will become clear later, it is important to put problem (4) in the more general framework of nonlinear optimisation and make use of other attractive and efficient algorithms. A number of algorithms exist to solve unconstrained as well as constrained minimisation problems. These include non-gradient methods such as direct and Newton-Raphson methods, and all gradient methods including steepest descent and projected/reduced gradient methods (Gill et al. 1981). Most of these algorithms share a similar property, namely the search for and movement along directions of descent. In various problems, however, the search for suitable step sizes (in line search) can be problematic particularly when the cost-function to be minimised is not quadratic, for which the algorithm can converge to the wrong local minima.

Other algorithms, such as simulated annealing, can be used to overcome problems related to convergence to (wrong) local minima. Simulated annealing (Kirkpatrick et al. 1983; Metropolis et al. 1993; Lin and Kernighan 1973) makes use of the fact that when a crystal is heated then slowly cooled, thermal mobility of the atoms is gradually lost and the crystal gets to the state of its minimum absolute energy. The

algorithm avoids local minima through a process of random shocks, see e.g. Hannachi and Legras (1995) for an atmospheric application. The algorithm is particularly useful when the feasible set is connected, but can easily fail when it is not connected.

Instead of straight line search directions, an alternative is to proceed from an initial condition to the solution by following a rather smooth curvilinear trajectory. To find the minima of $F(\mathbf{x})$, then a simple trajectory is provided by the continuous steepest descent trajectory of the ordinary differential equation (ODE)

$$\frac{d\mathbf{x}}{dt} = -\nabla F(\mathbf{x}) \quad (7)$$

with suitably chosen initial condition (Brown 1986). In fact, if \mathbf{x}^* is an isolated local minimum of $F(\mathbf{x})$, then \mathbf{x}^* is a stable fixed point of the dynamical system (9), see e.g. Hirsch and Smale (1974), and hence can be reached by integrating (9) from some suitable initial condition. Such methods have been around since the mid 1970 (Evtushenko 1974; Botsaris 1976, 1978) and can make use of efficient integration algorithms available for dynamical systems. Trajectories defined by second order differential equations have also been suggested (Snyman 1982).

When the optimisation is constrained, the solution also can be approached in a similar way using the projected or reduced gradient method. Here the continuous path can be obtained through projection of the gradient $\nabla F(\mathbf{x})$ onto the tangent space of the feasible set, i.e. the set or surface satisfying the constraints (Botsaris 1979, 1981; Evtushenko 1974; Evtushenko and Zhdan 1977). For a review of these methods as well as further references see Brown (1986).

For the EOF problem the successive variance maximisation and orthonormality conditions yield the following projected gradient formulation. For a given $k = 1, \dots, p$, designate by

$$\pi_k = I_p - \sum_{l=1}^{k-1} \mathbf{a}_l \mathbf{a}_l^T \quad (8)$$

the projection operator onto the orthogonal complement to the space spanned by the first $k-1$ EOFs ($\mathbf{a}_1, \dots, \mathbf{a}_{k-1}$), then EOF \mathbf{a}_k is obtained as the limit, when $t \rightarrow \infty$, of the solution to the following system of ODEs:

$$\frac{d}{dt} \mathbf{a}_k = \pi_k (I_p - \mathbf{a}_k \mathbf{a}_k^T) \nabla F(\mathbf{a}_k) = \pi_{k+1} \nabla F(\mathbf{a}_k). \quad (9)$$

It can be seen that the eigenvalues λ_k^2 , $k = 1, \dots, r$, of the covariance matrix S satisfy (see, e.g. Magnus and Neudecker 1995)

$$\lambda_k^2 = \max_{P_{k-1}^T \mathbf{x} = 0} \frac{\mathbf{x}^T S \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \quad (10)$$

where $P_k = (\mathbf{a}_1, \dots, \mathbf{a}_k)$ is the matrix consisting of the set of the k leading eigenvectors of S . Note that the projection operator π_{k+1} simply derives from orthogonality between \mathbf{a}_k and \mathbf{a}_l , $l < k$, plus the fact that the condition $\mathbf{a}^T \mathbf{a} = 1$ is equivalent to

$$\pi \mathbf{a} = (I_p - \mathbf{a}\mathbf{a}^T) \mathbf{a} = \mathbf{0}.$$

So π_k is precisely the projection onto the linear space of the feasible set $\{\mathbf{x}; P_{k-1}^T \mathbf{x} = 0\}$. The projection operator is in fact used to ensure that the gradient yields a feasible point at each step during the optimisation.

So as can be seen from (6), by construction the EOFs are orthogonal and the PCs uncorrelated. These useful geometric constraints can be, however, problematic when it comes to physical interpretation. To overcome some of these setbacks a number of methods have been proposed and are currently used in atmospheric science to aid interpretation. The most common method is rotation, which is discussed below.

2.2 Rotated EOFs

Horel (1981) and Richman (1981, 1986) argued that EOFs can be too non-local and dependent on the size and the shape of the spatial domain. As pointed out by Horel (1981) and others who applied rotated factors, invariance or constancy of a solution, e.g. factors or EOFs, when the domain changes is a fundamental necessity if the solution is to be physically meaningful. Also in his detailed review on rotated EOFs, Richman (1986) maintains that unrotated EOFs exhibit four characteristics that hamper their utility to isolate individual modes of variation. These are: domain dependence, subdomain instability, sampling problems, and inaccurate portrayal of the physical relationship embedded within the data.

The principal objective of rotated EOFs (REOFs) is to obtain patterns with simple structure, and thereby overcome some of the drawbacks of EOFs mentioned above. To aid interpretation the most common definition of simplicity is to drive the EOF coefficients (PC loadings) to have either small or large magnitude with few or no intermediate values. Rotation of EOFs attempts precisely to achieve this. Rotation acts on a pre-determined subset of EOFs (or PCs). Suppose that we decide to keep the leading m EOFs; $A_m = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$. Rotation of the $(p \times m)$ matrix A_m is normally achieved by finding a rotation matrix R such that the rotated loadings

$$B = A_m R \tag{11}$$

optimise a pre-determined "simplicity" criterion. The scaled loadings $A_m \Lambda_m$, where $\Lambda = (\lambda_1, \dots, \lambda_m)$ represents is the diagonal matrix of the leading m singular values

of X , can also be rotated. Depending on the orthogonality/non-orthogonality of the rotation matrix R , one obtains an orthogonal/oblique rotation respectively. Various simplicity criteria exist in the literature, such as the most widely used VARIMAX orthogonal rotation (Kaiser 1958; Richman 1986; Krzanowski and Marriott 1995; Jolliffe 2002). The VARIMAX criterion chooses R so as to maximise the function:

$$g(B) = \sum_{k=1}^m \left[\frac{1}{p} \sum_{j=1}^p b_{jk}^4 - \left(\frac{1}{p} \sum_{j=1}^p b_{jk}^2 \right)^2 \right] \quad (12)$$

where b_{jk} , $j = 1, \dots, p$, and $k = 1, \dots, m$ are the elements of the matrix B , i.e. $[B]_{jk} = b_{jk}$. VARIMAX therefore maximises the total variance of the squared loading coefficients within each column of $B = A_m R$. This operation drives the squared loading coefficients towards 0 or 1 and hence the rotated loading coefficients towards 0 or ± 1 . Another well known orthogonal rotation criteria is QUARTIMAX, which attempts to maximise instead the sum of the fourth power of the (rotated) loadings B . The structures obtained using QUARTIMAX tend in general to be less local than those obtained using VARIMAX.

In oblique rotation, one seeks a (non-orthogonal) rotation matrix R with unit length columns, such that the oblique rotated loadings:

$$B = A_m (R^T)^{-1} \quad (13)$$

minimise a certain criterion $f(B)$. Various oblique rotation criteria exist in the literature. A familiar example is the QUARTIMIN criterion (Carroll 1953; Harman 1976, p. 305) for which:

$$f(B) = \frac{1}{4} \sum_{r \neq s} \sum_i b_{ir}^2 b_{is}^2. \quad (14)$$

In this manuscript we have applied orthogonal and oblique rotation. Results from VARIMAX will be shown, but the discussion will also include results using other criteria. Although Richman (1986)'s results suggest that VARIMAX is among the least accurate, his findings were based on synthetic data, and as such should not be taken as a rigid rule. In general, results are often more sensitive to the choice of how many EOFs to rotate than to the choice of rotation criteria. Detailed comparison between existing rotation methods is beyond the scope of this paper and is left for future research.

3 Simplified EOF method: SCoTLASS

3.1 Background

REOFs have been introduced as a way to improve interpretation and yield simpler, more regional and physically believable patterns than EOFs. There is, however, a difficulty in building objective simplicity criteria. As noted by Jolliffe et al. (2002) concentrating the EOF coefficients close to 0 or ± 1 is not the only or best possible definition of simplicity. For example, a constant 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 drawbacks which make the method quite controversial (Richman 1986, 1987; Jolliffe 1987, 1995; Mestas-Nuñez 2000). First, to apply rotation, one has to decide the number of EOFs or PCs to be rotated. Technically, of course, all the EOFs can be rotated, but this defeats the objective of dimension reduction and is likely to lead to trivial solutions. The result of rotation will depend, in general, on the particular number chosen. Among the other issues that need to be addressed when using rotation are whether it should be orthogonal or oblique, and which of the large number of simplicity criteria should be used. The choice of normalisation constraint (Jolliffe 1995), although not unique to rotation, can make fundamental differences in REOFs whereas for EOFs it amounts simply to a rescaling. A simplification technique that can overcome most of these problems, and which in the meantime retains some of the useful properties of EOFs, is desirable. Such a technique is described next.

3.2 LASSO-based simplified EOFs

Chapter 11 of Jolliffe (2002) describes various simplification techniques. Most of these techniques attempt explicitly or implicitly to reduce the two stages of rotated PCA into just one step. Here we discuss a particularly interesting method of simplicity that is derived from a form of regression analysis. In multiple linear regression, a common problem that arises is instability of regression coefficients because of collinearity or high dimensionality. This problem has been investigated by Tibshirani (1996) who proposed a technique known as the least absolute shrinkage and selection operator (LASSO). The LASSO approach attempts to shrink some regression coefficients exactly to zero, hence implicitly selecting variables. The same idea was adapted by Jolliffe et al. (2003) for use in PCA in order to shrink some loadings to zero. Jolliffe et al. (2003) called the technique Simplified Component Technique-LASSO (SCoTLASS). Here we refer to the SCoTLASS EOF method as simplified EOFs (SEOFs),

but the reader should note the use of adjectives "simple" and "simplified" can also describe other different techniques in the literature.

The SEOF method makes use of the main principles behind EOFs and REOFs simultaneously. It achieves simultaneously the three objectives of successive variance maximisation, orthogonality and simplicity of the patterns. The objective of SEOFs is to seek directions $\mathbf{a}_k = (a_{k1}, a_{k2}, \dots, a_{kp})^T$, $k = 1, \dots, p$, that maximise

$$F(\mathbf{a}_k) = \mathbf{a}_k^T S \mathbf{a}_k \quad (15)$$

subject to the constraints

$$\mathbf{a}_k^T \mathbf{a}_l = \delta_{kl}, \quad (16)$$

as with EOFs, but to achieve simplicity the LASSO technique requires an extra constraint to be satisfied (Jolliffe et al. 2003):

$$\|\mathbf{a}_k\|_1 = \sum_{j=1}^p |a_{kj}| = \mathbf{a}_k^T \text{sign}(\mathbf{a}_k) \leq \tau, \quad (17)$$

where τ is a tunable parameter, and $\text{sign}(\mathbf{a}_k) = (\text{sign}(a_{k1}), \dots, \text{sign}(a_{kp}))^T$ is the sign of \mathbf{a}_k . Because of constraint (16), Eq. (17) is only possible for $\tau \geq 1$. Furthermore, since $\|\mathbf{a}\|_1$ reaches its maximum on the unit sphere only when all the components are equal we get $\|\mathbf{a}\|_1 \leq \sqrt{p}$, and hence for $\tau \geq \sqrt{p}$ we regain standard EOFs/PCs. Therefore EOFs can be regarded as a special case of SEOFs obtained when $\tau \geq \sqrt{p}$.

3.3 Numerical solution of SEOFs

The optimisation problem (15-17) is non-quadratic and nondifferentiable due to the LASSO condition. The solution can only be obtained numerically using a suitable descent algorithm. The nondifferentiable condition (17) is a particular nuisance for optimisation, and it is desirable to smooth it out. Trendafilov and Jolliffe (2005) used the fact that $\tanh(x) \sim \frac{|x|}{x} = \text{sign}(x)$ for large values of $|x|$ to transform (17) to the smooth constraint

$$\mathbf{a}_k^T \tanh(\gamma \mathbf{a}_k) - \tau = \sum_{j=1}^d a_{kj} \tanh(\gamma a_{kj}) - \tau \leq 0 \quad (18)$$

for some fixed sufficiently large number γ . Trendafilov and Jolliffe (2005) then solved this optimisation problem using the projected gradient approach and for completeness we briefly review the method here. To make it look like the standard EOF problem (4) or (9), condition (18) is incorporated into the function (15) as an exterior penalty function, see e.g. Gill et al (1981). This means that this condition will be explicitly

taken into account only if it is violated. This resulting cost function which is to be maximised reads:

$$F_\mu(\mathbf{a}_k) = \frac{1}{2}\mathbf{a}_k^T S \mathbf{a}_k - \mu H(\mathbf{a}_k^T \tanh(\gamma \mathbf{a}_k) - \tau) \quad (19)$$

where the exterior penalty function is given by

$$H(x) = \max(0, x),$$

and μ is a large positive number. Again to make (19) differentiable we use the fact that $\max(x, y) = \frac{1}{2}(x + y + |x - y|)$, and hence the exterior penalty function is replaced by $H(x) = \frac{1}{2}x(1 + \tanh\gamma x)$. Finally the optimisation problem (16-19) is solved using the projection operator π_k applied to the gradient to yield:

$$\frac{d\mathbf{a}_k}{dt} = \pi_k \left(I_p - \mathbf{a}_k \mathbf{a}_k^T \right) \nabla F_\mu(\mathbf{a}_k). \quad (20)$$

The k 'th SEOF \mathbf{a}_k is then obtained as the limit of the solution to the dynamical system (20) as $t \rightarrow \infty$.

In this high dimensional application, we have found it more efficient and substantially faster for $k \geq 2$ to slightly modify Eq. (20) as follows. To find the k 'th SEOF we first remove the effect of the previous $k - 1$ SEOFs by defining the "residual" data matrix:

$$Y_k = X \left(I_p - \sum_{l=0}^{k-1} \mathbf{a}_l \mathbf{a}_l^T \right) = X \pi_k \quad (21)$$

with the convention $\mathbf{a}_0 = \mathbf{0}$. We then compute the corresponding "residual" covariance matrix

$$S_k = \frac{1}{n} Y_k^T Y_k = \left(I_p - \sum_{l=0}^{k-1} \mathbf{a}_l \mathbf{a}_l^T \right) S \left(I_p - \sum_{l=0}^{k-1} \mathbf{a}_l \mathbf{a}_l^T \right). \quad (22)$$

The k 'th SEOF \mathbf{a}_k is then obtained as the stationary solution to the dynamical system:

$$\frac{d}{dt} \mathbf{a}_k = \left(I_p - \mathbf{a}_k \mathbf{a}_k^T \right) \nabla F_\mu^{(k)}(\mathbf{a}_k) \quad (23)$$

where $F_\mu^{(k)}$ is defined as in (19) except that S is replaced by S_k , i.e.

$$F_\mu^{(k)}(\mathbf{a}) = \frac{1}{2} \mathbf{a}_k^T S_k \mathbf{a} - \mu H(\mathbf{a}^T \tanh(\gamma \mathbf{a}) - \tau). \quad (24)$$

4 Application to winter monthly sea level pressure

4.1 Dataset

We illustrate the methods with the winter monthly mean sea level pressure (SLP) over the Northern Hemisphere (NH) north of 20°N . The dataset comes from the

the National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay et al. 1996; Kistler et al. 2001). The dataset is available on a $2.5^\circ \times 2.5^\circ$ regular grid and spans the period January 1948 to December 2000. The winter season is defined by December through to February (DJF). The mean annual cycle is first calculated by averaging the monthly data over the years and then subtracted from the data set to get SLP anomalies. An area weighting that consists of multiplying the SLP anomalies by the square root of the cosine of the corresponding latitude is applied.

4.2 Application

4.2.1 EOF results

Figure 1 shows the two leading patterns EOF1 and EOF2 of the winter monthly SLP. They explain respectively 21% and 13% of the total winter variance and are well separated according to the North et al. (1982) rule of thumb. EOF1 (Fig. 1a) has a low over the polar region and two high pressure centres over the Mediterranean/North East Atlantic and over the North Pacific respectively. It corresponds to the familiar Arctic Oscillation (AO) pattern (Thompson and Wallace 1998, 2000; Wallace 2000, Wallace and Thompson 2002). EOF2 in Fig. 1b shows two separated centres over the North East Atlantic and North Pacific respectively. Figure 2 shows the corresponding standardised PC time series. Note in particular the increasing trend after the mid-seventies in the first PC, a well known feature of the AO pattern.

4.2.2 REOF results

We have applied various rotation criteria to the winter monthly SLP EOFs including orthogonal rotation, e.g. VARIMAX, and QUARTIMAX, and also oblique rotation, e.g. QUARTIMIN. The discussion, however, will be centered around VARIMAX, but will also mention results using other criteria. For various values of m , rotation is applied to the leading m EOFs using (11) or (13). We have also investigated the rotation of both; the (unscaled) EOFs and the EOFs scaled by the square root of the corresponding eigenvalues. We therefore have four types of rotation; (i) orthogonal rotation of EOFs, (ii) orthogonal rotation of scaled EOFs, (iii) oblique rotation of EOFs, and (iv) oblique rotation of scaled EOFs.

In all the cases studied, using various rotation criteria and various number of EOFs chosen for rotation, we have found that (i) and (iii) give virtually the same result, which is discussed below. Figure 3 shows a scatter plot of loadings in REOFs using VARIMAX versus REOFs using QUARTIMIN for $m = 30$. Similar plots have been

obtained for other values of m and other criteria. The scatter with negative slope (Fig. 3) simply indicates that the corresponding REOFs have opposite sign. Because the order of REOFs is arbitrary, we have ordered them according to the variance of the associated time series although it should be noted that because the time series are correlated, the variances are not additive. For example, when 3 EOFs are rotated the first rotated pattern (not shown) comes out as the NAO with maximum variance whereas the Pacific pattern has the least variance of its time series. For $4 \leq m \leq 8$, however, the time series of the NAO rotated pattern loses its maximum variance in favour of a pattern that has most of its structure located around the polar region. Figure 4 shows an example for $m = 6$. The time series of the polar structure (Fig. 4a) has the maximum variance whereas time series of the NAO (Fig. 4b) and the North Pacific pattern (Fig. 4c) have respectively the third and fourth maximum variances. It is worth noting that when the NAO can be identified, the Pacific pattern time series always has less variance than that of the NAO.

As m increases further the scale of the rotated patterns becomes smaller and more concentrated. In particular the NAO pattern starts losing its structure. For example, when $m = 12$ (not shown) we get two NAO-like patterns with their northern centres located respectively over Hudson Bay and the west of Scotland. Figure 5 shows an example when $m = 20$, where we obtain a small scale AO-like (Fig. 5a) and NAO-like pattern (Fig. 5b) and another pattern looking like the NAO-jet (not shown). An example of the North Pacific pattern for $m = 20$ is shown in Fig. 5c. Finally for m larger than about 30 the rotated patterns clearly become single structured, with smaller and smaller spatial extension that become localised around selected individual grid points.

The case (ii), i.e. orthogonal rotation of scaled EOFs, turns out to be more robust regarding changes in the number m of chosen EOFs. This is due to the monotonic structure of the spectrum of the covariance matrix whereby successive ranked EOFs contribute less than the previous ones, and therefore cannot change substantially the structure of the leading REOFs. Note that the ranks of REOFs are now obtained from their squared norms. This robustness is found to apply for the two methods used here, i.e. VARIMAX and QUARTIMAX. The leading REOFs from both methods are similar, but this correspondence breaks down for higher REOFs. For example when $m = 30$, only the leading 15 REOFs using VARIMAX or QUARTIMAX are similar. Figure 6 shows the first three REOFs obtained with VARIMAX (with similar results for QUARTIMAX) using $m = 20$ EOFs. The first two REOFs show respectively the NAO (6a) and the North Pacific pattern (6b). The third REOF (6c) shows a structure similar to the Eurasian type I pattern of Barnston and Livezey (1987)

known also as the Scandinavian pattern. Note that the variances given by the squares of the REOFs norms are not additive because the associated time series XB are not uncorrelated. Any (non-trivial) orthogonal rotation of EOFs cannot conserve both the spatial orthogonality and temporal uncorrelatedness.

For case (iv), on the other hand, we found that the rotation algorithm runs into convergence problems. This problem is related to matrix inversion where the rotation matrices turn out to be badly conditioned. Although this may not happen with synthetic data, the problem seems more likely to arise with real data.

4.2.3 SEOF results

To obtain the winter monthly mean SLP SEOFs we numerically integrated (23) using the function ODE15S in the MATLAB software which solves stiff differential equations. The values of the constants γ and μ were fixed to 1000 and 800 respectively, but the results are invariant to changes in these constants. Figure 7 shows an example of the function $F_\mu(EOF1)$ versus γ for $\mu = 1000$. It is clear that the function becomes nearly independent of γ for γ larger than few hundred. The SEOF solutions are also found to be invariant to changes in μ for $\mu \geq 100$. The obtained solutions for the various values of μ are virtually identical. The integration of (23) is performed sequentially for a given number k of SEOFs to be computed. Because the algorithm is computationally intensive, we have used a coarser grid of $5^\circ \times 5^\circ$ by taking SLP values at every other grid point, and we have limited ourselves to computing and discussing the leading three SEOF patterns.

The SEOFs search can be made more efficient by running the algorithm for various values of τ starting from $\tau_0 = \sqrt{p}$ where the solution is known, then gradually decreasing τ , e.g. by 2 or 4, and using, for each τ , the previous solution, obtained e.g with $\tau + 4$ or $\tau + 2$, as initial condition. In this way the CPU time can be reduced by about one order of magnitude compared to the case with a random initial condition. This is because for small τ the feasible set being not-connected (see Trendafilov and Jolliffe 2005), the algorithm may jump between the different connected parts of the set when the initial condition is not close to the optimum.

The leading SEOFs have been computed for various values of the threshold τ , in the range $8 - 30$. Note that above $\tau = 32.8 (= \sqrt{1080})$ we get PCA. From about $\tau = 26$ upwards, the leading SEOFs do not change much and they start looking much like their EOF counterparts. For smaller values of the threshold τ , in the range $8 \leq \tau \leq 20$ we obtain the NAO as the leading pattern and the North Pacific pattern as the next SEOF. Figure 8 shows SEOF1 (8a) and SEOF2 (8b) for $\tau = 8$ showing a localised structure. The shaded area indicates the region where the loadings (or

coefficients) are "exactly" zero. Fig. 8a shows a clear dipolar NAO pattern with centres of opposite signs centred respectively over Iceland and west of the Iberian peninsula. The second pattern (8b) shows a clear monopolar structure centred over the Northern Mid-Pacific. Figure 9 is similar to Fig. 8 but is obtained with $\tau = 18$. Note how the shaded region shrinks and the pattern gets enlarged. Figure 10 shows the same patterns but for $\tau = 24$ with further reduction of the shaded area.

It is clear that the leading SEOF pattern is a NAO for a wide range of thresholds chosen in the middle of the interval $1 \leq \tau \leq \sqrt{1080}$ and that the second SEOF comes out as the North Pacific pattern. Figure 11 shows the time series of the simplified principal components (SPCs) associated with SEOF1 and SEOF2 patterns respectively for $\tau = 18$. The variance of the NAO time series (Fig. 11a) is $5.87 \times 10^3 \text{ hPa}^2$ whereas the variance of SPC2 (Fig. 11b) is $2.67 \times 10^3 \text{ hPa}^2$. For larger τ the difference between the two variances increases until they reach the respective variances of the PC's, i.e. $7.28 \times 10^3 \text{ hPa}^2$ and $2.6 \times 10^3 \text{ hPa}^2$ (a ratio of nearly 3) and this is obtained when τ is around 26. For smaller τ , however, the difference between both the variances becomes smaller. So, for thresholds less than $\frac{1}{2}\sqrt{p}$ the variances of both the time series are not too different and this may explain why the simple patterns found by SEOFs appear to be combined in a single pattern using EOF analysis (Fig. 1a). For example, when $\tau = 10$ the variance associated with the first pattern (NAO) is $2.61 \times 10^3 \text{ hPa}^2$ whereas for the second pattern (North Pacific Centre) one gets $2.33 \times 10^3 \text{ hPa}^2$, i.e. with variance ratio of the order 0.89. This also implies that the first two local maxima corresponding to SEOFs 1 and 2 are close to each other in state space, and therefore, it is quite difficult to separate them using EOFs alone. For comparison the third simplified EOF obtained for $\tau = 12$ (not shown) is similar the Scandinavian pattern (Fig. 6c).

The SEOF method has several advantages. In SEOFs simplicity, maximum variance, and orthogonality of spatial patterns are all addressed simultaneously. In addition, whereas rotated EOFs depend in a fundamental way on the normalisation chosen, there is no secondary operation, like rotation, on the SEOFs, so no real dependence on the normalisation.

The method, however, has, like any other technique, some potential disadvantages. The method relies on the choice of the threshold parameter τ . There is currently no objective way of making this choice. However, this is no more of a problem than choosing the number of EOFs to rotate, and the simplicity criterion, in REOFs. In our example, leading SEOFs are reasonably invariant to changes in the threshold parameters for a wide range of values, except that the spatial extention of the patterns gets reduced with the threshold parameter. Another disadvantage is the potential

loss of variance compared to the same number of EOFs or REOFs. However, our experience is that the simplicity gained is often worth the small loss of variance. Also, compared to EOFs, SEOFs lose one of the space and time orthogonality properties, but this is also true of REOFs.

Finally, the method is more expensive computationally compared, to EOFs or REOFs when rotating a relatively small number of EOFs. However, when rotating a large number of EOFs, using e.g. oblique rotation the computation can also be quite expensive, and other recently proposed dimension-reducing techniques, such as nonlinear PCA, are even more computationally expensive.

The non-differentiability of the L_1 -norm constraint does not pose a serious problem regarding the smoothness of SEOFs because the atmosphere contains already examples of sharp gradients such as fronts or jet streams. Note also that the problem we solve is a smoothed version of the original one. Furthermore, the L_1 -norm constraint is not of equality type, so the field can still be smooth while satisfying the constraint.

4.3 Summary

Rotation achieves simpler structures than EOFs. To apply rotation, however, one has to decide the type of rotation, the criterion to be minimised, and whether to scale EOFs. Rotation of (unscaled) EOFs seems to be robust regarding changes in the type of rotation and the rotation criteria, but not to changes in the number m of EOFs. On the other hand, scaled EOFs seem to be better off with orthogonal rotation, regarding consistency of the leading rotated patterns. It is clear, however, that such rotation of scaled EOFs excludes contribution to the leading modes from low ranked EOFs, with small eigenvalues, but which can be physically relevant. So here is the question: which way to go? We argue that simplified EOFs offer a possible answer.

In SEOFs the simplicity is controlled by the threshold parameter, which is the only parameter required. The other parameters, e.g. μ and γ , are introduced for numerical purposes, and the final solutions are almost invariant to changes in these parameters. In our example the structure of the leading SEOF patterns is also largely invariant vis-a-vis changes to the threshold parameter, as long as the latter is not too close to its upper bound. For example, we obtain the NAO, the North Pacific pattern, and the Scandinavian pattern for a wide range of values of τ . Furthermore, as this parameter gets smaller the spatial extension of the SEOFs gets smaller but the structure is similar. As the threshold parameter is decreased the variance of the leading SEOFs decreases as expected. This is illustrated in Fig. 12, which shows the ratio between the variance of SEOF1 to that of EOF1 versus the simplicity parameter τ . For this

example, the convergence toward EOFs is obtained about the value $\tau = \frac{2}{3}\sqrt{p}$. The loss in variance is justified by the increased simplicity. From a comparison between SEOFs and physically relevant teleconnection patterns, the results suggest the value $\tau = \frac{1}{3}\sqrt{p}$ to be a reasonably good choice.

The method is computationally intensive, but can still be very useful to gain insights if we are interested in the leading few patterns for interpretation. In our example, the method seems to give comparable results for the (simplified) leading modes of variability compared to orthogonal rotation of scaled EOFs. For higher order modes, however, since there is no consistent correspondence between the rotated modes of one orthogonal rotation with another orthogonal rotation, SEOFs can be used as an alternative to interpret high order (simplified) modes of variability.

5 Conclusions

Finding patterns from climate data with moderately large dimensions is, and will remain a challenging task. The task is even more challenging when it comes to physical interpretation. The EOF technique has been around for more than five decades in climate research and is widely used because it allows a useful space display and time display, and is easy to solve using a number of known matrix algorithms, such as the QR algorithm or SVD. The method, however, displays a number of drawbacks when it comes to interpretation because of the strong geometric constraints imposed upon EOFs, such as orthogonality, uncorrelatedness, and domain dependence.

Rotation of the loadings (EOFs) helps alleviate some of the problems encountered with EOFs and yields simpler structures. Rotation, however, raises more issues that need to be addressed, such as choice of the type, and criterion of rotation as well as deciding whether to scale the EOFs. The analyses show that orthogonal and oblique rotations of (unscaled) EOFs produce similar results. The obtained REOFs, however, are not robust to changes in the number of EOFs chosen for rotation. Orthogonal rotations of EOFs scaled by the square root of the corresponding eigenvalues produce leading rotated patterns that are almost invariant to changes in the rotation criteria or the number of EOFs chosen for rotation. Higher order REOFs, however, are not invariant. Also oblique rotation of scaled EOFs can easily run into convergence problems due to bad conditioning.

We have presented a simplified EOF (SEOF) technique that is able to identify patterns with simple structures. The method is formulated in a similar manner to EOFs, and thus keeps some of the EOF's useful properties, but also uses the simplicity property of REOFs simultaneously. Variance maximisation, spatial orthogonality,

and simplicity are all addressed simultaneously; hence the shortcoming in REOFs of needing to choose the number of EOFs to be rotated is overcome.

The SEOF method finds orthogonal patterns that maximise variance subject to an extra constraint that induces simplicity, namely that the sum of the absolute value of the loadings is to be less than a chosen threshold τ . This threshold parameter τ chosen from within $[1, \sqrt{p}]$, is essentially the only parameter for SEOFs. The remaining parameters are only introduced for numerical purposes, but do not change the structure of the final solutions. The results in our example are not overly sensitive to the choice of the threshold. The nondifferentiability of the L_1 -norm constraint does not pose a serious problem since there are in the atmosphere sharp gradients. The algorithm is more expensive computationally than EOFs and REOFs combined, when rotating a relatively small number of EOFs. We have applied the method to winter monthly mean SLP using NCAR/NCEP reanalyses from January 1948 to December 2000 using December through to February data and compared the results of the leading patterns with those obtained using REOFs. A recommended choice for the threshold parameter corresponds approximately to the first third of the interval, i.e. $\tau \approx \frac{1}{3}\sqrt{p}$. In general, however, the choice of τ may be a difficult problem, and it could be necessary to examine several values before giving a final answer. In our example, the SEOF patterns retain their structure, but with decreasing extent because of the increasing number of zero-grid points, as the threshold decreases. The NAO pattern is particularly robust to changes in the threshold. The Pacific pattern, a monopolar structure, and the Scandinavian pattern were obtained as second and third SEOFs respectively for most thresholds.

We have also shown that for large values of the simplicity parameter the variances of the leading simplified PC's tend toward those of their PCs counterpart. For smaller τ than about $\frac{1}{2}\sqrt{p}$, the variances explained by the first two SEOFs in our example are close to each other. The two patterns form two local maxima of the function to be maximised and are close to each other in state space. This may be one reason why conventional EOF analysis is unable to pick them individually, but combine them yielding the AO pattern observed in Fig. 1.

To conclude, we argue that the SCoTLASS method of simplicity provides a good alternative to rotated EOFs, and yields robust simplified EOFs. The SEOFs still explain a large amount of variance and yield simple versions of the NAO, the North Pacific pattern, and the Scandinavian pattern, as leading modes of variability. The increase in simplicity yields, as expected, a decrease in variance. We argue, however, that this loss of variance is justified by the increased simplicity.

Acknowledgements The authors wish to thank Prof Brian Hoskins for his positive feedback on the paper, and two anonymous reviewers for their constructive comments, and one of them for providing a particularly stimulating and thorough review that improved the manuscript. AH was funded by the Centre for Global Atmospheric Modelling (CGAM), and ITJ was funded by National Environmental Research Council (NERC) grant: Discipline Bridging Award NER/T/S/2003/00126.

References

- Ambaum, M.H.P., B.J. Hoskins, and D.B. Stephenson, 2001: Arctic Oscillation or North Atlantic Oscillation? *J. Climate*, **14**, 3495-3507.
- Ambaum, M.H.P., B.J. Hoskins, and D.B. Stephenson, 2002: Corrigendum: Arctic Oscillation or North Atlantic Oscillation? *J. Climate*, **15**, 553.
- Ambrizzi, T., B. J. Hoskins, and H.-H. Hsu, 1995: Rossby wave propagation and teleconnection patterns in the Austral winter. *J. Atmos. Sci.*, **52**, 3661-3672.
- Dommgenget, D., and M. Latif, 2002: A cautionary note on the interpretation of EOFs. *J. Climate*, **15**, 216-225.
- Barnston, A. G., and B. E. Livezey, 1987: Classification, seasonality, and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115**, 1083-1126.
- Botsaris, C.A., and D.H. Jacobson, 1976: A Newton-type curvilinear search method for optimisation. *J. Math. Anal. Appl.*, **54**, 217-229.
- Botsaris, C.A., 1978: A class of methods for unconstrained minimisation based on stable numerical integration techniques. *J. Math. Anal. Appl.*, **63**, 729-749.
- Botsaris, C.A., 1979: A Newton-type curvilinear search method for constrained optimisation. *J. Math. Anal. Appl.*, **69**, 372-397.
- Botsaris, C.A., 1981: Constrained optimisation along geodesics. *J. Math. Anal. Appl.*, **79**, 295-306.
- Brown, A.A., 1986: Optimisation methods involving the solution of ordinary differential equations. PhD thesis, the Hatfield polytechnic, available from the British library.
- Carroll, J. B., 1953: An analytical solution for approximating simple structure in factor analysis. *Psychometrika*, **18**, 23-38.
- Cheng, X., G. Nitsche, and J.M. Wallace, 1995: Robustness of low-frequency circulation patterns derived from EOF and rotated EOF analysis. *J. Climate*, **8**, 1709-1720.
- Evtushenko, J. G., 1974: Two numerical methods of solving non-linear programming problems. *Soviet Math. Dokl.*, **15**, 420-423.
- Evtushenko, J. G., and V.G. Zhadan, 1977: A relaxation method for solving problems of non-linear programming. *USSR Comput. Math. and Math. Phys.*, **17**, 73-87.

- Fukuoka, A., 1951: A study of 10-day forecast (A synthetic report). *The Geophysical Magazine, Tokyo*, Vol. XXII, 177-218.
- Gill, P.E., W. Murray, and M.H. Wright, 1981: *Practical Optimization*. Academic Press, London.
- Golub, G.H., and C.F. van Loan, 1996: *Matrix computation*. John Hopkins University Press, Baltimore, MD.
- Hannachi, A., and B. Legras, 1995: Simulated annealing and weather regimes classification. *Tellus, 47 A*, 955-973.
- Hannachi, A., and A. O'Neill, 2001: Atmospheric multiple equilibria and non-Gaussian behaviour in model simulations. *Q. J. R. Meteorol. Soc.*, **127**, 939-958.
- Harman, H. H., 1976: *Modern Factor Analysis* (3d edition). The University of Chicago Press, Chicago.
- Held, I. M., 1983: Stationary and quasi-stationary eddies in the extratropical troposphere: Theory. in *Large-Scale Dynamical Processes in the Atmosphere*, B. J. Hoskins and R. P. Pearce, Eds, Academic Press, 127-168.
- Hirsch, M.W., and S. Smale, 1974: *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press, London.
- Horel, J.D., 1981: A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 hPa height field. *Mon. Wea. Rev.*, **109**, 2080-2902.
- Hoskins, B. J., and D. J. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, **38**, 1179-1196.
- Hoskins, B. J., and T. Ambrizzi, 1993: Rossby wave propagation on a realistic longitudinally varying flow. *J. Atmos. Sci.*, **50**, 1661-1671.
- Jolliffe, I.T., 1987: Rotation of principal components: Some comments. *J. Climatol*, **7**, 507-510.
- Jolliffe, I.T., 1995: Rotation of principal components: Choice of normalization constraints. *J. Appl. Stat.* **22**, 29-35.
- Jolliffe, I. T., 2002: *Principal Component Analysis*. Springer, second Edition, New York.
- Jolliffe, I.T., M. Uddin, and S.K. Vines, 2002: Simplified EOFs—three alternatives to retain. *Clim. Res.*, **20**, 271-279.
- Jolliffe, I. T., 2003: A cautionary note on artificial examples of EOFs. *J. Climate*, **16**, 1084-1086.

- Jolliffe, I.T., N.T. Trendafilov, and M. Uddin, 2003: A Modified principal component technique based on the LASSO. *J. Comput. Graphical Stat.*. **12**, 531-547.
- Kaiser, H. F., 1958: The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187-200.
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetma, B. Reynolds, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.
- Killworth, P. D., and M. E. McIntyre, 1985: Do Rossby wave critical layers absorb, reflect, or over-reflect? *J. Fluid Mech.*, **161**, 449-492.
- Kirkpatrick, S., C. D. Gelatt Jr, and M. P. Vecchi, 1983: Optimization by simulated annealing. *Science*, **220**, 671-680.
- Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kouksy, H. van den Dool, R. Jenne, and M. Fiorino, 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247-267.
- Krzanowski, W.J., and F.H.C. Marriott, 1995: *Multivariate Analysis, Part 2. Classification, Covariance Structures and Repeated Measurements*. Arnold, London.
- Kutzbach, J. E., 1967: Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. *J. Appl. Meteor.*, **6**, 791-802.
- Lin, S., and Kernighan, B. W., 1973: An effective heuristic algorithm for the travelling salesman problem. *Operat. Res.*, **21**, 498-516.
- Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather prediction. Technical report, Statistical Forecast Project Report 1, Dept. of Meteor., MIT, 1956. 49pp.
- Magnus, J.R., and H. Neudecker, 1995: *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons, Chichester.
- Mardia, K.V., J.T. Kent, and J.M. Bibby, 1979: *Multivariate Analysis*. Academic Press, London.
- Mestas-Nuñez A.M., 2000: Orthogonality properties of rotated empirical

- modes. *Int. J. Climatol.*, **20**, 1509-1516.
- Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, 1953: Equation of state calculation by fast computing machines. *J. Chem. Phys.*, **21**, 1087-1092.
- Nigam, S., and I. M. Held, 1983: The influence of a critical latitude on topographically forced stationary waves in a barotropic model. *J. Atmos. Sci.*, **40**, 2610-2622.
- North, G.R., T.L., Bell, R.F. Cahalan, and F.J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Weather Rev.*, **110**, 699-706.
- Obukhov, A.M., 1947: Statistically homogeneous fields on a sphere. *Usp. Mat. Nauk.*, **2**, 196-198.
- Obukhov, A.M., 1960: The statistically orthogonal expansion of empirical functions. *Bull. Acad. Sci. USSR Geophys. Ser.* (English Transl.), 288-291.
- Preisendorfer, R.W., 1988: *Principal Component Analysis in Meteorology and oceanography*. C.D. Mobley, ed., Elsevier, Amsterdam.
- Richman, M. B., 1981: Obliquely rotated principal components: An improved meteorological map typing technique. *J. Appl. Meteor.*, **20**, 1145-1159.
- Richman, M. B., 1986: Rotation of principal components. *J. Climatol.*, **6**, 293-335.
- Richman, M. B., 1987: Rotation of principal components: A reply *J. Climatol.*, **7**, 511-520.
- Snyman, J.A., 1982: A new and dynamic method for unconstrained optimisation. *Appl. Math. Modelling*, **6**, 449-462.
- Simmons, A.J., J.M. Wallace, and G.W. Branstator, 1983: Barotropic wave propagation and instability, and atmospheric teleconnection patterns. *J. Atmos. Sci.*, **40**, 1363-1392.
- Stephenson, D. B., A. Hannachi, and A. O'Neill, 2004: On the existence of multiple climate regimes. *Q. J. R. Meteorol. Soc.*, **130**, 583-605.
- Thompson, D.W.J., and J.M. Wallace, 1998: The Arctic oscillation signature in wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, **25**, 1297-1300.
- Thompson, D.W.J., and J.M. Wallace, 2000: Annular modes in the extra-tropical circulation. Part I: Month-to-month variability. *J. Climate.*, **13**, 1000-1016.

- Tibshirani, R., 1996: Regression shrinkage and selection via the LASSO. *J. R. Stat. Soc. B*, **58**, 267-288.
- Trendafilov, N.T., and I.T. Jolliffe, 2005: Numerical solution of the SCoTLASS. *Comput. Stat. Data Anal.* In Press.
- von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge.
- Wallace, J.M., 2000: North Atlantic oscillation/annular mode: Two paradigms—one phenomenon. *Quart. J. Roy. Meteor. Soc.*, **126**, 791-805.
- Wallace, J.M., and D.W.J., Thompson, 2002: The Pacific Center of Action of the Northern Hemisphere annular mode: Real or artifact? *J. Climate*, **15**, 1987-1991.
- Ward, M.N., and C.K. Folland, 1991: Prediction of seasonal rainfall in the north Nordeste of Brazil using eigenvectors of sea-surface temperature. *Int. J. Climatol.*, **11**, 711-743.
- Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, San Diego.

Figure Caption

- Fig. 1. The leading two EOFs one (a) and two (b) of monthly mean winter (DJF) sea level pressure. Positive contours solid, and negative contours dashed. Note that the EOFs have been multiplied by 100.
- Fig. 2. The leading two scaled PCs corresponding to the leading two EOFs of Fig. 1.
- Fig. 3. Scatter plot of VARIMAX REOFs versus QUARTIMIN REOFs using $m = 30$ EOFs. The scatter with negative slopes corresponds to similar REOFs but with opposite sign.
- Fig. 4. Three rotated EOF patterns using the leading $m = 6$ EOFs, showing the first pattern (a), the NAO (b), and the North Pacific pattern (c). These rotated patterns are ranked first, third, and fourth respectively. The order is fixed according to the variance of the corresponding time series. Positive contours solid, and negative contours dashed. Values have been multiplied by 100 as in Fig. 1.
- Fig. 5. As in Fig. 4. but for $m = 20$. The order of the patterns are 1 (a), 6 (b), and 10 (c).
- Fig. 6. The leading three REOFs obtained using VARIMAX rotation of the first $m = 20$ EOFs scaled by the square root of the corresponding eigenvalues. The ranks of REOFs are given by their norm squared.
- Fig. 7. The function $F_\mu(EOF1)$ given by (19) versus the parameter γ for $\mu = 1000$.
- Fig. 8. The leading two simplified EOFs; SEOF1 (a) and SEOF2 (b) obtained for a threshold parameter $\tau = 8$. The shaded area shows the grid points with "exactly" zero loadings. As for EOFs and REOFs, the SEOFs have been multiplied by 100.
- Fig. 9. As in Fig. 8 but for $\tau = 18$.
- Fig. 10. As in Fig. 8 but for $\tau = 24$.
- Fig. 11. The scaled time series of the first two simplified PC1 (a), and PC2 (b) corresponding to Fig. 6., i.e. $\tau = 18$.
- Fig. 12. The ratio of variance of SPC1 to that of PC1 versus the simplicity parameter τ .

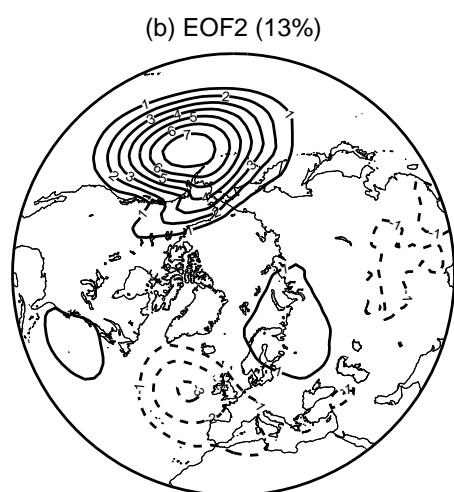
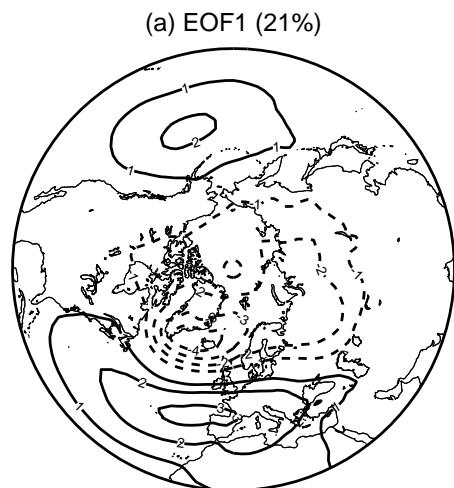


Figure 1: The leading two EOFs one (a) and two (b) of monthly mean winter (DJF) sea level pressure. Positive contours solid, and negative contours dashed. Note that the EOFs have been multiplied by 100.

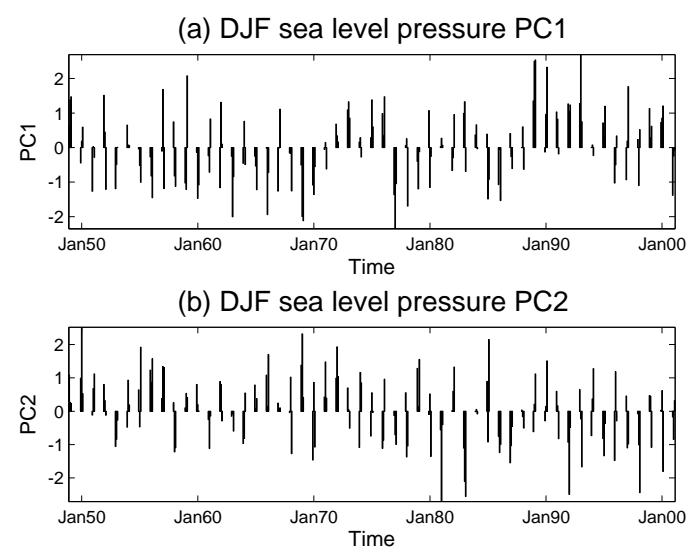


Figure 2: The leading two scaled PCs corresponding to the leading two EOFs of Fig. 1.

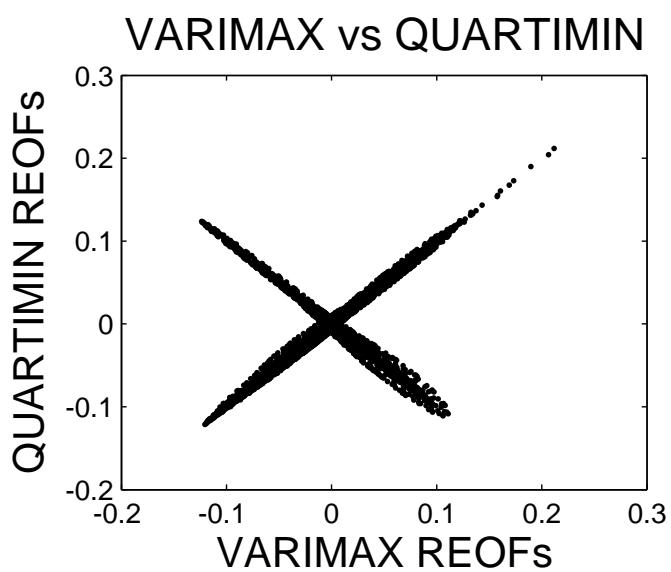


Figure 3: Scatter plot of VARIMAX REOFs versus QUARTIMIN REOFs using $m = 30$ EOFs. The scatter with negative slopes corresponds to similar REOFs but with opposite sign.

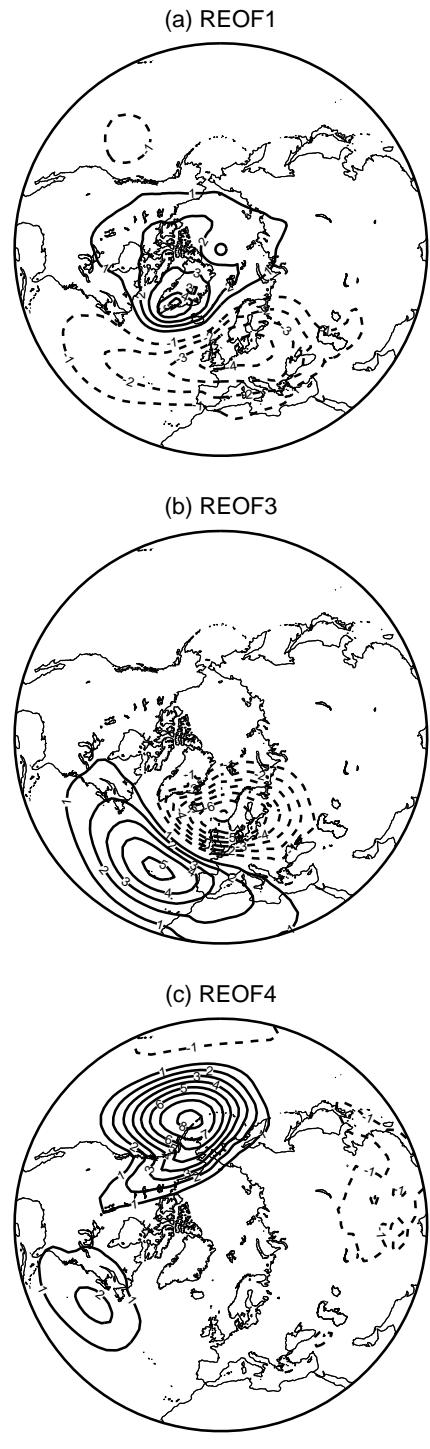


Figure 4: Three rotated EOF patterns using the leading $m = 6$ EOFs, showing the first pattern (a), the NAO (b), and the North Pacific pattern (c). These rotated patterns are ranked first, third, and fourth respectively. The order is fixed according to the variance of the corresponding time series. Positive contours solid, and negative contours dashed. Values have been multiplied by 100 as in Fig. 1.

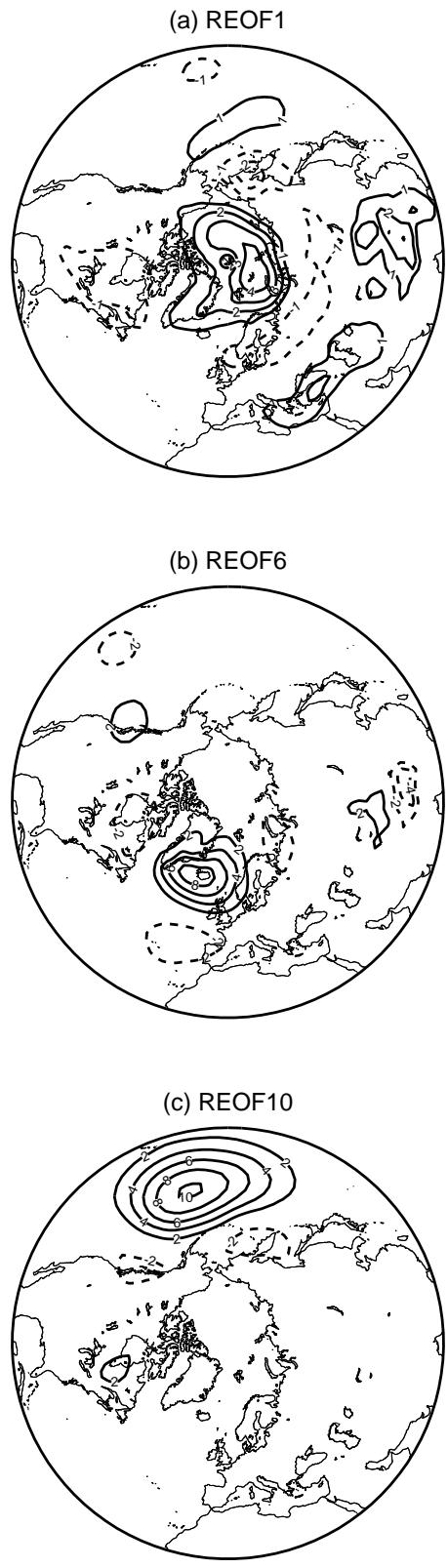
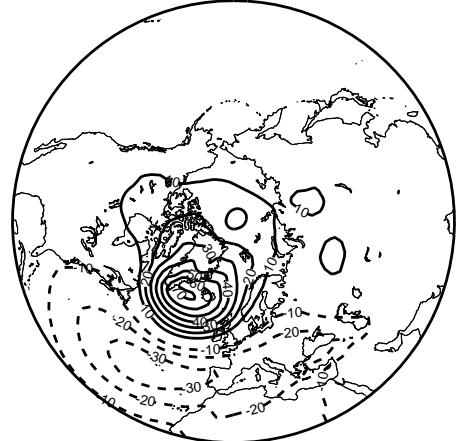
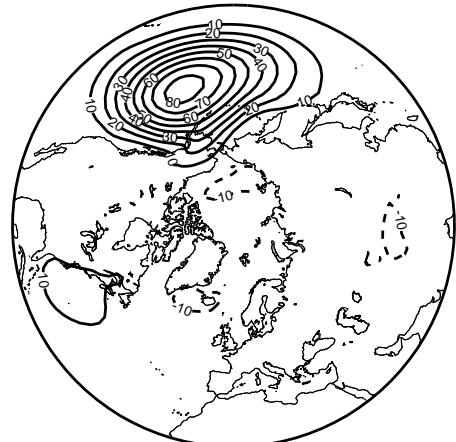


Figure 5: As in Fig. 4. but for $m = 20$. The order of the patterns are 1 (a), 6 (b), and 10 (c).

(a) REOF 1 (m=20)



(b) REEOF 2 (m=20)



(c) REEOF 3 (m=20)

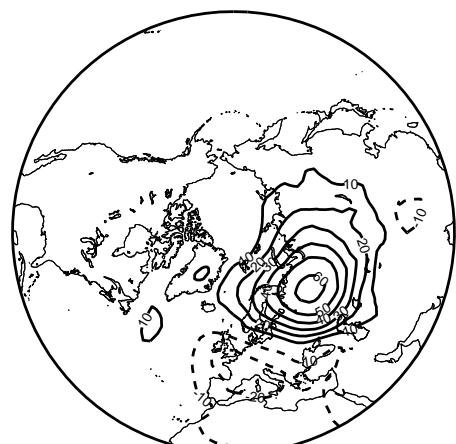


Figure 6: The leading three REEOFs obtained using VARIMAX rotation of the first $m = 20$ EOFs scaled by the square root of the corresponding eigenvalues. The ranks of REEOFs are given by their norm squared.

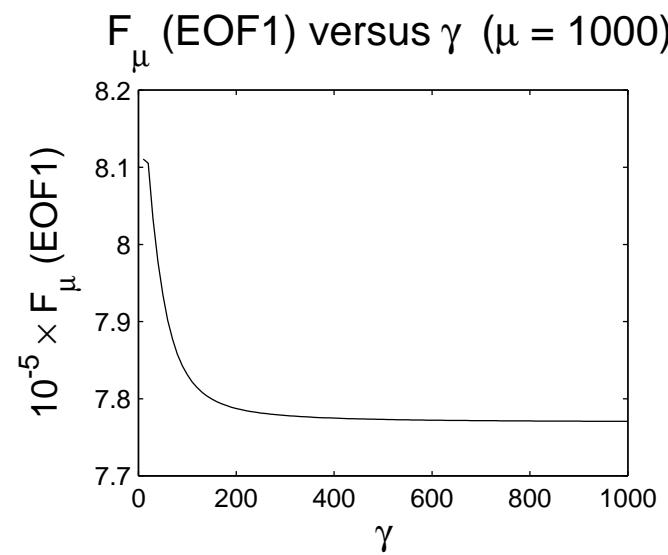
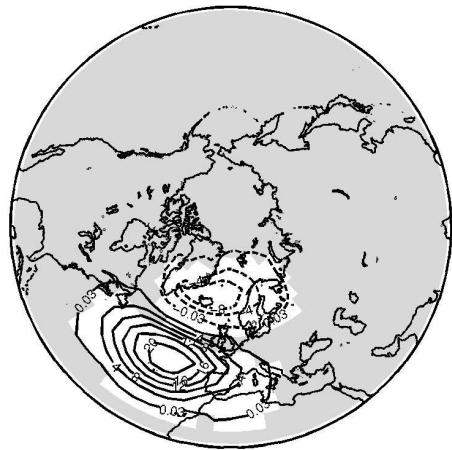


Figure 7: The function $F_\mu(\text{EOF1})$ given by (19) versus the parameter γ for $\mu = 1000$.

(a) SEOF1 ($\tau=8$)



(b) SEOF2 ($\tau=8$)

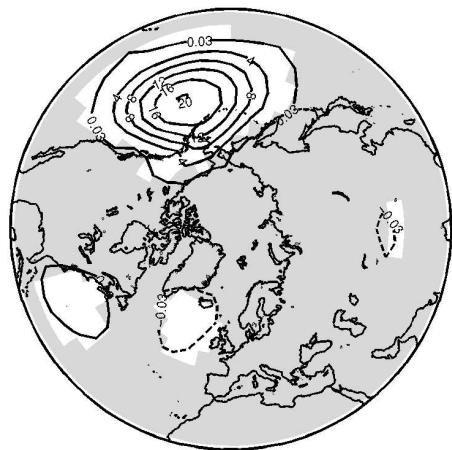
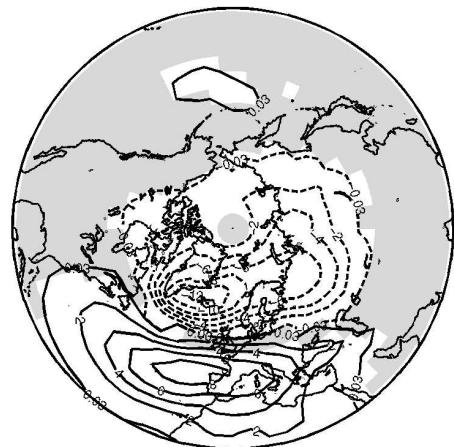


Figure 8: The leading two simplified EOFs; SEOF1 (a) and SEOF2 (b) obtained for a threshold parameter $\tau = 8$. The shaded area shows the grid points with "exactly" zero loadings. As for EOFs and REEOFs, the SEOFs have been multiplied by 100.

(a) SEOF1 ($\tau=18$)



(b) SEOF2 ($\tau=18$)

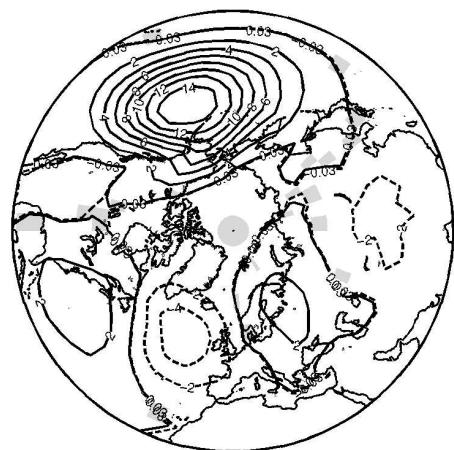
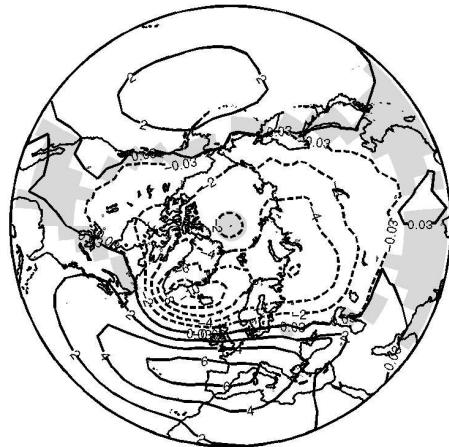


Figure 9: As in Fig. 8 but for $\tau = 18$.

(a) SEOF1 ($\tau=24$)



(b) SEOF2 ($\tau=24$)

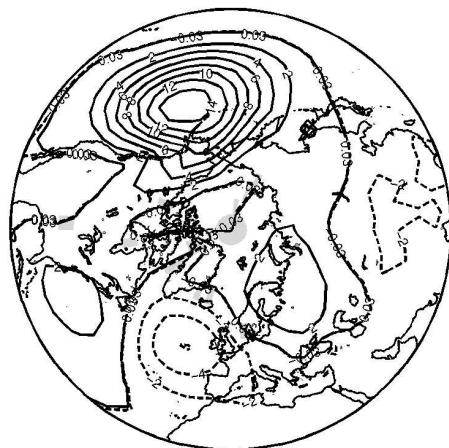


Figure 10: As in Fig. 8 but for $\tau = 24$.

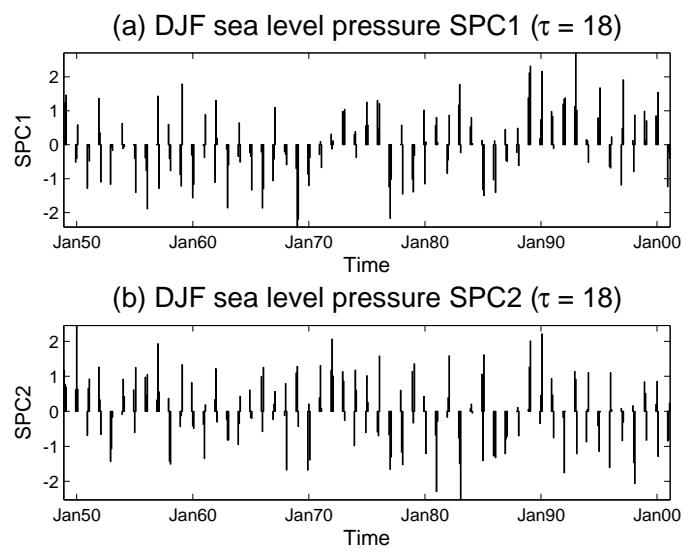


Figure 11: The scaled time series of the first two simplified PC1 (a), and PC2 (b) corresponding to Fig. 6., i.e. $\tau = 18$.

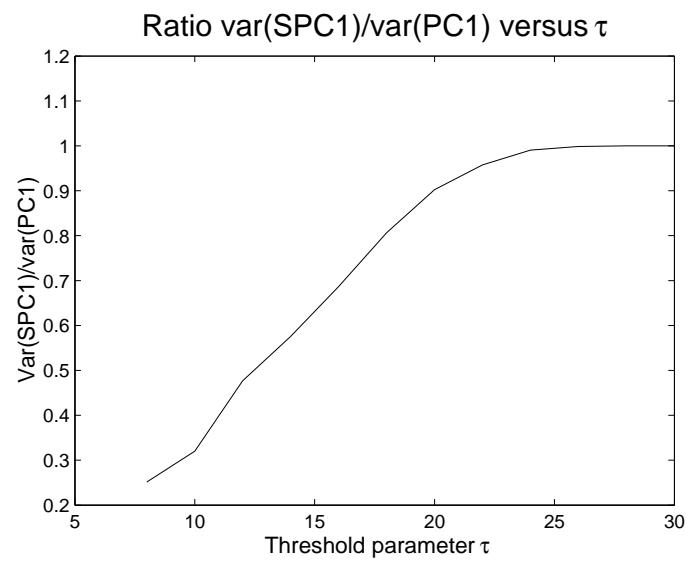


Figure 12: The ratio of variance of SPC1 to that of PC1 versus the simplicity parameter τ .