Empirical Orthogonal Functions: The Medium is the Message

Adam H. Monahan[1,2], John C. Fyfe[3], Maarten H.P. Ambaum [4], David B. Stephenson [5], and Gerald R. North [6]

[1] School of Earth and Ocean Sciences
University of Victoria P.O. Box 3065 STN CSC Victoria, BC, Canada, V8W 3V6
(email: monahana@uvic.ca)

[2] Earth System Evolution Program
Canadian Institute for Advanced Research, Toronto, ON, Canada

[3] Canadian Centre for Climate Modelling and Analysis
Environment Canada
University of Victoria P.O. Box 3065, STN CSC Victoria, BC, Canada, V8W 3V6
(email: John.Fyfe@ec.gc.ca)

[4] Department of Meteorology
University of Reading, PO Box 243 Reading RG6 6BB, United Kingdom
(email:m.h.p.ambaum@reading.ac.uk)

[5] Exeter Climate Systems, Mathematics Research Institute
University of Exeter, North Park Road Exeter EX4 4QF, United Kingdom
(email:d.b.stephenson@ex.ac.uk)

[6] Department of Atmospheric Sciences and the Department of Oceanography
Texas A&M University College Station, TX 77843-3146, USA
(email:g-north@tamu.edu)

Submitted to Journal of Climate

December 15, 2008

Abstract

Empirical Orthogonal Function (EOF) analysis is a powerful tool for data compression and dimensionality reduction broadly used in meteorology and oceanography. Often in the literature, EOF modes are interpreted individually, independent of other modes. In fact, it can be shown that no such attribution can generally be made. This review demonstrates that in general individual EOF modes (i) will not correspond to individual dynamical modes, (ii) will not correspond to individual kinematic degrees of freedom, (iii) will not be statistically independent of other EOF modes, and (iv) will be strongly influenced by the non-local requirement that modes maximise variance over the entire domain. The goal of this review is not to argue against the use of EOF analysis in meteorology and oceanography, but rather to demonstrate the care that must be taken in the interpretation of individual modes in order to distinguish the medium from the message.

1 Introduction

Since its introduction to meteorology by Edward Lorenz (Lorenz, 1956), Empirical Orthogonal Function (EOF) analysis (also known as Principal Component Analysis (PCA), the Karhunen-Loève transform, or Proper Orthogonal Decomposition) has become a fundamental statistical tool in atmosphere, ocean, and climate science for data diagnostics and dynamical mode reduction (e.g. the recent review by Hannachi et al., 2007). In particular, it has become common to use EOF analysis as a tool to probe the physics underlying the variability in a geophysical field of interest. There is no problem with the use of EOF analysis to identify structures in geophysical data (from observations or general circulation models (GCM)) which carry relatively large fractions of variance in the field under consideration. Problems can begin when these statistical structures are interpreted as being of individual dynamical, kinematic, or statistical meaning and are used to define the subsequent physical questions accordingly. Here we review the conditions under which EOFs are of individual physical or statistical meaning, and argue that these conditions are not generally satisfied in real observational or GCM fields.

Following a brief overview of EOF analysis in Section 2, the conditions under which individual EOF modes and dynamical modes correspond in a stable linear system driven by noise (a paradigmatic system for sustained variability around some basic state) are computed in Section 3; these conditions are the exception rather than the rule for geophysical flows. Section 4 further demonstrates that individual EOF modes cannot be expected to be related to individual kinematic degrees of freedom of the system, in the context of an idealised model of the extratropical eddy-driven jet. Furthermore, it is demonstrated in Section 5 that if the probability distribution of the field being analysed is non-Gaussian, then EOF

modes cannot be expected in general to be mutually independent (despite being uncorrelated by construction). Finally, the influence on the structure of EOFs imposed by the non-local requirement that they maximise variance over the entire spatial domain is discussed in Section 6. A discussion and conclusions follow in Section 7.

Note that in this review we follow convention and use the term *EOF mode* as shorthand for the pair of an EOF spatial structure and its associated time series (the Principal Component, or PC). In the present context, the word *mode* is used in the following sense (as given by the Oxford English Dictionary, `http://dictionary.oed.com`): "a particular form, manner, or variety in which some quality, phenomenon, or condition occurs or is manifested". This sense does not imply that a given mode be of individual significance (dynamically, kinematically, or statistically) independent of other modes; in fact, the following discussion is intended precisely to argue against such interpretations in general.

Empirical Orthogonal Function analysis is a powerful tool for data compression and dimensionality reduction in atmosphere, ocean and climate science: the purpose of this review is not to suggest that this technique is without value or to advocate its abandonment by meteorologists or oceanographers. Rather, the intent is to bring together several strands of argument demonstrating that some of the standard interpretations of EOF modes in terms of the underlying physics that are common in the literature have a very doubtful basis in general. That these interpretations can be problematic follows as a consequence of the fundamental mathematical structure of EOF analysis. The following critique addresses ways in which the mathematical details of EOFs (following from basic definitions) are imprinted upon the EOF modes independent of the dynamical, kinematic, or statistical features of the field under consideration. These mathematical details are not incidental to the results of EOF analysis, but rather - as the following discussion will show - they are central. It is to

this discussion that we now turn.

2 EOF Analysis: Some Relevant Facts

We present a brief review of those aspects of EOF analysis that are relevant to the present discussion; a more comprehensive discussion can be found in e.g. von Storch and Zwiers (1999) or the recent review by Hannachi et al. (2007). The following discussion also draws on basic results of liner algebra, an introduction to which can be found in Arfken (1985). This discussion will not consider issues arising around the statistical sampling of EOFs: it will be assumed that the underlying probability distributions (and therefore, all statistical moments) are known exactly. The issues considered in this review are not artifacts of sampling variability.

To begin, we consider an N-dimensional vector time series $\mathbf{x}(t)$, which we may think of as a continuous field (such as temperature or geopotential) sampled at N discrete points in space. Without loss of generality $\mathbf{x}$ can be assumed to be of zero mean (that is, any initially nonzero average has already been subtracted out). The covariance matrix of $\mathbf{x}$ is given by

$$C = \left\langle \mathbf{x}\mathbf{x}^T \right\rangle, \tag{1}$$

where the angle brackets denote probabilistic expectation (that is, the population mean): if $p(\mathbf{x})$ is the probability density function of $\mathbf{x}$, then

$$\langle f(\mathbf{x}) \rangle = \int f(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \tag{2}$$

for any function $f(\mathbf{x})$. The EOFs can be defined as the eigenvectors $\mathbf{e}_k$ of C:

$$C\mathbf{e}_k = \mu_k \mathbf{e}_k \tag{3}$$

with corresponding eigenvalues μ_k. Conventionally, the EOFs are ordered in decreasing magnitude of μ_k: that is $\mu_1 \geq \mu_2 \geq ... \geq \mu_N$. It is convenient to assume that these eigenvectors are unit norm:

$$||\mathbf{e}_k||^2 = \mathbf{e}_k \cdot \mathbf{e}_k = 1. \tag{4}$$

As C is a non-negative definite square matrix, the eigenvalues μ_k are all non-negative and the eigenvectors $\mathbf{e}_k$ form a complete orthonormal basis. That is,

$$\mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{5}$$

and any N-dimensional vector $\mathbf{z}$ can be expressed as a linear combination of the $\mathbf{e}_k$:

$$\mathbf{z} = \sum_{n=1}^{N} \zeta_n \mathbf{e}_n \tag{6}$$

such that the expansion coefficients are projections of $\mathbf{z}$ on $\mathbf{e}_n$:

$$\zeta_n = \mathbf{z} \cdot \mathbf{e}_n \tag{7}$$

In particular, we can write

$$\mathbf{x}(t) = \sum_{n=1}^{N} \alpha_n(t) \mathbf{e}_n \tag{8}$$

where the expansion coefficient time series $\alpha_n(t)$ are the principal components. It follows from the fact that the $\mathbf{e}_k$ are orthonormal eigenvectors of C that the PC time series are mutually uncorrelated:

$$\langle \alpha_j \alpha_k \rangle = \mu_k \delta_{jk} \tag{9}$$

and that the variance of $\alpha_j(t)$ is μ_j. Thus, EOF analysis expresses the (discretely sampled) field $\mathbf{x}$ as the superposition of N mutually orthogonal spatial patterns modulated by N mutually uncorrelated time series. The spatial patterns and time series occur in matched pairs, which are generally referred to as EOF *modes*.

We note the following points:

1. The EOF expansion can be interpreted geometrically as a change of coordinates in R^N through an orthogonal rotation to a basis in which C is diagonal. This emphasises that the EOF expansion is nothing more than another way of describing the time series $\mathbf{x}$ in terms of a new basis set in which this description is particularly simple (from the perspective of the distribution of variance).

2. The total variance of $\mathbf{x}$, $\sum_{n=1}^{N} \mathrm{var}(x_n)$, is simply the sum of the eigenvalues μ_n so the fraction of variance "accounted for" by the nth EOF mode is

$$\frac{\mathrm{var}(x_n)}{\sum_{i=1}^{N} \mathrm{var}(x_i)} = \frac{\mu_n}{\sum_{i=1}^{N} \mu_i} \tag{10}$$

3. EOFs of $\mathbf{x}$ can be obtained variationally in terms of a sequence of lower-dimensional linear approximations to $\mathbf{x}$ of successively decreasing (or non-increasing) mean squared error. In particular, $\mathbf{e}_1$ is the vector in R^N such that the difference between $\mathbf{x}$ and the projection of $\mathbf{x}$ along $\mathbf{e}_1$ is minimised in a mean-square sense: that is, the misfit

$$\epsilon^2 = \left\langle ||\mathbf{x} - (\mathbf{x} \cdot \mathbf{e}_1)\mathbf{e}_1||^2 \right\rangle \tag{11}$$

is minimised by $\mathbf{e}_1$ among all unit vectors in R^N. Having determined $\mathbf{e}_1$, $\mathbf{e}_2$ is defined such that $\mathbf{e}_2$ is orthogonal to $\mathbf{e}_1$ and the misfit

$$\epsilon^2 = \left\langle ||\mathbf{x} - \sum_{k=1}^{2} (\mathbf{x} \cdot \mathbf{e}_k)\mathbf{e}_k||^2 \right\rangle \tag{12}$$

is minimised, and so on for increasing $\mathbf{e}_k$ until all N EOFs are found. This variational problem admits an analytic solution such that the EOFs are given by the eigenvector problem (3), which is how the EOFs are generally found operationally. A particular benefit of the variational formulation is that it demonstrates that the truncated EOF expansion

$$\hat{\mathbf{x}} = \sum_{n=1}^{K} \alpha_n \mathbf{e}_n \tag{13}$$

where $K < N$ is an optimal K-dimensional approximation to $\mathbf{x}$ in the sense that the mean squared difference between the two is minimal. Furthermore, by Eqn.. (10) the fraction of variance variance accounted for by the K-dimensional approximation is

$$\frac{\sum_{k=1}^{K} \mu_k}{\sum_{n=1}^{N} \mu_n}.$$

The eigenvalues μ_k do not increase with k, so as the dimensionality of the approximation K increases the newly-included EOF modes are successively as or less important to the overall variance of the approximation. Truncated EOF expansions therefore provide a powerful tool for *data compression* or *dimensionality reduction* when it is desired to capture as much variance as possible in a lower-dimensional approximations. We note in passing that the fact that an EOF mode accounts for little variance does not necessarily imply that it is dynamically irrelevant (e.g. Crommelin and Majda, 2004), so a truncated EOF basis might not be optimal for representing the dynamics of a system.

There are various technical issues associated with the computation of EOFs that will not be considered in this study, including mixing of EOFs due to sampling-error induced degeneracy of the associated eigenvalues (e.g. North et al., 1982), the use of different norms (such as the total energy) to define the EOFs, or choices of spatial weighting (e.g. correlation-based EOFs, or latitude-dependent weights accounting for sphericity of the domain; c.f. Baldwin et al., in press). The focus of this review is the accuracy of common interpretations of individual EOF modes, which is distinct from these more technical concerns.

3 EOFs and Dynamical Modes

The most natural system in which the statistical modes produced by EOF analysis might be expected to have clear individual dynamical significance is one governed by linear dynamics for which the notion of "dynamical modes" as eigenvectors of the linear dynamical operator is straightforward. In fact, North (1984) demonstrated that the correspondence between EOFs and dynamical modes holds only in a very specialised class of linear dynamical systems which are expected to be the exception rather than the rule in the (linearised) dynamics of real geophysical flows. The original North (1984) argument considered the statistics and dynamics of a continuous field with dynamics described by a noise-forced partial differential equation; for the sake of simplicity, the following discussion considers the dynamics of a finite N-dimensional vector $\mathbf{x}$ (which might be considered as representing a continuous field sampled at a number of discrete points).

Many studies have shown that sustained small-amplitude variability in a broad range of physical situations in the atmosphere and ocean (e.g. Farrell and Ioannou, 1996; Kleeman, 2008; Zanna and Tziperman, 2008) can be described by linear dynamics subject to random (i.e. "turbulent") forcing representing the effects of unresolved physical scales:

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x} + B\dot{\mathbf{W}}, \tag{14}$$

where A is a constant $N \times N$ matrix, B is a constant $N \times M$ matrix, and $\dot{\mathbf{W}}$ is a vector of independent white noise processes:

$$\left\langle \dot{W}_i(t)\dot{W}_j(t') \right\rangle = \delta(t-t')\delta_{ij} \tag{15}$$

(uncorrelated in both space and time). The matrix A is the linear dynamical operator governing the "deterministic dynamics" (most naturally thought of as a linearisation of the

nonlinear dynamics around some basic state), while B characterises the spatial structure of the noise forcing. The deterministic dynamics will be stable (that is, perturbations will all decay asymptotically to zero in the absence of noise forcing) if the real parts of all eigenvalues of A are negative. In this situation, sustained variability of $\mathbf{x}$ is maintained by the turbulent noise forcing alone.

As a first statement about the connection between EOFs and dynamical modes, we can say immediately that the two sets of vectors will not correspond in the case that the linear operator A is non-normal; that is, if it does not commute with its adjoint $A^\dagger$: $AA^\dagger - A^\dagger A \neq 0$. In this case, the eigenvectors of A are in general not mutually orthogonal. As the EOFs are necessarily mutually orthogonal by construction, the EOFs and the dynamical modes will not generally be the same. In fact, studies such as Penland (1996) and Farrell and Ioannou (1996) demonstrate that the EOFs and the dynamical eigenvectors may be very different indeed. Non-normality of the dynamical matrix is the generic case for the linearised dynamics of geophysical systems, particularly in the presence of shear or coupling between systems with very different timescales (e.g. Farrell and Ioannou, 1996; Kleeman, 2008). It follows that we can say that, as a general rule, EOFs and dynamical modes will not coincide.

This argument does not rule out the possibility that EOFs correspond to dynamical modes in the special case that the linearised dynamics are governed by a normal operator. In this case, it can be shown (Appendix A) that the EOFs will only correspond to the dynamical eigenvectors of A if the noise has no spatial structure: that is, if it is spatially uncorrelated. If this is not the case, the structure of the driving noise will be imprinted on the covariance matrix of the damped, driven system so that the EOFs of $\mathbf{x}$ mix the structure of the noise with the structure of the linearised dynamics. In this case again, the EOFs and dynamical modes will not correspond.

The correspondence between EOFs and dynamical modes is even less clear in the case of systems governed by nonlinear dynamics, in which the concept of "dynamical mode" must be generalised to the more abstract notion of dynamically invariant subspaces (which, once entered by the system, will not be exited). Such subspaces will not in general even be planar, in contrast to the case of subspaces spanned by EOFs or linear dynamical modes. For such systems the EOFs will be of course determined by - but on an individual basis cannot be expected to bear any simple relationship to - the dynamics. Consider first an example drawn from Mo and Ghil (1987), in which the vector $\mathbf{x}(t)$ describes the perfectly elliptical orbit of a planetary body around a star (Figure 1). In the absence of external perturbations, this elliptical orbit will be "dynamically invariant". By symmetry, the leading EOF $\mathbf{e}_1$ will be aligned along the semimajor axis of the ellipse (which is longer and therefore carries the most variance) and the second EOF $\mathbf{e}_2$ along the semiminor axis. Both of these directions are dynamically significant, but neither is a "dynamically invariant subspace" on its own. Another illustrative example, also drawn from Mo and Ghil (1987), is that of the EOFs of the Lorenz (1963) model with the standard parameter values for which the system's trajectory settles down on to a strange attractor (Figure 2). Once again, the EOFs identify the orthogonal directions carrying the most variance but do not cleanly correspond to any invariant aspect of the dynamics. The EOFs of this highly nonlinear dynamical system do not have individual dynamical significance.

4 EOFs and Kinematic Degrees of Freedom

We have seen that EOFs will not generally be of individual dynamical significance. As statistical properties of a system are descriptions of variability and thus inherently kinematic,

we might ask if individual EOF modes will be simply related to natural kinematic descriptors of variability. That this cannot be expected to be the case in general will be illustrated by the example of a simple model of a fluctuating jet in zonal-mean zonal wind, for which the EOF problem is analytically solvable.

In both observations and atmospheric GCMs, the leading EOF of extratropical zonal-mean zonal wind is a dipole with a central zero-crossing at approximately the mean latitude of the eddy-driven jet (e.g. Fyfe and Lorenz, 2005; Monahan and Fyfe, 2006; Eichelberger and Hartmann, 2007). To address the kinematic significance of this EOF mode, we consider a jet in zonal-mean zonal wind with Gaussian profile and fluctuating in strength and position:

$$u(x,t) = U(t) \exp\left(-\frac{[x - x_c(t)]^2}{2\sigma_0^2}\right), \tag{16}$$

where x is a meridional coordinate. In this model, the jet strength $U(t)$ and position $x_c(t)$ are the natural kinematic variables of the jet - what we will call the kinematic degrees of freedom. For convenience, we will assume that fluctuations in $U(t)$ and $x_c(t)$ are independent and Gaussian. Based on observations of the extratropical zonal-mean eddy-driven jet (in either hemisphere), we will assume that both of $l = \text{std}(U)/\text{mean}(U)$ and $h = \text{std}(x_c)/\sigma_0$ are $<< 1$. With these assumptions, the covariance matrix of $u(x,t)$ can be computed analytically and expanded as a Taylor series in the small parameters l, h. From these expansions the leading EOFs can be determined in terms of the normalised basis vectors $f_1(x)$, $f_2(x)$, and $f_3(x)$ (Figure 3) corresponding respectively to a monopole, a dipole, and a tripole. By symmetry, the dipole is orthogonal to the monopole and tripole, but the monopole and tripole themselves are not mutually orthogonal (and therefore cannot simultaneously be EOFs).

In the case of pure fluctuations in jet strength, the only EOF with nonzero variance is the monopole. For fluctuations in position alone, the leading two EOFs are the dipole and the

tripole, respectively. When the jet fluctuates in both strength and position, if fluctuations in position are relatively large compared to those of strength (as is the case in observations) then the leading EOF is the dipole $f_1(x)$. The leading PC time series is given by (to leading order in the small parameter h)

$$\alpha_1(t) = \left(\frac{\sqrt{\pi}}{2\sigma_0}\right)^{1/2} U(t)x_c(t). \tag{17}$$

That is, while the dipole pattern arises due to the presence of fluctuations in position, the fluctuations in jet strength also project upon it and are therefore mixed into the associated PC time series. The first EOF mode bundles together both kinematic degrees of freedom and cannot be uniquely associated with position fluctuations. Furthermore, the spatial pattern of the second EOF is a monopole/tripole hybrid where the degree of hybridization is determined by the quantity

$$\delta = \frac{3h^4}{8l^2}. \tag{18}$$

When $\delta << 1$, $\mathbf{e}_2$ is a monopole and when $\delta >> 1$ it is a tripole: in between, it is a linear combination of the two. The monopole comes in from strength fluctuations and the tripole from position fluctuations, but as these are not orthogonal they cannot both simultaneously be EOFs. Spatial structures which are EOFs in the case of fluctuations in a single kinematic degree of freedom on its own will not necessarily be EOFs in the presence of multiple fluctuating degrees of freedom as a consequence of the requirement that the EOFs be mutually orthogonal. Not surprisingly, the time series associated with the second EOF, $\alpha_2(t)$, also mixes together variability in both strength and position.

The observed extratropical eddy-driven jet fluctuates in strength, position, and width (with the first and third of these correlated as a consequence of momentum conservation). The above arguments can be generalised to include fluctuations in jet width (Monahan

and Fyfe, 2006), to relax the Gaussian assumptions on jet profile and kinematic parameter probability distributions (Monahan and Fyfe, in press), and to consider the geopotential EOFs associated with the fluctuating jet (Monahan and Fyfe, 2008). The central conclusion remains unchanged: despite the fact that they directly reflect the kinematics of variability, the defining constraints on EOFs (orthogonal straight-line axes with uncorrelated time series) prevent them in general from being in simple one-to-one correspondence with kinematic descriptors of the field.

5 EOFs of Non-Gaussian Fields

We have seen that individual EOF modes cannot in general be expected to correspond to individual dynamical or kinematic modes. That is, EOFs will not generally partition the state space into dynamically or kinematically independent structures. One might at least hope that the EOF decomposition will partition the state space into *statistically* independent structures: after all, the PC time series are mutually uncorrelated. However, decorrelation does not imply independence: two uncorrelated variables can still be nonlinearly related (Appendix B). If the distribution of $\mathbf{x}$ is multivariate Gaussian, then the PC modes will in fact be both mutually uncorrelated and independent. If $\mathbf{x}$ is non-Gaussian, however, PC modes will not generally be mutually independent.

The idealised zonal jet model with fluctuations in both strength and position considered in Section 4 provides an example of a system in which the PC modes are not independent. When the fluctuations in position are taken to be Gaussian, the resulting field is non-Gaussian (as $u(x,t)$ is a nonlinear function of $x_c(t)$). A series of numerically simulated scatterplots of $\alpha_1(t)$ against $\alpha_2(t)$ are presented in Figure 4 for values of the parameter δ (Eqn. 18) between 0.1

and 10. A clear dependence between α_1 and α_2 is evident when δ is not very small (that is, when position fluctuations are sufficiently strong). In particular, strong negative and positive excursions of α_1 are both associated with strong positive excursions of α_2. The colours in Figure 4 correspond to those points for which the jet is anomalously strong ($U > \text{mean}(U)$, red dots) or anomalously weak ($U < \text{mean}(U)$, blue dots) The curvature of the interface between the two regions is another measure of the coupling between these two PC time series. That this coupling of EOF modes is not simply an artifact of this idealised system is illustrated by scatterplots of α_1 against α_2 for 500 hPa zonal-mean zonal winds from the Southern Hemisphere summertime in reanalyses data and from a dry primitive equation GCM (Figure 5). The idealised GCM is characterised by stronger fluctuations of the jet position than are the observations, and the two scatterplots in Figure 5 correspondingly resemble scatterplots in Figure 4 for larger and smaller values of δ, respectively.

Another field in which non-Gaussianity is manifest through statistical dependence of EOF modes is tropical Pacific sea surface temperature (SST), as was discussed in Monahan and Dai (2004). Maps of the leading EOF patters of SST as computed from the Hadley Centre Sea Ice and SST dataset (Rayner et al., 2003) are presented in Figure 6. Also presented are maps of the estimated standard deviation and skewness fields; the latter corresponds to the normalised third-order moment (measuring the asymmetry of a probability distribution around its mean):

$$\text{skew}(a) = \left\langle \left(\frac{a - \langle a \rangle}{\text{std}(a)} \right)^3 \right\rangle, \tag{19}$$

and vanishes if the distribution is Gaussian. Non-zero values of this statistic are therefore a measure of non-Gaussianity. The skewness field illustrated in Figure 6 indicates that the SST probability density tilts towards positive anomalies in the eastern equatorial Pacific and towards negative anomalies in a horseshoe-shaped band from the central subtropical

South Pacific through the western equatorial Pacific back up to the northern subtropics. The leading SST EOF, which carries the most variance, bears a strong resemblance to the standard deviation field. Also notable is the similarity between $\mathbf{e}_2$ and the skewness field, the reason for which becomes evident through an inspection of a scatterplot of α_1 with α_2 (Figure 7). From this plot it is evident that strong positive and negative anomalies of α_1 (corresponding respectively to El Niño and La Niña events) are both associated with strong positive anomalies of α_2. In other words, the second EOF mode makes a positive contribution to the SST field during both extreme phases of ENSO, so on average the strongest positive SST anomalies during El Niño are located further east than the strongest negative SST anomalies during La Niña. This asymmetry in the SST field between the opposing phases of ENSO is then manifest in the skewness field. Again we see a relationship between non-Gaussianity of the field and statistical dependence of the EOF modes.

We therefore see that individual EOFs will not in general correspond even to statistically independent structures. This fact is mitigated by the observation that for many geophysical fields deviations from Gaussianity are not strong. Nevertheless, in considering the results of an EOF analysis, it is worth bearing in mind that as a general rule an EOF decomposition will not generally partition state space into directions that are independent dynamically, kinematically, or statistically.

6 Non-locality of EOFs

We have seen that individual EOF modes are not in general mutually independent either dynamically, kinematically, or statistically: as an even weaker result, we might hope that when a field contains structures that are truly statistically independent then these will be

separated by the EOFs. In fact, this will not the case if the structures are localised or not orthogonal. These results follow from the non-locality of EOFs, which by definition maximise variance over the entire analysis domain, and from the requirement that by construction EOFs be mutually orthogonal.

A simple illustrative example of the general inability of EOFs to separate localised structures was presented in Ambaum et al. (2001). This example considers the three-dimensional vector $\mathbf{x} = (x, y, z)$ such that x and z are of unit variance and are mutually independent, and $y = -x - z$; each of these variables might be thought of as corresponding to, for example, the time series of a teleconnection pattern. It follows that the covariance matrix of this system is

$$C = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}. \tag{20}$$

Only two EOFs are of nonzero variance:

$$\mathbf{e}_1 = (1, -2, 1) \tag{21}$$

$$\mathbf{e}_2 = (-1, 0, 1), \tag{22}$$

carrying respectively 75% and 25% of the total variance. Despite the fact that x and z are independent, they are mixed together in the leading EOF because of their mutual relationships with the third variable y. Because the leading EOF by construction maximises the explained variance over the entire domain, the localised features are combined. As a general rule, independent localised features that share some variance in common with other parts of the field will not be separated by EOF analysis.

In the case of independent but non-orthogonal patterns, the argument is straightforward.

Suppose for example that

$$\mathbf{x}(t) = a_1(t)\mathbf{u}_1 + a_2(t)\mathbf{u}_2 \tag{23}$$

where $\mathbf{u}_1 \cdot \mathbf{u}_2 \neq 0$ and the time series a_1 and a_2 are truly independent. The expansion (23) has the form of a (truncated) EOF expansion, but as the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are not orthogonal, these cannot be the EOFs. A more detailed discussion of this point and its consequences is presented in Dommenget and Latif (2002).

In the end, can we at least hope that the structure of individual EOFs will be entirely determined by the statistical features of the field under consideration? In fact, the answer is no - this will not always be the case. When a field is characterised by spatially homogeneous statistics (that is, invariant from place to place) the EOFs will be strongly influenced by the size and shape of the domain (e.g. Buell, 1975, 1979; Richman, 1986; Dommenget, 2007). For example, if $\mathbf{x}$ represents the values of a field observed at points along a line such that neighbouring points are related by an AR(1) process:

$$x_j(t) = \frac{1}{d}x_{j-1}(t) + \alpha\epsilon_j(t) \tag{24}$$

where ϵ_j are independent and identically distributed random variables then the EOFs of $\mathbf{x}$ are sinusoids regardless of the magnitudes of d or α (Allen and Smith, 1994). This result can be shown to follow from the Wiener-Kinchin theorem, from which we know that the Fourier coefficients of a stationary field are uncorrelated random variables (e.g. Gardiner, 1997; Yaglom, 1961, in which it is also shown that the EOFs of a homogeneous random field on a sphere will be spherical harmonics); the spatial orthogonality of the Fourier modes then ensures that these are EOFs. The scales of the EOF spatial structures will be determined by the size of the domain. In particular, the leading EOF mode will be the gravest mode with a wavelength determined not by the properties of the field but by the size of the domain. If

interpreted as being of individual significance, these EOF modes will impose spatial structure where in fact there is none. Dommenget (2007) uses this fact to suggest a "stochastic null hypothesis" for determining if EOF structures more reflect the variability of the field or the geometry of the domain.

7 Conclusions

The study of atmospheric and oceanic fields (either observed or simulated by a GCM) requires consideration of datasets of very high dimensionality (not uncommonly of order 10^3). Such datasets obviously cannot be visualised in their raw form: some sort of statistical processing is required to reduce the dimensionality of the data to a size accessible to the human mind. This statistical processing imposes a filter on the data, so that the structures that emerge are a convolution of both the data itself *and the statistical technique used for the analysis*. That is, the statistics used leave their imprint on the output, so that the results obtained are influenced by the features of the filter. Empirical Orthogonal Function analysis is a powerful and versatile tool for dimensionality reduction, but it is not free from this "bias".

The examples presented in this review have demonstrated that in general EOF modes cannot be expected to be of individual dynamical, kinematic, or statistical meaning (independent of other EOF modes). These facts follow from the definition of EOF modes as spatially orthogonal, temporally uncorrelated (but not necessarily independent), and non-locally maximising variance over the analysis domain. Different generalisations of EOF analysis have been introduced to relax some of these constraints. A well-established technique which attempts to extract non-orthogonal or localised features from data is so-called "rotated principal component analysis" (e.g. Richman, 1986; Hannachi et al., 2007) in which either

the constraints of spatial or temporal orthogonality (or both) are relaxed. A special case of rotation known as Independent Component Analysis (ICA; Aires et al., 2002) attempts to determine spatial patterns such that the projection time series are statistically independent. Statistical techniques designed to diagnose low-dimensional nonlinear structure in multivariate datasets include Nonlinear Principal Component Analysis (NLPCA; e.g. Monahan et al., 2003), which determines a single globally nonlinear approximation which is optimal in the sense that it minimises the mean-squared error; and methods which approximate the data by a collection of distinct locally linear (but globally nonlinear) approximations (e.g. Horenko, 2008). While all of these methods have their utility for particular problems, they all share with EOF analysis the fundamental limitation that they do not make use of any physical understanding of the system under consideration.

It is possible to incorporate dynamical information into the construction of physical modes. For example, Brunet (1994) and Brunet and Vautard (1996) demonstrate that if the norm used to define the orthogonality of EOF modes is a globally conserved quantity such as wave activity, then (under certain approximations) the resulting Empirical Normal Modes (ENMs) can be expected to be in close correspondence to the eigenvectors of the linearised dynamics. Principal Interaction Patterns (PIPs), which are defined variationally to minimise the norm of the difference in tendencies between the full dynamics and its projection onto a lower-dimensional subspace, have proved useful for model reduction (e.g. Crommelin and Majda, 2004; Kwasniok, 2007). The explicit embedding of dynamics into the reduction techniques makes both ENMs and PIPs attractive from a physical perspective, although their application to observational datasets is nontrivial (e.g. Zadra et al., 2002; Kwasniok, 2007). Finally, somewhere between EOF analysis and diagnostic tools which encode explicit dynamical information lie techniques which make use of the temporal structure of the dataset,

such as Extended EOF analysis, Multichannel Singular Spectrum Analysis (MSSA), and complex EOF analysis (e.g. Hannachi et al., 2007).

As was stated in the Introduction, the purpose of this review is not to denounce the use of EOFs in the analysis of meteorological or oceanographic data. Empirical Orthogonal Function analysis is an extremely useful tool for data compression and dimensionality reduction. Empirical Orthogonal Functions lie along the principal axes of the attractor of the system in state space, and the knowledge of what spatial structures carry the most variance is valuable in guiding the development of physical understanding. For example, the recognition that planetary orbits (Figure 1) are elliptical was a crucial observation in the development of Newtonian mechanics. Furthermore, the fact that, taken together, the leading M EOFs provide an M-dimensional approximation to the attractor of the system which is optimal in terms of the fraction of variance explained can usefully be exploited in the construction of approximate low-dimensional dynamical models (although this too must be done carefully; e.g. Crommelin and Majda, 2004).

Statistical modes reflect the variability of a system driven by the underlying physics, but they will not in general be of individual significance: this is particularly true when the statistical modes are subject to constraints imposed by the method itself. The preceding discussion does not imply that for any particular system individual EOF modes are necessarily not of individual dynamical, kinematic, or statistical meaning. Rather, this analysis demonstrates that such an interpretation cannot *a priori* be expected to be valid, and must be justified by other lines of argument (dynamical, kinematic, or statistical). No statistical tool will ever replace a good mechanistic understanding of a system under consideration. In the analysis of atmospheric and oceanic variability, any statistical analysis must be interpreted with an eye towards structures imposed by both the underlying dynamics and the

statistical tool itself in order to distinguish the medium from the message.

Acknowledgements

AHM acknowledges support from the Natural Sciences and Engineering Research Council of Canada and from the Canadian Institute for Advanced Research Earth System Evolution Program. GRN wishes to thank H. J. Haynes for endowment funds. The authors acknowledge helpful comments from Joel Culina, Slava Kharin, and Nathan Gillett.

Appendix A: Covariance matrix of a stable linear stochastic-dynamical system

Consider a system with dynamics given by Eqn. (14), such that the linear operator A is stable (all eigenvalues have negative real part). Independent of the initial state of the system, a statistical equilibrium will eventually be reached between the fluctuating forcing and the damped response such that the mean of $\mathbf{x}$ is zero and the covariance matrix C satisfies:

$$AC + CA^T = -BB^T \tag{25}$$

(e.g. Penland, 1996). The statistical equilibrium represented by Eqn. (25) allows us to compute the statistics of the response to fluctuating forcing in terms of the deterministic dynamics and the character of the noise.

In the case that A is normal (that is, $AA^T - A^TA = 0$), Eqn. (25) can be solved for the covariance matrix C (Gardiner, 1997). In this case, the eigenvectors of A will form a complete orthogonal set and there will be a unitary matrix $\mathcal{U}$ (for which $\mathcal{U}\mathcal{U}^\dagger = \mathcal{U}^\dagger\mathcal{U} = I$, where I is the $N \times N$ unit matrix) such that

$$A = \mathcal{U}^\dagger \Lambda \mathcal{U}, \tag{26}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_N)$ is a diagonal matrix of the eigenvalues of A. It follows then that

$$(\mathcal{U}A\mathcal{U}^\dagger)(\mathcal{U}C\mathcal{U}^\dagger) + (\mathcal{U}C\mathcal{U}^\dagger)(\mathcal{U}A^\dagger\mathcal{U}^\dagger) = -\mathcal{U}BB^T\mathcal{U}^\dagger \tag{27}$$

and so

$$\sum_k \lambda_i \delta_{ik} F_{kj} + \sum_k F_{ij} \lambda_k^* \delta_{kj} = -(\mathcal{U}BB^T\mathcal{U}^\dagger)_{ij}, \tag{28}$$

where

$$F = \mathcal{U}C\mathcal{U}^\dagger. \tag{29}$$

Solving for C, we obtain

$$C = \mathcal{U}^\dagger F \mathcal{U}, \tag{30}$$

where F is the matrix with components

$$F_{ij} = -\frac{(\mathcal{U} B B^T \mathcal{U}^\dagger)_{ij}}{\lambda_i + \lambda_j^*}. \tag{31}$$

If F is diagonal, then we can conclude that C and A are diagonalised by the same unitary matrix and therefore have the same eigenvectors, in which case the dynamical modes and EOFs coincide. In general, F will not be diagonal unless B has the same eigenvectors as A (a highly unlikely situation) or $BB^T = \beta I$ so that the noise is spatially uncorrelated with spatially uniform variance: that is, if the noise is white in space.

We note that it is often possible to make a change of coordinates from $\mathbf{x}$ to $\mathbf{y}$ such that the dynamics of the new variable is governed by a normal operator (Farrell and Ioannou, 1996). However, this fact is of little help if we are interested in the statistics of the original field $\mathbf{x}$ itself.

Appendix B: Uncorrelated vs. Independent Variables

Two variables a and b are statistically independent only if their joint probability distribution function factors as the product of its marginals: $p(a, b) = p(a)p(b)$, so the state of a is irrelevant to the state of b (and vice versa). Independence implies vanishing correlation, but the reverse implication does not hold in general. This can be demonstrated with a simple example: if a is Gaussian distributed with mean zero and unit variance and $b = a^2$, then $\langle ab \rangle = 0$ so these two variables are uncorrelated. However, as the value of b is determined exactly by that of a these are clearly not independent.

References

Aires, F., W. B. Rossow, and A. Chédin, 2002: Rotation of EOFs by the Independent Component Analysis: Toward a solution of the mixing problem in the decomposition of geophysical time series. *J. Atmos. Sci.*, **59**, 111–123.

Allen, M. R. and L. A. Smith, 1994: Investigating the origins and significance of low-frequency modes of climate variability. *Geophysical Research Letters*, **21**, 883–886.

Ambaum, M. H., B. J. Hoskins, and D. B. Stephenson, 2001: Arctic Oscillation or North Atlantic Oscillation? *J. Climate*, **14**, 3495–3507.

Arfken, G., 1985: *Mathematical Methods for Physicists*. Academic Press, San Diego, 985 pp.

Baldwin, M. P., D. B. Stephenson, and I. T. Jolliffe, in press: Spatial weighting and iterative projection methods for EOFs. *J. Clim.*.

Brunet, G., 1994: Empirical normal-mode analysis of atmospheric data. *J. Atmos. Sci.*, **51**, 932–952.

Brunet, G. and R. Vautard, 1996: Empirical normal modes versus Empirical Orthogonal Functions for statistical prediction. *J. Atmos. Sci.*, **53**, 3468–3489.

Buell, C. E.: 1975, The topography of the empirical orthogonal functions. *Fourth Conf. on Probability and Statistics in Atmospheric Sciences,* Tallahassee, FL, Amer. Meteor. Soc., 188–193.

— 1979, On the physical interpretation of empirical orthogonal functions. *Sixth Conf. on Probability and Statistics in Atmospheric Sciences,* Banff, AB, Canada, Amer. Meteor. Soc., 112–117.

Crommelin, D. T. and A. J. Majda, 2004: Strategies for model reduction: Comparing different optimal bases. *J. Atmos. Sci.*, **61**, 2206–2217.

Dommenget, D., 2007: Evaluating EOF modes against a stochastic null hypothesis. *Clim. Dyn.*, **28**, 517–531.

Dommenget, D. and M. Latif, 2002: A cautionary note on the interpretation of EOFs. *J. Climate*, **15**, 216–225.

Eichelberger, S. J. and D. L. Hartmann, 2007: Zonal jet structure and the leading mode of variability. *J. Climate*, **20**, 5149–5163.

Farrell, B. F. and P. J. Ioannou, 1996: Generalized stability theory. Part I: autonomous operators. *J. Atmos. Sci.*, **53**, 2025–2040.

Fyfe, J. C. and D. J. Lorenz, 2005: Characterizing zonal wind variability: Lessons from a simple GCM. *J. Climate*, **18**, 3400–3404.

Gardiner, C. W., 1997: *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer, 442 pp.

Hannachi, A., I. Joliffe, and D. Stephenson, 2007: Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.*, **27**, 1119–1152; DOI:10.1002/joc.1499.

Horenko, I., 2008: On simultaneous data-based dimension reduction and hidden phase identification. *J. Atmos. Sci.*, **65**, 1941–1954.

Kleeman, R., 2008: Stochastic theories for the irregularity of ENSO. *Phil. Trans. R. Soc. A*, **366**, 2511–2526; doi:10.1098/rsta.2008.0048.

Kwasniok, F., 2007: Reduced atmospheric models using dynamically motivated basis functions. *J. Atmos. Sci.*, **64**, 3452–3474.

Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather prediction. Technical report, Department of Meteorology, MIT, science Report 1.

— 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.

Mo, K. C. and M. Ghil, 1987: Statistics and dynamics of persistent anomalies. *J. Atmos. Sci.*, 877–901.

Monahan, A. H. and A. Dai, 2004: The spatial and temporal structure of ENSO nonlinearity. *J. Climate*, **17**, 3026–3036.

Monahan, A. H. and J. C. Fyfe, 2006: On the nature of zonal jet EOFs. *J. Climate*, **19**, 6409–6424.

— 2008: On annular modes and zonal jets. *J. Climate*, **21**, 1964–1979.

— in press: How generic are zonal jet EOFs? *J. Atmos. Sci.*.

Monahan, A. H., J. C. Fyfe, and L. Pandolfo, 2003: The vertical structure of wintertime climate regimes of the Northern Hemisphere extratropical atmosphere. *J. Climate*, **16**, 2005–2021.

North, G. R., 1984: Empirical orthogonal functions and normal modes. *J. Atmos. Sci.*, **41**, 879–887.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling errors in the estimation of Empirical Orthogonal Functions. *Mon. Weath. Rev.*, **110**, 699–706.

Penland, C., 1996: A stochastic model of IndoPacific sea surface temperature anomalies. *Physica D*, **98**, 534–558.

Rayner, N., D. Parker, E. Horton, C. Folland, L. Alexander, D. Rowell, E. Kent, and A. Kaplan, 2003: Global analyses of SST, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 10.1029/2002JD002670.

Richman, M. B., 1986: Rotation of principal components. *Int. J. Climatology*, **6**, 293–335.

von Storch, H. and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge.

Yaglom, A.: 1961, Second-order homogeneous random fields. *Proc. 4th Berkeley Symp. Math. Stat. Probab.*, Univ. California Press, volume 2, 593–622.

Zadra, A., G. Brunet, and J. Derome, 2002: An empirical normal mode diagnostic algorithm applied to NCEP renalyses. *J. Atmos. Sci.*, **59**, 2811–2829.

Zanna, L. and E. Tziperman, 2008: Optimal surface excitation of the thermohaline circulation. *J. Phys. Oceanogr.*, **38**, 1820–1830.

Figure Captions

Figure 1: First and second EOFs ($\mathbf{e}_1$ and $\mathbf{e}_2$ respectively) of an elliptical planetary orbit. Redrafted following Mo and Ghil (1987).

Figure 2: Lorenz (1963) attractor $\mathbf{x}(t)$ for standard parameters producing a strange attractor. The coloured lines are the projections of $\mathbf{x}(t)$ onto the three EOF modes: $\mathbf{p}_j(t) = (\mathbf{x}(t) \cdot \mathbf{e}_j)\mathbf{e}_j$. Redrafted following Mo and Ghil (1987).

Figure 3: Normalised basis vectors $f_1(x)$, $f_2(x)$, $f_3(x)$ (respectively the monopole, dipole, and tripole) from which are constructed the EOFs of the idealised jet in zonal-mean zonal wind.

Figure 4: Scatterplots of numerically calculated $\alpha_1(t)/U_0$ vs $\alpha_2(t)/U_0$ for $h = 0.26$ and $\delta = 0.1$, 0.25, 1, 2.5, and 10. Blue dots denote those points for which $U < U_0$ and red dots denote those points for which $U > U_0$.

Figure 5: As in Figure 4 for 500 hPa zonal-mean zonal wind data from Southern Hemisphere summertime reanalyses (left panel) and from a dry primitive equation GCM. Adapted from Fyfe and Lorenz (2005).

Figure 6: Maps of the leading two EOF patterns of tropical Pacific SST, along with the standard deviation and skewness (Eqn. 19) fields. Redrawn from Monahan and Dai (2004).

Figure 7: Scatterplot of leading two PC modes of tropical Pacific SST (each normalised to unit variance). Redrawn from Monahan and Dai (2004).

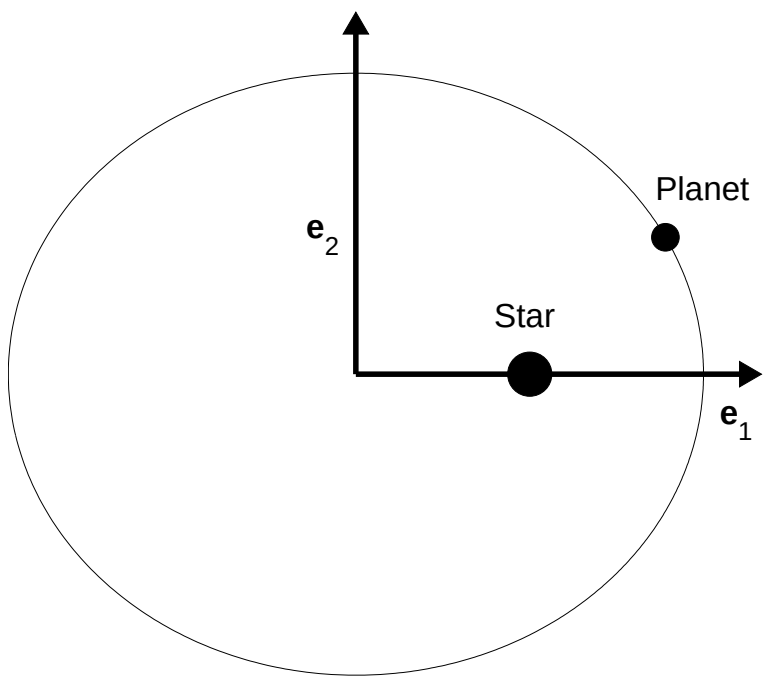

Figure 1: First and second EOFs ($\mathbf{e}_1$ and $\mathbf{e}_2$ respectively) of an elliptical planetary orbit. Redrafted following Mo and Ghil (1987).

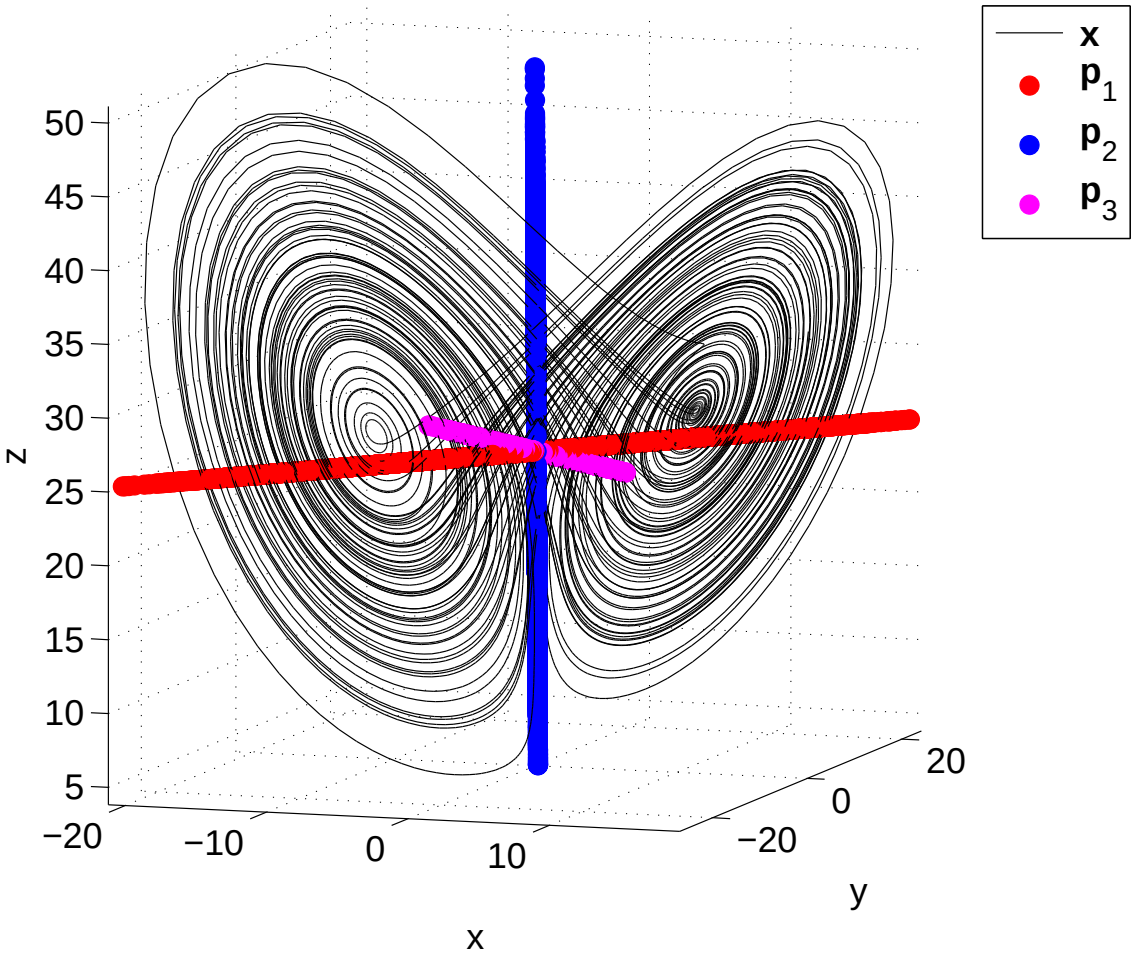

Figure 2: Lorenz (1963) attractor $\mathbf{x}(t)$ for standard parameters producing a strange attractor. The coloured lines are the projections of $\mathbf{x}(t)$ onto the three EOF modes: $\mathbf{p}_j(t) = (\mathbf{x}(t) \cdot \mathbf{e}_j)\mathbf{e}_j$. Redrafted following Mo and Ghil (1987)

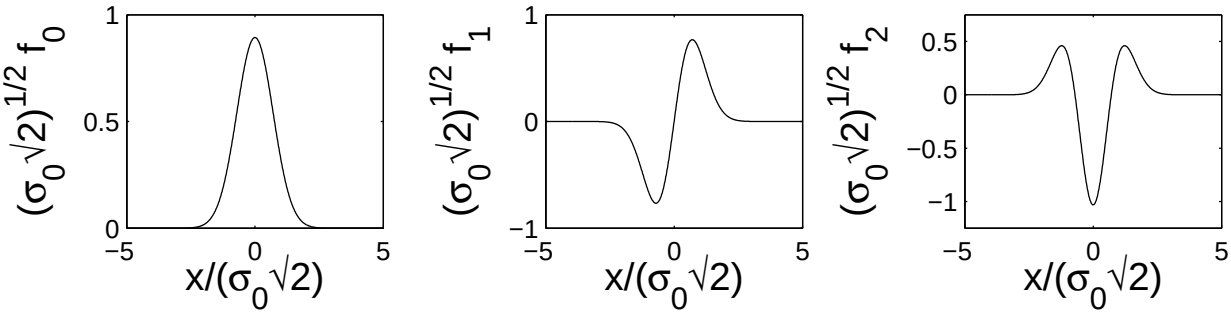

Figure 3: Normalised basis vectors $f_1(x)$, $f_2(x)$, $f_3(x)$ (respectively the monopole, dipole, and tripole) from which are constructed the EOFs of the idealised jet in zonal-mean zonal wind.

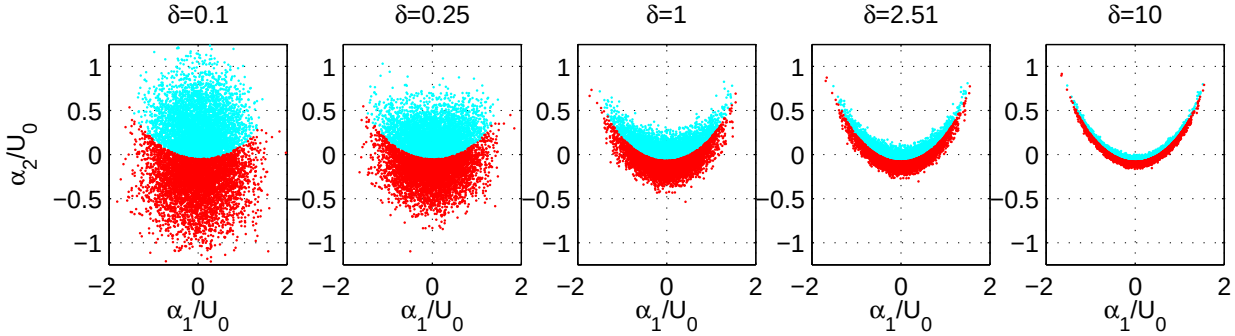

Figure 4: Scatterplots of numerically calculated $\alpha_1(t)/U_0$ vs $\alpha_2(t)/U_0$ for $h = 0.26$ and $\delta = 0.1$, 0.25, 1, 2.5, and 10. Blue dots denote those points for which $U < U_0$ and red dots denote those points for which $U > U_0$.

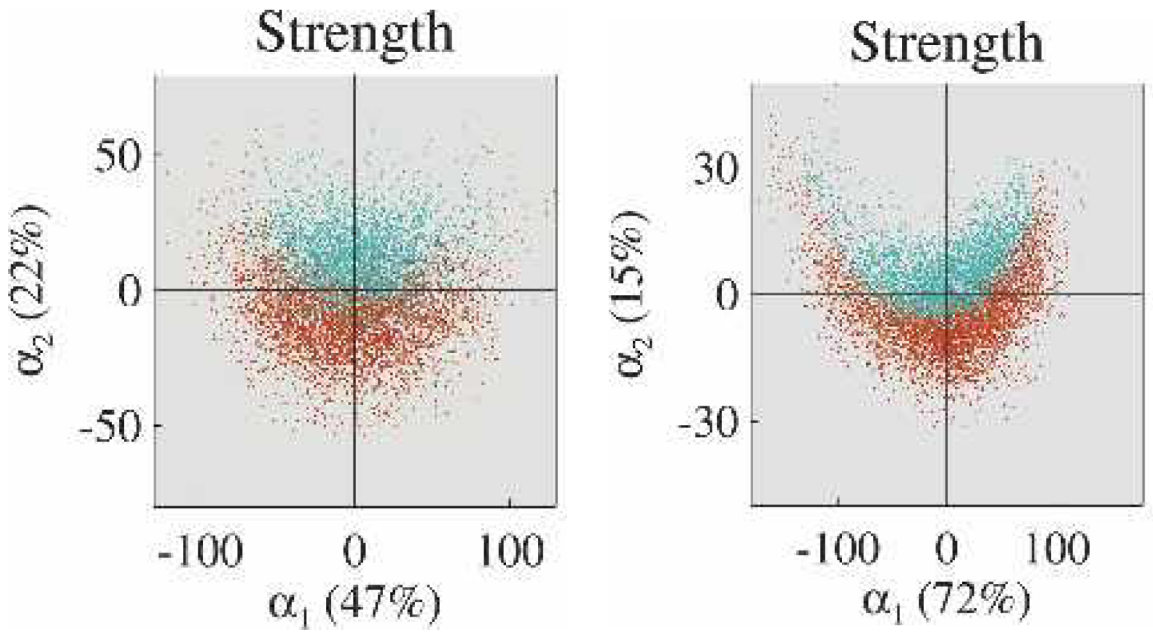

Figure 5: As in Figure 4 for 500 hPa zonal-mean zonal wind data from Southern Hemisphere summertime reanalyses (left panel) and from a dry primitive equation GCM. Adapted from Fyfe and Lorenz (2005).

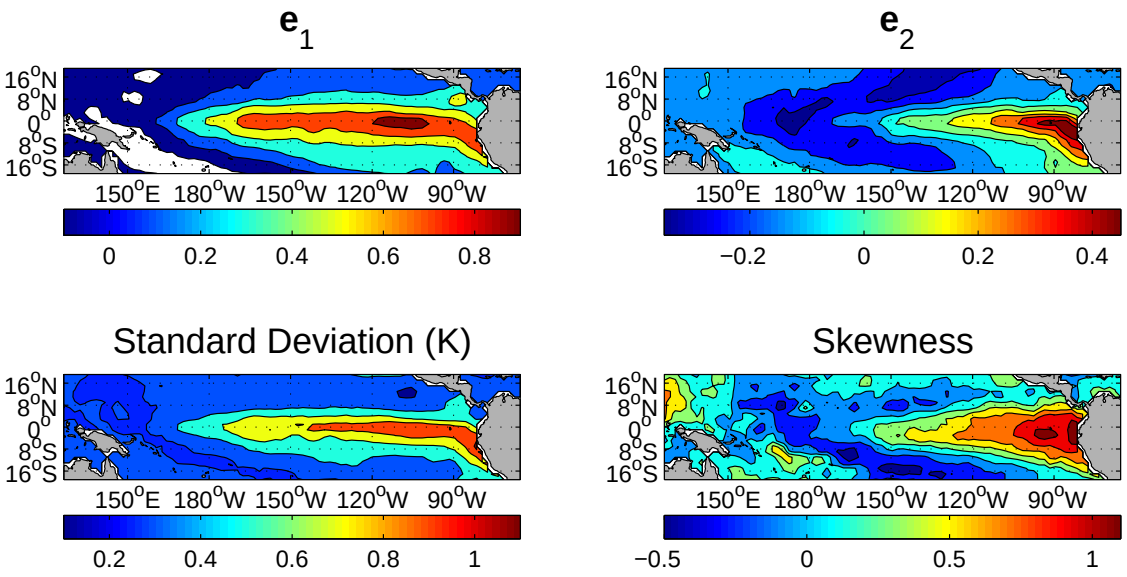

Figure 6: Maps of the leading two EOF patterns of tropical Pacific SST, along with the standard deviation and skewness (Eqn. 19) fields. Redrawn from Monahan and Dai (2004).

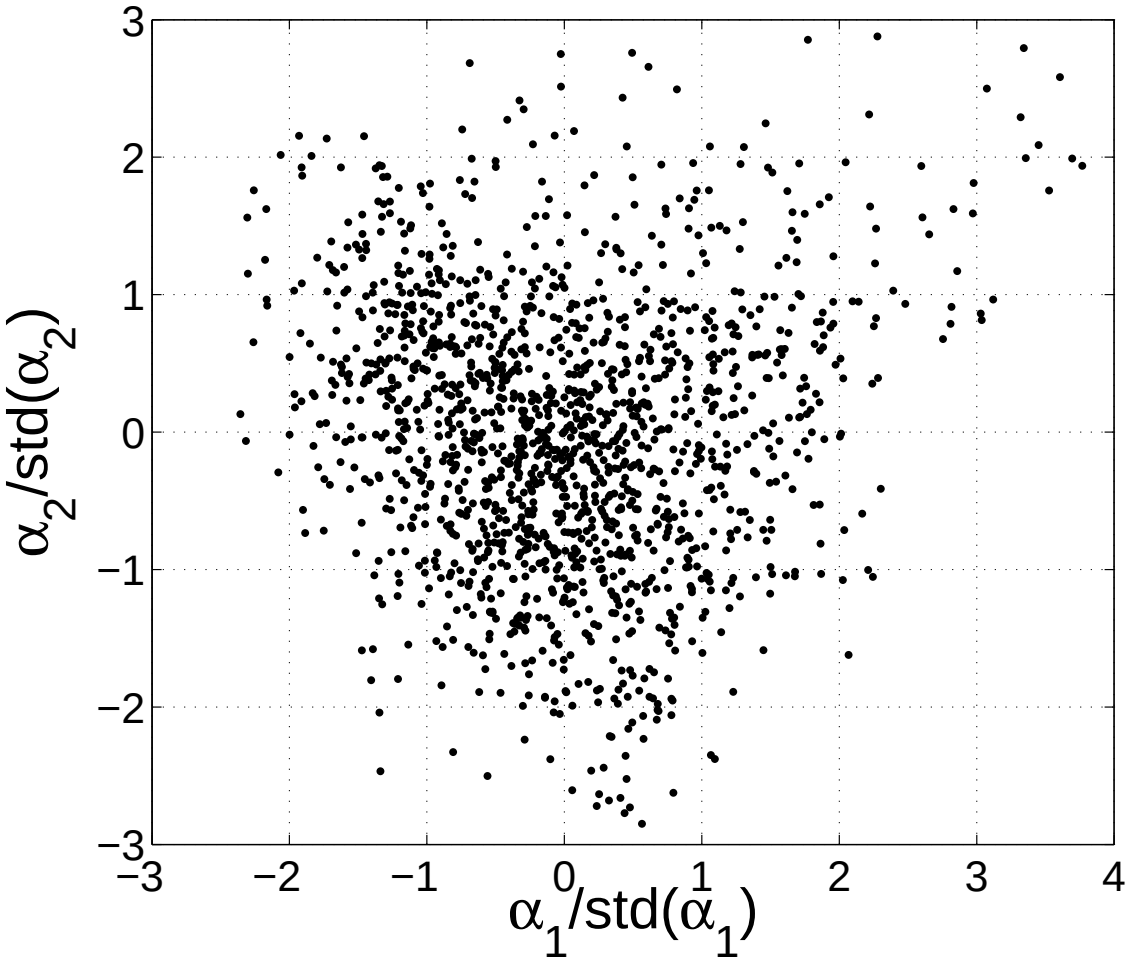

Figure 7: Scatterplot of leading two PC modes of tropical Pacific SST (each normalised to unit variance). Redrawn from Monahan and Dai (2004).