Tellus (2001), 53A, 469–480

TELLUS
ISSN 0280–6495

Identifying signals from intermittent low-frequency behaving systems

By A. HANNACHI* and M. R. ALLEN, *Atmospheric, Oceanic and Planetary Physics, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK*

(Manuscript received 17 February 2000; in final form 18 December 2000)

ABSTRACT

An approach is presented to recover a true signal from an intermittent regime-like behaving noisy system given an estimate of the noise covariance $\mathbf{C}_N$. The method is based on minimizing the noise probability distribution on isosurfaces of the data probability distribution and involves the spectrum of the matrix $\mathbf{C}_D\mathbf{C}_N^{-1}$ where $\mathbf{C}_D$ represents an estimate of the data covariance within a given regime. In single channel, i.e., one-dimensional, case the intermittent low-frequency time series is split according to the regime level. For each regime the obtained sub-sampled time series is first centred by removing its time mean and then projected onto a higher dimensional space using the method of the delay coordinates. The oscillation for each regime level is then obtained as the leading eigenvector of $\mathbf{C}_D\mathbf{C}_N^{-1}$. In the multi-channel case the signal pattern is obtained in the same way locally in each cluster forming the data. The approach is first tested with the Lorenz system yielding the correct oscillation. The method is then applied to the Pacific 500-hPa geopotential height from a set of multidecadal integrations with the General Circulation Model (GCM) of the Hadley Centre, forced with observed Sea Surface Temperature (SST), in order to identify the nonlinear atmospheric response associated with El Niño Southern Oscillation (ENSO).

1. Introduction

Several techniques can be found in the literature to identify regular signals from noisy data such as maximum entropy method (Ulrych and Bishop, 1975), singular spectrum analysis, SSA, (Fraedrich, 1986; Broomhead and King, 1986a,b; Vautard et al., 1992), and signal-to-noise (S/N) ratio maximizing approach (Allen and Smith, 1997). Most of these techniques, however, attempt to identify oscillations or regular signals embedded in data that behave "nicely". For instance it can be shown that SSA, for example, is justified only if the hypothetical noise in the dataset is white (Allen and Smith, 1997), but that it can fail to identify genuine and detectable oscillations in the presence of coloured noise. Allen and Smith (1997) proposed a refinement based on maximizing the S/N ratio. However, the S/N ratio itself fails when, for example, the data present a non-Gaussian, such as regime, behaviour (Hannachi 2000).

Now, many physical and natural phenomena, such as the climate system, display a complicated behaviour involving most of the time high- and intermittent low-frequency variability leading to some sort of regime behaviour (Lorenz, 1963; Namias, 1964; Charney and DeVore, 1979, Ghil and Childress, 1987; Palmer, 1999; Corti et al., 1999). The latter feature necessarily implies that the system is expected to display a non-Gaussian behaviour from which one wishes to extract a true signal.

In this manuscript we attempt to recover a true

* Corresponding author. Department of Meteorology, University of Reading, Earley Gate PO Box 243, Reading RG6 6BB, UK.
email: han@met.rdg.ac.uk

high frequency signal, such as oscillations in single channel, embedded in an intermittent low-frequency, or non-Gaussian, behaving system given an estimate of the noise covariance. The knowledge of the noise covariance is, in general, a strong hypothesis. In most climate studies, however, this hypothesis can be met due to the use of climate models that can provide the noise estimate. For instance when we perform an ensemble of integrations of a climate model with the same boundary, but with different initial, conditions the signal from the obtained ensemble can be understood as the reproducible component whereas the noise can be represented by the scatter among the different members. This question is addressed in detail in Section 4.

The manuscript is organised as follows. In Section 2, we briefly introduce the approach for a Gaussian system then we extend it to the non-Gaussian case (theoretical details on the approach can be found in Hannachi (2000)). In Section 3, we apply the approach to the Lorenz (1963) system. The application to the 500-hPa geopotential height from the General Circulation Model (GCM) is presented in Section 4. Summary and conclusion are presented in the last Section.

2. Signal detection from a regime-like behaving system

Let us suppose that we are in the presence of a noisy data embedding a signal that we would like to recover given an estimate of the noise covariance. The data can be either unidimensional, in which case the signal is simply an oscillation, or multi-dimensional (multi-channel), in which case the signal is a pattern (Section 4). In the single-channel case the time series $(x_i)_{1\leqslant i\leqslant n} = (x(t_i))_{1\leqslant i\leqslant n}$ can be projected onto an m-dimensional space using the delay coordinates $(\mathbf{X}_i)_{1\leqslant i\leqslant n-m+1}$:

$$\mathbf{X}_i = (x_i, \ldots, x_{i+m-1}) \quad \text{for } i \leqslant 1 \leqslant n-m+1. \qquad (1)$$

(Takens, 1981) to yield a multi-channel data whose covariance matrices can be calculated. Next, we designate by $\mathbf{C}_\mathrm{D}$ and $\mathbf{C}_\mathrm{N}$ the respective estimates of the covariance matrices of the data and the noise, and we propose to find the signal embedded in the noisy data given $\mathbf{C}_\mathrm{D}$ and $\mathbf{C}_\mathrm{N}$.

To solve this problem, let us first suppose that the data, and also the noise, can be approximated by Gaussian distributions with respective probability distribution functions (p.d.fs.) p_D and p_N given by:

$$p_\mathrm{D}(\mathbf{x}) = (2\pi)^{-m/2}|\mathbf{C}_\mathrm{D}|^{-1/2}\exp[-\tfrac{1}{2}\mathbf{x}^\mathrm{T}\mathbf{C}_\mathrm{D}^{-1}\mathbf{x}], \qquad (2)$$

$$p_\mathrm{N}(\mathbf{x}) = (2\pi)^{-m/2}|\mathbf{C}_\mathrm{N}|^{-1/2}\exp[-\tfrac{1}{2}\mathbf{x}^\mathrm{T}\mathbf{C}_\mathrm{N}^{-1}\mathbf{x}], \qquad (3)$$

where $|C_\mathrm{D}|$ stands for the determinant of C_D. Naturally the signal orientation is more likely to be easily captured in regions of the state space where p_N is at a minimum. Since the data contain the signal (and also the noise) and because we do not want any bias in the estimation, we choose equiprobable subsets of the data for this operation. Therefore to optimally filter the signal out, we define this latter as being the pattern $\mathbf{e}$ that minimizes p_N on iso-surfaces of p_D (see Hannachi (2000) for more details). Note that the same result is obtained by maximizing p_D on isosurfaces of p_N. This definition of identifying the signal can be seen as a surrogate to simultaneously minimizing p_N and maximizing p_D. Since the iso-surfaces of p_D are defined by constant values of the quadratic form $\mathbf{x}^\mathrm{T}\mathbf{C}_\mathrm{D}^{-1}\mathbf{x}$, i.e., $\mathbf{x}^\mathrm{T}\mathbf{C}_\mathrm{D}^{-1}\mathbf{x} = \alpha$ (where α is a constant), the signal pattern is obtained by solving:

$$\max \mathbf{x}^\mathrm{T}\mathbf{C}_\mathrm{N}^{-1}\mathbf{x} - \lambda(\mathbf{x}^\mathrm{T}\mathbf{C}_\mathrm{D}^{-1}\mathbf{x} - \alpha), \qquad (4)$$

where λ is a Lagrange multiplier. Eq. (4) eventually leads to the generalized eigenproblem:

$$\mathbf{C}_\mathrm{D}\mathbf{C}_\mathrm{N}^{-1}\mathbf{x} = \lambda\mathbf{x}. \qquad (5)$$

Therefore the signal pattern $\mathbf{e}$ is provided by the eigenvector corresponding to the leading eigenvalue of the *detection* matrix $\mathbf{C}_\mathrm{D}\mathbf{C}_\mathrm{N}^{-1}$ (Hannachi, 2000). This formulation via the p.d.fs. is useful because it can be extended to the case where the data, for example, displays a non-Gaussian behaviour as presented below.

Now we suppose that our system intermittently resides in two (or more) different states (or regimes) and therefore the signal can be regime-dependent. In each regime, the system can well be approximated by a Gaussian p.d.f. and hence the resulting p.d.f. of the system can be approximated by a mixture of two (or more) Gaussians (Everitt and Hand, 1981). For instance, for two regimes we have:

$$\begin{aligned}
&(2\pi)^{m/2}p_\mathrm{D}(\mathbf{x}) \\
&= \alpha_1|\mathbf{C}_{\mathrm{D}_1}|^{-1/2} \\
&\quad\times \exp[-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_1)^\mathrm{T}\mathbf{C}_{\mathrm{D}_1}^{-1}(\mathbf{x}-\boldsymbol{\mu}_1)] \\
&\quad+ \alpha_2|\mathbf{C}_{\mathrm{D}_2}|^{-1/2} \\
&\quad\times \exp[-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_2)^\mathrm{T}\mathbf{C}_{\mathrm{D}_2}^{-1}(\mathbf{x}-\boldsymbol{\mu}_2)], \qquad (6)
\end{aligned}$$

where $\boldsymbol{\mu}_i$ and $\mathbf{C}_{D_i}$ $(i=1,2)$ are the centre and the covariance matrix of the ith regime respectively. In eq. (6), α_1 is the proportion of the data in the first cluster $(\alpha_1+\alpha_2=1)$.

The previous decomposition into Gaussians is in fact made possible via the fundamental result that any p.d.f. can be approximated as closely as desired by a finite mixture of Gaussians (Anderson and Moore, 1979). The application of the full mixture algorithm to find Gaussian components from data is beyond the scope of this paper and we can only refer the reader, for example, to Everitt and Hand (1981) for details or to Haines and Hannachi (1995) and Hannachi (1997) who applied the mixture model to find regimes from a GCM. In this manuscript we adopt a rather simple approach to get the regimes (see below). Now it can be shown, after some approximations, that for more general p.d.f. p_D given by (6) the optimization problem (4) can be transformed to yield an optimization over each Gaussian component of the p.d.f. p_D (6) separately (Hannachi, 2000). Therefore the signal $\mathbf{e}_i$, i.e., the oscillation or the signal pattern, in the ith regime emerges as the eigenvector corresponding to the leading eigenvalue of the ith detection matrix $\mathbf{C}_{D_i}\mathbf{C}_N^{-1}$.

3. Application to the Lorenz system

We consider here the classical Lorenz system (Lorenz, 1963) which exhibits high frequency oscillations with intermittent low frequency, or regime, behaviour. The Lorenz system has been extensively used in studies of climate dynamics (Marshall and Molteni, 1993; Molteni et al., 1993, Palmer, 1999; and Corti et al., 1999). Palmer (1999), for example, used the same system as a paradigm to interpret the recent climate change in terms of changes in the frequency of the ocurrence of natural atmospheric circulation regimes. Here we would like to identify the high frequency oscillation in the system.

The model equations for the Lorenz system are:

$$\begin{aligned}\dot{x}&=-\sigma(x-y),\\ \dot{y}&=-xz+rx-y,\\ \dot{z}&=xy-bz.\end{aligned}\tag{7}$$

The parameters are, as in the original paper, $\sigma=10$, $r=30$, and $b=8/3$. System (7) is integrated forward in time using an Euler scheme with time step $dt=10^{-2}$. Here we use the time series $x(t)$, that we contaminate by a red-noise, to address the question of identifying the oscillation frequency in the system. In Fig. 1 we show the variable $x(t)$ (1a) and the contaminated time series $x'(t)$ (1b). The time series $x'(t)$ is projected onto an m-dimensional space using the method of the delay coordinates (Takens, 1981) with m fixed to 30 and the covariance matrices $\mathbf{C}_D$ and $\mathbf{C}_N$ of the (contaminated) time series and the noise (in delay coordinates) respectively are then computed. Fig. 2 shows the spectrum of $\mathbf{C}_D\mathbf{C}_N^{-1}$ (2a) along with the

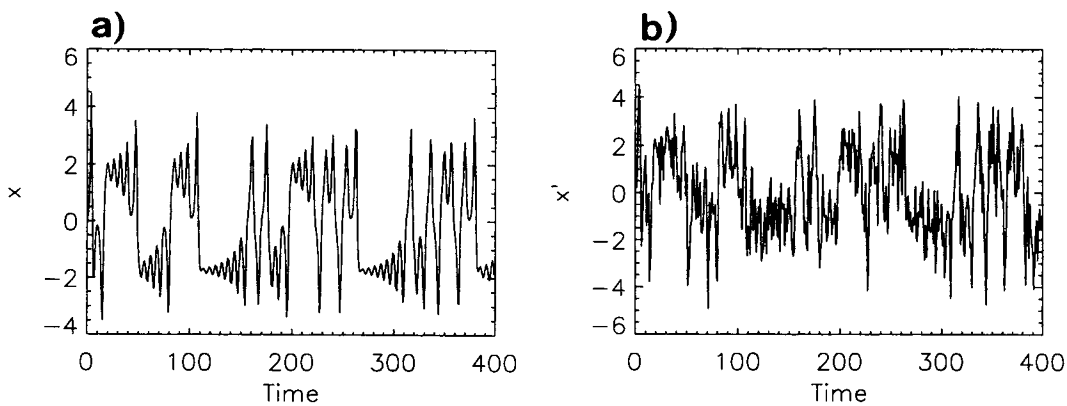

Fig. 1. The time series $x(t)$ of the Lorenz system sampled every tenth time step (a), and the contaminated test series, $x'(t)$, obtained by adding a red-noise to the (resampled and rescaled) trajectory $x(t)$ (b). An Euler numerical scheme with time step $dt=10^{-2}$ is used for the time integration.

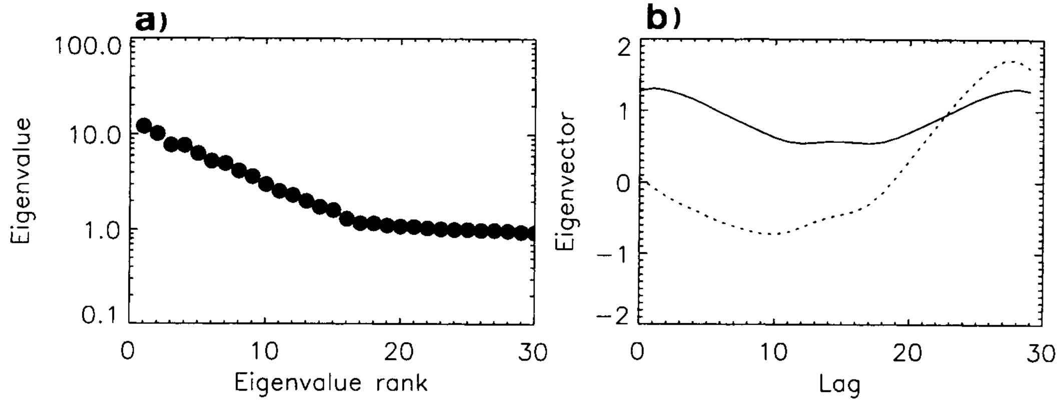

Fig. 2. The eigenspectrum of the detection matrix $\mathbf{C}_D\mathbf{C}_N^{-1}$ in decreasing order corresponding to the Lorenz system shown in Fig. 1b (a), and the eigenvectors corresponding to the leading two eigenvalues (b).

eigenvectors corresponding to the two leading eigenvalues (2b). As Fig. 2 shows, the procedure fails to identify the oscillation frequency in the model because we did not take into consideration the regime, or the intermittent low-frequency, behaviour of the system.

Now to take the previous intermittent low-frequency behaviour into account we split the contaminated time series $x'(t)$ simply by considering the subsample corresponding to the part of the time series where $x'(t) \geqslant 0$ which is obtained by concatenating together the different segments of $x'(t)$ above zero, to yield a new time series whose covariance matrix is computed in the usual way. This is just a simplified way to discriminate between the regimes of our test series. The subsampled time series is first centred by removing its time mean since we are interested in the variability about its regime center. The covariance matrix $\mathbf{C}_{D_1}$ of the obtained subsample is then computed and the spectrum of $\mathbf{C}_{D_1}\mathbf{C}_N^{-1}$ obtained. This spectrum is displayed in Fig. 3a showing a pair of leading equal eigenvalues well separated from the rest of the spectrum indicating therefore the existence of an oscillation contained in the time series which can be revealed by the eigenvectors. In fact, the eigenvectors corresponding to the leading pair of eigenvalues are displayed in Fig. 3b. They clearly show a pair of sine waves in quadrature whose oscillation frequency represents the oscillation contained in the time series $x'(t)$.

The same oscillation is also obtained from the other subsample corresponding to the part of the time series below zero (not shown) since the system contains the same oscillation in both regimes. However, we carried out similar experiments (not shown) with generated test series where the signal has two different periods (one period over each regime) and switches randomly between two regime centers, i.e., the regime-dependent signal case. After the system was noise-contaminated, the same procedure was applied as in the Lorenz system and the true signals recovered.

The previous result is found to be quite robust to reasonable changes in the level of split (Hannachi, 2000). For instance, we found that we do not need to cut the time series through the origin but we can use any value between -1 and 1 and we still obtain the same signal. The previous result is concerned with oscillations embedded in noisy unidimensional time series. The same approach can also be applied to multichannel cases as presented next with data taken from an atmospheric General Circulation Model (GCM).

4. Application to an atmospheric GCM data

We consider in this Section the application of the approach presented above to data from a GCM. Precisely we would like to show how the previous approach can be used to address the question related to the nonlinear atmospheric response associated with El Niño Southern Oscillation (ENSO) using state-of-the-art GCMs. We use here the Hadley Centre atmospheric

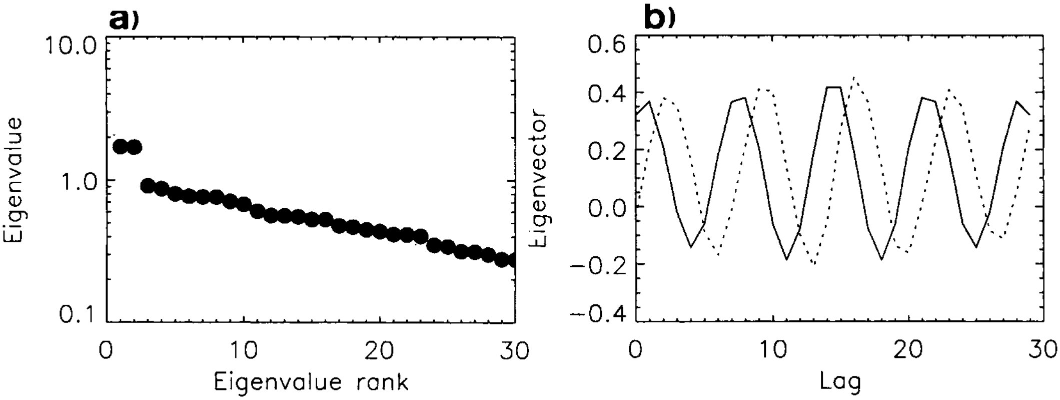

Fig. 3. (a) As in Fig. 2a but for $\mathbf{C}_{\mathrm{D}_1}\mathbf{C}_{\mathrm{N}}^{-1}$ corresponding to the part of the time series $x'(t)$ above the origin (note the leading pair of eigenvalues corresponding to the oscillation), and (b) the eigenvectors corresponding to the leading pair of eigenvalues showing the right oscillation.

model, HadAM1, of the UK Meteorological Office. It is a grid point model with 3.75° × 2.5° horizontal resolution and 19 vertical levels. It has a gravity wave drag and radiation schemes with radiative fluxes computed in 4 long-wave and 6 short-wave bands, a penetrative convection scheme and a land surface scheme. Details on the model can be found for example in Davies et al. (1997).

We consider a set of 6 members of multidecadal integrations of HadAM1 GCM forced by observed Sea Surface Temperature (SST) and integrated for 45 years from October 1948 to December 1993 with different initial condition for each member. In this study we focus attention on the inter-annual winter 500-hPa geopotential height field over the Pacific sector (2.5°N–75°N, 123.5°E–60°W). The winter fields are obtained by averaging the December through to February (DJF) anomalies. A detailed analysis of the application of the same approach to the North Pacific and the North Atlantic geopotential heights from the same model is presented in Hannachi (2001), and for convenience, we choose the North Pacific data to illustrate the theory presented in Section 2.

Since the different members have the same boundary condition the low frequency signal from the ensemble can be understood as the reproducible component whereas the scatter among the different members represents the noise. The noise fields are therefore estimated as the combined residuals obtained by subtracting the ensemble mean from each ensemble member. In Fig. 4 we display the first two Empirical Orthogonal Functions (EOFs) of the noise fields. The first mode (4a) shows a north-south seesaw and explains 28% of the inter-annual variability whereas the second mode (4b) shows a Pacific–North American (PNA) pattern with 20% explained variability. These two modes are not degenerate as indicated by their respective standard errors (not shown).

In Fig. 5 we show the first two EOFs of the ensemble mean. The first mode (5a) shows the PNA pattern with 47% explained variance whereas the second mode (5b) shows a pattern reminiscent of the North Pacific Oscillation (NPO) similar to the first mode of the internal noise (4a) and explains 20% of the interannual variance. The ensemble mean is generally accepted to represent an approximation of the model response to the forcing pattern. This is particularly true, at least theoretically, for an infinite sample size, but for a finite sample the ensemble mean is always contaminated by the internal variability noise and we therefore make use of the previous approach to filter the signal out.

In Fig. 6a we show the signal pattern, indicating a PNA response, obtained as the leading eigenvector of $\mathbf{C}_{\mathrm{D}}\mathbf{C}_{\mathrm{N}}^{-1}$ where $\mathbf{C}_{\mathrm{D}}$ and $\mathbf{C}_{\mathrm{N}}$ represent respectively estimates of the covariance matrices of the ensemble mean and the internal variability noise. Now since the operator $\mathbf{C}_{\mathrm{D}}\mathbf{C}_{\mathrm{N}}^{-1}$ is in general

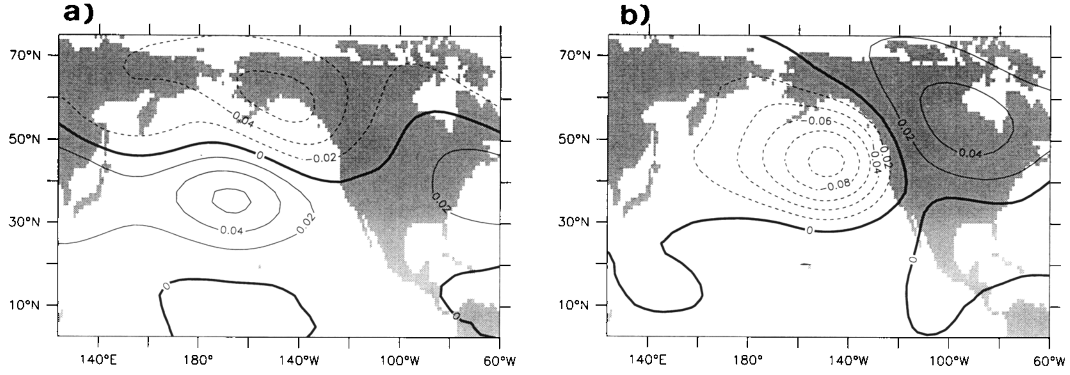

Fig. 4. The first (a) and the second (b) North Pacific EOFs of the internal variability noise of the winter geopotential height at 500-hPa.

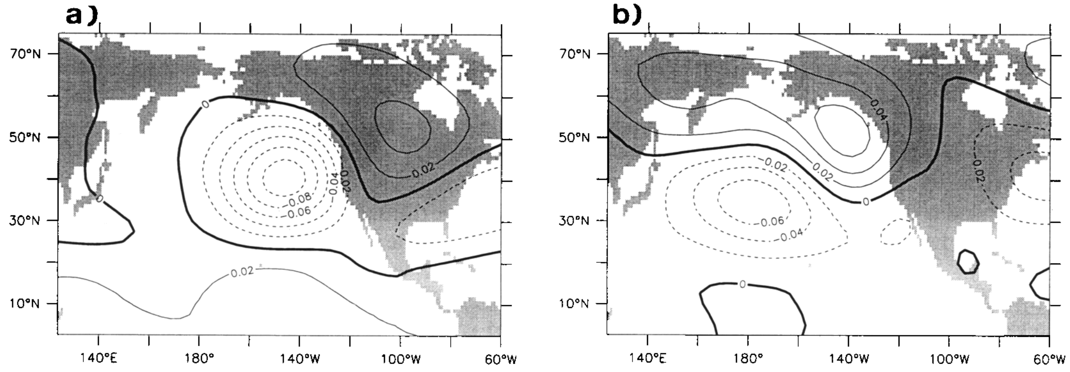

Fig. 5. As in Fig. 4 but for the ensemble mean winter geopotential height.

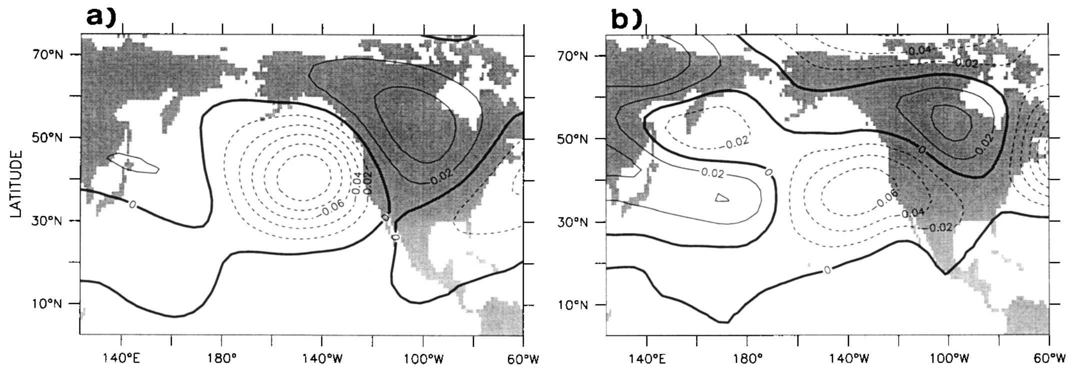

Fig. 6. The response pattern obtained as the eigenvector corresponding to the leading eigenvalue of the detection matrix $\mathbf{C_D C_N^{-1}}$ (a), and the adjoint mode (b) estimated from the winter 500-hPa geopotential height of the ensemble mean and the noise fields.

nonsymmetric its eigenvectors and those of its adjoint are different. Precisely, the eigenvectors of the adjoint operator represent the weight patterns as in Thacker (1996) or optimal filters as in Venzke et al. (1999). The leading adjoint mode (Fig. 6b) indicates the pattern onto which the different members of the ensemble have a (nearly) common projection corresponding to a common response (Fig. 6a). The adjoint mode can be shown to maximize the S/N ratio, $(\mathbf{u}^T\mathbf{C}_D\mathbf{u})/(\mathbf{u}^T\mathbf{C}_N\mathbf{u})$, and therefore plays the role of a filter and is not to be interpreted like the response pattern because it is in a "different" space. In fact, as shown in Hannachi (2000, 2001), the adjoint mode and the response pattern are in duality.

In Fig. 7, we display, in fact, the time series of the projection of each ensemble member onto the leading weight pattern (7a) as well as the ENSO index (7b) obtained by averaging the winter SST anomaly over the NINO3 region in the east tropical Pacific (5°S–5°N, 150°W–90°W). Note in particular that the two plots (Figs 7a, b) are very similar with a correlation of the order of 0.9. Now the first panel (7a) clearly indicates that the response pattern (Fig. 6) is the common signal to all members, while the second (7b), and also the first (7a), panels along with Fig. 6 clearly indicate a −PNA pattern associated with positive phase of ENSO (La Niña) and a +PNA associated with opposite phase (El Niño). This analysis clearly indicates a linear response associated with the phase of ENSO. Now, given the relationship between the nonlinear nature associated with the thermodynamic control on deep convection over the tropical Pacific and the teleconnection response over the northern midlatitudes, the linear response clearly cannot explain this nonlinearity.

In order to address the question related to the nonlinear response we resort to the analysis presented in Section 2. It is well documented that atmospheric low-frequency variability can display regime, or intransitive (Lorenz, 1968), behaviour (Lorenz, 1963; Namias 1964; Leith, 1973; Charney and DeVore 1979; Wallace and Gutzler, 1981; and Cheng and Wallace, 1993, etc.). The regime behaviour is very important in climate dynamics. For instance there is a trend among the scientific community who interpret the climate in terms of a dynamical system switching between nonlinear circulation regimes and that external forcing can only act to change the residence time in these same regimes (Palmer, 1999; Corti et al., 1999; and Hasselmann, 1999). This necessarily leads to a non-Gaussian behaviour where the mixture analysis presented in Section 2 can be applied, see for example Haines and Hannachi (1995) and Hannachi (1997) for an application of the mixture model to data from a GCM.

The probability distribution of the ensemble mean geopotential heights within the first two principal components (PCs) is displayed in Fig. 8a showing a signature of bimodality. This p.d.f. has been computed using the kernel method (Silverman, 1994). A mixture of two Gaussians has also been fitted to the same data (Fig. 8b). The corresponding centres are indicated by small

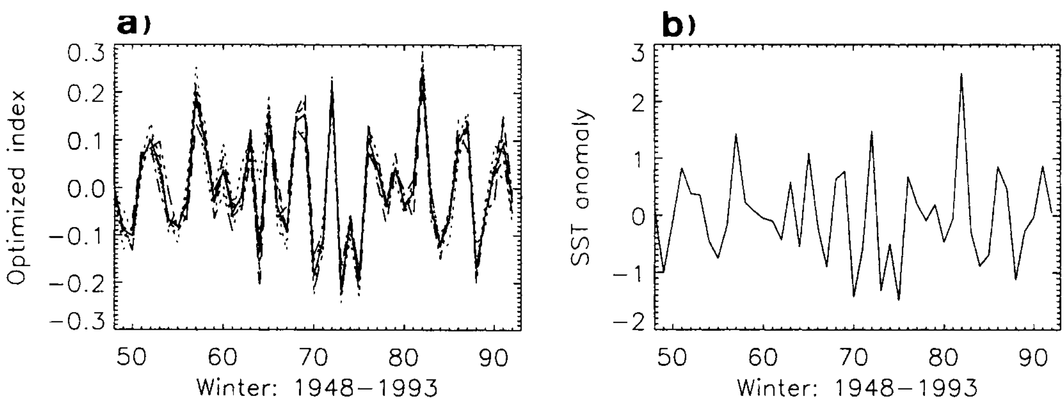

Fig. 7. The projection of the winter geopotential height fields of each ensemble member onto the leading weight pattern (a), and the winter ENSO index described in the text (b). Units: 1500 m (a), and °C (b).

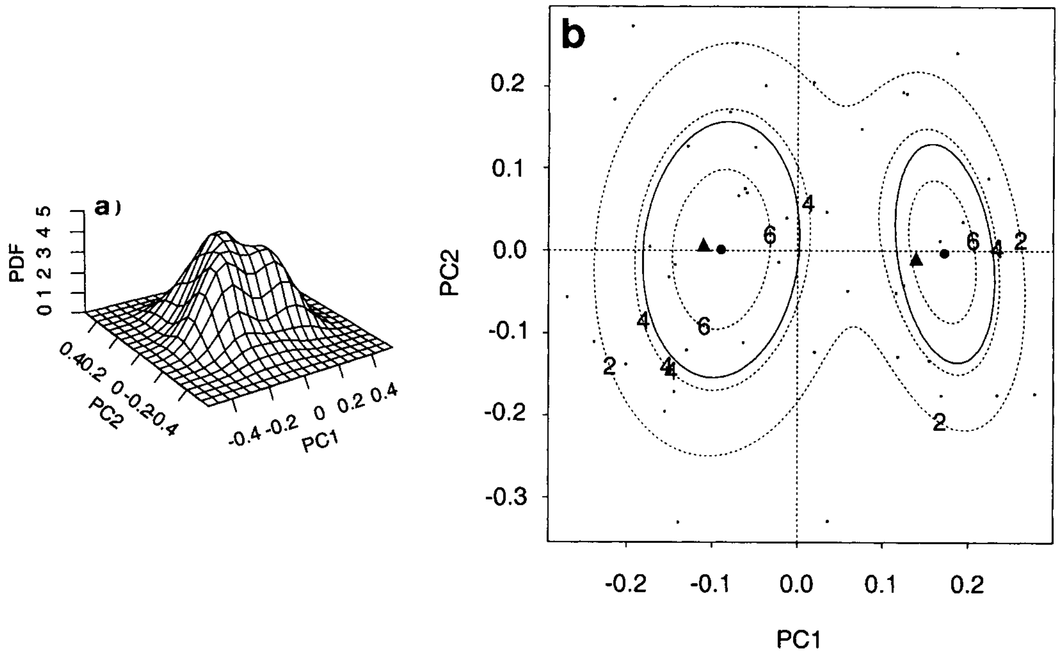

Fig. 8. (a) The probability distribution of the ensemble mean heights within the first two principal components. (b) The best fit of a mixture of two Gaussians to the ensemble mean (dotted contours) along with the covariance ellipse of each Gaussian (solid contour) whose centres are shown as small filled circles. The filled triangles in (b) represent the centres of the data obtained by splitting the ensemble mean according to the sign of the ENSO index.

filled circles while the regime centres obtained by splitting the ensemble mean according to the sign of the ENSO index are shown as filled triangles. Note the alignment of these centres along the PC1 axis (8a,b) representing respectively $\pm$PNA.

Although Fig. 8a shows signature of bimodality, one is inclined not to emphasise much on this feature primarily because of the small sample size at hand. Rather the important issue here is the non-normality. Analyses of the height fields from the same model (different integration) show that the corresponding p.d.f. is best fitted with two Gaussians indicating the $\pm$PNA patterns respectively (Hannachi and O'Neill, 2001, *Quart. J. Roy. Meteorol. Soc.*, in press). Furthermore, recently Burgers and Stephenson (1999) have shown that the ENSO time series is strongly non-Gaussian.

The centres obtained from the mixture model and those obtained according to the sign of the ENSO index are not obviously identical but not very different (Fig. 8b), they are rather close to one another. Now because we are primarily interested in the atmospheric response associated with positive and negative phases of ENSO, and to mimic the analysis presented in the previous Section, for simplicity, we decide to split the ensemble mean field according to the sign of the ENSO index (7b). The corresponding two regime centres, $\pm$PNA, have been subtracted from their respective subsamples prior to computing the spectra. The signal pattern for each regime is then obtained as outlined in Section 2 and also as in the example of the previous Section.

In Fig. 9a we display the spectrum of $\mathbf{C}_{\mathrm{D}_1}\mathbf{C}_{\mathrm{N}}^{-1}$ corresponding to the first regime, i.e., the positive phase of ENSO (La Niña). Note that only the first half of the spectrum is displayed. The error bars in panel 9a are calculated according to the rule of thumb of North et al. (1982). The eigenvector corresponding to the leading, non-degenerate, eigenvalue is shown in panel 9b. Fig. 9 therefore indicates that the model response to positive phase of ENSO, La Niña, looks like a $-$PNA.

The same thing is performed with the other half of the ensemble mean data corresponding to negative phases of ENSO, El Niño. Surprisingly,

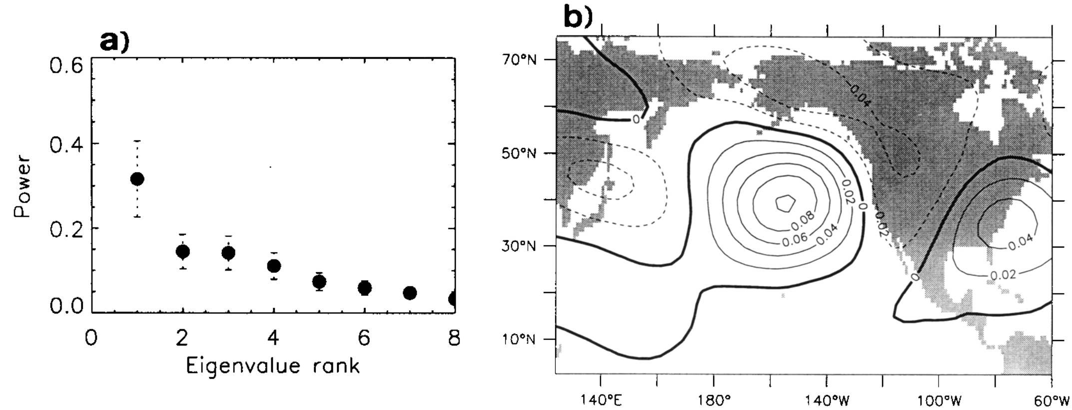

Fig. 9. The spectrum of $\mathbf{C}_{D_1}\mathbf{C}_N^{-1}$ relative to the first regime corresponding to the positive phase of ENSO, La Niña, (a), and the eigenvector corresponding to the leading eigenvalue (b).

however, the results are quite different. The spectrum (Fig. 10a) indicates that the leading eigenvalue is degenerate where the first two eigenvalues overlap. The two eigenvectors corresponding to the degenerate leading pair of eigenvalues are displayed in Figs. 10b,c.

The first one (10b) shows a high over North America and a zonally stretched low over the

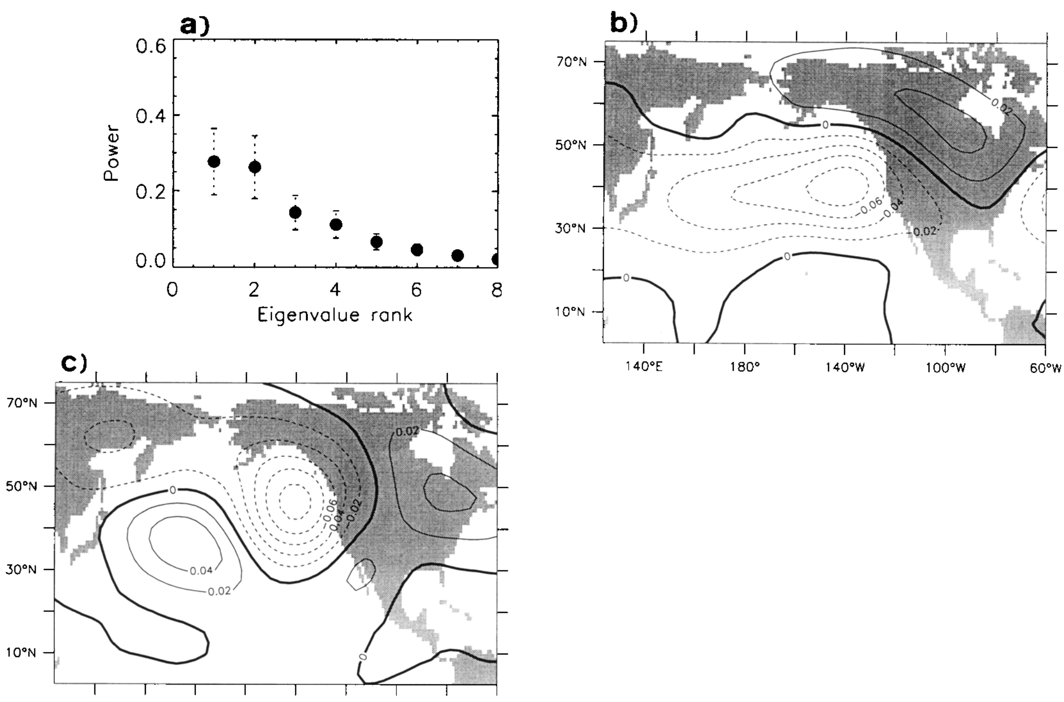

Fig. 10. The spectrum (a) and the first two eigenvectors (b, c) corresponding to the leading pair of eigenvalues of $\mathbf{C}_{D_2}\mathbf{C}_N^{-1}$ for the second regime corresponding to the negative phase of ENSO (El Niño).

North Pacific ocean. It bears some similarities to a PNA but incorporates also some signatures of NPO particularly over the northern North Pacific. Note in particular the nonsymmetry between this pattern and the previous one (9b) where the center of the zonally stretched low (10b) is shifted eastward with respect to the corresponding (positive) center of Fig. 9b and therefore the two patterns are not precisely reverse of each other. The second pattern (10c) is slightly zonally orientated with its main centre located over the gulf of Alaska and shifted north-eastward compared to the corresponding centre of a PNA. This pattern is nearly in quadrature with the patterns of Fig. 10b and also Fig. 9b. Note that the patterns of Fig. 10 are not in exact quadrature since the detection matrix $\mathbf{C}_{\mathrm{D}_2}\mathbf{C}_{\mathrm{N}}^{-1}$ is not symmetric.

A whole range of (winter) northern extratropics teleconnection patterns are identified in the literature (Wallace and Gutzler, 1981; Barnston and Livezey, 1987; and Bell and Halpert, 1995), but not all of them arise from SST variability. Precisely, only few of them, the PNA, the West Pacific (WP), and the Tropical Northern Hemisphere (TNH), are found to be associated with tropical Pacific SST variability or ENSO (Trenberth et al., 1998). Fig. 10b, for example, is the equivalent of the PNA shown in Trenberth et al. (1998), while Fig. 10c bears some similarities to the −TNH and incorporates some signature from the WP particularly over the western North Pacific. A close examination shows that Fig. 10c is reminiscent of the winter regime PAC4 of Kimoto and Ghil (1993) except that the strongest anomaly in their pattern lies over eastern Canada while it is over western Canada here.

This analysis clearly indicates the nonlinear nature of the model response where La Niña seems to produce a −PNA-like pattern while during the opposite phase, El Niño, two response patterns are obtained. The nonlinear atmospheric response associated with the phase of ENSO has also been noted, although in a different context, in other investigations (Livezey et al., 1997; Hoerling et al., 1997). For instance Hoerling et al. (1997) found, using observations as well as model simulations, that the atmospheric response to extreme phases of the tropical Pacific SST anomalies exhibit appreciable nonlinearity. In particular, they found nonlinearity from GCMs simulations even when these latter are forced by tropical Pacific SST anomalies, associated with El Niño and La Niña, that undergo identical evolution but with opposite phase.

The nonlinear nature of the response related particularly to the existence of a non-unique atmospheric response to El Niño forcing, the PNA- and the −TNH-like patterns, is particularly interesting since it indicates that not all El Niño events systematically trigger the PNA. There are in fact occasions where the PNA was not observed during some, although strong, El Niño events. Among the reasons that can lead to this kind of behaviour is the fact that the PNA is above all an internal mode that does not get systematically locked to the El Niño phase. Local effects such as local midlatitude SST anomalies and baroclinic adjustment can play a role in triggering the pattern.

5. Summary and conclusion

In this manuscript, we presented an approach to recover a true signal from an intermittent regime-like behaving system given an estimate of the noise covariance. The approach can be applied to both single- and multi-channel systems. In the single-channel, i.e., unidimensional, case the signal is simply an oscillation whereas in the muti-channel case the signal can be a spatial pattern.

The method uses the respective p.d.fs. of the data and the noise and aims at minimizing the noise p.d.f. on iso-surfaces of the data p.d.f.. When the system does not display regime behaviour, i.e., approximately Gaussian, the signal is obtained as the eigenvector corresponding to the leading eigenvalue of the detection matrix $\mathbf{C}_{\mathrm{D}}\mathbf{C}_{\mathrm{N}}^{-1}$ where $\mathbf{C}_{\mathrm{D}}$ and $\mathbf{C}_{\mathrm{N}}$ represent respectively estimates of the covariance matrices of the data and the noise. In case of regime behaviour, the signal is obtained over each regime in the same way and where $\mathbf{C}_{\mathrm{D}}$ becomes the covariance matrix of the subsample (or cluster) from the data corresponding to the regime under consideration.

The method has been successfully applied to the Lorenz system. The time series of the first component of the system, contaminated by a red-noise, has been split into two regimes according to its sign. For each subsample, or regime, the obtained time series has been first centred, by removing its time mean, then projected onto a multidimensional space using the method of the delay coordinates, the covariances computed and the signal obtained. In particular it has been shown that the signal cannot by recovered without regime separation.

Next a set of multidecadal GCM integrations of the Hadley Centre has been considered in order to investigate the atmospheric response to ENSO over the North Pacific sector. We used for this purpose the 500-hPa geopotential height fields obtained from a set of six independent integrations submitted all to the same observed SST boundary forcing. The previous approach has been applied to the ensemble mean and the noise which is obtained as the combined departure of each ensemble member from the ensemble mean. Without regime separation, the analysis results in a linear response as a function of the ENSO phase, where a +PNA (−PNA) is obtained during El Niño (La Niña) phase.

This linear response, however, is not in accord with the nonlinear nature involved in the tropical Pacific convection and its relationship to teleconnection patterns. To recover this nonlinear signal, the ensemble mean has been separated into two regimes according to the sign of the ENSO index, and the signal response for each subsample computed. It is found that the obtained signals revealed in fact nonlinearity associated with opposite phases of ENSO.

During the positive phase of ENSO, i.e., negative NINO3 SST anomaly, the response looks more like a −PNA. During the opposite phase, however, the model seems to generate two response patterns. The first pattern indicates a PNA-like structure while the second one is more equivalent to the −TNH with some signature of WP. Both patterns are not precisely the opposite of the pattern associated with positive phase of ENSO.

This nonlinear behaviour, related to the asymmetry between the response patterns associated with opposite phases of ENSO, is found to be in agreement with previous investigations where similar observations have been noted from both observations and model simulations. The fact that not all El Niño events systematically trigger a PNA seems to be in agreement with our findings. Local SST anomaly and also baroclinic adjustment can play a role in triggering, or not, the PNA during El Niño events.

6. Acknowledgment

This work was funded by the UK Universities Global Atmospheric Modelling Project (UGAMP). Computing was performed on local sun stations at AOPP. The HadAM1 data was provided by the Hadley Centre. Particular thanks are due to D. Rowell and to S. Brown for his help in getting the data. We also thank two anonymous reviewers for their constructive comments that greatly improved the manuscript.

REFERENCES

Allen, M. R. and L. A. Smith, 1997. Optimal filtering in singular spectrum analysis. *Phys. lett. A* **234**, 419–428.

Anderson, B. D. and J. B. Moore, 1979. *Optimal filtering*. Prentice Hall. New Jersey.

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

Bell, G. D. and M. S. Halpert, 1995. Interseasonal and interannual variability: 1986 to 1993, *NOAA Atlas* **12**, 256 pp. U.S. Dept. Commer., Washington, D.C.

Broomhead, D. S. and G. King, 1986a. Extracting qualitative dynamics from experimental data. *Physica D* **20**, 217–236.

Broomhead, D. S. and G. King, 1986b. On the qualitative analysis of experimental dynamical systems. *Nonlinear phenomena and chaos*, S. Sarkar, Ed. Adam Hilger, pp. 113–144.

Burgers, G. and D. B. Stephenson, 1999: The "Normality" of El Nino. *Geophys. Res. Lett.* **26**, 1027–1030.

Charney, J. G. and J. G. DeVore, 1979. Multiple flow equilibria in the atmosphere and blocking. *J. Atmos. Sci.* **36**, 1205–1216.

Cheng, X. and J. M. Wallace, 1993. Cluster analysis of the Northern Hemisphere wintertime 500-hPa height field: Spatial patterns. *J. Atmos. Sci.* **50**, 2674–2696.

Corti, S., F. Molteni and T. N. Palmer, 1999. Signature of recent climate change in frequencies of natural atmospheric circulation regimes. *Nature* **398**, 799–802.

Davies, J. R., D. P. Rowell and C. K. Folland, 1997. North Atlantic and European seasonal predictability using an ensemble of multi-decadal AGCM simulations. *Int. J. Climatology* **17**, 1263–1284.

Everitt, B. S. and D. J. Hand, 1981. *Finite mixture distributions*. Chapmann and Hall. London.

Fraedrich, K. 1986. Estimating the dimension of weather and climate attractors. *J. Atmos. Sci.* **43**, 419–432.

Ghil, M. and S. Childress, 1987. *Topics in geophysical fluid dynamics: atmospheric dynamics, dynamo theory and climate dynamics*. Springer-Verlag. New York.

Haines, K. and A. Hannachi, 1995. Weather regimes in the Pacific from a GCM. *J. Atmos. Sci.* **52**, 2444–2462.

Hannachi, A. 1997. Low-frequency variability in a GCM: three-dimensional flow regimes and their dynamics. *J. Climate* **10**, 1357–1379.

Hannachi, A. 2000. Probabilistic-based approach to optimal filtering. *Phys. Rev. E* **61**, 3610–3619.

Hannachi, A. 2001. Toward a nonlinear identification of the atmospheric response to ENSO. *J. Climate* **14**, 2138–2149.

Hasselmann, K. 1999. Climate change: Linear and nonlinear signature. *Nature* **398**, 755–756.

Hoerling, M. P., A. Kumar and M. Zhong, 1997. El Niño, La Niña, and the nonlinearity of their teleconnections. *J. Climate* **10**, 1769–1786.

Kimoto M. and M. Ghil, 1993. Multiple flow regimes in the Northern Hemisphere winter: Part II: Sectorial regimes and preferred transitions. *J. Atmos. Sci.* **50**, 2645–2673.

Leith, C. E. 1973. The standard error of time-average estimates of climatic means. *J. Appl. Meteor.* **12**, 1066–1069.

Livezey, R. E., A. Leetmaa, M. Mautani, H. Rui, N. Ji and A. Kumar, 1997. Teleconnective response of the Pacific-North American region atmosphere to large central equatorial Pacific SST anomalies. *J. Climate* **10**, 1787–1820.

Lorenz, E. N. 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141.

Lorenz, E. N. 1968. Climatic determinism. *Meteor. Monogr.* **8**, 1–3.

Marshall, J. C. and F. Molteni, 1993. Towards a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.* **50**, 1792–1818.

Molteni, F., L. Ferranti, T. N. Palmer and P. Viterbo, 1993. A dynamical interpretation of the global response to equatorial Pacific SST anomalies. *J. Climate* **6**, 777–795.

Namias, J. 1964. Seasonal persistence and recurrence of European blocking during 1958–1960. *Tellus* **16**, 394–407.

North, G. R., T. L. Bell, R. F. Cahalan, and F. G. Moeng, 1982: sampling errors in the estimation of the empirical orthogonal functions. *Mon. Wea. Rev.* **110**, 699–706.

Palmer, T. N. 1999. A nonlinear dynamical perspective on climate prediction. *J. Climate* **12**, 575–591.

Silverman, B. W. 1994. *Density estimation for statistics and data analysis.* Chapman and hall. London.

Takens, F. 1981. *Detecting strange attractors in turbulence.* Lecture notes in mathematics, D. A. Rand and L.-S. Young, eds. Springer Verlag. Berlin.

Thacker, W. C. 1996. Metric-based principal components: data uncertainties. *Tellus* **48A**, 584–592.

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998. Progress during TOGA in understanding and modelling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.* **103**, 14,291–14,324.

Ulrych, T. J. and T. N. Bishop, 1975. Maximum entropy spectral analysis and autoregressive decomposition. *Rev. Geophys. and Space Phys.* **13**, 183–200.

Vautard, R., P. Yiou and M. Ghil, 1992. Singular spectrum analysis: a toolkit for short, noisy and chaotic signals. *Physica D* **58**, 95–126.

Venzke, S., M. R. Allen, R. T. Sutton and D. P. Rowell, 1999. The atmospheric response over the North Atlantic to decadal changes in sea surface temperature. *J. Climate* **12**, 2562–2584.

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