Advanced Spectral Methods

for Climatic Time Series

M. Ghil[1] [2], M. R. Allen[3], M. D. Dettinger[4], K. Ide[1], D. Kondrashov[1], M. E. Mann[5], A. W. Robertson[1], A. Saunders[1], Y. Tian[1], F. Varadi[1] and P. Yiou[6] [7]

Revised version

June 29, 2001

---

[1]Department of Atmospheric Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90095-1565, USA

[2]Corresponding author

[3]Space Science and Technology Department, Rutherford Appleton Laboratory, Chilton, Didcot, OX11 0QX, UK

[4]U. S. Geological Survey, San Diego, CA 92123, USA

[5]Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22903, USA

[6]Laboratoire des Sciences du Climat et de l'Environnement, UMR CEA–CNRS, 91191 Gif-sur-Yvette, France

[7]The authors, besides the corresponding one, are in alphabetical order.

Abstract

The analysis of uni- or multivariate time series provides crucial information to describe, understand, and predict climatic variability. The discovery and implementation of a number of novel methods for extracting useful information from time series has recently revitalized this classical field of study. Considerable progress has also been made in interpreting the information so obtained in terms of dynamical systems theory. In this review, we describe the connections between time-series analysis and nonlinear dynamics, discuss signal-to-noise enhancement, and present some of the novel methods for spectral analysis. The various steps, as well as the advantages and disadvantages of these methods, are illustrated by their application to an important climatic time series, the Southern Oscillation Index. This index captures major features of interannual climate variability and is used extensively in its prediction. Regional and global sea-surface temperature data sets are used to illustrate multivariate spectral methods. Open questions and further prospects conclude the review.

Keywords: spectral analysis, time series, prediction, climate

Contents

1 Introduction and Motivation

A time series provides useful information about the physical, biological or socio-economic system that produced it. The purpose of time-series analysis is to determine some of the system's key properties by quantifying certain features of the time series. These properties can then help understand and predict the system's future behavior.

To illustrate the ideas and methods reviewed here, we shall turn to one of the best-known climatic time series. This time series is made up of monthly values of the Southern Oscillation Index (SOI). It will be introduced in Sec. 2.2 and is shown in Fig. 2 there.

At this point we merely notice that physical processes usually operate in continuous time. Most measurements though are done and recorded in discrete time. Thus the SOI time series, as well as most climatic and other geophysical time series, are available in discrete time.

1.1 Analysis in the time domain vs. the spectral domain

Two basic approaches to time series analysis are associated with the time domain or the spectral domain. We present them at first in the linear context in which the physical sciences have operated for most of the last two centuries. In this context, the physical system can be described by a linear ordinary differential equation (ODE), or a system of such equations, subject to additive random forcing.

It goes well beyond the scope of this review paper to introduce the concepts of random variables, stochastic processes and stochastic differential equations. We refer the interested reader to *Feller* [1968, 1970] for the former two concepts and to *Arnold* [1974] and *Schuss* [1980] for the latter. Many of the standard books on classical spectral methods that are cited in Secs. 3.1 and 3.2 also contain good elementary introductions to stochastic processes in discrete and, sometimes, continuous time.

We concentrate here on time series in discrete time and consider therefore at first the

simple case of a scalar, linear ordinary difference equation with random forcing,

$$X(t+1) = \sum_{j=1}^{M} a_j X(t - M + j) + \sigma \xi(t). \tag{1.1}$$

Its constant coefficients a_j determine the solutions $X(t)$ at discrete times $t = 0, 1, \ldots, n, \ldots$ In Eq. (1.1) the random forcing $\xi(t)$ is assumed to be white in time, i.e., uncorrelated from t to $t+1$, and Gaussian at each t, with constant variance equal to unity. In the *Yule* [1927] – *Walker* [1931] method for the time-domain approach one computes the coefficients a_j and the variance σ^2 from a realization of X having length N, $\{X(t) : 1 \leq t \leq N\}$.

This method is discussed further in Sec. 3.3 below, where (1.1) is treated as an autoregressive (AR) process of order M, which we denote by AR(M). The notation used in the present paper is summarized in Table 1. The table lists the main symbols in alphabetical order and indicates the section where each symbol is introduced first. This facilitates the comparison between the methods we review, since the literature of each method tends to use its own notation.

Table 1: *Glossary of the principal symbols*

The spectral-domain approach is motivated by the observation that the most regular, and hence predictable, behavior of a time series is to be periodic. This approach then proceeds to determine the periodic components embedded in the time series by computing the associated periods, amplitudes and phases, in this order.

The classical implementation of the spectral-domain approach is based on the Bochner-Khinchin-Wiener theorem [*Box and Jenkins*, 1970] which states that the lag-autocorrelation function of a time series and its *spectral density* are Fourier transforms of each other. *Hannan*'s [1960] introduction to this approach and its implementation excels by its brevity and clarity; the so-called Blackman-Tukey implementation is presented in Sec. 3.2 below. We shall use here both the more mathematical term of spectral density and the term of *power*

spectrum, often used in the scientific and engineering literature, according to whether the context is more theoretical or more applied.

The remainder of this review is organized as follows. Section 2 deals mainly with signal-to-noise (S/N) ratio enhancement and introduces singular-spectrum analysis (SSA) as an important and flexible tool for this enhancement. Connections between SSA and empirical orthogonal functions (EOFs) are outlined in Appendix A. Statistical tests for the reliability of SSA results are also discussed in this section, along with connections to wavelet analysis.

In Sec. 3, we present, in succession, three methods of spectral analysis: Fourier-transform based, maximum entropy, and multi-taper. Both Secs. 2 and 3 use the SOI time series for the purposes of illustrating the methods "in action." In Sec. 4, the multivariate extensions of the maximum-entropy method and of single-channel SSA are introduced and a few additional applications are mentioned or illustrated. The review concludes with a section on open questions, from the point of view of both the methodology and its applications.

1.2   Time series and nonlinear dynamics

Before proceeding with the technical details, we give in this subsection a quick perspective on the "nonlinear revolution" in time-series analysis. In the 1960s and '70s, the scientific community found out that much of the irregularity in observed time series, which had traditionally been attributed to the above-mentioned random "pumping" of a *linear* system by infinitely many (independent) degrees of freedom (*d-o-f*), could be generated by the *nonlinear* interaction of a few *d-o-f* [*Lorenz*, 1963; *Smale*, 1967; *Ruelle and Takens*, 1971]. This realization of the possibility of deterministic aperiodicity or "chaos" [*Gleick*, 1987] created quite a stir.

The purpose of this review is to describe briefly some of the implications of this change in outlook for time-series analysis, with a special emphasis on climatic time series. Many general aspects of nonlinear time-series analysis are reviewed by *Drazin and King* [1992], *Ott*

et al. [1994], and *Abarbanel* [1996]. We concentrate here on those aspects that deal with regularities and have proven most useful in studying climatic variability.

A connection between deterministically chaotic time series and the nonlinear dynamics generating them was attempted fairly early in the young history of "chaos theory." The basic idea was to consider specifically a scalar, or univariate, time series with apparently irregular behavior, generated by a deterministic or stochastic system. This time series could be exploited — so the thinking went — in order to ascertain, first, whether the underlying system has a finite number of *d-o-f*. An upper bound on this number would imply that the system is deterministic, rather than stochastic, in nature. Next, we might be able to verify that the observed irregularity arises from the fractal nature of the deterministic system's invariant set, which would yield a fractional, rather than integer, value of this set's dimension. Finally, one could maybe reconstruct the invariant set or even the equations governing the dynamics from the data.

This ambitious program [*Packard et al.*, 1980; *Roux et al.*, 1980; *Ruelle*, 1981] relied essentially on the *method of delays*, based in turn on the *Whitney* [1936] embedding lemma and the *Mañé* [1981]–*Takens* [1981] theorems. We first describe an easy connection between a univariate and a multivariate time series.

Let us assume that the univariate time series is the solution of a scalar nonlinear ODE of order p,

$$X^{(p)} = G(X^{(p-1)}, \ldots, X). \tag{1.2}$$

This scalar higher-dimensional equation is equivalent to the following system of first-order ODEs,

$$\dot{X}_i = F_i(X_1, \ldots, X_j, \ldots, X_p), \quad 1 \le i, j \le p; \tag{1.3}$$

7

here $\dot{X} \equiv dX/dt \equiv X^{(1)}$ and $X^{(p)} \equiv d^p X/dt^p$. It suffices to write

$$X \equiv X_1, \ \dot{X}_1 = X_2, \ \dots, \ \dot{X}_{p-1} = X_p, \ \dot{X}_p = G(X_1, \dots, X_p), \qquad (1.4)$$

so that $F_1 = X_2$, $F_2 = X_3, \dots$, $F_p = G$. In other words, the succesive derivatives of $X(t)$ can be thought of as the component of a vector $\mathbf{X} \equiv (X_1, \dots, X_p)$. The Euclidean space $\mathbb{R}^p$ in which the vector $\mathbf{X} = \mathbf{X}(t)$ evolves is called the *phase space* of the first-order system (1.3).

Let us now change the point of view and consider system (1.3) for arbitrary right-hand sides $F_i(\mathbf{X})$, rather than for the specific F_i's given by (1.4). Such an ODE system represents a fairly general description of a differentiable dynamical system in continuous time [*Arnold*, 1973, 1983] . We are interested at first in the case in which only a single time series $\hat{X}(t)$ is known, where $\hat{X} = X_i$ for some fixed component $i = i_0$ or $\hat{X}$ is some sufficiently smooth function of all the components X_i. For the solutions of such a system to be irregular, i.e., other than (asymptotically) steady or periodic, three or more *d-o-f* are necessary. Can one then go from (1.3) to (1.2) just as easily as in the opposite direction? The answer, in general is "no"; hence, a slightly more sophisticated procedure needs to be applied.

This procedure is called the method of delays and it tries, in some sense, to imitate the Yule-Walker inference of (1.1) from the time series $\{X(t) : t = 1, \dots, N\}$. First of all, one acknowledges that the data $X(t)$ are typically given at discrete times $t = n\Delta t$ only. Next, one admits that it is hard to actually get the right-hand sides F_i; instead one attempts to reconstruct the invariant set on which the solutions of (1.3) that satisfy certain constraints lie.

In the case of conservative, Hamiltonian systems [*Lichtenberg and Lieberman*, 1992], there are typically unique solutions through every point in phase space. The irregularity in the solutions' behavior is associated with the intricate structure of *cantori* [*Wiggins*, 1988], complicated sets of folded tori characterized by a given energy of the solutions lying on them.

These cantori have, in particular, finite and fractional dimension, being self-similar *fractals* [*Mandelbrot*, 1982].

Hamiltonian systems, however, are — mathematically speaking — structurally unstable [*Smale*, 1967] in the function space of all differentiable dynamical systems. Physically speaking, on the other hand, "open" systems in which energy is gained externally and dissipated internally, abound in nature. Therefore, climatic time series, as well as most other time series from nature or the laboratory, are more likely to be generated by forced dissipative systems [*Lorenz*, 1963; *Ghil and Childress*, 1987, Ch. 5]. The invariant sets associated with irregularity here are "strange attractors" [*Ruelle and Takens*, 1971], towards which all solutions tend asymptotically, i.e., long-term irregular behavior in such systems is associated with these attractors. These objects are also fractal, although rigorous proofs to this effect have been much harder to give than in the case of Hamiltonian cantori [*Guckenheimer and Holmes*, 1983; *Lasota and Mackey*, 1994].

Mañé [1981], *Ruelle* [1981], and *Takens* [1981] had the idea, developed further by *Sauer et al.* [1991], that a single observed time series $\hat{X}(t)$ — where $\hat{X} = X_{i_0}$ or, more generally, $\hat{X} = \phi(X_1(t), \ldots, X_p(t))$ — could be used to reconstruct the attractor of a forced dissipative system. The basis for this reconstruction idea is essentially the fact that such a solution covers the attractor densely, i.e., as time increases, it will pass arbitrarily close to any point on the attractor. Time series observed in the natural environment, however, have finite length and sampling rate, as well as significant measurement noise.

The embedding idea has been applied, therefore, most successfully to time series generated numerically or by laboratory experiments in which sufficiently long series could be obtained and noise was controlled better than in nature. *Broomhead and King* [1986a], for instance, applied successfully SSA to the reconstruction of the *Lorenz* [1963] attractor. As we shall see, for climate and other geophysical time series, it might be possible to attain a more modest goal: to describe merely a "skeleton" of the attractor that is formed by a few

9

robust periodic orbits.

In the climate context, *Lorenz* [1969] had already pointed out a major stumbling block for applying the attractor reconstruction idea to large-scale atmospheric motions. While using more classical statistical methods, he showed that the recurrence time of sufficiently good analogs for weather maps was of the order of hundreds of years, at the spatial resolution of the observational network then available for the Northern Hemisphere.

The next-best target for demonstrating from an observed time series the deterministic cause of its irregularity was to show that the presumed system's attractor had a finite and fractional dimension. Various dimensions, metric and topological, can be defined [*Kaplan and Yorke*, 1979; *Farmer et al.*, 1983]; among them, the one that became the most popular, since easiest to compute, was the correlation dimension [*Grassberger and Procaccia*, 1983]. In several applications its computation proved rather reliable, and hence useful. Climatic time series, however, tended again to be rather too short and noisy for comfort (see for instance, *Ruelle* [1990] and *Ghil et al.* [1991] for a review of this controversial topic).

A more robust connection between classical spectral analysis and nonlinear dynamics seems to be provided by the concept of "ghost limit cycles." The road to chaos [*Eckmann*, 1981] proceeds from stable equilibria, or *fixed points*, through stable periodic solutions, or *limit cycles*, and on through quasi-periodic solutions lying on *tori*, to *strange attractors*. The fixed points and limit cycles are road posts on this highway from the simple to the complex. That is, even after having lost their stability to successively more complex and realistic solutions, these simple attractors still play a role in the observed spatial patterns and the time series generated by the system.

A "ghost fixed point" is a fixed point that has become unstable in one or a few directions in phase space. Still the system's trajectories will linger near it for extended time intervals [*Legras and Ghil*, 1985]. Likewise, a ghost limit cycle is a closed orbit that has become slightly unstable but is visited, again and again, by the system's trajectories [*Kimoto and*

Ghil, 1993].

Consider the periodic solution shown in Fig. 1a as embedded in Euclidean three-dimensional phase space. It is neutrally stable in the direction tangent to itself, while in the plane perpendicular to this tangent it is asymptotically stable in one direction and unstable in the other, as shown in the Poincaré section of Fig. 1b. In a multi-dimensional phase space, it is plausible that the directions of stability are numerous or even infinite in number. The directions of instability, however, would still be few in number, for parameter values not too far from those at which the Hopf bifurcation that gave rise to the limit cycle in the first place occurs. Hence solutions of the full system would easily be attracted to this barely unstable limit cycle, follow it closely for one or a few turns, be ejected from its neighborhood, only to return later, again and again.

The analogous picture for a ghost fixed point was illustrated in detail for an atmospheric model with 25 *d-o-f* by *Legras and Ghil* [1985]; see also *Ghil and Childress* [1987; Figs. 6.12 and 6.18 there]. Thus the "ghosts" of fixed points and limit cycles leave their imprint on the system's observed spatio-temporal behavior.

Figure 1: *The schematic diagram of a ghost limit cycle.*

The episodes during which the system trajectory circles near a ghost limit cycle result in nearly periodic segments of the time series and hence contribute to a spectral peak with that period. This concept was illustrated using 40 years of an atmospheric multivariate time series by *Kimoto and Ghil* [1993] for the so-called intraseasonal oscillations of the Northern Hemisphere (see also *Ghil and Mo* [1991a]). We shall show in the subsequent sections of the present review, in particular Sec. 2.2, how this concept can be generalized to associate multiple spectral peaks with a robust skeleton of the attractor, as proposed by *Vautard and Ghil* [1989].

2 Enhancing the Signal-to-Noise (S/N) Ratio

2.1 Motivation for singular spectrum analysis (SSA)

SSA is designed to extract information from short and noisy time series, and thus provide insight into the — unknown or only partially known — dynamics of the underlying system that generated the series [*Broomhead and King*, 1986*a*; *Fraedrich*, 1986; *Vautard and Ghil*, 1989]. We outline here the method for univariate time series and generalize for multivariate ones in Sec. 4.2.

The analogies between SSA and spatial EOFs are summarized in Appendix A, along with the basis of both in the Karhunen–Loève theory of random fields and of stationary random processes. Multi-channel SSA (see Sec. 4.2) is numerically analogous to the Extended EOF (EEOF) algorithm of *Weare and Nasstrom* [1982]. The two different names arise from the origin of the former in the dynamical-systems analysis of univariate time series, while the latter had its origins in the principal component analysis of meteorological fields. The two approaches lead to different methods for the choice of key parameters — such as the, fixed or variable, window width — and hence to differences in the way of interpreting results.

The starting point of SSA is to *embed* a time series $\{X(t) : t = 1, \ldots, N\}$ in a vector space of dimension M, i.e., to represent it as a trajectory in the phase space of the hypothetical system that generated $\{X(t)\}$. In concrete terms, this is equivalent to representing the behavior of the system by a succession of overlapping "views" of the series through a sliding M-point window.

Let us assume, for the moment, that $X(t)$ is an observable function $\hat{X}(t)$ of a noise-free system's dependent variables $X_i(t)$, as defined in Eq. (1.4), and that the function ϕ that maps the p variables $\{X_i(t) : i = 1, \ldots, p\}$ into the single variable $X(t)$ has certain properties that make it generic in the dynamical systems sense of *Smale* [1967]. Assume, moreover, that $M > 2d + 1$ — where d is the dimension of the underlying attractor on which the system

evolves — and that d is known and finite. If so, then the representation of the system in the "delay coordinates" described in Eq. (2.1) below will share key topological properties with a representation in *any* coordinate system. This is a consequence of *Whitney*'s [1936] embedding lemma and indicates the potential value of SSA in the qualitative analysis of the dynamics of nonlinear systems [*Broomhead and King*, 1986*a, b*; *Sauer et al.*, 1991]. The quantitative interpretation of SSA results in terms of attractor dimensions is fraught with difficulties, however, as pointed out by a number of authors [*Broomhead et al.*, 1987; *Vautard and Ghil*, 1989; *Paluš and Dvořák*, 1992].

We use SSA here, therefore, mainly (i) for data-adaptive signal-to-noise (S/N) enhancement and associated data compression; and (ii) to find the attractor's *skeleton*, given by its least-unstable limit cycles (see again Fig. 1 above). The embedding procedure applied to do so constructs a sequence $\{\tilde{\mathbf{X}}(t)\}$ of M-dimensional vectors from the original time series X, by using lagged copies of the scalar data $\{X(t) : 1 \leq t \leq N\}$,

$$\tilde{\mathbf{X}}(t) = (X(t), X(t+1), \ldots, X(t+M-1)); \qquad (2.1)$$

the vectors $\tilde{\mathbf{X}}(t)$ are indexed by $t = 1, \ldots, N'$, where $N' = N - M + 1$.

SSA allows one to unravel the information embedded in the delay-coordinate phase space by decomposing the sequence of augmented vectors thus obtained into elementary patterns of behavior. It does so by providing data-adaptive filters that help separate the time series into components that are statistically independent, at zero lag, in the augmented vector space of interest. These components can be classified essentially into trend(s), oscillatory patterns, and noise. As we shall see, it is an important feature of SSA that the trend(s) need not be linear and that the oscillations can be amplitude- and phase-modulated.

SSA has been applied extensively to the study of climate variability, as well as to other areas in the physical and life sciences. The climatic applications include the analysis of pale-

oclimatic time series [*Vautard and Ghil*, 1989; *Yiou et al.*, 1994, 1995], interdecadal climate variability [*Ghil and Vautard*, 1991; *Allen and Smith*, 1994; *Plaut et al.*, 1995; *Robertson and Mechoso*, 1998], as well as interannual [*Rasmusson et al.*, 1990; *Keppenne and Ghil*, 1992] and intraseasonal [*Ghil and Mo*, 1991*a, b*] oscillations. SSA algorithms and their properties have been investigated further by *Penland et al.* [1991], *Allen* [1992], *Vautard et al.* [1992], and *Yiou et al.* [2000]. The SSA-MTM Toolkit first documented by *Dettinger et al.* [1995*a*] was built, largely but not exclusively, around this technique.

2.2   Decomposition and reconstruction

In this section, we illustrate the fundamental SSA formulae with the classical example of a climatic time series, the Southern Oscillation Index (SOI). SOI is a climatic index connected with the recurring El Niño conditions in the tropical Pacific. It is defined usually as the difference between the monthly means of the sea-level pressures at Tahiti and at Darwin (Australia).

We use this definition and the monthly data given in the archive

http://tao.atmos.washington.edu/pacs/additional_analyses/soi.html.

The SOI data in this archive are based on the time series at each of the two stations being deseasonalized and normalized [*Ropelewski and Jones*, 1987]. The seasonal cycle is removed by subtracting the average of the values for each calendar month over a reference interval, in this case 1951–1980. The residues from this operation, called monthly *anomalies* in the climatological literature, are then normalized with respect to the standard deviation computed over the same interval.

The archived SOI data for 1866–1997 are from the Climate Research Unit of the University of East Anglia and those for 1998-1999 are from the Climate Prediction Center of the U.S. National Centers for Environmental Prediction (NCEP). They are obtained by taking

14

the difference between the anomalies at Tahiti and those at Darwin and dividing by the standard deviation of this difference over the same 30-year reference interval. The time interval we consider here goes from January 1942 to June 1999, during which no observations are missing at either station; this yields $N = 690$ raw data points. Note that this raw SOI time series is centered and normalized over the reference interval 1951–1980, but not over the entire interval of interest. We show in Fig. 2 the SOI obtained from this raw data set. It actually has mean -0.0761 and standard deviation equal to 1.0677. All subsequent figures, however, use a version of the series that has been correctly centered over the interval January 1942–June 1999.

Figure 2: *Variations of the Southern Oscillation Index (SOI).*

SSA is based on calculating the principal directions of extension of the sequence of augmented vectors $\{\tilde{\mathbf{X}}(t) : t = 1, \ldots, N'\}$ in phase space. The $M \times M$ covariance matrix $\mathbf{C}_X$ can be estimated directly from the data as a *Toeplitz matrix* with constant diagonals, i.e., its entries c_{ij} depend only on the lag $|i - j|$, cf. *Vautard and Ghil* [1989]:

$$c_{ij} = \frac{1}{N - |i - j|} \sum_{t=1}^{N - |i - j|} X(t)X(t + |i - j|). \tag{2.2}$$

The eigenelements $\{(\lambda_k, \rho_k) : k = 1, \ldots, M\}$ of $\mathbf{C}_X$ are then obtained by solving

$$\mathbf{C}_X \rho_k = \lambda_k \rho_k. \tag{2.3}$$

The eigenvalue λ_k equals the partial variance in the direction ρ_k and the sum of the λ_k, i.e., the trace of $\mathbf{C}_X$, gives the total variance of the original time series $X(t)$.

An equivalent formulation of Eq. (2.3), which will prove useful further on, is given by forming the $M \times M$ matrix $\mathbf{E}_X$ that has the eigenvectors ρ_k as its columns and the diagonal

15

matrix $\boldsymbol{\Lambda}_X$ whose elements are the eigenvalues λ_k, in decreasing order:

$$\mathbf{E}_X^{\mathrm{t}} \mathbf{C}_X \mathbf{E}_X = \boldsymbol{\Lambda}_X; \qquad (2.4)$$

here $\mathbf{E}_X^{\mathrm{t}}$ is the transpose of $\mathbf{E}_X$. Each eigenvalue λ_k gives the degree of extension, and hence the variance, of the time series in the direction of the orthogonal eigenvector ρ_k.

A slightly different approach to computing the eigenelements of $\mathbf{C}_X$ was originally proposed by *Broomhead and King* [1986a]. They constructed the $N' \times M$ *trajectory matrix* $\mathbf{D}$ that has the N' augmented vectors $\tilde{\mathbf{X}}(t)$ as its rows and used singular-value decomposition (SVD; see, for instance, *Golub and Van Loan* [1996]) of

$$\mathbf{C}_X = \frac{1}{N'} \mathbf{D}^{\mathrm{t}} \mathbf{D} \qquad (2.5)$$

to obtain the square roots of λ_k. The latter are called the *singular values* of $\mathbf{D}$ and have given SSA its name.

Allen and Smith [1996] and *Ghil and Taricco* [1997] have discussed the similarities and differences between the approaches of *Broomhead and King* [1986a] and *Vautard and Ghil* [1989] in computing the eigenelements associated in SSA with a given time series $X(t)$. Both the Toeplitz estimate (2.2) and the SVD estimate (2.5) lead to a symmetric covariance matrix $\mathbf{C}_X$. In addition, the eigenvectors ρ_k of a Toeplitz matrix are necessarily odd and even, like the sines and cosines of classical Fourier analysis. The Toeplitz approach has the advantage of better noise reduction when applied to short time series, as compared to the SVD approach. This advantage comes at the price of a slightly larger bias when the time series is strongly nonstationary over the interval of observation $1 \leq t \leq N$ [*Allen*, 1992]. Such bias-vs.-variance trade-offs are common in estimation problems.

To obtain S/N separation one plots the eigenvalue spectrum illustrated in Fig. 3. In this plot, an initial plateau that contains most of the signal is separated by a steep slope from

the noise; the latter is characterized by much lower values that form a flat floor or a mild slope [*Kumaresan and Tufts*, 1980; *Pike et al.*, 1984; *Vautard and Ghil*, 1989].

Figure 3: *Singular spectrum of the SOI time series.*

As the $M \times M$ matrix $\mathbf{C}_X$ is symmetric, standard algorithms [*Press et al.*, 1988] will perform its spectral decomposition efficiently, as long as M is not too large. The choice of M is based on a trade-off between two considerations: quantity of information extracted vs. the degree of statistical confidence in that information. The former requires as wide a window as possible, i.e., a large M, while the latter requires as many repetitions of the features of interest as possible, i.e., as large a ratio N/M as possible. The choice of $M = 60$ in Fig. 3 allows us to capture periodicities as long as 5 years, while $N/M = 11$ is fairly safe and the diagonalization of $\mathbf{C}_X$ does not introduce large numerical errors either.

In Fig. 3, there is a clear grouping of the first 5 eigenvalues, followed by a very steep slope of three additional eigenvalues. The latter are well separated from the first 5, as well as from the remaining 52 eigenvalues, which form the mildly sloping and flattening out "tail" of the SSA spectrum.

The S/N separation obtained by merely inspecting the slope break in a "scree diagram" of eigenvalues λ_k or singular values $\lambda_k^{1/2}$ vs. k works well when the *intrinsic* noise that perturbs the underlying deterministic system and the *extrinsic* noise that affects the observations are both white, i.e. uncorrelated from one time step to the next [see definition of $\xi(t)$ in Eq. (1.1)]. This rudimentary separation works less well when either noise is red, i.e., when it is given by an AR(1) process (see Sec. 3.3) or is otherwise correlated between time steps [*Vautard and Ghil*, 1989]. The difficulties that arise with correlated noise led *Allen* [1992] and *Allen and Smith* [1994] to develop Monte Carlo SSA (see Sec. 2.3).

When the noise properties can be estimated reliably from the available data, the application of a so-called "pre-whitening operator" can significantly enhance the signal-separation

capabilities of SSA [*Allen and Smith*, 1997]. The idea is to pre-process the time series itself — or, equivalently but often more efficiently, the lag-covariance matrix $\mathbf{C}_X$ — such that the noise becomes uncorrelated in this new representation. SSA is then performed on the transformed data or covariance matrix and the results are tranformed back to the original representation for inspection.

By analogy with the meteorological literature, the eigenvectors ρ_k of the lag-covariance matrix $\mathbf{C}_X$ have been called empirical orthogonal functions (EOFs; see *Preisendorfer* [1988] and references there) by *Fraedrich* [1986] and by *Vautard and Ghil* [1989]. The EOFs corresponding to the first five eigenvalues are shown in Fig. 4. Note that the two EOFs in each one of the two leading pairs — i.e., EOFs 1 and 2, as well as EOFs 3 and 4 — are in quadrature and that each pair of EOFs corresponds in Fig. 3 to a pair of eigenvalues that are approximately equal and whose error bars overlap. *Vautard and Ghil* [1989] argued that, subject to certain statistical significance tests discussed further below, such pairs correspond to the nonlinear counterpart of a sine–cosine pair in the standard Fourier analysis of linear problems.

Figure 4: *First five EOFs of the SOI time series.*

In the terminology of our Sec. 1 here, such a pair gives a handy representation of a ghost limit cycle. The advantage over sines and cosines is that the EOFs obtained from SSA are not necessarily harmonic functions and, being data-adaptive, can capture highly *anharmonic* oscillation shapes. Indeed, relaxation oscillations [*Van der Pol*, 1940] and other types of nonlinear oscillations [*Stoker*, 1950], albeit purely periodic, are usually not sinusoidal, i.e., they are anharmonic. Such nonlinear oscillations often require, therefore, the use of many harmonics or subharmonics of the fundamental period when carrying out classical Fourier analysis, while a single pair of SSA eigenmodes might suffice. Capturing the shape of an anharmonic oscillation — such as a seesaw or boxcar, albeit slightly rounded or smoothed

— is easiest when the SSA window is exactly equal to the single period being analyzed.

Projecting the time series onto each EOF yields the corresponding principal components (PCs) $\mathbf{A}_k$:

$$A_k(t) = \sum_{j=1}^{M} X(t + j - 1)\rho_k(j). \tag{2.6}$$

Figure 5 shows the variations of the five leading PCs. Again, the two PCs in each of the pairs $(1, 2)$ and $(3, 4)$ are in quadrature, two by two (see Figs. 4a, b). They strongly suggest periodic variability at two different periods, of about 4 and 2 years, respectively. Substantial amplitude modulation at both periodicities is present, too.

Figure 5: *First five PCs of the SOI time series.*

The fifth PC, shown in Fig. 5c, contains both a long-term, highly nonlinear trend and an oscillatory component. We shall discuss the trend of the SOI series in connection with Figs. 6b and 16a further below.

We can reconstruct that part of a time series that is associated with a single EOF or several by combining the associated PCs:

$$R_{\mathcal{K}}(t) = \frac{1}{M_t} \sum_{k \in \mathcal{K}} \sum_{j=L_t}^{U_t} A_k(t - j + 1)\rho_k(j); \tag{2.7}$$

here $\mathcal{K}$ is the set of EOFs on which the reconstruction is based. The values of the normalization factor M_t, as well as of the lower and upper bound of summation L_t and U_t, differ between the central part of the time series and its endpoints [*Ghil and Vautard*, 1991; *Vautard et al.*, 1992]:

$$(M_t, L_t, U_t) = \begin{cases} \left(\frac{1}{t}, 1, t\right), & 1 \le t \le M - 1, \\ \left(\frac{1}{M}, 1, M\right), & M \le t \le N', \\ \left(\frac{1}{N - t + 1}, t - N + M, M\right), & N' + 1 \le t \le N. \end{cases} \tag{2.8}$$

The reconstructed components (RCs) have the property of capturing the phase of the time series in a well-defined least-square sense, so that $X(t)$ and $R_\mathcal{K}(t)$ can be superimposed on the same time scale, $1 \leq t \leq N$. This is an advantage of the RCs over the PCs, which have length $N - M$ and do not contain direct phase information within the window width M.

No information is lost in the reconstruction process, since the sum of all individual RCs gives back the original time series. Partial reconstruction is illustrated in Fig. 6 by summing the variability of PCs 1–4, associated with the two leading pairs of eigenelements; it is common to refer to such a reconstruction (2.7), with $\mathcal{K} = \{1, 2, 3, 4\}$, as RCs 1–4. The portion of the SOI variability thus reconstructed contains 43% of the total variance. It captures the quasi-oscillatory behavior isolated by these two leading pairs, with its two distinct near-periodicities.

Figure 6: *Partial reconstructions of the SOI time series.*

It is clear that the partial SOI reconstruction (heavy solid curve in Fig. 6a) is smooth and represents the essential part of the interannual variability in the monthly SOI data (light solid). Each of the two pairs of RCs, 1-2 and 3-4, can be thought of as retracing a ghost limit cycle in the phase space of the tropical climate system. These two limit cycles can then be said to form the robust skeleton of the attractor. It is unlikely that a time series of a few hundred points, like the SOI or other typical climatic time series, will suffice to capture the attractor's fine structure [*Vautard and Ghil*, 1989].

This robust skeleton, however, provides sufficient information for most practical purposes. In particular, warm events (El Niños) and cold ones (La Niñas) over the eastern Tropical Pacific are captured quite well, during the 57.5 years of record, as minima and maxima of the partially reconstructed SOI. We check this statement by comparing the heavy solid curve in Fig. 6a with the vertical arrows along the figure's upper and lower abscissae.

These arrows correspond to strong (heavy arrows) or moderate (light arrows) El-Niño/ Southern-Oscillation (ENSO) events. The events are defined subjectively, calendar quarter by calendar quarter, from reanalyzed surface temperature data produced at NCEP and the U. K. Meteorological Office for the years 1950–1999 (see

http://www.cpc.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.html).

Large positive peaks in the partial SOI reconstruction, i.e., those that exceed one standard deviation, match the strong La Niñas quite well (downward-pointing arrows on the upper abscissa). The only exceptions are the 1950 and 1971 cold events, which were of moderate strength, and the weak 1996 La Niña.

The same good match obtains between the large negative peaks in the figure's heavy solid curve, i.e., those that exceed one standard deviation, and the strong El Niños (upward-pointing arrows on the lower abscissa). The only notable exception is the large peak in 1977–78, which was classified subjectively as a weak warm event. The 1957–58 and 1991–92 events appear as moderate-size minima in the partial SOI reconstruction. They are included in the NCEP list as strong El Niños for one (January–March 1958) or two (January–June 1992) seasons, but neither was strong during the second half of the calendar year. Thus the only discrepancies among the oscillatory part of the SOI, based on RCs 1–4, and the subjective NCEP classification is in the intensity — moderate vs. strong, or vice-versa — of a few events.

Earlier SSA results do support the present emphasis on the doubly-periodic character of ENSO phenomena. They include the analyses of *Rasmusson et al.* [1990] for sea-surface temperatures and near-surface zonal winds in the tropical Indo-Pacific belt, those of *Keppenne and Ghil* [1992] for a slightly different treatment of the SOI, as well as those of *Jiang et al.* [1995a] for sea-surface temperatures and of *Unal and Ghil* [1995] for sea-level heights in the tropical Pacific. In all these data sets and SSA analyses, a quasi-biennial and a lower-frequency, quasi-quadrennial oscillatory pair were reliably identified among the leading SSA

eigenelements.

Shown in Fig. 6b is also a filtered version of RC-5, which captures well the small but significant long-term trend of the SOI time series in Fig. 2. To eliminate the oscillatory component apparent in PC-5 (Fig. 5c), we applied SSA with the same 60-month window to the full RC-5. The two leading eigenmodes correspond to a pure trend, shown in Fig. 6b, while the second eigenpair corresponds to a quasi-biennial oscillation (not shown). The SOI trend in Fig. 6b agrees, up to a point, with the one captured by the multi-taper reconstruction in Sec. 3.4.2 (see Fig.16a there). Given the recent interest in the interdecadal variability of ENSO, we postpone a further discussion of this result for the moment when its multi-taper version is also in hand.

Reliable S/N separation and identification of oscillatory pairs is not always as easy as in the case of interannual climate variability in the tropical Pacific. Global surface-air temperatures, for instance, present a considerably more difficult challenge for identifying interannual and interdecadal oscillations. *Elsner and Tsonis*'s [1991] excessive reliance on eigenvalue rank-order as a criterion of significance in SSA has led to considerable confusion in this case (see *Allen et al.* [1992a, b]).

Reliable identification of the true signal conveyed by a short, noisy time series and of the oscillatory components within this signal requires effective criteria for statistical significance, which are treated in the next subsection. Subject to these caveats, a clean signal — obtained by partial reconstruction over the correct set of indices $\mathcal{K}$ — provides very useful information on the underlying system, which is often poorly or incompletely known.

Such a signal can then be analyzed further, both visually and by using other spectral-analysis tools that are described in Sec. 3. The Maximum Entropy Method (MEM), which we describe in Sec. 3.3, works particularly well on signals so enhanced by SSA [*Penland et al.*, 1991].

2.3 Monte Carlo SSA

In the process of developing a methodology for applying SSA to climatic time series, a number of heuristic [*Vautard and Ghil*, 1989; *Ghil and Mo*, 1991a; *Unal and Ghil*, 1995] or Monte Carlo [*Ghil and Vautard*, 1991; *Vautard et al.*, 1992] methods have been devised for S/N separation or the reliable identification of oscillatory pairs of eigenelements. They are all essentially attempts to discriminate between the significant signal as a whole, or individual pairs, and *white noise*, which has a flat spectrum. A more stringent "null hypothesis" [*Allen*, 1992] is that of *red noise*, since most climatic and other geophysical time series tend to have larger power at lower frequencies [*Hasselmann*, 1976; *Mitchell*, 1976; *Ghil and Childress*, 1987].

For definiteness, we shall use here the term of red noise exclusively in its narrow sense, of an AR(1) process given by Eq. (1.1) with $M = 1$ and $a_1 < 0$, as required by weak or *wide-sense stationarity* (see Appendix A for an exact definition). Other stochastic processes that have a continuous spectral density $S(f)$ which decreases monotonically with frequency f will be called "warm-colored."

The power spectrum $S(f)$ of the AR(1) process is given by (e.g., *Chatfield* [1984])

$$S(f) = S_0 \frac{1 - r^2}{1 - 2r \cos(2\pi f/f_N) + r^2}. \tag{2.9}$$

Here $0 < S_0 < \infty$ is the average value of the power spectrum, related to the white-noise variance σ^2 in Eq. (1.1) by

$$S_0 = \frac{\sigma^2}{1 - r^2}, \tag{2.10}$$

while r is the lag-one autocorrelation, $0 < r < 1$, and the *Nyquist frequency* $f_N = 2/\Delta t$ is the highest frequency that can be resolved for the sampling rate Δt. Note that in Eqs. (1.1) and (2.1) we have used $\Delta t = 1$ for simplicity and without loss of generality, since Δt can always be redefined as the time unit. It is useful, at this point, to recall — for clarity's sake

— that it not necessary to do so. The characteristic decay time scale τ of the AR(1) noise can be estimated by

$$\tau = -\frac{\Delta t}{\log r}. \tag{2.11}$$

In general, straightforward tests can be devised to compare a given time series with an idealized noise process: the continuous spectrum of such a process is known to have a particular shape, and if a particular feature of the data spectrum lies well above this theoretical noise spectrum, it is often considered to be statistically "significant." A single realization of a noise process can, however, have a spectrum that differs greatly from the theoretical one, even when the number of data points is large. It is only the (suitably weighted) average of such sample spectra over many realizations that will tend to the theoretical spectrum of the ideal noise process. Indeed, the Fourier transform of a single realization of a red-noise process can yield arbitrarily high peaks at arbitrarily low frequencies; such peaks could be attributed, quite erroneously, to periodic components.

More stringent tests have to be used, therefore, to establish whether a time series can be distinguished from red noise or not. *Allen* [1992] devised such a test that compares the statistics of simulated red-noise time series with those of a given climatic time series. The principle of this test is implicit in some of *Broomhead and King*'s [1986a] ideas. The application of SSA in combination with this particular Monte Carlo test against red noise has become known as "Monte Carlo SSA" (MC-SSA; see *Allen and Smith* [1994, 1996]).

MC-SSA can be used, more generally, to establish whether a given time series can be distinguished from other well-defined processes. We only present here, for the sake of brevity and clarity, the original test against an AR(1) process. *Allen* [1992] proposes, in fact, to estimate the mean X_0 of the process at the same time as the other parameters. We rewrite, therefore, Eq. (1.1) here for the particular case at hand as

$$X(t) = a_1[X(t-1) - X_0] + \sigma\xi(t) + X_0; \tag{2.12}$$

here, as in (1.1), $\xi(t)$ is a Gaussian-distributed white-noise process with zero mean and unit variance.

When testing against the process (2.12), the first step in MC-SSA is to estimate the mean X_0 and the coefficients a_1 and σ from the time series $X(t)$ by using a maximum-likelihood criterion. *Allen and Smith* [1996] provide low-bias estimators that are asymptotically unbiased in the limit of large N and close to unbiased for series whose length N is at least an order of magnitude longer than the decorrelation time $\tau = -1/\log r$. *Cochrane and Orcutt* [1949] have shown that, when τ is not very small relative to N, the use of a crude estimate for the mean X_0 — i.e., "centering" the time series first — can lead to severe biases in the subsequent estimation of a_1. This is not the case for the SOI time series used here, as $\tau \ll N$ for it. Hence we have used an SOI time series, based on the data in Fig. 2, that has been centered.

Based on the estimated values $\hat{X}_0$, $\hat{a}_1$ and $\hat{\sigma}$ of these parameters, an ensemble of simulated red-noise data is generated and, for each realization, a covariance matrix $\mathbf{C}_R$ is computed. In the nonlinear dynamics literature, such simulated realizations of a noise process are often called *surrogate data* [*Drazin and King*, 1992; *Ott et al.*, 1994].

The covariance matrices of the surrogate data are then projected onto the eigenvector basis $\mathbf{E}_X$ of the original data by using Eq. (2.4) for their SVD,

$$\mathbf{\Lambda}_R = \mathbf{E}_X^{\mathrm{t}} \mathbf{C}_R \mathbf{E}_X. \tag{2.13}$$

Since (2.13) is not the SVD of the particular realization $\mathbf{C}_R$, the matrix $\mathbf{\Lambda}_R$ is not necessarily diagonal, as it is in Eq. (2.4). Instead, $\mathbf{\Lambda}_R$ measures the resemblance of a given surrogate set with the original data set of interest. The degree of resemblance can be quantified by computing the statistics of the diagonal elements of $\mathbf{\Lambda}_R$. The statistical distribution of these elements, determined from the ensemble of Monte Carlo simulations, gives confidence

intervals outside which a time series can be considered to be significantly different from a random realization of the process (2.12). For instance, if an eigenvalue λ_k lies outside a 90% noise percentile, then the red-noise null hypothesis for the associated EOF (and PC) can be rejected with this level of confidence. Otherwise, that particular SSA component of the time series cannot be considered as significantly different from red noise. Additional problems posed by the multiplicity of SSA eigenvalues and other finer points are also discussed by *Allen* [1992] and by *Allen and Smith* [1996].

As the next step in the analysis of our SOI time series, we apply an MC-SSA noise test to it. In order to enhance the readability of the diagrams for the SSA spectra in the presence of MC-SSA error bars, we associate a dominant frequency with each EOF detected by SSA, as suggested by *Vautard et al.* [1992], and plot in Fig. 7 the eigenvalues (diamonds) vs. frequency, following *Allen and Smith* [1996].

Figure 7: *Monte Carlo singular spectrum of the SOI time series.*

Such a plot is often easier to interpret, with respect to the MC-SSA error bars, than plotting vs. the eigenvalue's rank k as in Fig. 3. Care needs to be exercised, however, since the dominant-frequency estimate may be ambiguous or uncertain, due to the possible anharmonicity of the EOFs, especially for low frequencies. This is patently the case for the 5th mode in Figs. 3–5. This mode appears at zero frequency in Fig. 7, while we know very well that it has a quasi-biennial oscillatory component besides its capturing the SOI's nonlinear trend.

The error bars shown in Fig. 7 for each EOF represent 90% of the range of variance found in the state-space direction defined by that EOF in an ensemble of 1000 red-noise realizations, i.e., it denotes the interval between the 5th and 95th percentile. Hence, eigenvalues lying outside this interval are relatively unlikely (at the 10% level) to be due merely to the red-noise process (2.12) against which they are being tested.

The high values in Fig. 7 exhibit a significant quasi-biennial oscillation and an oscillatory component with a period of about 50 months. The low values near 1 cycle yr^{-1} are due to the fact that the seasonal cycle has been removed prior to the analysis, and the consequent suppression of power near annual-cycle periods has not been taken into account in the noise-parameter estimation. *Allen and Smith* [1996] recommend that, as far as possible, seasonal cycles should not be removed prior to the analysis, but that their presence should be taken into account explicitly in the parameter estimation (see also *Jiang et al.* [1995a] and *Unal and Ghil* [1995]). This recommendation has to be weighted against two related considerations. First, the reliable identification of a given periodicity becomes harder as the number of periodicities to be estimated increases and, second, the seasonal cycle and its physical causes are fairly well known and understood.

The MC-SSA algorithm described above can be adapted to eliminate known periodic components and test the residual against noise. This adaptation can provide better insight into the dynamics captured by the data. Indeed, known periodicities — like orbital forcing on the Quaternary time scale or seasonal forcing on the intraseasonal-to-interannual one — often generate much of the variance at the lower frequencies manifest in a time series and alter the rest of the spectrum. *Allen* [1992] and *Allen and Smith* [1996] describe this refinement of MC-SSA which consists in restricting the projections given by Eq. (2.13) to the EOFs that do *not* account for known periodic behavior.

Monte Carlo simulation is a robust, flexible and nonparametric approach to assessing the significance of individual eigenmodes in SSA. Since it can be computationally intensive, *Allen and Smith* [1996] suggest a much faster, albeit parametric alternative. In an appendix, they also clarify the relationship between confidence intervals from MC-SSA and earlier heuristic approaches (e.g., *Vautard and Ghil* [1989], *Ghil and Mo* [1991a] and *Unal and Ghil* [1995]).

MC-SSA provides, furthermore, the means of evaluating the significance of frequency separation between apparently distinct spectral peaks, without appealing to the application

of additional spectral methods. The associated "bandwith" is $1/M$ and is discussed further in a multi-channel context in Sec. 4.2. Given the fact that the assignment of frequency to an eigenpair is not entirely unique (see discussion of the zero-frequency component in Fig. 7), we still recommend the application of other spectral methods for the detailed study of spectral content, in addition to SSA and MC-SSA.

2.4   Multi-scale SSA and wavelet analysis

Wavelet analysis has become a basic tool for the study of intermittent, complex and self-similar signals, because it works as a mathematical microscope that can focus on a specific part of the signal to extract *local* structures and singularities [*Strang*, 1989; *Meyer*, 1992, 1993; *Daubechies*, 1992]. In climate dynamics [*Meyers et al.*, 1993; *Weng and Lau*, 1994; *Torrence and Compo*, 1998] and geophysics [*Kumar and Foufoula-Georgiou*, 1997] wavelets have been used mostly to follow changes in frequency of one or more periodic signals. While SSA follows amplitude and phase modulation of a signal easily (see Sec. 2.2 here, as well as *Plaut and Vautard* [1994] and *Moron et al.* [1998]), small changes in frequency are subsumed typically in a single pair of SSA eigenmodes.

A wavelet transform requires the choice of an analyzing function or "mother wavelet" ψ that has general admissibility properties [*Meyer*, 1992; *Daubechies*, 1992], as well as the more specific property of time and frequency localization, i.e., ψ and its Fourier transform $\mathcal{F}\psi$ must decay rapidly outside a given interval. Functions ψ based on a Gaussian, $\psi(x) = \exp(-x^2)$, first proposed in this context by *Gabor* [1946], possess the localization property even though they do not verify the admissibility condition that their integral over the real line $\mathbb{R}$ vanish [*Delprat et al.*, 1992].

A ψ-wavelet transform W_ψ in continuous time and frequency is simply a projection of a

signal $X(t)$, $-\infty < t < \infty$, onto b-translated and a-dilated versions of ψ:

$$W_\psi(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} X(t)\psi\left(\frac{t-b}{a}\right) dt. \qquad (2.14)$$

If most of ψ is concentrated between $[-1, 1]$, say (up to a rescaling), then Eq. (2.14) is clearly an analysis of X in the interval $[b-a, b+a]$. Using the successive derivatives $\psi^{(n)}$ of a given mother wavelet ψ in Eq. (2.14) is equivalent (up to a normalization factor) to a ψ-analysis of the successive derivatives of the time series X; this is easy to see through an integration by parts.

The original signal, or a filtered version of it, can be reconstructed from the family of wavelet transforms. Hence, for scale values a in an interval I, a reconstructed version X_I of the signal $X(t)$ is:

$$X_I(t) = A_\psi \int_{a \in I} \int_{b=-\infty}^{\infty} W_\psi(a, b)\psi\left(\frac{t-b}{a}\right) \frac{dadb}{a^2}; \qquad (2.15)$$

A_ψ is a normalization factor which only depends on the mother wavelet ψ. This formulation is essentially a bandpass filter of X through I; if I is the positive real line, $I = \mathbb{R}^+$, then $X_I(t) = X(t)$. Note that the Gaussian $\psi(x) = \exp(-x^2/2)$ itself cannot be used in the reconstruction formula (2.15), because it does not satisfy $\int_{-\infty}^{\infty} \psi(x)dx = 0$, although its first derivative does. The forward transform of Eq. (2.14), however, is well-defined and the Gaussian can be used as an analyzing tool [*Arneodo et al.*, 1993; *Yiou et al.*, 2000].

A large number of wavelet bases have been introduced to satisfy the conflicting requirements of completeness, localization in both time and frequency, and orthogonality or — for nonorthogonal bases — limited redundancy. To provide an optimal multi-scale decomposition of a given signal, an automatic, time-varying adjustment of the mother wavelet's shape may be desirable. This could replace the current practice of searching through extensive "libraries" of mother wavelets (e.g., http://www.mathsoft.com/wavelets.html). To

provide such a data-adaptive variation in basis with time, *Yiou et al.* [2000] have introduced *multi-scale SSA*.

Systematic comparisons between SSA, the wavelet transform, and other spectral-analysis methods have been carried out by *Yiou et al.* [1996] and in Table 1 of *Ghil and Taricco* [1997]. Further analogies between certain mathematical features of SSA and wavelet analysis were mentioned by *Yiou* [1994]. Table 2 here summarizes the most useful mathematical parallels between the two time-series analysis methods.

Table 2: *Analogy between SSA and wavelet analysis.*

In SSA, the largest scale at which the signal X is analyzed in Eq. (2.6) is approximately $N\Delta t$, the length of the time series, and the largest period is the window width $M\Delta t$. As a consequence, the EOFs ρ_k contain information from the whole time series, as in the Fourier transform.

In order to define a local SSA, the SSA methodology was extended by using a time-frequency analysis within a running time window whose size W is proportional to the order M of the covariance matrix. Varying M, and thus W in proportion, a multi-scale representation of the data is obtained. A *local* SSA is performed by sliding windows of length $W \leq N\Delta t$, centered on times $b = W/2, \ldots, N\Delta t - W/2$, along the time series. This method is useful when the local variability — assumed to be the sum of a trend, statistically significant variability, and noise — changes in time.

A priori, the two scales W and M can vary independently, as long as W is larger than $M\Delta t$, $W/(M\Delta t) \geq \alpha > 1$, and α is large enough [*Vautard et al.*, 1992]. In the wavelet transform, however, the number of oscillations of the mother wavelet is fixed and independent of the scale (width) of the analyzing wavelet. In this spirit, *Yiou et al.* [2000] fixed the ratio $W/M = \alpha\Delta t$ and relied therewith on the oscillation property of the EOFs to provide a fixed number of zeroes for the data-adaptive "wavelet" ρ_k of each local SSA analysis. They

used $\alpha = 3$ in most of their calculations, as well as only the one or two leading EOFs, ρ_1 and/or ρ_2, on each W-interval. This provides an analysis at a fixed scale W (see Table 2). Sampling a set of W values that follow a geometrical sequence, for instance in powers of 2 or 3, provides a multi-scale analysis very similar to the wavelet transform.

For a given position b and fixed W, we thus obtain local EOFs that are the direct analogs of analyzing wavelet functions. The number of EOF oscillations increases roughly with order and the zeroes of ρ_{k+1} separate those of ρ_k; this emulates an important property of successive analyzing wavelets. The first EOF thus corresponds approximately to an analyzing wavelet function with a single extremum and no zero inside the window, for instance the Gaussian or the "Mexican hat"; such a basic wavelet is denoted by $\psi = \psi^{(0)}$ in Table 2. The second EOF has a single zero and is reminiscent of the first derivative of the Gaussian, denoted by $\psi^{(1)}$ in Table 2, and so on. *Vautard and Ghil* [1989] demonstrated this oscillation property of the EOFs for red noise in continuous time and *Allen* [1992] did so for the same type of noise process in discrete time. Appendix B of *Yiou et al.* [2000] provides a more general proof that is based on the concept of total positivity for lag-covariance matrices.

For each b and each EOF ρ_k, it is possible to obtain local PCs $\mathbf{A}_k$ and RCs $\mathbf{R}_k$ [see Eqs. (2.6) and (2.7)]. The kth PC at time b is

$$A_k^b(t) = \sum_{j=1}^{M} X(t+j-1)\rho_k^b(j), \tag{2.16}$$

and the corresponding RC is

$$R_k^b(t) = \frac{1}{M_t} \sum_{j=1}^{M} A_k^b(t-j)\rho_k^b(j), \tag{2.17}$$

with $b - W/2 \leq t \leq b + W/2$. The crucial difference between this local version and global SSA is that the RCs are obtained here from local lag-covariance matrices. As b varies from $W/2$ to $N\Delta t - W/2$, this implies that the RCs will be truncated near the edges of the time

series.

We thus see that the local SSA method provides simultaneous "wavelet transforms" of the data by a set of analyzing wavelet functions, corresponding to the M different EOFs of the lag-covariance matrix. When $W = \alpha M \Delta t$ is small, local SSA provides a small-scale analysis of the signal with a few distinct analyzing functions, i.e., a small subset of EOFs indexed by k. This is reasonable as there are not many possible structures at scales that approach the sampling time scale. On the other hand, at large scales, local SSA can also provide the simultaneous analysis by many different analyzing mother wavelet functions, $\{\rho_k : 1 \leq k \leq M\}$, and thus reflect the large complexity of the structures that is possible over the entire time series.

The most important property of this local SSA analysis is that the analyzing functions are *data adaptive.* In other words, the shape of these analyzing functions is not imposed a priori, like in a wavelet analysis, but explicitly depends on the time series itself. For instance, an oscillatory behavior could be followed in a given time series by white or colored noise and then by deterministically intermittent behavior. These changes in behavior could indicate regime transitions that the system which generates the signal underwent while under observation. If so, an analyzing wavelet which is adapted to each section of the signal will definitely help follow such regime transitions in time.

Yiou et al. [2000] performed multi-scale SSA on the monthly SOI data for the years 1933–1996 (see Sec. 2.2). The parameters were $\alpha = 3$ and geometric scale increments of 2. They computed, moreover, an "instantaneous" frequency by least-square fitting a sine wave to each local EOF of interest, as done in Monte-Carlo SSA for the global EOFs. The instantaneous frequency can also be obtained from a complex wavelet transform [*Delprat et al.*, 1992; *Farge*, 1992], by using information on the phase of the transform.

The analysis of Yiou and colleagues did not reveal any evidence of self-similarity or fractality in the SOI. Instead, they find a preferred scale of variability between 3 and 5 years

(not shown), which corresponds to ENSO's low-frequency mode (see Sec. 2.2 of the present paper and references there). The first two local EOFs are consistently paired and in phase quadrature, which shows that the nonlinear oscillation associated with this mode is robust and persists throughout the sixty-odd years being examined.

The computation of the instantaneous frequency in multi-scale SSA allows one to detect an abrupt frequency shift of ENSO's low-frequency mode near 1960 (Fig. 8). The characteristic periodicity goes from 57 months (between 1943 and 1961) to 39 months (between 1963 and 1980). A decrease in period in the early 1960s was observed already by *Moron et al.* [1998] in Tropical Pacific sea-surface temperatures, by using multi-channel (global) SSA, and by *Wang and Wang* [1996] in a sea-level pressure record at Darwin, using wavelet methods. *Mann and Park* [1996b; Fig. 9 there] also observed a "pinch-out" in the amplitude of the quasi-biennial oscillation in the 1960s.

Figure 8: *Instantaneous SOI frequencies.*

Moron et al. [1998] noticed, on the one hand, a change in the low-frequency mode's periodicity in the early 1960s, by using multi-channel SSA (see Sec. 4.2 here) with different window widths (from $M = 72$ months to 168 months) on sea-surface temperature fields for 1901–1994 (their Fig. 2 and Table 4). On the other hand, these authors found that the trend of the sea-surface temperatures in the Tropical Pacific exhibited an increase from 1950 on (their Fig. 4). They related this surface trend to a change in the parameters — such as the thermocline depth along the Equator — of the coupled ocean-atmosphere oscillator responsible for ENSO [*Ghil et al.*, 1991; *Neelin et al.*, 1994, 1998].

The frequency of a linear oscillator always changes smoothly as a function of the coefficients in the linear ODE that governs it. That is typically the case for nonlinear oscillators that depend on a single parameter as well [*Stoker*, 1950]. Two notable exceptions involve period doubling [*Feigenbaum*, 1978; *Eckmann*, 1981] and the Devil's staircase [*Feigenbaum*

et al., 1982; *Bak*, 1986]. In both these routes to chaos, frequencies change abruptly as one parameter, in the former, or two, in the latter [*Schuster*, 1988], change gradually.

Moron et al. [1998] hypothesized, therefore, that — as one of the parameters of the ENSO oscillator crosses a threshold value in the early 1960s — the period of the seasonally forced climatic oscillator jumps from one step of the Devil's staircase to another. This would confirm an important aspect of recent theoretical work on ENSO's oscillatory regularity, as well as on the irregular occurences of major warm and cold events [*Chang et al.*, 1994; *Jin et al.*, 1994; *Tziperman et al.*, 1994].

Yiou et al. [2000] compared the results in Fig. 8 here with those of *Moron et al.* [1998], using multi-channel SSA (see Sec. 4.2 in the present paper), and those of *Wang and Wang* [1996]. The latter authors used both the standard wavelet transform with a Morlet mother wavelet and *Mallat and Zhang*'s [1993] *waveform transform* to analyze the Darwin sea-level pressures for 1872–1995. The frequency jump in Fig. 8 is much sharper than when using either global SSA or wavelet-type methods (not shown here). It is also sharper than the changes in SOI variability described by *Torrence and Compo* [1998], who used both Morlet and Mexican-hat basis functions for their analysis. Multi-scale SSA can thus provide sharper regime transitions in the evolution of a nonlinear system than either global SSA or a wavelet analysis that uses a fixed set of basis functions.

3   Spectral Analysis Methods

3.1   Generalities and definitions

Both deterministic [*Eckmann and Ruelle*, 1985] and stochastic [*Hannan*, 1960] processes can, in principle, be characterized by a function of *frequency f* (instead of *time t*). This function $S(f)$ is called the power spectrum in the engineering literature or the spectral density in the mathematical one (see Sec. 1.1 for our use of the terms). Thus, a very irregular motion

possesses a smooth and continuous spectrum, which indicates that *all* frequencies in a given band are excited by such a process. On the other hand, a purely periodic or quasi-periodic process is described by a single line or a (finite) number of lines in the frequency domain. Between these two extremes, nonlinear deterministic but "chaotic" processes can have spectral peaks superimposed on a continuous and wiggly background [*Ghil and Childress*, 1987, Sec. 12.6; *Ghil and Jiang*, 1998].

In theory, for a spectral density $S(f)$ to exist and be well defined, the dynamics generating the time series has to be ergodic and allow the definition of an invariant measure, with respect to which the first and second moment of the generating process are computed as an ensemble average. The reason for the spectral theory of linear random processes being more familiar is simply that the construction of such an invariant measure is easier for such processes [*Hannan*, 1960; *Priestley*, 1981a]. In practice, the distinction between deterministically chaotic and truly random processes via spectral analysis can be as tricky as the attempted distinctions based on the dimension of the invariant set (see Sec. 1.2 and references therein). In both cases, the difficulty is due to the shortness and noisiness of climatic time series.

Given the existence and invariance of an appropriate measure, the rigorous definition of the process's lag-correlation function involves a Lebesgue integral over that measure. On the spectral side of the Wiener-Khinchin equality [see Eqs. (3.2) and (3.3) in the next subsection], one then has to use a Stieltjes integral — rather than the more familiar Riemann integral — with respect to the usual measure on $\mathbb{R}$ to compute the cumulative power $P(f)$ of the process over all frequencies between $-\infty$ and the given frequency f (e.g., *Hannan* [1960]). The spectrum is continuous at all points f where the left and right limits of this integral are equal, i.e. $P(f-) = P(f+)$. At such points $S(f) = dP/df$. Where this is not the case, an upward jump in $P(f)$ occurs, i.e., $P(f+) = P(f-) + p(f)$, where $p(f) > 0$. Such a discontinuity in $P(f)$ corresponds, loosely speaking, to a Dirac δ-function in the "spectral

density." The *line spectrum* is the sum of these δ-functions that contributes to $S(f)$, along with the previously described *continuous spectrum* (e.g., *Priestley* [1981a]).

In practice, spectral analysis methods attempt to estimate either the continuous part of the spectrum or the "lines" or both. The lines are often estimated from discrete and noisy data as more-or-less sharp "peaks." The estimation and dynamical interpretation of the latter, when present, is often more robust and easier to understand than the nature of the processes that might generate the broad-band background, whether deterministic of stochastic.

The numerical computation of the power spectrum of a random process is an ill-posed inverse problem [*Jenkins and Watts*, 1968; *Thomson*, 1982]. For example, a straightforward calculation of the discrete Fourier transform of a random time series — which has a continuous spectral density — will provide a spectral estimate whose variance is equal to the estimate itself [*Jenkins and Watts*, 1968; *Box and Jenkins*, 1970]. In the remainder of this section, we outline three techniques to reduce this variance that are commonly used for the spectral analysis of climatic time series and point out their respective properties, advantages and failings.

Each one of the three techniques outlined in Secs. 3.2–3.4 below — Blackman-Tukey, maximum entropy, and multi-taper — can provide confidence intervals for the estimates it produces. Still, these "error bars" are based on certain assumptions about the process generating the time series that is being analyzed. These assumptions are rarely, if ever, met in practice by the physical processes one wishes to examine. Therefore, we highly recommend to apply several independent techniques to a given time series before drawing any conclusion about its spectrum.

3.2 Classical spectral estimates

A generic problem of time series analysis is the finiteness of the time interval on which the series is known. Estimating the spectrum of $\{X(t) : t = 1, \ldots, N\}$ by a discrete Fourier transform yields its periodogram, which is the Fourier transform squared. Doing so corresponds to convolving the true spectrum with the spectrum of a box-car function; this induces power leakage, due to the lobes of the Fourier transform of the box car. The resulting systematic distortion of the spectrum adds to the problem of the variance of the spectral estimate.

3.2.1 The periodogram and the correlogram

The so-called *periodogram* estimate $\hat{S}_X(f)$ of $S(f)$ is given by

$$\hat{S}_X(f) = \frac{\Delta t}{N} \left| \sum_{t=1}^{N} X(t) e^{-2\pi i f t \Delta t} \right|^2 ; \tag{3.1}$$

the use of $i = \sqrt{-1}$ in this section should not be confused with the time index i used in Sec. 2. The estimate $\hat{S}_X(f)$ is itself a random-valued function. Its mean leads to the correct S_X as the length $N\Delta t$ of the sample series tends to infinity, but the standard deviation in this limit $N \to \infty$ is as large as the mean $S_X(f)$ being estimated [*Hannan*, 1960]. An estimate for which the variance does not tend to zero as the sample size tends to infinity is called statistically *inconsistent*. Moreover, for any finite N, the sample periodogram $\hat{S}_X$ is a *biased* estimate of S_X, i.e., its mean over distinct realizations of the same process does not equal S_X in general.

Blackman and Tukey [1958] gave an alternative method to estimate the power spectrum of a given time series $\{X(t) : t = 1, \ldots, N\}$. Their method helps reduce the estimate's variance and bias, and attenuate the leakage effects of the periodogram [*Chatfield*, 1984]. The method's starting point is the so-called Wiener-Khinchin or Bochner-Khinchin-Wiener theorem, which states that the spectral density S_X is equal to the Fourier transform of the

auto-correlation function ϕ_X [*Jenkins and Watts*, 1968; see also Secs. 1.1 and 3.1 here]:

$$\phi_X(k) = \frac{\mathcal{E}\{X(t+k)X(t)\}}{\mathcal{E}\{X^2(t)\}}, \tag{3.2}$$

where $\mathcal{E}$ is the expectation operator.

Hence the power spectrum $S_X(f)$ of $X(t)$ can be estimated by the Fourier transform $\tilde{S}_X(f)$ of an estimate $\hat{\phi}_X$ of the true ϕ_X, called the *correlogram*:

$$\tilde{S}_X(f) = \sum_{k=-(N-1)}^{N-1} \hat{\phi}_X(k)e^{-2\pi ifk}. \tag{3.3}$$

This estimate is also biased and inconsistent, as it stands. Many methods have been devised to reduce the bias, variance and leakage in the *direct* estimate (3.1) or the *indirect* estimate (3.3). A succinct summary is provided by *Ghil and Taricco* [1997] and details are given in a number of text books and reference books [*Chatfield*, 1984; *Kay*, 1988].

A simple way to reduce the bias and variance is to average the estimate over non-overlapping bins in the frequency domain, in order to smooth the periodogram:

$$\bar{S}_X(f_k) = \sum_{j=-M}^{M} g_j \hat{S}_X(f_{k-j}), \tag{3.4}$$

where $\{f_k : k = 1, \ldots, N_f\}$ is a discrete set of frequencies, with $N_f \leq N$, and $\{g_j\}$ is a set of $2M + 1$ smoothing weights, with $M < N/2$. With well-chosen weights, the estimate $\bar{S}_X$ becomes *consistent*, i.e., its variance converges to 0 as M and N tend to infinity [*Chatfield*, 1984; *Percival and Walden*, 1993]. In (3.4), the choice of M is dictated by a trade-off between frequency resolution — the larger M the better — and the estimated variance, which is proportional to M/N — hence the smaller M the better [*Kay*, 1988; see also the discussion of window width for SSA in Sec. 2.2 here]. Therefore, a rule of thumb is to take M no larger than $N/5$ or $N/10$, to avoid spurious results from high-variance estimates.

The counterpart of this approach in the indirect estimation procedure is to use a *lag window* $\{W_\mu(k) : k = -(N-1), \ldots, N-1\}$, with a smoothing parameter μ. An example of such a window is given in Eq. (3.7) below (see Sec. 3.2.2). The power spectrum estimate then becomes:

$$\bar{\bar{S}}_X(f) = \sum_{k=-(N-1)}^{N-1} W_\mu(k)\hat{\phi}_X(k)e^{-2\pi i f k}. \tag{3.5}$$

With a careful choice of the shape of the window $\{W_\mu(k)\}$ for fixed μ and of the parameter value μ, the estimate $\bar{\bar{S}}_X$ enjoys the same statistical properties as $\bar{S}_X$ in Eq. (3.4), i.e., consistency and lower bias. It turns out that the two procedures are formally equivalent and that a lag window $\{W_\mu(k)\}$ can be deduced from weights $\{g_j\}$ in Eq. (3.4), and vice-versa [*Percival and Walden*, 1993]. Computationally, it is more efficient to calculate (3.2) and (3.5) than (3.1) and (3.4). Hence the indirect or correlogram method — associated with the names of *Blackman and Tukey* [1958] — is most often the algorithm of choice.

The windows traditionally have a compact support of length $m < N$, so that Eq. (3.5) becomes, with $\mu = m$:

$$\tilde{S}_X(f) = \sum_{k=-(m-1)}^{m-1} W_m(k)\hat{\phi}_X(k)e^{-2\pi i f k}; \tag{3.6}$$

m is called the *truncation point* or *window width* in this standard formula. Such lag windows are heuristically chosen as modified cosine functions, cubic functions or tent functions [*Chatfield*, 1984]. Classical Blackman-Tukey estimation deals largely with the choice of window shape, often called "window carpentry," and of the window width m, often called "opening and closing" the window.

39

3.2.2 Bias and variance reduction

We illustrate the variance reduction properties of the lag window estimate in Fig. 9 with a Bartlett window of varying width m. The Bartlett window is simply:

$$W_m(k) = 1 - \frac{|k|}{m}, \tag{3.7}$$

with $k = 1, \ldots, m$. The smoothed estimate (heavy solid) clearly has a much smaller variance than the raw-periodogram estimate (light solid). Confidence levels relative to a pure red noise — or AR(1) process, which has a spectral slope that behaves like $[1 + (2\pi f)^2]^{-1}$, cf. Eq. (2.9) — can be calculated for each window shape [*Jenkins and Watts*, 1968], but the spectral resolution is generally poor if the number N of data points, and hence m, is low.

Figure 9: *Classical spectral estimates of the SOI time series.*

It turns out that the Blackman-Tukey, or windowed-correlogram, method is quite efficient for estimating the continuous part of the spectrum, but is less useful for the detection of components of the signal that are purely sinusoidal or nearly so. The reason for this is twofold: the low resolution of this method and the fact that its estimated erros bars are still essentially proportional to the estimated mean $\tilde{S}_X(f)$ at each frequency f. We shall see in Sec. 3.4 how to choose, according to the *multi-taper* method (MTM), a set of optimal tapers that maximize the resolution. Moreover, MTM also allows the formulation of statistical-significance tests that are independent, in principle, of the amplitude of the sinusoidal component that is being estimated.

Vautard and Ghil [1989] showed the (data-adaptive) bandpass filtering effects of reconstructed components (RCs). In Fig. 10 we illustrate the filtering properties of SSA on the SOI time series, by removing the trend and "noise" components with SSA, and computing the smoothed correlogram of the residual time series. This filter enhances the components

with periodicities around 3.5 years and 5.6 months. These approximate periodicities will be confirmed and refined using the *maximum-entropy* method (MEM) and MTM in the next two subsections. Even with this SSA prefiltering, though, the lag-windowed correlogram is not able to distinguish a quasi-biennial from a quasi-quadrennial peak, as the MEM and MTM are able to do.

Figure 10: *Blackman-Tukey correlogram estimate of the SSA-filtered SOI.*

3.3   Maximum-entropy method (MEM)

This method is based on approximating the time series under study by a linear AR process (1.1) of order M, AR(M). It thus performs best when estimating line frequencies for a time series that is actually generated by such a process. Details can be found in *Burg* [1967] or *Childers* [1978].

Given a time series $\{X(t) : t = 1, \ldots, N\}$ that is assumed to be generated by a wide-sense stationary process with zero mean and variance σ^2, $M' + 1$ estimated auto-correlation coefficients $\{\hat{\phi}_X(j) : j = 0, \ldots, M'\}$ are computed from it:

$$\hat{\phi}_X(j) = \frac{1}{N + 1 - j} \sum_{t=1}^{N-j} X(t)X(t+j). \tag{3.8}$$

In the absence of prior knowledge about the process that generates the time series $X(t)$, M' is arbitrary and has to be optimized. The purpose of evaluating (3.8) is to determine the spectral density $\hat{S}_X$ that is associated with the most random, or least predictable, process that has the same auto-correlation coefficients $\hat{\phi}$. In terms of information theory [*Shannon*, 1949], this corresponds to the concept of *maximal entropy*, hence the name of the method.

In practice, one obtains estimates $\{\hat{a}_j : j = 0, \ldots, M'\}$ of the regression coefficients $a_j : j = 0, \ldots, M\}$ from the time series $X(t)$ by assuming that the latter is generated by an AR(M) process and that its order M equals M'. The autocorrelation coefficients $\hat{\phi}_X(j)$ are

41

computed according to Eq. (3.8) and used to form the same Toeplitz matrix $\mathbf{C}_X$ as in SSA [*Vautard and Ghil*, 1989; *Penland et al.*, 1991; *Vautard et al.*, 1992]. This matrix is then inverted using standard numerical schemes [*Press et al.*, 1988] to yield the estimated $\{\hat{a}_j\}$. The spectral density S_X of the true AR process with coefficients $\{a_j : j = 0 \ldots, M\}$ is given by

$$S_X(f) = \frac{a_0}{\left| 1 + \sum_{j=1}^{M} a_j e^{2\pi i j f} \right|^2},$$
(3.9)

where $a_0 = \sigma^2$ is the variance of the residual noise ξ in Eq. (1.1). Therefore, the knowledge of the coefficients $\{\hat{a}_j : j = 0, \ldots, M'\}$, determined from the time series $X(t)$, also yields an estimate $\hat{S}_X$ of the power spectrum.

An example of MEM estimates is given for the SOI time series in Fig. 11, using a number of lags $M' = 10$, 20 and 40. It is clear that the number of peaks increases with M'. Two distinct interannual peaks appear only for $M' = 40$. This separation between a quasi-biennial and a low-frequency peak is accompanied, unfortunately, by many spurious peaks at higher frequencies.

Figure 11: *MEM spectral estimates of the SOI time series.*

In general, if the time series is not stationary or otherwise not well approximated by an AR process — e.g., if it is generated by nonlinear dynamics — it is necessary to exercise great care in applying MEM. In such cases, cross-testing with the application of other techniques is especially important.

As the number of peaks in a MEM spectrum increases with M', regardless of the spectral content of the time series, an upper bound for M' is generally taken as $N/2$. Heuristic criteria have been devised to refine the choice of a reasonable M' [*Haykin and Kessler*, 1983; *Benoist*, 1986], based on minimizing the residual of a least-square fit between the AR approximation and the original time series [*Akaike*, 1969, 1974; *Haykin and Kesler* 1983].

Such "information-content" criteria, however, often tend to either underestimate [*Benoist*, 1986] or overestimate [*Penland et al.*, 1991] the order of regression of a time series, depending on its intrinsic characteristics.

The effects of the S/N enhancement performed by SSA decomposition (cf. Sec. 2) on MEM analysis are illustrated in Fig. 12, where the noise components identified by SSA were filtered out prior to MEM analysis. In this example, the power spectrum is much smoother than in Fig. 11. The regression order $M' = 10$ suffices to separate a quasi-biennial peak at about 2.4 years from a quasi-quadrennial peak at roughly 4.8 years, while no spurious peaks appear at high frequencies. By contrast, Blackman-Tukey spectra of the same time series (see Fig. 10 here and also *Rasmusson et al.* [1990]) fail to separate these two peaks with sufficient statistical confidence.

Figure 12: *MEM analysis of the SOI time series after SSA prefiltering.*

The subannual peak at 5.6 months, which is associated with the EOFs 10-11, is also much sharper here than in Fig. 10. It can be distinguished therewith reliably from the known semi-annual peak of tropical variability; in fact, the latter had been removed from the time series along with the seasonal cycle. A coupled-mode solution with spatial features that resemble an oceanic Kelvin wave and a period close to 6 months has been obtained in a hybrid ocean-atmosphere model with no seasonal cycle by *Neelin* [1990]. It provides a plausible physical mechanism for our 5.6-month peak.

3.4   Multi-taper method (MTM)

This method provides useful tools for the spectral estimation [*Thomson*, 1982; *Percival and Walden*, 1993] and signal reconstruction (e.g., *Park* [1992]) of a time series whose spectrum may contain both broad-band and line components. Like the classical periodogram and correlogram methods of Sec. 3.2, MTM is nonparametric, since it does not use an a priori,

parameter-dependent model of the process that generated the time series under analysis, like MEM does. MTM reduces the variance of spectral estimates by using a small set of tapers [*Thomson*, 1982; *Percival and Walden*, 1993] rather than the unique data taper or spectral window used by the classical methods. The data are premultiplied by orthogonal tapers constructed to minimize the spectral leakage due to the finite length of the time series and a set of independent estimates of the power spectrum is computed.

The optimal tapers or "eigentapers" belong to a family of functions known as discrete prolate spheroidal sequences (DPSS). They are defined as the eigenvectors of a suitable Rayleigh-Ritz minimization problem and were extensively studied by *Slepian* [1978]. More precisely, the tapers are the discrete set of eigenfunctions that solve the variational problem of minimizing leakage outside of a frequency band with half-bandwidth equal to pf_R, where $f_R = 1/(N\Delta t)$ is the *Rayleigh frequency*, Δt is the sampling interval, and p is a suitably chosen integer. Because the windowing functions or eigentapers are the specific solution to an appropriate variational problem, this method is less heuristic than traditional nonparametric techniques (see Sec. 3.2 here, as well as *Box and Jenkins* [1970] and *Jenkins and Watts* [1968]). Averaging over the (small) ensemble of spectra obtained by this procedure yields a better and more stable estimate — i.e., one with lower variance — than do single-taper methods [*Thomson*, 1990a].

Detailed algorithms for the calculation of the eigentapers are readily available [*Thomson*, 1982; *Percival and Walden*, 1993]. In practice, only the first $2p - 1$ tapers provide usefully small spectral leakage [*Slepian*, 1978; *Thomson*, 1982; *Park et al.*, 1987]. Thus, the number K of tapers used should be less than $2p - 1$ in any application of MTM.

The choice of the bandwidth $2pf_R$ and number of tapers K thus represents the classical trade-off between spectral resolution and the stability or variance-reduction properties of the spectral estimate [*Thomson*, 1982]. The case $p = 1$ and $K = 1$ is simply the single-tapered discrete Fourier transform (DFT) of *Blackman and Tukey* [1958]. For instrumental climate

records, with a typical length of a few hundred points, the choice $p = 2$ and $K = 3$ offers a good compromise between the required frequency resolution for resolving distinct climate signals (e.g., ENSO and decadal-scale variability) and the benefit of multiple spectral degrees of freedom, i.e., of reduced variance [e.g., *Mann and Park*, 1993]. Longer data sets permit the use of a greater number K of tapers while maintaining a desired frequency resolution. The optimal choice of p and K depends in general on the length and other properties of the time series under study.

We show in Fig. 13 the $K = 3$ leakage-resistant Slepian tapers for bandwidth parameter $p = 2$. In this example, the effective half-bandwidth is $2f_R$ for the spectral estimate centered at any particular frequency f_0. This half-bandwidth is a measure of spectral resolution and it equals, in the present case, twice that of an unsmoothed DFT. Thus a spectral estimate at the five-year periodicity of $f = 0.2$ cycle/yr for a time series with $N = 690$ months (57.5 years) averages the raw periodogram over the frequency interval between $f \approx 0.165$ and $f \approx 0.235$ cycle/yr. The variance of the spectral estimate at f_0 is decreased, on the other hand, threefold by averaging the three independent eigenspectra. This variance-resolution trade-off is clearly superior to conventional periodogram smoothing. The process of MTM spectral estimation is described in greater detail in the following subsection.

Figure 13: *The first $K = 3$ eigentapers for the SOI time series in Fig.2.*

Because of the improvements in its spectral estimation properties over the classical methods of Sec. 3.2, the MTM method has been widely applied to problems in geophysical signal analysis, including analyses of instrumental data on the atmosphere and oceans [*Kuo et al.*, 1990; *Ghil and Vautard*, 1991; *Mann and Park*, 1993, 1994, 1996*a, b*; *Lall and Mann*, 1995; *Mann et al.*, 1995*b*; *Thomson*, 1995], paleoclimate proxy data [*Chappellaz et al.*, 1990; *Thomson*, 1990*a, b*; *Berger et al.*, 1991; *Mann et al.*, 1995*a*; *Mann and Lees*, 1996; *Mommersteeg et al.*, 1995; *Park and Maasch*, 1993; *Yiou et al.*, 1991, 1994, 1995, 1997], geochemical tracer

data [*Koch and Mann*, 1996], and seismological data [*Park et al.*, 1987; *Lees*, 1995]. Time-frequency "evolutive" analyses based on moving-window adaptations of MTM have also been applied to paleoclimatic records and model simulations [*Yiou et al.*, 1991; *Birchfield and Ghil*, 1993; *Mann et al.*, 1995a; *Mann and Park*, 1996b].

3.4.1   Spectral estimation

MTM can provide estimates of both the line components and the continuous background of the spectrum. Once the tapers $w_k(t)$ are computed for a chosen frequency bandwidth, the total power spectrum S_X can be estimated by averaging the individual spectra given by each tapered version of the data set. We call $\hat{S}_k(f) \equiv |Y_k(f)|^2$ the kth eigenspectrum, where Y_k is the discrete Fourier transform (DFT) of $\{X(t)w_k(t) : t = 1, \ldots, N\}$. The *high-resolution* multi-taper spectrum is a weighted sum of the K eigenspectra,

$$S_r(f) = \frac{\sum_{k=1}^{K} \mu_k |Y_k(f)|^2}{\sum_{k=1}^{K} \mu_k}; \qquad (3.10)$$

for the choice of weights μ_k and other details, see *Percival and Walden* [1993].

This spectral estimate's frequency resolution is $\pm p f_R$, which means that line components will actually be detected as peaks or bumps of width $2p f_R$. The situation is thus similar, in principle, to that for the classical spectral estimate of Sec. 3.2, except that the peaks can be identified with a higher resolution and greater confidence in MTM. For a white-noise process — or even one that has a locally flat spectrum near the line of interest — the high-resolution spectrum is chi-square distributed with $2K$ degrees of freedom [*Thomson*, 1982].

The relative weights on the contributions from each of the K eigenspectra can be adjusted further to obtain a more leakage-resistant spectral estimate, termed the *adaptively weighted*

multi-taper spectrum,

$$S_w(f) = \frac{\sum_{k=1}^{K} b_k^2(f) \mu_k |Y_k(f)|^2}{\sum_{k=1}^{K} b_k^2(f) \mu_k}. \tag{3.11}$$

The weighting functions $b_k(f)$ further guard against broad-band leakage for a "warm-colored" process (see Sec. 2.3) that is locally white, i.e., that has a fairly flat spectrum in the frequency range of interest. The adaptive spectrum estimate has an effective number of degrees of freedom ν that generally departs only slightly from the nominal value $2K$ of the high-resolution multi-taper spectrum [*Thomson*, 1982].

The purpose of *harmonic analysis* is to determine the line components in the spectrum that correspond to a purely periodic or multiply periodic signal in terms of their frequency, amplitude, and phase. The Fourier transform of a clean periodic signal in continuous time and of infinite length yields a *Dirac function* at the frequency of the signal, viz., a line (or peak of zero width) with infinite magnitude. As described in Sec. 3.1, it is the jump in the cumulative power at that frequency that is proportional to the periodic signal's amplitude squared.

A spectral estimate based on the methods discussed so far gives indirect information on such a signal's amplitude at all frequencies. For a periodic signal sampled at discrete times over a finite time interval, the area under the peak centered at its true frequency is proportional to the signal's amplitude squared, while the peak's width is, roughly speaking, inversely proportional to the length N of the time series. The area under the peak is nearly constant, since the peak's height is also proportional to N.

Harmonic analysis attempts to determine directly the (finite) amplitude of a (pure) line in the spectrum of a time series of finite length. We explain next how this is done within MTM. *MacDonald* [1989] describes pure-line estimation using the maximum-likelihood approach for single-window periodogram estimators [*Schuster*, 1898; *Whittle*, 1952]. *Foias et al.* [1988]

47

proved rigorous mathematical results on maximum-likelihood estimation of sinusoids in noise of arbitrary color. The drawback of these single-window results is that they work well only when the S/N ratio is fairly high. Thus, SSA prefiltering, which enhances the S/N ratio, may help "classical" pure-line estimation, as suggested by the results presented in Secs. 3.2 and 3.3 for standard peak detection. The MTM approach described below, on the other hand, can also ascertain the presence of pure sinusoids in a fairly high-noise background.

Assume the time series $X(t)$ is the sum of a sinusoid of frequency f_0 and amplitude B, plus a "noise" $\eta(t)$ which is the sum of other sinusoids and white noise. One can then write

$$X(t) = Be^{2\pi i f_0 t} + \eta(t). \tag{3.12}$$

If $\{w_k(t) : k = 0, \ldots, K - 1\}$ are the first K eigentapers and $U_k(f)$ is the DFT of w_k, a least-squares fit in the frequency domain yields an estimate $\hat{B}$ of the amplitude B:

$$\hat{B}(f_0) = \frac{\displaystyle\sum_{k=1}^{K} U_k^\star(0)\, Y_k(f_0)}{\displaystyle\sum_{k=1}^{K} |U_k(0)|^2}, \tag{3.13}$$

where the asterisk denotes complex conjugation.

A statistical confidence interval for $\hat{B}$ can be given by a Fisher-Snedecor test or *F-test* [*Kendall and Stuart*, 1979]. This test is roughly based on the ratio of the variance captured by the filtered portion of the time series $X(t)$, using K eigentapers, to the residual variance. By expanding the variance of the model one finds that it is the sum of two terms,

$$\theta = |\hat{B}(f_0)|^2 \sum_{k=1}^{K} |U_k(0)|^2 \tag{3.14}$$

48

and

$$\psi = \sum_{k=1}^{K} |Y_k(f_0) - \hat{B}(f_0)U_k(0)|^2, \tag{3.15}$$

that are respectively the "explained" and "unexplained" contributions to the variance.

The random variable $F(f)$ of the unknown frequency f is defined by

$$F(f) = (K-1)\frac{\theta}{\psi}. \tag{3.16}$$

If the time series $X(t)$ were a pure white-noise realization, F would obey a Fisher-Snedecor law with 2 and $2K-2$ degrees of freedom. One can interpret its numerical value for given data by assuming that $B = 0$ — i.e., that $X(t)$ is white — and trying to reject the white-noise null hypothesis. In practice, the spectrum need only be "locally white" in the sense that the K eigenspectra which describe the local characteristics of the spectrum should be distributed as they would be for white noise [*Thomson*, 1982].

This harmonic-analysis application of MTM is able to detect low-amplitude harmonic oscillations in a relatively short time series with a high degree of statistical significance or to reject a large amplitude if it failed the F-test, because the F-value $F(f)$ does not depend — to first order — on the magnitude of $\hat{B}(f)$. This feature is an important advantage of MTM over the classical methods of Sec. 3.2, in which the error bars scale with the amplitude of a peak (e.g., *Jenkins and Watts* [1968]).

The key assumption in this harmonic-analysis appproach, however, is that the time series is produced by a process that consists of a finite superposition of separate, purely periodic, fixed-amplitude components and white or, at least, locally white noise. If not, a continuous spectrum — in the case of a colored noise or a chaotic system — will be broken down into spurious lines with arbitrary frequencies and possibly high F-values. In essence, the above procedure assumes that the "signal" is represented by lines in the spectrum corresponding to phase-coherent harmonic oscillations, while the "noise" is represented by the continuous

component of the spectrum.

In geophysical applications, signals are most often associated with narrow-band variability that is not strictly harmonic, i.e., sinusoidal, nor even purely periodic. Truly harmonic signals are, in fact, rarely detected in climatic and most other geophysical time series [*Ghil and Childress*, 1987; *Park*, 1992]. This is often due to the fact that the period, amplitude and phase of a limit cycle like that illustrated in Fig. 1 are typically each a function of slowly changing parameters. Such signals are well captured by the SSA's pairs of eigenelements (see Sec. 2.2). To analyze the precise spectral content of an SSA oscillatory pair requires, however, the application of additional spectral estimation tools (see Secs. 3.2 and 3.3).

The complex nature of geophysical signals and noise outlined in the preceding paragraph has motivated *Mann and Lees* [1996] to modify the conventional MTM approach to line estimation. Their algorithm combines the harmonic-signal detection procedure described above with a criterion for detecting significant narrow-band, "quasi-oscillatory" signals which may exhibit phase and amplitude modulation as well as intermittently oscillatory behavior. It thus provides for the detection of both harmonic and narrow-band signals while making use of a robust estimate of the background noise.

Mann and Lees [1996] retain the pure-line test of the traditional MTM procedure, as given by Eqs. (3.13)—(3.16). All peaks, however, whether purely harmonic or narrow-band, are tested for significance relative to the null hypothesis of a red-noise background. The spectral density of this background noise is assumed to be given by Eqs. (2.9)–(2.11) and the parameters S_0 and r — the noise process's variance and the lag-one autocorrelation — are estimated empirically from the data.

Discrimination against a red-noise background is particularly important in climate studies, where the system under investigation always contains longer time scales than those of immediate interest. This leads to greater power at lower frequencies and greater likelihood of prominent peaks in the spectrum there, even in the absence of any signals (e.g., *Hasselmann*

50

[1976]; *Mitchell* [1976]). *Thomson* [1990a] and *Thomson and Chave* [1990] have developed therefore quadratic-inverse spectrum estimates aimed more particularly at the shape and properties of continuous spectra and cross-spectra.

Mann and Lees [1996] provide a robust estimate of the spectral background (2.9) by minimizing, as a function of r, the misfit between an analytical AR(1) red-noise spectrum and the adaptively weighted multi-taper spectrum convolved with a median smoother. The median smoothing operation insures that the globally estimated noise background is insensitive to "outliers." These outliers may be due to peaks associated with significant signals or to spurious ones. The median smoother guards against inflated estimates of noise variance and noise autocorrelation; the latter arise in a conventional Box-Jenkins approach due to the contamination of noise parameters by contributions from the signal, e.g., from a significant trend or oscillatory component of the series. A median smoothing width of $\Delta f = \min(f_N/4, 2pf_R)$ is a good compromise between describing the full variation of the background spectrum over the Nyquist interval, on the one hand, and insensitivity to narrow spectral features, on the other.

Significance levels for harmonic or narrow-band spectral features relative to the estimated noise background can be determined from the appropriate quantiles of the chi-square distribution, by assuming that the spectrum is distributed with $\nu = 2K$ degrees of freedom [*Mann and Lees*, 1996]. A reshaped spectrum is determined in which the contributions from harmonic signals are removed [*Thomson*, 1982], based on their passing a significance threshold for the F variance-ratio test described above. In this way, noise background, harmonic, and narrow-band signals are isolated in two steps. The harmonic-peak detection procedure provides information as to whether the signals are best approximated as harmonic or narrowband, i.e., as phase-coherent sinusoidal oscillations or as amplitude-and-phase–modulated, and possibly intermittent, oscillations. In either case, they must be found to be significant relative to a specified noise hypothesis, such as that of red noise (2.9) used above.

To illustrate the revised MTM procedure of *Mann and Lees* [1996], we apply the approach to the SOI series discussed earlier. Consistent with the SSA-MEM results of Sec. 3.3 (see Fig. 12 there), the MTM analysis in Fig. 14 recognizes two highly significant interannual peaks, one centered at $f = 0.18$ cycle/year (roughly a 5.5-year period) and another centered at $f = 0.41$ cycle/year (roughly a 2.5-year period).

Figure 14: *Adaptively weighted MTM spectrum of the SOI time series.*

These two signals are significant well above the 99% level: for $p = 2$ and $N = 690$ monthly samples, there are about 7 statistically independent bandwidths in the spectral estimate within the subannual band ($f < 0.5$ cycle/year); hence not even one interannual or lower-frequency peak is expected to arise by pure chance at the 99% level. A low-frequency variation that cannot be distinguished from a trend for the given data set is also isolated as significant at the 99% level relative to the estimated red-noise background.

We associate the two interannual peaks with the low-frequency and quasi-biennial ENSO signals, respectively, although a slight discrepancy exists with respect to the SSA-MEM results in Fig. 12. The difference between the 2.5-year period here and the 2.4 years in Fig. 12 is negligible. That between 5.5 years here and 4.8 years there is probably due to the relative shortness of the record and the width of the peak, both here and in Fig. 12. SSA-MEM emphasizes the middle of the peak, while MTM selects a harmonic feature that might lie close to one of the peak's sides.

It is interesting to note that the AR(1) model does *not* provide a good description of the noise background in a neighborhood of the spectrum surrounding $f = 1.0$ cycle/year (see also Fig. 7 and its discussion in Sec. 2.3). This discrepancy could be due, at least in part, to the prior removal of the seasonal cycle (see also *Thomson* [1995] and Sec. 2.3 here). Away from $f = 1.0$ cycle/year, the robustly estimated red-noise spectrum provides a good visual fit to the data set's spectral background. *Mann and Lees*'s [1996] noise-background

estimation procedure thus appears to be fairly insensitive to the anomalous behavior of the spectrum near the annual cycle.

Higher-frequency peaks between $f = 0.3$ cycle/month and $f = 0.5$ cycle/month (periods of 2–3 months) may be associated with intraseasonal oscillations in the tropical atmosphere but raise some concerns through their large number and relatively poor match with results obtained using other atmospheric data sets. The estimate of the noise background in this intraseasonal band is probably lowered unduly by the attempt to fit a single red-noise spectrum to both this band and the interannual one. *Gilman et al.* [1963] already had recognized that multiple decorrelation time scales may characterize certain atmospheric and climate phenomena. Thus one might be better off in a case like this, when applying MTM, to use a more complex null hypothesis for significance testing. Such a hypothesis of a fairly general colored-noise background is available in the SSA-MTM Toolkit (http://www.atmos.ucla.edu/tcd/ssa), where it is described as "locally white." This composite noise model, however, has less of a physical justification than the straightforward red noise of Eqs. (2.9)–(2.12) and raises questions of parsimony in estimating its parameters.

To better understand the distinction between the inferences from the conventional MTM procedure of *Thomson* [1982] and the revised procedure of *Mann and Lees* [1996] we compare the results of the two in Fig. 15. The F-test criterion for harmonic signals in Fig. 15a yields 7 peaks at the 99% confidence level, and 27 peaks at the 95% confidence level. Many of these peaks are associated with very weak power in the spectrum. Note that the harmonic-peak test has the higher Rayleigh frequency resolution $\pm f_R$, and not the broader bandwidth $\pm p f_R$ of the adaptive multi-taper spectral estimate (3.11). In a series of $N = 690$ data points, we would expect about 7 and 35 peaks at the 99% and 95% confidence levels respectively from chance alone. This does not differ significantly from the F-test results for the time series in Fig. 2. The harmonic-analysis approach alone does, therefore, not distinguish any significant features from random noise in our SOI data.

Figure 15: *Comparison between MTM spectra.*

The results of this approach are nonetheless useful in the reshaping procedure used by *Mann and Lees* [1996]. Figure 15b also shows the reshaped (solid) and adaptively weighted (dashed) multi-taper spectrum along with the significance levels relative to the robustly estimated red-noise background. The dashed peaks satisfy the harmonic detection test at the 90% level and are also significant relative to red noise at the same level. There are 8 such peaks, including both a lower-frequency and a quasi-biennial ENSO peak. The inference in this case is that the two distinct ENSO-band signals are both significant well above the 99% level against red noise; furthermore, each of these two narrow-band signals is likely to contain a harmonic, phase-coherent oscillation with a 90% level of confidence. If a 99% threshold were required in the F-test for the reshaping procedure, neither of these two peaks would pass the harmonic test. Nonetheless, they are still highly significant narrow-band features of the spectrum. The true ENSO signal is almost certainly associated with amplitude and phase modulation over time (see *Jiang et al.* [1995a] and Fig. 6 here) and these features are not appropriately modeled in the harmonic-signal test alone.

3.4.2 Signal reconstruction

Once significant peaks have been isolated in the spectrum, relative to the specified null hypothesis, the associated signals can be reconstructed in the time domain using the information from the multi-taper decomposition. These reconstructions of oscillatory signals are analogous to the SSA reconstructed components (RCs) described in Sec. 2.2, except that information from a frequency-domain decomposition, rather than a lag-domain decomposition, is used to reconstruct the partial signal of interest. As in the lag-domain case of SSA, the reconstruction becomes more delicate near the endpoints of the time interval over which the data are provided.

The reconstructed signal that corresponds to a peak centered at frequency f_0 is written

in continuous time as

$$\tilde{X}(t) = \Re \left\{ A(t)e^{-2\pi i f_0 t} \right\} \qquad (3.17)$$

or, for the discrete-time case at hand,

$$\tilde{X}(n\Delta t) = \Re \left\{ A_n e^{-2\pi i f_0 n\Delta t} \right\} . \qquad (3.18)$$

We determine the envelope function $A(t)$ from a time-domain inversion of the spectral-domain information contained in the K complex eigenspectra [*Park*, 1992; *Park and Maasch*, 1993; *Mann and Park*, 1994, 1996*b*], while $\Re$ is the real part of the Fourier term in curly brackets. The envelope $A(t)$ has K complex degrees of freedom, and allows for phase and amplitude variations in the time reconstruction of the signal centered at frequency f_0. It thus represents a considerable refinement of the classical, single-taper *complex demodulation* method [*Bloomfield*, 1976; *Hasan*, 1983].

The discrete-time sequence describing the complex envelope A_n is determined from a discrete inverse problem that uses the complex amplitudes of each of the K eigenspectra and appropriate boundary conditions [*Park*, 1992; *Park and Maasch*, 1993]. The three lowest-order boundary constraints in this inversion involve minimizing the envelope A_n itself or its slope near the endpoints, or maximizing its smoothness there. Supplementary information regarding the signal (see the discussion in *Park* [1992]) might favor any one of these three choices that involve the sequence A_n and its first or second (divided) differences. The amplitude of the seasonal cycle in surface temperature, for example, is nearly constant in time, and a minimum-slope constraint is thus most appropriate (see *Mann and Park* [1996*b*]). In the absence of such a priori information, one can obtain a nearly optimal reconstruction through seeking the weighted linear combination of these three constraints that minimizes the mean-square difference between the reconstructed signal and the raw data series [*Mann and Park*, 1996*b*].

In Fig. 16 we show the reconstruction, based on the latter recipe, of the SOI that corresponds to the three interannual signals that are significant at the 99% level in Figure 14 — the trend, low-frequency, and quasi-biennial components. The two oscillatory signals each exhibit roughly one (nondimensional, normalized) unit of peak-to-peak variation. The amplitude modulations of both oscillatory components, low-frequency (Fig. 16b) and quasi-biennial (Fig. 16c), are similar, with minimal amplitude in the early 1960s (compare Fig. 8 and its discussion in Sec. 2.4).

Figure 16: *Multi-taper reconstruction of the significant interannual components in SOI.*

The long-term trend in the SOI (Fig. 16a), while weak, still exhibits a peak-to-peak variation of about 0.5 units or one half of the amplitude of the two oscillatory signals. The oscillatory components of SOI are well-known to be anticorrelated with sea-surface temperatures in the eastern Tropical Pacific. The trend in Fig. 16a also seems to be anticorrelated with the temperature trend for the Equatorial Pacific shown in Fig. 4b of *Moron et al.* [1998], although the number of degrees of freedom in these two trends does not allow us to establish the statistical significance of this anticorrelation. The filtered RC-5 in Fig. 6b shares the overall features of the reconstructed trend in Fig. 16a, although it also displays some interdecadal variability not present in the latter.

In Fig. 17 we plot the partial reconstruction of SOI (heavy solid) obtained by summing the trend and two oscillatory components shown separately in Fig. 16. This reconstruction captures 28% of the total variance in the raw monthly data (light solid) vs. the 43% captured by RCs 1–4 in Fig. 6a. The overall match with strong and moderate El Niños and La Niñas is also fairly good, although somewhat less compelling than in Sec. 2.2. The MTM trend in Fig 16a, on the other hand, seems to provide a cleaner picture of the SOI's nonoscillatory evolution over the latter half of the 20th century than the SSA trend in Fig. 6b. Both SSA and MTM reconstructions have much, therefore, to contribute to our description and

56

understanding of climate variability.

Figure 17: *Partial MTM reconstruction of SOI.*

4 Multivariate Methods

So far, we have discussed the spectral analysis of a single, *scalar* time series. In many geophysical applications — and to a lesser extent in astrophysics and space physics — one is presented with a sequence of maps or other time-discrete sample of a *vector*-valued process. The L components at each time may be values of a field at discrete grid points — in one, two or three space dimensions — or values of the field's projections onto certain (spatial) basis functions imposed by the geometry of the domain. It is useful to think of the values of such a component at successive times as a (scalar) time series and it is common to refer to each of these time series as a *channel*. The methods discussed in this section are thus referred to sometimes as *multi-channel*. It is also common in the statistical literature to refer to a scalar time series as *univariate* and to a vector time series as *multivariate*.

Most of the univariate methods of Secs. 2 and 3 — SSA, wavelets, MEM and MTM — have been extended to multivariate time series. The computation of cross-correlations and cross-spectra by classical methods is presented by *Hannan* [1970] and *Priestley* [1981b], among others. *Mann and Park* [1999] provide an up-to-date review of multivariate MTM methods and their applications. There exists an extensive literature on applying wavelets to the evolution of turbulent flow fields [*Farge*, 1992] and related multivariate problems [*Torrence and Compo*, 1998]. For the sake of brevity, we shall restrict ourselves therefore to a succint presentation of multivariate MEM and SSA.

4.1 Principal oscillation patterns (POPs)

In principle, one can extend MEM to multivariate time series via the inverse modeling of vector-valued AR processes [*Ooms*, 1988]. In practice, if both the number L of channels and the order M of the AR model are high, the issues of computational and statistical stability of the LM coefficients involved become rapidly prohibitive. Thus, it is only vector-valued AR(1) processes that have been widely applied to geophysical problems under the name of Principal Oscillation Patterns (POPs) [*Hasselmann*, 1988; *Von Storch et al.*, 1988] or Empirical Normal Modes (ENMs) [*Penland*, 1989; *Penland and Ghil*, 1993].

Penland [1996] provides an excellent review of vector-valued stochastic processes $\mathbf{X}(t)$ in continuous time and of their connection to discrete-time AR processes. Her review is well illustrated by the application of the basic concepts to the spectral analysis and AR prediction of sea-surface temperatures in the tropical Pacific. The details of the methodology are well beyond the scope of this brief review, but the basic idea is to use the data in order to estimate two matrices, $\mathbf{B}$ and $\mathbf{Q}$, in the model

$$d\mathbf{X} = \mathbf{B}\mathbf{X}dt + d\xi(t). \tag{4.1}$$

Here $\mathbf{B}$ is the (constant) dynamics matrix and $\mathbf{Q}$ the lag-zero covariance of the vector white-noise process $d\xi(t)$, considered here in continuous time. For the subtleties of time-continuous stochastic processes we refer to *Penland* [1996], and further references there; $\xi(t)$ is a continuous, nowhere differentiable Wiener process. For the present purposes, it suffices to recall that — by the wide-sense stationarity assumption (see Appendix A) — $\mathbf{B}$ has to be stable, i.e., its eigenmodes have to decay in time. The latter property is the counterpart, in the present vector-matrix context, of the requirement that $a_1 < 0$ for the scalar AR(1) process of Eqs. (1.1) and (2.12).

The two matrices $\mathbf{B}$ and $\mathbf{Q}$ are related by a fluctuation-dissipation relation [*Leith*, 1975]:

$$\mathbf{B}\mathbf{C}(0) + \mathbf{C}(0)\mathbf{B}^{\mathrm{t}} + \mathbf{Q} = 0, \tag{4.2}$$

where $\mathbf{C}(\tau) = \mathcal{E}\{\mathbf{X}(t+\tau)\mathbf{X}^{\mathrm{t}}(t)\}$ is the lag-covariance matrix of the process $\mathbf{X}(t)$ being modeled and $(.)^{\mathrm{t}}$ is the transpose of a vector or matrix. One proceeds to estimate the Green's function $\mathbf{G}(\tau) = \exp(\mathbf{B}\tau)$ at a given lag τ_0 from the sample of $\mathbf{C}(\tau)$ by

$$\mathbf{G}(\tau_0) = \mathbf{C}(\tau_0)\mathbf{C}^{-1}(0). \tag{4.3}$$

Diagonalizing $\mathbf{G}(\tau_0)$ allows one to obtain $\mathbf{B} = \mathbf{B}_{\tau_0}$, via its eigenvalues and (left and right) eigenvectors. If Eq. (4.1) is the correct model for the process $\mathbf{X}(t)$, then the matrix $\mathbf{B}_\tau$ obtained for a certain $\tau = \tau_0$ should be, in fact, independent of τ [*Penland and Ghil*, 1993]. The ENM approach [*Penland*, 1989] differs from the original POP approach [*Hasselmann*, 1988] by: i) building on the theory of continuous-time stochastic processes; and ii) considering the whole spectrum of $\mathbf{B}$, rather than pre-selected (pairs of) eigenvalues only.

There are various reasons for which $\mathbf{G}(\tau)$, and hence $\mathbf{B}_\tau$, might be substantially different from one value of τ to another. These include problems and trade-offs in the choice of τ, as well as nonlinearity, instability or nonstationarity of the true process being sampled. The vector AR(1) method, whether in its POP or ENM form, works best when the spectrum is dominated by a single peak (see *Von Storch et al.* [1988], *Penland* [1996], and further references there), which results in a well-separated pair of POPs or ENMs. The "tau-test" of recomputing the eigendecomposition of $\mathbf{G}(\tau)$ and, therewith, of $\mathbf{B}_\tau$ for various τ values is particularly useful in determining the stability of the results, since error bars are — like for single-channel MEM and even more so — fairly difficult to come by.

Recently, *Egger* [1999] has extended the POP approach to higher-order vector AR processes. He calls such a vector AR(M) process a MOP-model. MOPs can capture, in principle,

a multiplicity of oscillating patterns that elude POPs. Still, they encounter difficulties when the underlying system that produces the signal is nonlinear, as is the case with POPs.

4.2   Multi-channel SSA

Multi-channel SSA (or M-SSA) is a natural extension of SSA to a time series of vectors or maps, such as time-varying temperature or pressure distributions over the globe. The use of (M-)SSA for such multivariate time series was proposed theoretically, in the context of nonlinear dynamics, by *Broomhead and King* [1986*b*]. SSA can be applied — as explained in Secs. 1.2 and 2.1 — in the context of systems of ordinary differential equations and the reconstruction of (the skeleton of) their attractor.

M-SSA is a generalization of this approach to systems of partial differential equations and the study of the spatio-temporal structures that characterize the behavior of solutions on their attractor [*Temam*, 1997]. It has been applied to intraseasonal variability of large-scale atmospheric fields by *Kimoto et al.* [1991], *Keppenne and Ghil* [1993], and *Plaut and Vautard* [1994], as well as to ENSO; for the latter, both observed data [*Jiang et al.*, 1995*a*; *Unal and Ghil*, 1995] and coupled GCM model simulations [*Robertson et al.*, 1995*a, b*] were used.

In the meteorological literature, extended EOF (EEOF) analysis is often assumed to be synonymous with M-SSA [*Von Storch and Zwiers*, 1999]. The two methods are both extensions of classical principal component analysis (PCA) [*Preisendorfer*, 1988] but they differ in emphasis: EEOF analysis [*Barnett and Hasselmann*, 1979; *Weare and Nasstrom*, 1982; *Lau and Chan*, 1985] typically utilizes a number L of spatial channels much greater than the number M of temporal lags, thus limiting the temporal and spectral information. In M-SSA, on the other hand, based on the single-channel experience reviewed in Secs. 2 and 3.3, one usually chooses $L \leq M$ (see also Appendix A). Often M-SSA is applied to a few leading PCA components of the spatial data, with M chosen large enough to extract detailed temporal and spectral information from the multivariate time series.

Let $\{X_l(t) : l = 1, \ldots, L; t = 1, \ldots, N\}$ be an L-channel time series with N data points given at equally spaced intervals $n\Delta t$. We assume that each channel l of the vector $\mathbf{X}(n\Delta t)$ is centered and stationary in the weak sense.

The generalization of SSA to a multivariate time series requires the construction of a "grand" block-matrix $\tilde{\mathbf{T}}_{\mathbf{X}}$ for the covariances that has the form

$$
\tilde{\mathbf{T}}_{\mathbf{X}} =
\begin{pmatrix}
\mathbf{T}_{1,1} & \mathbf{T}_{1,2} & . & . & & . & & \mathbf{T}_{1,L} \\
\mathbf{T}_{2,1} & \mathbf{T}_{2,2} & . & & & & & . \\
 & . & . & . & . & & & . \\
 & . & & . & . & . & \mathbf{T}_{l,l'} & & . \\
 & . & & . & . & . & & & . \\
 & . & & & & . & & . & \mathbf{T}_{L-1,L} \\
\mathbf{T}_{L,1} & & . & & . & . & \mathbf{T}_{L,L-1} & \mathbf{T}_{L,L}
\end{pmatrix} .
\tag{4.4}
$$

Each block $\mathbf{T}_{l,l'}$ is a matrix that contains estimates of the lag covariance between channels l and l'. Extending the approach of *Vautard and Ghil* [1989] (see Sec. 2.2), *Plaut and Vautard* [1994] proposed to obtain a "least-biased" estimator by using, for each lag m ($m = 0, \ldots, M - 1$), the longest-possible segment of each channel. Thus, the entries (j, j') of each block $\mathbf{T}_{l,l'}$ can be written as

$$
(T_{l,l'})_{j,j'} = \frac{1}{\tilde{N}} \sum_{n=\max(1,1+j-j')}^{\min(N,N+j-j')} X_l(t) X_{l'}(t + j - j'),
\tag{4.5}
$$

where $\tilde{N}$ is a factor that depends on the range of summation, to wit,

$$
\tilde{N} = \min(N, N + j - j') - \max(1, 1 + j - j') + 1
\tag{4.6}
$$

[compare Eq. (2.2)].

Note that, unlike in the single-channel case, here $\mathbf{T}_{l,l'}$ is Toeplitz but not symmetric. Its main diagonal contains the *Vautard and Ghil* [1989] estimate of the lag-zero covariance of channels l and l'. The diagonals in the lower-left triangle of $\mathbf{T}_{l,l'}$ contain the lag-m covariance of channels l and l', with l' leading l, while the diagonals in the upper-right triangle contain the covariances with l leading l'. Equation (4.5) ensures that $\mathbf{T}_{l',l} = (\mathbf{T}_{l,l'})^{\mathrm{t}}$ so that $\tilde{\mathbf{T}}_{\mathbf{X}}$ is symmetric, but it is *not* Toeplitz. Note that the original formula in *Plaut and Vautard* [1994]—their Eq. (2.5)—does *not*, in fact, yield a symmetric matrix. More generally, we only expect to obtain a grand matrix $\tilde{\mathbf{T}}_{\mathbf{X}}$ that is both symmetric and of Toeplitz form provided the spatial field being analyzed is statistically homogeneous, i.e., its statistics are invariant with respect to arbitrary translations and rotations.

An alternative approach [*Broomhead and King*, 1986a, b; *Allen and Robertson*, 1996] to computing the lagged cross-covariances is to form the multi-channel trajectory matrix $\tilde{\mathbf{X}}$ by first augmenting each channel $\{X_l(t) : t = 1, \ldots, N\}, 1 \leq l \leq L$, of $\mathbf{X}$ with M lagged copies of itself,

$$
\tilde{\mathbf{X}}_l = \begin{pmatrix} X_l(1) & X_l(2) & . & . & X_l(M) \\ X_l(2) & X_l(3) & . & . & X_l(M+1) \\ . & . & . & . & . \\ X_l(N'-1) & . & . & . & X_l(N-1) \\ X_l(N') & X_l(N'+1) & . & . & X_l(N) \end{pmatrix}, \quad 1 \leq l \leq L, \tag{4.7}
$$

and then forming the full augmented *trajectory matrix*:

$$
\tilde{\mathbf{D}} = \left(\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2, \ldots, \tilde{\mathbf{X}}_L \right). \tag{4.8}
$$

These two steps generalize Eqs. (2.1) and (2.5). The grand lag-covariance matrix $\tilde{\mathbf{C}}_{\mathbf{X}}$

that is analogous to $\mathbf{C}_X$ in Eq. (2.5) is given by

$$\tilde{\mathbf{C}}_{\mathbf{X}} = \frac{1}{N'}\tilde{\mathbf{X}}^{\mathrm{t}}\tilde{\mathbf{X}} = \begin{pmatrix} \mathbf{C}_{1,1} & \mathbf{C}_{1,2} & . & . & . & \mathbf{C}_{1,L} \\ . & \mathbf{C}_{2,2} & . & . & . & . \\ . & & . & . & . & . \\ . & & & . & . \mathbf{C}_{l,l'} & . \\ . & & . & . & . & . \\ \mathbf{C}_{L,1} & \mathbf{C}_{L,2} & . & . & . & \mathbf{C}_{L,L} \end{pmatrix}. \tag{4.9}$$

The blocks of $\tilde{\mathbf{C}}_{\mathbf{X}}$ are given by

$$\mathbf{C}_{l,l'} = \frac{1}{N'}\tilde{\mathbf{X}}_l^{\mathrm{t}}\tilde{\mathbf{X}}_{l'} \tag{4.10}$$

with entries

$$(C_{l,l'})_{j,j'} = \tilde{C}_{IJ} = \frac{1}{N'}\sum_{t=1}^{N'} X_l(t+j-1)X_{l'}(t+j'-1), \tag{4.11}$$

where $I = j + M(l-1)$ and $J = j' + M(l'-1)$ [*Allen and Robertson*, 1996]. Here $\mathbf{C}_{l,l'}$ is neither Toeplitz nor symmetric, but again $\mathbf{C}_{l',l} = (\mathbf{C}_{l,l'})^{\mathrm{t}}$ and hence $\tilde{\mathbf{C}}_{\mathbf{X}}$ is symmetric.

Each block $\mathbf{C}_{l',l}$ contains, like $\mathbf{T}_{l,l'}$ in Eq. (4.4), an estimate of the lag covariances of the two channels l and l'. The difference between the Toeplitz method of *Vautard and Ghil* [1989] and the trajectory-matrix method of *Broomhead and King* [1986 a, b] in estimating the lag-covariance matrix, when applied to M-SSA, consists in calculating Eqs. (4.4) and (4.5) vs. Eqs. (4.9)–(4.11). The Toeplitz method extracts, according to the fixed lag m that is being considered at the moment, the two longest matching segments from the two channels l and l'; the matching is determined by the requirement that, for each entry in the designated trailing channel, there exist an entry in the designated leading channel that is m time steps "ahead." It then uses the same estimated lag covariance for all entries along the appropriate diagnoal in the block and thus yields a Toeplitz-form submatrix.

The trajectory-matrix method gradually slides, regardless of m, the same $N'-$long win-

dows over the two channels. To compute the first entry in the diagonal that contains the lag-m covariances, the two matching windows are situated so that one starts at the first time point of the trailing channel and at the $(m+1)$st point of the leading channel. Both windows are then slid forward by one point in time to produce the second entry. This results in slightly different values, from entry to entry, until the last point of the leading channel is covered and the $(M - m)$th entry of the diagonal, which is the last one, is calculated. This latter method thus retains detailed information on the variation of the lag-covariance estimates from one pair of segments of the channels l and l' to another, while the Toeplitz method produces a single, and smoother, global estimate for each lag m.

Diagonalizing the $LM \times LM$ matrix $\tilde{\mathbf{C}}_{\mathbf{X}}$ or $\tilde{\mathbf{T}}_{\mathbf{X}}$ yields LM eigenvectors $\{\mathbf{E}^k : 1 < k \leq LM\}$ that are not necessarily distinct. The extent to which the eigenpairs $(\lambda_k, \mathbf{E}^k)$ obtained by diagonalizing $\tilde{\mathbf{C}}_{\mathbf{X}}$ equal those obtained from $\tilde{\mathbf{T}}_{\mathbf{X}}$ is a good indication of the robustness of the M-SSA results. Each eigenvector $\mathbf{E}^k$ is composed of L consecutive M-long segments, with its elements denoted by $E_l^k(j)$. The associated space-time PCs $\mathbf{A}^k$ are single-channel time series that are computed by projecting $\tilde{\mathbf{X}}$ onto the EOFs:

$$A^k(t) = \sum_{j=1}^{M} \sum_{l=1}^{L} X_l(t+j-1)E_l^k(j), \tag{4.12}$$

where t varies from 1 to N'. This is the multi-channel counterpart of Eq. (2.6).

For a given set of indices $\mathcal{K}$, RCs are obtained by convolving the corresponding PCs with the EOFs. Thus, the kth RC at time t for channel l is given by:

$$R_l^k(t) = \frac{1}{M_t} \sum_{j=L_t}^{U_t} A^k(t-j+1)E_l^k(j). \tag{4.13}$$

As in Eq. (2.7), the normalization factor M_t equals M, except near the ends of the time series, and the sum of all the RCs recovers the original time series [*Plaut and Vautard*, 1994], as it does in the single-channel case. The factor M_t, as well as the lower and upper bounds

of summation, L_t and U_t, are given by the same Eq. (2.8) as in the single-channel case.

As an example, we apply M-SSA to the near-global data set of monthly sea-surface temperatures from the Global Sea-Ice and Sea-Surface Temperature (GISST) data set [*Rayner et al.*, 1995] for 1950–94, from 30°S to 60°N, on a 4°-latitude by 5°-longitude grid. The mean seasonal cycle was subtracted at the outset, and the data prefiltered with standard PCA [*Preisendorfer*, 1988] to retain the 10 leading spatial PCs that describe 55.2% of the variance. This favors the association of larger decorrelation times with larger spatial scales, as expected for climatic [*Fraedrich and Boettger*, 1978] and other geophysical fields, and the channels are uncorrelated at zero lag. We use therewith $L = 10$ channels and $N = 540$ months; no other filtering or detrending was applied. The SST anomalies of this data set were averaged over the Niño-3 area in the eastern Tropical Pacific (5°S–5°N, 90°–150°W), and the evolution of these averaged anomalies is shown in Fig. 18 (light solid line).

Figure 18: *SST anomalies averaged over the Niño-3 region.*

For comparison with previous ENSO studies we choose $M = 60$ months, i.e., a five-year M-window. To demonstrate the usage of N'-windows vs. M-windows [see Appendix A, especially Eq. (A.1) and Fig. 21, for the definition of "complementary windows"], we also set $M = 480$, which yields an N'-window of 5 years, and $M = 270$, which yields the M- and the N'-windows equal both to $N/2$.

We use here the lag-covariance matrices $\tilde{\mathbf{C}}_{\mathbf{X}}$ that are based on the trajectory-matrix approach of *Broomhead and King* [1986*a, b*] for ease of comparison. In all three cases $ML > N'$ so that it is more efficient to diagonalize the *reduced* $(N' \times N')$ *covariance matrix* with elements given by Eq. (A.1), rather than the $(ML \times ML)$ matrix whose elements are given by Eq. (4.9).

The leading oscillatory pair over the entire domain has a quasi-quadrennial period for all three values of M, as illustrated by the spatial average of the respective RCs over the

Niño-3 region (Fig. 18). The two reconstructions for $M = 60$ (heavy solid) and $N' = 60$ (dotted) isolate a very similar frequency-modulated mode, with a slightly longer period before 1965 (see Fig. 8 and discussion there). The quasi-quadrennial pair (modes 2 and 3 in these two cases), however, only accounts for 20.0% variance with $M = 60$ against 30.8% for $N' = 60$. This difference arises because the reduced covariance matrix has rank 60 in the latter case — which maximizes data compression — compared with rank $ML = 480$ for $M = 60$. The smaller one of M and N' determines the approximate spectral resolution $1/N'$ or $1/M$. Choosing $M = N' = 270$ yields the maximum spectral resolution but captures less variance for the oscillatory pair of interest: only 12.1% variance for the quasi-quadriennial pair, formed in this case by modes 3 and 4.

Figure 19 shows the spatio-temporal evolution of the quasi-quadrennial mode (RCs 2–3) for $N' = 60$, using the phase-compositing procedure of *Plaut and Vautard* [1994]. The maps (not shown) are almost indistinguishable for $M = 60$ (same pair) and $M = N' = 270$ (RCs 3–4). The succession of panels in Fig. 19 illustrates the progression of the oscillation, keyed to the phase index constructed by using the leading spatial PC of RCs 2–3 [*Moron et al.*, 1998].

Figure 19: *Phase-composites for RCs 2-3 with $N' = 60$.*

The maps are dominated by the classic El Niño signature over the tropical Pacific, together with the well-known "horse-shoe" pattern over the extratropical Pacific (see *Chao et al.* [2000] and references there). Amplitudes elsewhere are weak: In the northwestern Indian Ocean, though, there is some suggestion of a warming during El Niño and cooling during La Niña events that lags the eastern equatorial Pacific by 1/8 cycle ($\approx$ 6 months). Over the tropical Atlantic, there is also a small warming concurrent with the peak phase of El Niño and cooling simultaneously with La Niña.

As in the single-channel case, a test of statistical significance is needed to avoid spatially

smooth-looking but spurious oscillations that might arise from M-SSA of finitely sampled noise processes. The complementary N'-window approach of Eq. (A.1) (see Appendix A) allows the univariate SSA Monte Carlo test of *Allen* [1992] (see also *Allen and Smith* [1996]) to be extended to M-SSA in a straightforward manner, provided $ML > N'$, so that the reduced covariance matrix is completely determined and has full rank. As shown in our example above, N' can always be chosen sufficiently small, so that the complementary window N' used in Eq. (A.1) determines the spectral resolution.

Details of this essentially univariate test are given in *Allen and Robertson* [1996]. The usefulness of the test depends in an essential way on the channels being uncorrelated at zero lag or very nearly so. In the example at hand, the decorrelation condition holds exactly, since we use the PCs of spatial EOF analysis. When using time series from grid points that are sufficiently far from each other for decorrelation to be near-perfect, the test can still be useful.

In this test, the data series together with a large ensemble of red-noise surrogates are projected onto the eigenmodes of the reduced covariance matrix of either the data or the noise. The statistical significance of the projections is estimated as in the single-channel test described in Sec. 2.3. The noise surrogates are constructed to consist of univariate AR(1) segments, one per channel, that match the data in autocovariance at lag 0 and lag 1, channel-by-channel. The reason an essentially univariate test can be applied is because the eigenmodes do not depend on cross-channel lag-covariance, provided N' is interpreted as the spectral window used in Eq. (A.1). Since the test is biased if we project onto the data eigenmodes, we project onto the eigenmodes provided by the covariance matrix of the AR(1) noise.

Figure 20: *Significance test for the GISST data.*

Figure 20 shows the result of the test applied for $M = N' = 270$, the maximum ef-

fective resolution. In Fig. 20, the projections are plotted against the dominant frequencies associated with each noise eigenvector. Since the latter are near-sinusoidal in this case, the resulting spectrum is closely related to a traditional Fourier power spectrum. Both the quasi-quadrennial and the quasi-biennial modes pass the test at the 95% level. They are well separated in frequency by about $1/(20 \text{ months})$, which far exceeds the spectral resolution of $1/M = 1/N' = 1/(270 \text{ months}) \approx 1/(22 \text{ years})$. The two modes are thus significantly distinct from each other spectrally, in agreement with the univariate SOI results of Figs. 12 and 14, using MEM and MTM, respectively.

5 Summary and Perspectives

5.1 Summary

We have reviewed a number of recent developments in the spectral analysis of climatic and other geophysical time series. Basic ideas and their connections to dynamical systems theory were emphasized. The methods arising from these ideas and their numerical implementation were outlined and illustrated.

The key features of the methods discussed in this review are summarized in Table 3 (see also Table I in *Ghil and Taricco* [1997]). We have tried to provide up-to-date information on the most refined and robust statistical significance tests available for each one of the three methods that were discussed in depth: SSA, MEM and MTM. Still, this group of authors feels that none of these methods by itself can provide entirely reliable results, since every statistical test is based on certain probabilistic assumptions about the nature of the physical process that generates the time series of interest. Such mathematical assumptions are rarely, if ever, met in practice.

Table 3: *Comparison table of spectral methods.*

To establish higher and higher confidence in a spectral result, such as the existence of an oscillatory mode, a number of steps can be taken. First, the mode's manifestation is verified for a given data set by the best battery of tests available for a particular spectral method. Second, additional methods are brought to bear, along with their significance tests, on the given time series. This process was illustrated by the SOI time series in the present review. *Vautard et al.* [1992] and *Yiou et al.* [1996] provided similar illustrations for a number of synthetic time series, as well as for the IPCC time series of global surface-air temperatures [*IPCC*, 1990] and for the deuterium isotopic ratio [*Jouzel et al.*, 1993] that reflects local air temperatures at Vostok station in the Antarctica.

The application of the different univariate methods described here and of their respective batteries of significance tests to a given time series is facilitated by the SSA-MTM Toolkit, which was originally developed by *Dettinger et al.* [1995a]. The Toolkit has evolved as freeware over the last six years to become more effective, reliable, and versatile. Its latest version is available at http://www.atmos.ucla.edu/tcd/ssa.

The third step on the road to genuine confidence in a given oscillatory mode is to obtain additional time series that might exhibit the mode of interest, if indeed it is present. For ENSO, this includes, besides SOI, the Niño-3 sea-surface temperatures shown in Fig. 18, as well as the additional data sets discussed in connection with Fig. 6 in Sec. 2.2 here (see also *Rasmusson et al.* [1990], *Keppenne and Ghil* [1992], *Jiang et al.* [1995a] and *Unal and Ghil* [1995] for instrumental data and *Mann and Park* [1999] and further references there for proxy records). The existence of two oscillatory climate modes with interannual periodicities — one quasi-biennial, the other with a period of 4–6 years — while not present in every record and for every method, has been established with great confidence. These two modes are thus definitely part of the skeleton of the attractor for climate dynamics on interannual time scales.

The fourth and most difficult step on the road of confidence building is that of providing

a convincing physical explanation of an oscillation, once we have full statistical confirmation for its existence. This step consists in building and validating a hierarchy of models for the oscillation of interest (see *Ghil* [1994] and *Ghil and Robertson* [2000]). The modeling step is distinct from and thus fairly independent of the statistical-analysis steps discussed up to this point. It can be carried out before, after or in parallel with the other three steps.

Model validation requires specifically the theoretical prediction of an oscillatory feature not yet obtained by time-series analysis. The prediction might involve a spatial feature not yet diagnosed from existing observations or as-yet-unobserved spectral features: harmonics, subharmonics or combination tones. The subsequent detection in new data sets of such a theoretically predicted spatial or spectral feature, while never definitive, tends to greatly bolster the case for the proposed model. It thus reinforces more than anything else the confidence in the existence and importance of the mode under discussion.

We turn therewith to the implications that the discovery by spectral methods of oscillatory behavior can have for the deeper understanding of climatic variability and for its real-time prediction.

5.2   Implications for understanding

We have introduced in Sec. 1 the concept of a *ghost limit cycle* (Fig. 1) as the underpinning for the dynamic usefulness of a spectral peak. In a number of instances, from intraseasonal oscillations [*Ghil et al.*, 1991] through paleoclimatic variability [*Ghil*, 1994], the robust identification of an oscillatory mode in observations or proxy records has gone hand in hand with the description of a Hopf bifurcation in a single model or a full hierarchy of models [*Ghil and Robertson*, 2000]. Such a bifurcation gives rise — via an oscillatory instability and its nonlinear saturation — to a stable limit cycle. This limit cycle is destabilized in turn, as the forcing on the system increases or the constraints on it (such as viscous or eddy dissipation) decrease; it thus becomes a "ghost limit cycle." The process of successive bifurcations

engenders additional ghost limit cycles and thus leads to multiply periodic, and eventually irregular, behavior [*Guckenheimer and Holmes*, 1983; *Ghil and Childress*, 1987; *Schuster*, 1988].

In the present paper, we used seasonal-to-interannual climate variability as the Ariadne's thread to help guide the interested reader through the labyrinth of methodological exposition. As mentioned in Secs. 2.2 and 4.2, the data-adaptive signal-to-noise (S/N) enhancement obtained by applying SSA to univariate and multivariate climatic time series covering this variability led to the robust identification of separate 4/1-, 4/2- and 4/3-year peaks. These subharmonics of the seasonal cycle characterize the Devil's staircase mechanism of interaction between the El-Niño/Southern Oscillation's (ENSO's) intrinsic, self-sustained oscillation in the coupled ocean-atmosphere system and the seasonal forcing. This mechanism involves partial phase locking between the two oscillations and has been studied in detail in simple and intermediate ENSO models [*Chang et al.*, 1994; *Jin et al.*, 1994; *Tziperman et al.*, 1994; *Saunders and Ghil*, 2001], in which the period of the intrinsic ENSO cycle varies between 2 and 3 years, depending on models and parameters.

The separation of two or three peaks (quasi-quadrennial, quasi-biennial and, sometimes but not always, 16–17 months) and the spatio-temporal variability associated with them have been ascertained in various data sets [*Rasmusson et al.*, 1990; *Jiang et al.*, 1995a], as well as in coupled general circulation models [*Robertson et al.*, 1995a, b; Ghil and Robertson, 2000]. The Devil's staircase thus provides the most plausible explanation so far for the warm events peaking typically in boreal winter, as well as for the irregularity of the El Niños' spacing from one warm event to the next (see also discussion at the end of Sec. 2.4 and further below).

A similar dialog between spectral studies of (inter)decadal variability and dynamical studies of oscillatory model behavior is taking shape. Interdecadal modes of variability at 10–11, 14–15 and 25–27 years were initially identified — by combining the application of various spectral methods described in Secs. 2 and 3, to wit SSA, MEM, and MTM — in

both global [*Ghil and Vautard*, 1991; *Vautard et al.*, 1992] and regional [*Plaut et al.*, 1995] time series. *Allen and Smith* [1994, 1996] and *Mann and Lees* [1996] examined further the level of significance of those spectral features against a red-noise background. *Allen and Smith* [1994] described the spatial patterns associated with the variability at 10–11 and 25–27 years by using a method related to M-SSA.

The spectral peak at 14–15 years was confirmed by SSA of coral-reef proxy records from the tropical South Pacific by *Quinn et al.* [1993]. *Mann and Park* [1994] described this mode's spatial pattern on a global scale by using MTM-related multi-channel methods applied to instrumental temperature data (see also *Mann and Park* [1999]). Variability in the South Atlantic basin with a near-periodicity of 14–16 years has been described by *Venegas et al.* [1997] and *Robertson and Mechoso* [2000].

On the other hand, a plethora of oscillatory mechanisms is emerging in two-dimensional [*Quon and Ghil*, 1995], as well as in fully three-dimensional [*Weaver et al.*, 1993; *Chen and Ghil*, 1995; *Rahmstorf*, 1995] models of the ocean's thermohaline circulation (see summary table, with further references, in *Ghil* [1994]). Coupled ocean-atmosphere models are also producing decadal and interdecadal oscillations that seem to resemble the purely oceanic ones [*Delworth et al.*, 1993; *Chen and Ghil*, 1996] or differ from them in substantial ways [*Latif and Barnett*, 1994; *Robertson*, 1996].

As in the case of ENSO variability, a full hierarchy of models will probably be needed to clarify the connections between the (dimly) observed and the (imperfectly) modeled oscillations on the decadal and interdecadal time scales. A particularly interesting set of questions deals with the 14–15-year oscillation, its pattern and physical causes. The details of its spatio-temporal evolution in the North Atlantic have been described by *Moron et al.* [1998] and their implications for prediction have been outlined by *Sutton and Allen* [1997]. These details seem to point to the intrinsic variability of the North Atlantic's wind-driven circulation as a likely source [*Jiang et al.*, 1995c; *Speich et al.*, 1995; *Dijkstra and Molemaker*,

1999]. *Chao et al.* [2000], on the other hand, have described a 14–17-year oscillation encompassing the entire Pacific basin's sea-surface temperatures and roughly symmetric about the Equator. Its subharmonic, with a period of 31–35 years, appears in coral records of both δ^{18}O and δ^{13}C [*Quinn et al.*, 1993; *Boiseau et al.*, 1999], as well as in 2300 years of tree-ring records from Tasmania [*Cook et al.*, 1995].

Are the three oscillations — in the North Atlantic, South Atlantic, and the Pacific — phase-locked, as suggested by *Mann and Park* [1994]? If so, do they have a common origin that is purely oceanic or is it due to coupled ocean-atmosphere processes? If not, what are their separate origins?

An important aspect of nonlinear oscillations in a complex system with multiple time scales, like the climate system, is their modulation in amplitude, phase and frequency (see Table 3). We showed that SSA can follow well, via reconstructed components (RCs), variations in signal amplitude and phase that are associated with a fairly broad spectral peak [*Ghil and Mo*, 1991a; *Plaut and Vautard*, 1994; *Dettinger and Ghil*, 1998]. MTM-based harmonic analysis can be adapted, by using a sliding window, to follow the slow shift in time of a sharp line's frequency [*Yiou et al.*, 1991; *Birchfield and Ghil*, 1993]. Wavelet analysis provides an even more natural way of following quasi-adiabatic or gradual changes in the natural frequency of a climatic oscillator [*Meyers et al.*, 1993; *Yiou et al.*, 1995]. Suitable significance criteria for wavelet analysis results [*Flandrin*, 1999] need to be adapted further to the problems of following slow changes in frequency, as well as in amplitude and phase, of climate oscillations.

We showed that multi-scale SSA — which can be viewed as a data-adaptive form of wavelet analysis — is a novel and useful tool for the reliable identification of sharp, relatively sudden changes in frequency [*Yiou et al.*, 2000]. Its results, when applied to SOI, support further the Devil's Staircase theory of ENSO, by highlighting the sudden jump in the low-frequency oscillation's periodicity in the early 1960s (see also the discussion of secu-

lar warming in the Tropical Pacific and its effects on the evolution of ENSO variability over the 20th century in *Park and Mann* [2000]).

5.3 Implications for prediction

In the previous subsection, we have discussed the need for validating physical models of a spectrally detected oscillatory mode by formulating and verifying theoretical predictions of the oscillation's features, spatial or temporal, that have not yet been observed. There is, in climate dynamics and other geosciences, another type of prediction that can greatly help increase confidence in the description and explanation of the various time series generated by a given process, namely the forecasting of the process's future evolution in time. In fact, this is also one of the main purposes of time series analysis [*Wiener*, 1949; *Box and Jenkins*, 1970; *Priestley*, 1981*b*; *Ghil*, 1997].

The combination of SSA with MEM provides an immediate application of the spectral methods discussed herein to prediction. The key idea, as explained in Sec. 3.3, is that each significant RC obtained from an SSA analysis is a narrow-band time series. It can therefore be predicted fairly robustly by computing a low-order AR(M) process that fits this RC over the time interval available, up to the present epoch. The separate RC predictions can then be combined to form a partial reconstruction of the entire time series' future evolution [*Keppenne and Ghil*, 1992; *Vautard et al.*, 1992].

The decision about which RCs to use for this SSA-MEM prediction is based on a trade-off between the amount of variance one hopes to predict and the reliability of the prediction. The former consideration suggests using all significant components (see discussion in Secs. 2.2 and 2.3), the latter suggests restricting oneself to the oscillatory ones [*Plaut et al.*, 1995]. The accuracy of the prediction depends, in either case, on the extent to which the available data from the past do determine the regular — i.e., periodic and multiply periodic — behavior of the time series. Instrumental temperature data over the last few centuries do not seem,

for instance, to determine sufficiently well the behavior of global [*Ghil and Vautard*, 1991] or local [*Plaut et al.*, 1995] temperatures to permit a reliable climate forecast on the decadal time scale by this SSA-MEM method [*Ghil*, 1997].

On the seasonal-to-interannual time scale, it appears that the doubly-periodic variability associated with the low-frequency and quasi-biennial ENSO modes is sufficiently large and regular to help prediction substantially. We saw in Sec. 2.2 that these two modes capture 43% of the variance in the SOI time series and that their constructive interference accounts essentially for all strong and moderate El Niños and La Niñas (see Fig. 6). Table 4 shows that the SSA-MEM method performs as well as, or better than, two other statistical methods and three dynamical models, including a fully coupled GCM. This does not mean that dynamical ENSO models ought to be abandoned. It does mean, though, that they have to meet higher standards of forecast skill than beating the classical benchmark of "damped persistence," i.e., persistence gradually regressing to climatology.

Table 4: *Comparison of 6-month-lead skill of six ENSO forecast models.*

More importantly, the relatively high accuracy of the SSA-MEM forecast supports the essential role played by the two oscillatory modes, low-frequency and quasi-biennial, in ENSO variability. It emphasizes therewith the necessity for dynamical models to simulate the correct period, amplitude and phase of these two modes. Once a physical model does that, it can proceed to capture other aspects of ENSO dynamics, which are probably crucial in predicting the magnitude of extreme events, like the 1997-98 El Niño, or the extratropical aspects of seasonal-to-interannual climate variability.

To conclude, spectral methods of time series analysis and prediction are intimately connected with the nonlinear dynamics of climate. Advances in the methodology have benefitted from the deeper insights into the study of time series that dynamical-systems theory provides. These advances are contributing in turn to a better understanding and prediction of

climate phenomena and processes.

Acknowledgments

Appendix A. SSA, Spatial EOFs, and the Karhunen-Loève Theorems

A.1 Spatial and temporal EOFs

Table 5 outlines the formal analogy between SSA and the principal component analysis (PCA) of variability in spatial fields. The key idea in both cases is to identify the main patterns of this variability, in decreasing order of the associated variance. This idea is carried out by diagonalizing the appropriate covariance or correlation matrix, as shown in the table below.

Table 5: *Analogy between Karhunen–Loève decomposition in space and time.*

The eigenvectors of this matrix are called empirical orthogonal functions (EOFs) in the meteorological literature. We refer to *Preisendorfer* [1988] for the long history of the meteorological and oceanographic applications, and for references to the early work of I. Holmström, J.E. Kutzbach, E.N. Lorenz, and W.D. Sellers. In the fluid-mechanics literature, the whole analysis of turbulent fields by the principal-component approach has become known as proper orthogonal decomposition (POD). An excellent discussion and many more references are given in Ch. 3 of *Holmes et al.* [1996].

The rigorous mathematical foundation for the spectral representation of random fields is analogous to that of orthonormal bases for the solutions of the classical equations of mathematical physics [*Courant and Hilbert*, 1953, 1962]. In the case of random fields that possess finite first- and second-order moments, the counterpart of the linear partial differential operator being diagonalized is the two-point covariance operator. The appropriate theorems have been proven in the 1940s by *Kosambi* [1943], *Loève* [1945] and *Karhunen* [1946]. Mathematically rigorous results in the time domain, rather than the spatial domain, assume that the univariate stochastic process is *weakly stationary* or stationary in the wide sense (see *Hannan*

[1960], *Ghil and Taricco* [1997], and *Mallat* [1998]). *Wide-sense stationarity* requires the process to have the same finite mean and finite variance at each time t and a lag-autocorrelation function that is finite for all lags and depends only on lag.

The covariance matrices obtained when discretizing the spatially continuous random fields or time-continuous processes of Karhunen-Loève theory are often quite large. Because of their symmetry, the singular value decomposition (SVD) of these matrices is especially simple [*Golub and Van Loan*, 1996, Sec. 2.5] and hence it is the method of choice for their diagonalization. SVD is, in fact, the most widely used algorithm in the whole of numerical linear algebra.

It seems that *Colebrook* [1978] was the first one to observe in the oceanographic literature that the eigenvectors of the lag-correlation matrix are data-adaptive band-pass filters in the time domain. He used this observation to improve the spectral analysis of zooplankton records in the North Atlantic, grouped either by species or by area into "spatial" eigenvectors. *Vautard and Ghil* [1989] noticed the special role played by pairs of nearly equal eigenvalues that are associated with temporal EOFs in phase quadrature, and their possible connection to an efficient representation of anharmonic limit cycles.

Barnett and Hasselmann [1979] proposed the use of EOFs that incorporate a given number of lags for the prediction of atmospheric and oceanic fields in the Tropical Pacific. *Weare and Nasstrom* [1982] coined the term *extended EOF* (EEOF) analysis for their use of PCA to study variability in a sliding time window of meteorological and oceanographic fields. *Lau and Chan* [1985] applied this analysis to outgoing longwave radiation data over the Indian and tropical Pacific oceans.

All these applications used a small and fixed number of lags: two in the case of *Lau and Chan* [1985], three for *Weare and Nasstrom* [1982] and 10 or 12 in *Barnett and Hasselmann*'s [1979] work. This restriction was essential because of these authors' retaining the full extent of the spatial information provided.

Kimoto et al. [1991] and *Plaut and Vautard* [1994], on the other hand, were fully aware of the need to use variable windows in univariate SSA to ensure the statistical significance of the results, while capturing the main periods of interest. To permit the use of large and variable lag windows in multi-channel SSA (M-SSA), they recommended prior compression of the spatial information. This compression can be carried out either by spatial averaging [*Dettinger et al.*, 1995*b*; *Jiang et al.*, 1995*a*], or by prefiltering that uses spatial EOF analysis [*Plaut and Vautard*, 1994; *Moron et al.*, 1998].

EEOF analysis and M-SSA are thus both extensions of the classical Karhunen-Loève approach to the analysis of random fields and stationary processes. They both analyze, in one fell swoop, spatio-temporal variability of climatic fields and both use, more often than not, the classical SVD algorithm of numerical linear algebra. EEOF analysis has its roots in and stays closer to spatial EOF analysis. M-SSA has its roots in univariate SSA and has been developing a battery of statistical significance tests for spatio-temporal variability inspired by those incorporated in the SSA-MTM Toolkit [*Plaut and Vautard*, 1994; *Allen and Robertson*, 1996].

A.2   Complementary windows for M-SSA

Both univariate SSA and conventional PCA analysis are special cases of EEOF or M-SSA analysis: in SSA the number L of channels equals 1, while in PCA the number M of lags is 1. Following *Allen and Robertson* [1996], both algorithms can be understood in practice in terms of a "window." These windows are illustrated schematically in Fig. 21.

Figure 21: *Comparison between PCA, SSA, and EEOFs–M-SSA.*

Standard PCA slides a flat and narrow window, of length 1 and width L, over the data set's N fields, each of which contains L data points. PCA thus identifies the spatial patterns, i.e., the EOFs, which account for a high proportion of the variance in the N views of the data

set thus obtained. Equivalently, PCA can be described as sliding a long and narrow, $N \times 1$ window across the L input channels and identifying high-variance temporal patterns, i.e. the PCs, in the corresponding L views. In Fig. 21, the former view of things corresponds to sliding the $1 \times L$ window parallel to the x-axis along the t-axis. In the latter view, one slides an $N \times 1$ window that starts out by lying along the t-axis parallel to the x-axis. This latter case is also somewhat analogous to the trajectory-matrix version, Eq. (2.5), of single-channel SSA, except that in SSA we *reduce* the length of this long and narrow window to M, and consider only one channel.

These two different window interpretations carry over into M-SSA. In the first case one proceeds from spatial PCA to EEOFs. To do so, we extend our $1 \times L$ window by M lags to form an $M \times L$ window that lies in the horizontal (x, s)-plane of Fig. 21. By moving this window along the t-axis, we search for spatio-temporal patterns, i.e. the EEOFs, that maximize the variance in the $N' = N - M + 1$ overlapping views of the time series thus obtained. The EEOFs are the eigenvectors of the $LM \times LM$ lag-covariance matrix given by Eqs. (4.9)–(4.11).

The second conceptual route leads from single-channel SSA to M-SSA. To follow this route, we reduce the length of the $N \times 1$ window to $N' \times 1$. This window still lies initially along the t-axis and we search for temporal patterns, i.e. the N'-long PCs, that maximize the variance in the $M \times L$ views of the time series [*Allen and Robertson*, 1996; *Robertson*, 1996]. This is equivalent to SSA of the univariate time series formed by stringing together all the channels of the original multi-channel time series end-to-end, with the *complementary window N'* playing the role of M in SSA. The (reduced) PCs are the eigenvectors of the $N' \times N'$ *reduced covariance matrix* with the elements

$$C_{jj'}^{(R)} = \frac{1}{LM} \sum_{l=1}^{L} \left[\sum_{t=1}^{M} X_l(t + j - 1) X_l(t + j' - 1) \right]. \tag{A.1}$$

This matrix, unlike the $LM \times LM$ matrix of Eq. (4.9), is given by

$$\mathbf{C}^{(R)} = \frac{1}{LM} \tilde{\mathbf{X}} \tilde{\mathbf{X}}^{\mathrm{t}}. \tag{A.2}$$

In typical climate applications, $N' < LM$; hence $\mathbf{C}^{(R)}$ is smaller than $\tilde{\mathbf{C}}_{\mathbf{X}}$ and the latter has a null space of dimension $LM - N'$. The eigenvectors of $\tilde{\mathbf{C}}_{\mathbf{X}}$ that lie outside this null space, i.e., that correspond to nonzero eigenvalues, are identical to those of $\mathbf{C}^{(R)}$.

References

Abarbanel H. D. I. (1996) *Analysis of Observed Chaotic Data.* Springer-Verlag, New York, 272 pp

Akaike H. (1969) Fitting autoregressive models for prediction. *Ann. Inst. Statist. Math. 21*, 243–247

Akaike H. (1974) A new look at the statistical model identification. *IEEE Trans. Autom. Control 19*, 716–723

Allen M. (1992) Interactions Between the Atmosphere and Oceans on Time Scales of Weeks to Years. PhD thesis University of Oxford, 202 pp

Allen M., Smith L. A. (1994) Investigating the origins and significance of low-frequency modes of climate variability. *Geophys. Res. Lett. 21*, 883–886

Allen M. R., Robertson A. W. (1996) Distinguishing modulated oscillations from coloured noise in multivariate datasets. *Clim. Dyn. 12*, 775–784

Allen M. R., Smith L. A. (1996) Monte Carlo SSA: Detecting irregular oscillations in the presence of coloured noise. *J. Clim. 9*, 3373–3404

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

Allen M. R., Read P., Smith L. (1992*a*) Temperature time-series? *Nature 355*, 686

Allen M. R., Read P., Smith L. (1992*b*) Temperature oscillations. *Nature 359*, 679

Arneodo A., Argoul F., Muzy J. F., Tabard M. (1993) Beyond classical multifractal analysis using wavelets: Uncovering a multiplicative process hidden in the geometrical complexity of diffusion limited aggregates. *Fractals 1*, 629–649

Arnold L. (1974) *Stochastic Differential Equations: Theory and Applications*, J. Wiley, New York, 228 pp

Arnold V. I. (1973) *Ordinary Differential Equations.* MIT Press, Cambridge MA London England, 280 pp

Arnold V. I. (1983) *Geometrical Methods in the Theory of Ordinary Differential Equations.* Springer-Verlag, New York Heidelberg Berlin

Bak P. (1986) The Devil's staircase. *Physics Today 39*, 38–45

Barnett T. P., Hasselmann K. (1979) Techniques of linear prediction, with application to oceanic and atmospheric fields in the tropical Pacific. *Rev. Geophys. Space Phys. 17*, 949–968

Barnett T. P., Latif M, Graham N., Flügel M., Pazan S., White W. (1993) ENSO and ENSO-related predictability. Part I: Prediction of equatorial Pacific sea surface temperature with a hybrid coupled ocean-atmosphere model. *J. Climate 6*, 1545–1566

Barnston A. G., Ropelewski C. F. (1992) Prediction of ENSO episodes using canonical correlation analysis. *J. Climate 5*, 1316–1345

Barnston, A. G., Van den Dool H., Zebiak S., Barnett T. P., Jo M., Rodenhuis D., Cane M., Leetmaa A., Graham N., Ropelewski C., Kousky V., O'Lenic E., Livezey T. (1994) Long-lead seasonal forecasts - Where do we stand? *Bull. Amer. Met. Soc. 75*, 2097–2114

Benoist J. P. (1986) Analyse spectrale de signaux glaciologiques: Etude des glaces sédimentaires déposées à Dome C, Morphologie du lit d'un glacier. Thèse d'Etat USMT Grenoble

Berger A. L., Melice J. L., Hinnov L. (1991) A strategy for frequency spectra of Quaternary climate records. *Clim. Dyn. 5*, 227–240

Birchfield G. E., Ghil M. (1993) Climate evolution in the Pliocene-Pleistocene as seen in deep sea δ^{18}O records and in simulations: internal variability versus orbital forcing. *J. Geophys. Res. 98*(D6), 10385–10399

Blackman R. B., Tukey J. W. (1958) *The Measurement of Power Spectra From the Point of View of Communication Engineering.* Dover, New York

Bloomfield P. (1976) *Fourier Analysis of Time Series: An Introduction.* J. Wiley and Sons, New York

Boiseau M., Ghil M., Juillet-Leclerc A. (1999) Trends and interdecadal variability from South-Central Pacific coral records. *Geophys. Res. Lett. 26*, 2881–2884

Box G. E. P., Jenkins G. M. (1970) *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco

Broomhead D. S., King G. P. (1986*a*) Extracting qualitative dynamics from experimental data. *Physica D20*, 217–236

Broomhead D. S., King G. P. (1986*b*) On the qualitative analysis of experimental dynamical systems. In: Sarkar S. (ed.) Nonlinear Phenomena and Chaos. Adam Hilger, Bristol, pp 113–144

Broomhead D. S., Jones R., King G. P. (1987) Topological dimension and local coordinates from time series data. *J. Phys. A 20*, 563–569

Burg J. P. (1967) Maximum entropy spectral analysis. In: 37th Ann. Intern. Meeting. Soc. Explor. Geophys., Oklahoma City, Oklahoma

Chang P., Wang B., Li T., Ji L. (1994) Interactions between the seasonal cycle and the southern oscillation: Frequency entrainment and chaos in an intermediate coupled ocean-atmosphere model. *Geophys. Res. Lett. 21*, 2817–2820

Chao, Y., Ghil M., Li X., McWilliams J. C. 2000: Pacific interdecadal variability in this century's sea surface temperatures. *Geophys. Res. Lett. 27*, 2261–2264

Chappellaz J., Barnola J. M., Raynaud D., Korotkevitch Y. S., Lorius C. (1990) Ice-core record of atmospheric methane over the past 160,000 years. *Nature 345*, 127–131

Chatfield C. (1984) *The Analysis of Time Series: An Introduction.* 3rd edn, Chapman and Hall, New York, 286 pp

Chen F., Ghil M. (1995) Interdecadal variablity of the thermohaline circulation and high-latitude surface fluxes. *J. Phys. Oceanogr. 25*(11), 2547–2568

Chen F., Ghil M. (1996) Interdecadal variability in a hybrid coupled ocean-atmosphere model. *J. Phys. Oceanogr. 26*(8), 1561–1578

Childers D. G. (ed.) (1978) *Modern Spectrum Analysis.* IEEE Press, New York, 331 pp

Cochrane D., Orcutt G. H. (1949) Application of least squares regression to relationships containing autocorrelated error terms. *J. Amer. Stat. Assoc. 44*, 32–61

Colebrook J. M. (1978) Continuous plankton records: zooplankton and environment, North-East Atlantic and North Sea. *Oceanol. Acta 1*, 9–23

Constantin P., Foias C., Nicolaenko B., Temam R. (1989) *Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations.* Springer-Verlag, New York, 122 pp

Cook E. R., Buckley B. M., D'Arrigo R. D. (1995) Interdecadal temperature oscillations in the Southern Hemisphere: Evidence from Tasmanian tree rings since 300 B. C. In: Martinson D. G. et al. (eds.) National Natural Climate Variability on Decade-to-Century Time-Scales. National Academy Press, Washington, D.C., pp 523–532

Courant R., Hilbert D. (1953) *Methods of Mathematical Physics*. Vol I Wiley-Interscience, New York/London/Sidney, 560 pp

Courant R., Hilbert D. (1962) *Methods of Mathematical Physics*. Vol II Wiley-Interscience, New York/London/Sidney, 830 pp

Daubechies I. (1992) *Ten Lectures on Wavelets*. SIAM, Philadelphia PA

Delprat N., Escudé B., Guillemain P., Kronland-Martinet R., Tchamitchian P., Torrésani B. (1992) Asymptotic wavelet and gabor analysis: Extraction of instantaneous frequencies. *IEEE Trans. Inform. Theory 38*, 644–664

Delworth T., Manabe S., Stouffer R. (1993) Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. *J. Clim. 6*, 1993–2011

Dettinger M. D., Ghil M. (1998) Seasonal and interannual variations of atmospheric CO_2 and climate. *Tellus 50B*, 1–24

Dettinger M. D., Ghil M., Strong C. M., Weibel W., Yiou P. (1995*a*) Software expedites singular-spectrum analysis of noisy time series. *Eos Trans. AGU 76*(2), 12, 14, 21 (version 4.0 available as freeware on the World Wide Web at `http://www.atmos.ucla.edu/tcd`)

Dettinger M. D., Ghil M., Keppenne C. (1995*b*) Interannual and interdecadal variability in United States surface-air temperatures, 1910–87. *Clim. Change 31*, 35–66

Dijkstra and Molemaker (1999) Imperfections of the North-Atlantic wind-driven ocean circulation: Continental geometry and wind stress. *J. Mar. Res. 57*, 1–28

Drazin P. G., King G. P. (1992) Introduction. *Physica D 58*, 7–11

Eckmann J.-P. (1981) Roads to turbulence in dissipative dynamical systems. *Rev. Mod. Phys. 53*, 643–654

Eckmann J.-P., Ruelle D. (1985) Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys. 57*, 617–656 and *57*, 1115

Egger J. (1999) POPs and MOPs. *Clim. Dyn. 15*, 561–568

Elsner J. B., Tsonis A. A. (1991) Do bidecadal oscillations exist in the global temperature record? *Nature 353*, 551–553

Farge M. (1992) Wavelet transforms and their applications to turbulence. *Ann. Rev. Fluid Mech. 24*, 395–457

Farmer J. D., Ott E., Yorke J. A. (1983) The dimension of chaotic attractors. *Physica D 7*, 153–180

Feigenbaum M. J., (1978) Quantitative universality for a class of non-linear transformations. *J. Stat. Phys. 19*, 25–52

Feigenbaum M. J., Kadanoff L. P., Shenker S. J. (1982) Quasiperiodicity in dissipative systems: A renormalization group analysis. *Physica D 5*, 370–386

Feller W. (1968) *An Introduction to Probability Theory and its Applications.* 3rd edn, vol. I, Wiley, New York, 509 pp

Feller W. (1970) *An Introduction to Probability Theory and its Applications.* 2nd edn, vol. II, Wiley, New York, 699 pp

Flandrin P. (1999) *Time-Frequency/Time-Scale Analysis.* Academic Press, San Diego, 386 pp

Foias C., Frazho A. E., Sherman P. J. (1988) A geometric approach to the maximum likelihood spectral estimator of sinusoids in noise. *IEEE Trans. Info. Theory 34*, 1066–1070

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

Fraedrich K., Boettger H. (1978) A wavenumber-frequency analysis of the 500 mb geopotential at 50N. *J. Atmos. Sci., 35*, 745–750

Gabor D. (1946) Theory of communication. *J. Inst. Electr. Eng. 93*, 429–457

Ghil M. (1994) Cryothermodynamics: the chaotic dynamics of paleoclimate. *Physica D77*, 130–159

Ghil M. (1995) Atmospheric modeling. In: Martinson D. G. et al. (eds.) National Natural Climate Variability on Decade-to-Century Time-Scales. National Academy Press, Washington, D.C., pp 164–168

Ghil M. (1997) The SSA-MTM Toolkit: Applications to analysis and prediction of time series. *Proc. SPIE 3165*, 216–230

Ghil M., Childress S. (1987) *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory and Climate Dynamics.* Springer-Verlag, New York, 485 pp

Ghil M., Mo K. C. (1991*a*) Intraseasonal oscillations in the global atmosphere. Part I: Northern hemisphere and tropics. *J. Atmos. Sci. 48*(5), 752–779

Ghil M., Mo K. C. (1991*b*) Intraseasonal oscillations in the global atmosphere. Part II: Southern hemisphere. *J. Atmos. Sci. 48*(5), 780–790

Ghil M., Vautard R. (1991) Interdecadal oscillations and the warming trend in global temperature time series. *Nature 350*(6316), 324–327

Ghil M., Yiou P. (1996) Spectral methods: What they can and cannot do for climatic time series. In: Anderson D., Willebrand J. (eds.) Decadal Climate Variability: Dynamics and Predictability. Springer-Verlag, Berlin/Heidelberg, pp 445–481

Ghil M., Taricco C. (1997) Advanced spectral analysis methods. In: Castagnoli G. C., Provenzale A. (eds.) Past and Present Variability of the Solar-Terrestrial System: Measurement, Data Analysis and Theoretical Models. Societa Italiana di Fisica, Bologna, & IOS Press, Amsterdam pp 137–159

Ghil M., Jiang N. (1998) Recent forecast skill for the El Niño/Southern Oscillation. *Geophys. Res. Lett. 25*, 171–174

Ghil M., Robertson A. W. (2000) Solving problems with GCMs: General circulation models and their role in the climate modeling hierarchy. In: Randall D. (ed.) General Circulation Model Development: Past, Present and Future. Academic Press, San Diego, pp 285–325

Ghil M., Kimoto M., Neelin J. D. (1991) Nonlinear dynamics and predictability in the atmospheric sciences. *Rev. Geophys. 36 (Supplement)*, 46–55 (U. S. National Report IUGG, 1987–91)

Gilman D. L., Fuglister, F. J., Mitchell, J. J. (1963) On the power spectrum of "red noise." *J. Atmos. Sci. 20*, 182–184

Gleick J. (1987) *Chaos: Making a New Science.* Viking, New York, 354 pp

Golub G. H., Van Loan C. F. (1996) *Matrix Computations.* 3rd edn, The Johns Hopkins University Press, Baltimore London, 642 pp

Grassberger P., Procaccia I. (1983) Measuring the strangeness of strange attractors. *Physica D9*, 189–208

Guckenheimer J., Holmes P. (1983) *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.* Springer-Verlag, New York, 459 pp

Hannan E. J. (1960) *Time Series Analysis.* Methuen, London, 152 pp

Hannan E. J. (1970) *Multiple Time Series.* Wiley, New York, 536 pp

Hasan T.(1983) Complex demodulation: Some theory and applications. In: Brillinger D. R. and Krishnaiah P. R. (eds.) Handbook of Statistics 3: Time Series in the Frequency Domain. North Holland, Amsterdam, pp 125–156

Hasselmann K. (1976) Stochastic climate models. *Tellus 6*, 473–485

Hasselmann K. (1988) PIPs and POPs: The reduction of complex dynamical systems using Principal Interaction and Principal Oscillation Patterns. *J. Geophys. Res. 93D*, 11015–11021

Haykin S., Kessler S. (1983) Prediction-error filtering and maximum-entropy spectral estimation. In: Haykin S. (ed.) Nonlinear Methods of Spectral Analysis. Springer-Verlag, Berlin, pp 9–72

Holmes P., Lumley J. L., Berkooz G. (1996) *Turbulence, Coherent Structures, Dynamical Systems and Symmetry.* Cambridge University Press

Intergovernmental Panel on Climate Change (1990) *Climate Change, The IPCC Scientific Assessment,* J. T. Houghton, G. J. Jenkins, and J. J. Ephraums (eds.). Cambridge Univ. Press, Cambridge, MA, 365 pp

Jenkins G. M., Watts D. G. (1968) *Spectral Analysis and its Applications.* Holden-Day, San Francisco

Ji M., Kumar A., Leetmaa A. (1994) An experimental coupled forecast system at the National Meteorological Center: Some early results. *Tellus 46A*, 398–418

Jiang N., Neelin D., Ghil M. (1995*a*) Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific. *Clim. Dyn. 12*, 101–112

Jiang N., Ghil M., Neelin D. (1995*b*) Forecasts of equatorial Pacific SST anomalies by using an autoregressive process and singular spectrum analysis. In: Experimental Long-Lead Forecast Bulletin Vol. 4, No. 1 (Jan. 1995), pp. 24–27 and subsequent quaterly issues (1995–1997), National Meteorological Center, NOAA, U.S. Department of Commerce

Jiang S., Jin F.-F., Ghil M. (1995*c*) Multiple equilibria, periodic, and aperiodic solutions in a wind-driven, double-gyre, shallow-water model. *J. Phys. Oceanogr. 25*, 764–786

Jin F.-F., Neelin J. D., Ghil M. (1994) El Niño on the Devil's staircase: Annual subharmonic steps to chaos. *Science 264*, 70–72

Jin F.-F., Neelin J. D., Ghil M. (1996) El Niño/Southern Oscillation and the annual cycle: Subharmonic frequency-locking and aperiodicity. *Physica D 98*, 442–465

Jouzel, J., Barkov N. I., Barkov J. M., Barnola M., Bender J., Chappellaz C., Genthon V. M., Kotlyakov V., Lipenkov C., Lorius J. R., Petit D., Raynaud G., Raisbeck C., Ritz T., Sowers M., Stievenard F., Yiou P. (1993) Extending the Vostok ice-core record of paleoclimate to the penultimate glacial period. *Nature 364*, 407–412

Kaplan J. L., Yorke J. A. (1979) Chaotic behavior of multidimensional difference equations. In: Peitgen H. O., Walther H. O. (eds.) Functional Differential Equations and Approximation of Fixed Points Vol 730 of *Lecture Notes in Mathematics*. Springer-Verlag, New York, pp 228–237

Karhunen K. (1946) Zur Spektraltheorie stochastischer Prozesse. *Ann. Acad. Sci. Fenicae A1*, 34

Kay S. M. (1988) *Modern Spectral Analysis: Theory and Applications*. Prentice-Hall, Englewood Cliffs, NJ, 543 pp

Kendall M., Stuart A. (1977) *The Advanced Theory of Statistics*. Vol 2, 4th edn, Macmillan, New York

Keppenne C. L., Ghil M. (1992) Adaptive filtering and prediction of the Southern Oscillation Index. *J. Geophys. Res. 97*, 20449–20454

Keppenne C. L., Ghil M. (1993) Adaptive filtering and prediction of noisy multivariate signals: An application to subannual variability in atmospheric angular momentum. *Intl. J. Bifurcation & Chaos 3*, 625–634

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

Kimoto M., Ghil M., Mo K. C. (1991) Spatial structure of the extratropical 40-day oscillation. In: Proc. 8th Conf. Atmos. Oceanic Waves and Stability. Amer. Meteorol. Soc. , Boston, MA, pp 115–116

Koch D. M., Mann M. E. (1996) Spatial and temporal variability of ^{7}Be surface concentration. *Tellus 48B*, 387–398

Kosambi D. D. (1943) Statistics in function space. *J. Indian Math. Soc. 7*, 76–88

Kumar P., Foufoula-Georgiou E. (1997) Wavelet analysis for geophysical applications. *Rev. Geophys. 35*, 385–412

Kumaresan R., Tufts D. W. (1980) Data-adaptive principal component signal processing. In: Proc. Conf. Decision and Control IEEE , Albuquerque, pp 949–954

Kuo C., Lindberg C., Thomson D. J. (1990) Coherence established between atmospheric carbon dioxide and global temperature. *Nature 343*, 709–713

Lall U., Mann M. (1995) The Great Salt Lake: A barometer of low-frequency climatic variability. *Wat. Res. Res. 31*, 2503–2515

Lasota A., Mackey M. C. (1994) *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics.* 2nd edn, Springer-Verlag, New York

Latif M., Barnett T. (1994) Causes of decadal climate variability over the North Pacific and North America *Science* 266, 634–637

Lau K. M., Chan P. H. (1985) Aspects of the 40-50 day oscillation during the northern winter as inferred from outgoing longwave radiation. *Mon. Wea. Rev. 113*, 1889–1909

Lees J. M. (1995) Reshaping spectrum estimates by removing periodic noise: application to seismic spectral rations. *Geophys. Res. Lett. 21*, 199–236

Legras B., Ghil M. (1985) Persistent anomalies, blocking and variations in atmospheric predictability. *J. Atmos. Sci. 42*, 433–471

Leith C. E. (1975) Climate response and fluctuation dissipation. *J. Atmos. Sci. 32*, 2022–2026

Lichtenberg A. J., Lieberman M. A. (1991) *Regular and Chaotic Dynamics*. 2nd edn, Springer-Verlag, New York, 692 pp

Loève M. (1945) Fonctions aléatoires de second ordre. *C. R. Acad. Sci. Paris 220*

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

Lorenz E. N. (1969) Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci. 26*, 636–646

MacDonald G. J. (1989) Spectral analysis of time series generated by nonlinear processes. *Rev. Geophys. 27*, 449–469

Mallat S. (1998) *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 577 pp

Mallat S., Zhang Z. (1993) Matching pursuit with time-frequency dictionaries. *IEEE Trans. Signal Process. 41*, 3397–3415

Mandelbrot B. B. (1982) *The Fractal Geometry of Nature*. 2nd edn, Freeman, San Francisco, 460 pp

Mañé R. (1981) On the dimension of the compact invariant sets of certain non-linear maps. In: Rand D. A., Young L.-S. (eds.) Dynamical Systems and Turbulence Vol 898 of *Lecture Notes in Mathematics.* Springer, Berlin, pp 230–242

Mann M. E., Park J. (1993) Spatial correlations of interdecadal variation in global surface temperatures. *Geophys. Res. Lett. 20*, 1055–1058

Mann M. E., Park J. (1994) Global scale modes of surface temperature variability on inter-annual to century time scales. *J. Geophys. Res. 99*, 25819–25933

Mann M. E., Park J. (1996*a*) Greenhouse warming and changes in the seasonal cycle of temperature: model versus observations. *Geophys. Res. Lett. 23*, 1111–1114

Mann M. E., Park J. (1996*b*) Joint spatiotemporal modes of surface temperature and sea level pressure variablity in the Northern Hemisphere during the last century. *J. Clim. 9*, 2137–2162

Mann M. E., Lees J. M. (1996) Robust estimation of background noise and signal detection in climatic time series. *Clim. Change 33*, 409–445

Mann M. E., Park J. (1999) Oscillatory spatiotemporal signal detection in climate studies: A multiple-taper spectral domain approach. *Adv. Geophys. 41*, 1–131

Mann M. E., Lall U., Saltzman B. (1995*a*) Decadal-to-centennial-scale climate variability: Insights into the rise and fall of the Great Salt Lake. *Geophys. Res. Lett. 22*(8), 937–940

Mann M. E., Park J., Bradley R. S. (1995*b*) Global interdecadal and century-scale climate oscillations during the past five centuries. *Nature 378*, 266–270

Meyer Y. (1992) *Wavelets and Operators.* Cambridge University Press, Cambridge (England), 223 pp

Meyer Y. (1993) *Wavelets: Algorithms and Applications.* SIAM, Philadelphia PA, 133 pp

Meyers S. D., Kelly B. G., O'Brien J. J. (1993) An introduction to wavelet analysis in oceanography and meteorology: with application to the dispersion of Yanai waves. *Mon. Wea. Rev. 121*, 2858–2866

Mitchell J. M. (1976) An overview of climatic variability and its causal mechanisms. *Quatern. Res. 6*, 481–493

Mommersteeg H., Loutre M. F., Young R., Wijmstra T. A., Hooghiemstra H. (1995) Orbital forced frequencies in the 975000 year pollen record from Tenagi Philippon (Greece). *Clim. Dyn. 11*, 4–24

Moron V., Vautard R., Ghil M. (1998) Trends, interdecadal and interannual oscillations in global sea-surface temperatures. *Clim. Dyn. 14*, 545–569

Neelin J. D. (1990) A hybrid coupled general circulation model for El Niño studies. *J. Atmos. Sci. 47*, 674–693

Neelin J. D., Latif M., Jin F.-F. (1994) Dynamics of coupled ocean-atmosphere models: the tropical problem. *Annu. Rev. Fluid Mech. 26*, 617–659

Neelin J. D., Battisti D. S., Hirst A. C., Jin F.-F., Wakata Y., Yamagata T., Zebiak S. E. (1998) ENSO theory. *J. Geophys. Res. 103*, 14261–14290

Ooms M. (1988) *Empirical Vector Autoregressive Modeling.* Springer-Verlag

Ott E., Sauer T., Yorke J. A. (eds.) (1994) *Coping with Chaos, Analysis of Chaotic Data and the Exploitation of Chaotic Systems.* J. Wiley & Sons, 418 pp

Packard N. H., Crutchfield J. P., Farmer J. D., Shaw R. S. (1980) Geometry from a time series. *Phys. Rev. Lett. 45*, 712–716

Paluš M., Dvořák I. (1992) Singular value decomposition in attractor reconstruction: pitfalls and precautions. *Physica D55*, 221–234

Park J. (1992) Envelope estimation for quasi-periodic geophysical signals in noise: A multi-taper approach. In: Walden A. T., Guttorp P. (eds.) Statistics in the Environmental & Earth Sciences. Edward Arnold, pp 189–219

Park J., Maasch K. A. (1993) Plio-Pleistocene time evolution of the 100-kyr cycle in marine paleoclimate records. *J. Geophys. Res. 98*, 447–461

Park J., Mann M. E. (2000) Interannual temperature events and shifts in global temperature: A multiwavelet correlation approach. *Earth Interactions 4*, online paper 4-001

Park J., Lindberg C. R., Vernon F. L. I. (1987) Multitaper spectral analysis of high frequency seismograms. *J. Geophys. Res. 92*, 12675–12684

Penland C. (1989) Random forcing and forecasting using Principal Oscillation Pattern analysis. *Mon. Weather Rev. 117*, 2165–2185

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

Penland C., Ghil M. (1993) Forecasting Northern Hemisphere 700-mb geopotential height anomalies using empirical normal modes. *Mon. Wea. Rev. 121*, 2355–2372

Penland C., Ghil M., Weickmann K. (1991) Adaptive filtering and maximum entropy spectra, with application to changes in atmospheric angular momentum. *J. Geophys. Res. 96*(D12), 22659–22671

Percival D. B., Walden A. T. (1993) *Spectral Analysis for Physical Applications*. Cambridge University Press, Cambridge UK, 583 pp

Pike E. R., McWhirter J. G., Bertero M., de Mol C. (1984) Generalized information theory for inverse problems in signal processing. *IEE Proc. 131*, 660–667

Plaut G., Vautard R. (1994) Spells of low-frequency oscillations and weather regimes in the northern hemisphere. *J. Atmos. Sci. 51*(2), 210–236

Plaut G., Ghil M., Vautard R. (1995) Interannual and interdecadal variability in 335 years of Central England temperatures. *Science 268*, 710–713

Preisendorfer R. W. (1988) *Principal Component Analysis in Meteorology and Oceanography.* Elsevier, Amsterdam, 425 pp

Press W. H., Flannery B. P., Teukolski S. A., Vettering W. T. (1988) *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, 818 pp

Priestley, M. B. (1981*a*) *Spectral Analysis and Time Series, vol. 1 (Univariate Series).* Academic Press San Diego, 653 pp

Priestley, M. B. (1981*b*) *Spectral Analysis and Time Series, vol. 2 (Multivariate Series, Prediction and Control).* Academic Press San Diego, pp 654–890

Quinn T. M., Taylor F. W., Crowley T. J. (1993) A 173 year stable isotope record from a tropical south pacific coral. *Quat. Sci. Rev. 12*, 407–418

Quon C., Ghil M. (1995) Multiple equilibria and stable oscillations in thermosolutal convection at small aspect ratio. *J. Fluid Mech. 291*, 33–56

Rahmstorf S. (1995) Bifurcation of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. *Nature 378*, 145–149

Rasmusson E. M., Wang X., Ropelewski C. F. (1990) The biennial component of ENSO variability. *J. Mar. Syst. 1*, 71–96

Rayner N. A., Folland C. K., Parker D. E., Horton E. B. (1995) A new global sea-ice and sea surface temperature (GISST) data set for 1903–1994 for forcing climate models. *Hadley Centre Internal Note No. 69*, pp 13

Robertson A. W. (1996) Interdecadal variability over the North Pacific in a multi-century climate simulation. *Clim. Dyn. 12*, 227–241

Robertson A. W., Ma C.-C., Mechoso C. R., Ghil M. (1995*a*) Simulation of the Tropical-Pacific climate with a coupled ocean-atmosphere general circulation model. Part I: The seasonal cycle. *J. Clim. 8*, 1178–1198

Robertson A. W., Ma C.-C., Ghil M., Mechoso C. R. (1995*b*) Simulation of the Tropical-Pacific climate with a coupled ocean-atmosphere general circulation model. Part II: Interannual variability. *J. Clim. 8*, 1199–1216

Robertson A. W., Mechoso C. R. (1998) Interannual and decadal cycles in river flows of southeastern South America. *J. Climate 11*, 2570–2581

Robertson A. W., Mechoso C. R. (2000) Interannual and interdecadal variability of the South Atlantic Convergence Zone. *Mon. Wea. Rev. 128*, 2947–2957

Ropelewski C. F., Jones P. D. (1987) An extension of the Tahiti–Darwin Southern Oscillation index. *Mon. Wea. Rev. 115*, 2161–2165

Roux J. C., Rossi A., Bachelart S., Vidal C. (1980) Representation of a strange attractor from an experimental study of chemical turbulence. *Phys. Lett. A 77*, 391–393

Ruelle D. (1981) Small random perturbations of dynamical systems and the definition of attractors. *Commun. Math. Phys. 82*, 137–151

Ruelle D. (1990) Deterministic chaos: the science and the fiction. *Proc. Roy Soc. London, Ser. A 427*, 241–248

Ruelle D., Takens F. (1971) On the nature of turbulence. *Commun. Math. Phys. 20*, 167–192 and *23*:343–344

Sauer T., Yorke J. A., Casdagli M. (1991) Embedology. *J. Stat. Phys. 65*, 579–616

Saunders A., Ghil M. (2001) A Boolean delay equation model of ENSO variability. *Physica D, sub judice*

Saunders A., Ghil M., Neelin J.D. (1997) Forecasts of Niño-3 SST anomalies and SOI based on singular spectrum analysis combined with the maximum entropy method (1997-2000), *Experimental Long-Lead Forecasts Bulletin* Vol. 6, No. 3 (Sept. 1997), pp 43–44; subsequent quarterly issues published also electronically, http://grads.igs.org/ellfb (1997-2000)

Schuss Z. (1980) *Theory and Applications of Stochastic Differential Equations*, J. Wiley, New York, 321 pp

Schuster A. (1898) On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena. *J. Geophys. Res. 3*, 13–41

Schuster H. G. (1988) *Deterministic Chaos: An Introduction.* 2nd edn, Physik-Verlag, Weinheim, Germany, 270 pp

Shannon C. E. (1949) Communication in the presence of noise. *Proc. I.R.E. 37*, 10–21

Slepian S. (1978) Prolate spheroidal wave functions, Fourier analysis and uncertainty-V: The discrete case. *Bell. Sys. Tech. J. 57*, 1371–1430

Smale S. (1967) Differentiable dynamical systems. *Bull. Amer. Math. Soc. 73*, 199–206

Speich S., Dijkstra H., Ghil M. (1995) Successive bifurcations in a shallow-water model, applied to the wind-driven ocean circulation. *Nonlin. Proc. Geophys. 2*, 241–268

Stoker J. J. (1950) *Nonlinear Vibrations in Mechanical and Electrical Systems*, Wiley-Interscience, New York, 273 pp

Strang G. (1989) Wavelets and dilation equations: a brief introduction. *SIAM Rev. 31*(4), 614–627

Sutton R. T., Allen M. R. (1997) Decadal predictability of North Atlantic sea surface temperature and climate. *Nature 388* 563–567

Takens F. (1981) Detecting strange attractors in turbulence. In: Rand D. A., Young L.-S. (eds.) Dynamical Systems and Turbulence Vol 898 of *Lecture Notes in Mathematics*. Springer, Berlin, pp 366–381

Temam R. (1997) *Infinite-Dimensional Dynamical Systems in Mechanics and Physics.* 2nd edn, Springer-Verlag, New York, 648 pp

Thomson D. J. (1982) Spectrum estimation and harmonic analysis. *Proc. IEEE 70*(9), 1055–1096

Thomson D. J. (1990*a*) Quadratic-inverse spectrum estimates: applications to palaeoclimatology. *Phil. Trans. R. Soc. Lond. A 332*, 539–597

Thomson D. J. (1990*b*) Time series analysis of Holocene climate data. *Phil. Trans. R. Soc. Lond. A 330*, 601–616

Thomson D. J. (1995) The seasons, global temperature, and precession. *Science 268*, 59–68

Thomson D. J., Chave A. D. (1990) Jackknifed error estimates for spectra, coherences, and transfer functions, Ch. 2 in Haykin, S. (Ed.), *Advances in Spectrum Analysis*. Prentice-Hall

Torrence C., Compo G. P. (1998) A practical guide to wavelet analysis. *Bull. Amer. Meteorol. Soc. 79*, 61–78

Tziperman E., Stone L., Cane M. A., Jarosh H. (1994) El Niño chaos: overlapping of resonances between the seasonal cycle and the Pacific ocean-atmosphere oscillator. *Science 264*, 72–74

Unal Y. S., Ghil M. (1995) Interannual and interdecadal oscillation patterns in sea level. *Clim. Dyn. 11*, 255–278

Van den Dool, H. M. (1994) Searching for analogues, how long must we wait? em Tellus *46A*, 314–324

Van der Pol B. (1940) Biological rhythms considered as relaxation oscillations. *Acta Med. Scand. Suppl. 108*, 76–87

Vautard R., Ghil M. (1989) Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D35*, 395–424

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

Venegas S. A., Mysak L. A., Straub, D. N. (1997) Atmosphere-ocean coupled variability in the South Atlantic. *J. Climate 10*, 2904–2920

Von Storch H., Zwiers F.W. (1999) *Statistical Analysis in Climate Research*, Cambridge University Press, Cambridge, UK, 484 pp

Von Storch H., Bruns T., Fischer-Bruns I., Hasselmann K. (1988) Principal oscillation pattern analysis of the 30-to-60-day oscillation in general circulation model equatorial troposphere. *J. Geophys. Res. 93D*, 11022–11036

Walker G. (1931) On periodicity in series of related terms. *Proc. Roy. Soc. (London) Ser. A 131*, 518–532

Wang B., Wang Y. (1996) Temporal structure of the southern oscillation as revealed by waveform and wavelet analysis. *J. Clim. 9*, 1586–1598

Weare B. C., Nasstrom J. N. (1982) Examples of extended empirical orthogonal function analyses. *Mon. Wea. Rev. 110*, 784–812

Weaver A. J., Marotzke J., Cummins P. F., Sarachik E. S. (1993) Stability and variability of the thermohaline circulation. *J. Phys. Oceanogr. 23*, 39–60

Weng H.-Y., Lau K.-M. (1994) Wavelets, period-doubling and time-frequency localization with application to satellite data analysis. *J. Atmos. Sci. 51*, 2523–2541

Whitney B. (1936) Differentiable manifolds. *Ann. Math. 37*, 645–680

Whittle P. (1952) The simultaneous estimation of a time series harmonic component and covariance structure. *Trab. Estadist 3*, 43–57

Wiener N. (1949) *Extrapolation, Interpolation and Smoothing of Stationary Time Series, with Engineering Applications.* M.I.T. Press, Cambridge, MA, 163 pp

Wiggins S. (1988) *Global Bifurcations and Chaos (Analytical Methods).* Springer-Verlag, New York, 494 pp.

Yiou P. (1994) Dynamique du Paléoclimat: Des Données et des Modèles. PhD thesis Université Pierre et Marie Curie Paris 6, 183 pp

Yiou P., Genthon C., Jouzel J., Ghil M., Le Treut H., Barnola J. M., Lorius C., Korotkevitch Y. N. (1991) High-frequency paleovariability in climate and in CO_2 levels from Vostok ice-core records. *J. Geophys. Res. 96*(B12), 20365–20378

Yiou P., Ghil M., Jouzel J., Paillard D., Vautard R. (1994) Nonlinear variability of the climatic system, from singular and power spectra of Late Quaternary records. *Clim. Dyn. 9*, 371–389

Yiou P., Jouzel J., Johnsen S., Rögnvaldsson Ö. E. (1995) Rapid oscillations in Vostok and GRIP ice cores. *Geophys. Res. Lett. 22*(16), 2179–2182

Yiou P., Loutre M. F., Baert E. (1996) Spectral analysis of climate data. *Surveys Geophys. 17*, 619–663

Yiou P., Fuhrer K., Meeker L. D., Jouzel J., Johnsen S., Mayewski P. A. (1997) Paleoclimatic variability inferred from the spectral analysis of Greenland and Antarctic ice core data. *J. Geophys. Res. 102*(C12), 26441–26454

Yiou P., Sornette D., Ghil M. (2000) Data-adaptive wavelets and multi-scale SSA. *Physica D 142*, 254–290

Yule G. U. (1927) On a method of investigating periodicities in disturbed series. *Phil. Trans. Roy. Soc. (London) Ser. A 226*, 267–298

Zebiak, S. E., Cane M. A. (1987) A model El Nio Southern Oscillation. *Mon. Wea. Rev. 115*, 2262–2278

Tables

Table 1. Glossary of the principal symbols. Method and section number correspond to where the symbol appears first or is used differently from previous sections.

Symbols	Definition	Method	(Section)
$A_k(t)$	k-th principal component (PC) of $\{X(t)\}$	**SSA**	(Sec. 2.2)
$A_k^b(t)$	k-th local PC of $\{X(t)\}$ at time b	**SSA**	(Sec. 2.4)
$A(t)$ & A_n	continuous- & discrete-time envelope function	**MTM**	(Sec. 3.4)
$\mathbf{A}^k$	k-th multi-channel PC	**SSA**	(Sec. 4.2)
a	dilation factor of $\psi(x) \to \psi(x/a)$	**WLTs**	(Sec. 2.4)
a_1	lag-one auto-regression coefficient	**SSA**	(Sec. 2.3)
$\{a_j\}$	true regression coefficients with order M		(Sec. 1)
$\{\hat{a}_j\}$	estimates of a_j with order M'	**MEM**	(Sec. 3.3)
B & $\hat{B}(f_0)$	true & estimated amplitude of oscillation (at frequency f_0)	**MTM**	(Sec. 3.4)
$\mathbf{B}$ & $\mathbf{B}_\tau$	true & estimated dynamics matrix (with lag τ)	**POP**	(Sec. 4.1)
b	translation of $\psi(x) \to \psi(x-b)$	**WLTs**	(Sec. 2.4)
$b_k(f)$	weighting function	**MTM**	(Sec. 3.4)
$\mathbf{C}_R$	covariance of surrogate data for $\{\mathbf{X}(t)\}$	**SSA**	(Sec. 2.3)
$\mathbf{C}^{(\mathbf{R})}$	reduced covariance	**SSA**	(Appendix)
$\mathbf{C}(\tau)$	lag-covariance matrix of $\{\mathbf{X}(t)\}$ with lag τ	**POP**	(Sec. 4.1)
$\mathbf{C}_X$	covariance of $\{\tilde{\mathbf{X}}(t)\}$	**SSA**	(Sec. 2.2)
$\mathbf{C}_{\mathbf{X}}$	grand covariance matrix	**SSA**	(Sec. 4.2)

$\mathbf{D}$	trajectory matrix of $\{\tilde{\mathbf{X}}(t)\}$ in the univariate case		(Sec. 2.2)
$\tilde{\mathbf{D}}$	trajectory matrix of $\{\tilde{\mathbf{X}}_l(t)\}$ in the multivariate case	**SSA**	(Sec. 4.2)
d	dimension of underlying attractor		(Sec. 2.1)
$d\xi(t)$	white-noise vector in continuous time	**POP**	(Sec. 4.1)
$\mathbf{E}^k$	eigenvector matrix of $\tilde{\mathbf{C}}_{\mathbf{X}}$ or $\tilde{\mathbf{T}}_{\mathbf{X}}$	**SSA**	(Sec. 4.2)
F_i	right-hand sides of ODEs		(Sec. 2.1)
$F(f)$	F-test ratio	**MTM**	(Sec. 3.4)
f	frequency	**SSA**	(Sec. 2.3)
$\{f_k\}$	discrete sequence of frequencies	**BT**	(Sec. 3.2)
f_N	Nyquist frequency		(Sec. 2.3)
f_R	Rayleigh frequency	**MTM**	(Sec. 3.4)
f_0	fixed frequency of pure sinusoid	**MTM**	(Sec. 3.4)
$\mathbf{G}(\tau)$	Green's function associated with $\mathbf{B}$ at lag τ	**POP**	(Sec. 4.1)
$\{g_j\}$	smoothing weights for $\bar{S}_X(f_k)$	**BT**	(Sec. 3.2)
I	interval for dilations a	**WLTs**	(Sec. 2.4)
i	time index	**SSA**	(Sec. 2.4)
i	imaginary unit	**BT**	(Sec. 3.4)
K	number of tapers	**MTM**	(Sec. 3.4)
$\mathcal{K}$	a set of indices used in reconstruction	**SSA**	(Sec. 2.2)
L	number of channels	**POP**	(Sec. 4.1)
M	order of auto-regression		(Sec. 1)
M	embedding dimension		(Sec. 2.1)
$M\Delta t$	window width	**SSA**	(Sec. 2.2)
M_t	normalization factor for $R_{\mathcal{K}}(t)$	**SSA**	(Sec. 2.2)
M'	order of MEM	**MEM**	(Sec. 3.3)

N	length of $\{X(n\Delta t)\}$		(Sec. 1)
N'	length of $\{\tilde{\mathbf{X}}(n\Delta t)\}$		(Sec. 2.1)
$\bar{N}$	normalization factor for $\mathbf{T_X}$	**SSA**	(Sec. 4.2)
$P(f)$	cumulative power spectrum		(Sec. 3.1)
p	integer band-width parameter	**MTM**	(Sec. 3.4)
$\mathbf{Q}$	lag-zero covariance matrix of $d\xi(t)$	**POP**	(Sec. 4.1)
$R_{\mathcal{K}}(t)$	reconstructed component (RC) of $\{X(t)\}$ for a set $\mathcal{K}$	**SSA**	(Sec. 2.2)
$R_{\mathcal{K}}^b(t)$	$\mathcal{K}$-th local RC of $\{X(t)\}$ at time b	**SSA**	(Sec. 2.4)
$\mathbf{R}^k$	k-th multi-channel RC	**SSA**	(Sec. 4.2)
r	lag-one auto-correlation	**SSA**	(Sec. 2.3)
$S(f)$	power spectrum of AR(1) process	**SSA**	(Sec. 2.3)
S_0	average of $S(f)$	**SSA**	(Sec. 2.3)
$S_X(f)$ & $\hat{S}_X(f)$	true & estimated periodogram	**BT**	(Sec. 3.2)
$\tilde{S}_X(f)$	correlogram	**BT**	(Sec. 3.2)
$\bar{S}_X(f_k)$ & $\bar{\bar{S}}_X(f)$	direct & indirect estimate of $S_X(f)$	**BT**	(Sec. 3.2)
$S_X(f)$ & $\hat{S}_X(f)$	true & estimated (by $\{\hat{a}_j\}$) power spectrum	**BT**	(Sec. 3.2)
$\hat{S}_k(f)$	k-th eigenspectrum	**MTM**	(Sec. 3.4)
$S_r(f)$ & $S_w(f)$	high-resolution & adaptively-weighted multi-taper spectrum	**MTM**	(Sec. 3.4)
$\mathbf{T_X}$	grand block-matrix for covariances	**SSA**	(Sec. 4.2)
t	continuous time ($t \in \mathbb{R}$) or discrete time ($t \in \mathbb{Z}^+$)		(Sec. 1)
Δt	sampling rate		(Sec. 1)
$n\Delta t$	discretely sampled time		(Sec. 1)
$U_k(f)$	discrete Fourier transform of $w_k(t)$	**MTM**	(Sec. 3.4)

W	sliding-window length of wavelet	**SSA**	(Sec. 2.4)
$W_\psi(a,b)$	wavelet transform of $\{X(t)\}$ using b-translated and a-dilated $\psi((t-b)/a)$	**SSA**	(Sec. 2.4)
$W_\mu(k)$	lag-window for $\bar{\bar{S}}_X(f)$ with smoothing parameter μ	**BT**	(Sec. 3.2)
$W_m(k)$	Barlett window for $\bar{\bar{S}}_X(f)$ with window length m	**BT**	(Sec. 3.2)
$w_k(t)$	k-th taper	**MTM**	(Sec. 3.4)
$\hat{X}(t)$	observed time series		(Sec. 1)
X_n	$X(n\Delta t)$		(Sec. 1)
$X^{(p)}$	p-th order differentiation of $X(t)$ with respect to time t, $d^p X/dt^p$		(Sec. 1)
$\{X(t)\}$	univariate time series in $t \in \mathbb{R}$ or $t \in \mathbb{Z}^+$		(Sec. 1)
$\tilde{\mathbf{X}}(t)$	M-dimensional augmented vector of $X(t)$		(Sec. 2.1)
X_0	mean of $\{\mathbf{X}(t)\}$	**SSA**	(Sec. 2.3)
$X_I(t)$	reconstructed $X(t)$	**WLTs**	(Sec. 2.4)
$\tilde{X}(t)$ & $\tilde{X}(n\Delta t)$	continuous- & discrete-time reconstructed signal	**MTM**	(Sec. 3.4)
$\{\mathbf{X}(t)\}$	multivariate time series in $t \in \mathbb{R}$ or $t \in \mathbb{Z}^+$	**POP**	(Sec. 4.1)
$\{\tilde{\mathbf{X}}_l\}$	multi-channel augmented vector of $\{\tilde{\mathbf{X}}(t)\}$	**SSA**	(Sec. 4.2)
$\hat{Y}_k(f)$	discrete Fourier transform of $\{X(t)w_k(t)\}$	**MTM**	(Sec. 3.4)
α	time-scale ratio	**SSA**	(Sec. 2.4)
$\eta(t)$	additive noise	**MTM**	(Sec. 3.4)
θ	explained contribution to variance	**MTM**	(Sec. 3.4)
λ_k & $\mathbf{\Lambda}_X$	k-th eigenvalue & eigenvalue matrix of $\mathbf{C}_X$	**SSA**	(Sec. 2.2)
$\mathbf{\Lambda}_R$	projection of $\mathbf{C}_R$ onto $\mathbf{E}_X$	**SSA**	(Sec. 2.3)
μ_k	weight of k-th taper	**MTM**	(Sec. 3.4)
ν	number of degrees of freedom of spectrum	**MTM**	(Sec. 3.4)
$\xi(t)$	random forcing		(Sec. 1)

ρ_k & $\mathbf{E}_X$	k-th eigenvector & eigenvector matrix of $\mathbf{C}_X$	**SSA**	(Sec. 2.2)
σ^2	variance of random forcing		(Sec. 1)
τ	characteristic delay time	**SSA**	(Sec. 2.3)
$\phi_X(k)$ & $\hat{\phi}_X(k)$	true & estimated lag k auto-correlation function of $X(t)$	**BT**	(Sec. 3.2)
$\psi(x)$	mother wavelet of variable x	**WLTs**	(Sec. 2.4)
ψ	unexplained contribution to variance	**MTM**	(Sec. 3.4)

Table 2. Analogy between SSA and wavelet analysis. See text for notation.

Method	SSA	Wavelet transform
Analyzing function	EOF ρ_k	Mother wavelet ψ
Basic facts	ρ_k eigenvectors of $\mathbf{C}_X$	ψ chosen *a priori*
Decomposition	$\displaystyle\sum_{t'=1}^{M} X(t+t')\rho_k(t')$	$\displaystyle\int X(t)\psi\left(\frac{t-b}{a}\right)dt$
Scale	$W = \alpha M \Delta t$	a
Epoch	t	b
Average & trend	ρ_1	$\psi^{(0)}$
Derivative	ρ_2	$\psi^{(1)}$

Table 3. Comparison table of spectral methods. Footnotes are given in the table in parentheses; key references are grouped under abbreviations in square brackets.

Method	Para-metric	Lines	Cont. spectrum	Modulation	Error bars	Remarks
BT Sec. 3.2	No	Low resolution	Variance reduction	Partially (1)	Confidence intervals (2)	Simple
MEM/AR Sec. 3.3	Yes	High AR order (3)	Low AR order	Yes (4)	Confidence intervals	
POP & ENM Sec. 4.1	Yes	In space & time (5)	No	No	Robustness τ-test	Inverse-model dynamics: linear & stable
MTM Sec. 3.4	No	Harmonic analysis	Variance reduction	Amplitude & phase; frequency via ESA (6)	F-test (7); confidence intervals (8)	Cost of DPSS computation
SSA Secs. 2 and 4.2	No	Low resolution	Yes (9)	Amplitude & phase; some frequency shifts	Heuristic error bars; Monte Carlo confidence intervals (8)	Detects nonlinear oscillations
Wavelets Sec. 2.4	No	Variable resolution (10)	??	Amplitude & phase & frequency	??	Nonstationary signals, local in time & frequency

1. Only one degree of freedom in envelope, via complex demodulation.
2. Confidence intervals are obtained with respect to AR(0) and AR(1) null hypotheses.
3. High AR orders can produce very good resolution, but spurious peaks too.
4. Underlying AR model *can* describe modulated oscillation.
5. POP is used to detect spatio-temporal patterns on an AR(1) background.
6. Evolutive spectral analysis.
7. F-test for harmonic analysis has a high resolution but produces spurious peaks when the spectrum is continuous.
8. Confidence intervals are obtained from the null hypothesis of AR(0) or AR(1) noise.
9. Approximated by the near-constant slopes in the spectral "scree diagram."
10. The choice of the wavelet basis is delicate and involves a trade-off between time resolution, scale resolution

Table 4. Comparison of 6-month-lead skill of ENSO forecast models (after *Ghil and Jiang* [1998]); the first three models are dynamical, while the last three are statistical.

	Dynamical			Statistical		
Authors	Zebiak & Cane (1987)[*]	Barnett *et al.* (1993)[*]	Ji *et al.* (1994)[*]	Barnston & Ropelewski (1992)[*]	Van den Dool (1994)[*]	Jiang *et al.* (1995b)
Model	Physical: simple coupled	Hybrid: physical ocean model coupled to a statistical atmosphere	Physical: coupled ocean-atmosphere GCM	Statistical: canonical correlation analysis	Empirical: constructed analogs	Statistical: SSA and MEM
Predicted SST region **5°N-5°S**	Niño-3 90°-150°W	Central Pacific 140°-180°W	Niño-3.4 120°-170°W	Niño-3.4 120°-170°W	Niño-3.4 120°-170°W	Niño-3 90°-150°W
Period of record	1970-93	1966-93	1984-93	1956-93	1956-93	1984-93
Skill **(1982-93)**	Corr 0.62 RMSE 0.95 SD 1.08	Corr 0.65 RMSE 0.97 SD 1.10	Corr 0.69 RMSE 0.83 SD 1.00	Corr 0.66 RMSE 0.89 SD 1.11	Corr 0.66 RMSE 0.89 SD 1.11	Corr 0.74 RMSE 0.50 SD 1.00

[*]The first five columns are based on information from *Barnston et al.* [1994]; see further description of models and validation methodology in their paper and in *Ghil and Jiang* [1998]. SSA-MEM forecasts (last column) have continued to be equally skillful in real-time forecasting since their early prediction of the 1997–98 El Niño [*Saunders et al.*, 1997].

Table 5: Analogy between Karhunen–Loève decomposition in space (traditional or spatial EOFs) and time (temporal EOFs).

Spatial EOFs	SSA
$\phi(x,t) = \sum A_k(t)\rho_k(x)$	$X(t+s) = \sum A_k(t)\rho_k(s)$
x - space	s - lag
$C_\phi(x,y) = \mathcal{E}\left\{\phi(x,\omega)\phi(y,\omega)\right\}$ [a] $\cong \frac{1}{T}\int_0^T \phi(x,t)\phi(y,t)dt$	$C_X(s) = \mathcal{E}\left\{X(\cdot,\omega)X(s,\omega)\right\}$ [a] $\cong \frac{1}{T}\int_0^T X(t)X(t+s)dt$
$\mathbf{C}_\phi\rho_k(x) = \lambda_k\rho_k(x)$ [b]	$\mathbf{C}_X\rho_k(s) = \lambda_k\rho_k(s)$ [b]

[a] The expectation operator $\mathcal{E}$ represents averaging over a theoretical ensemble of realizations denoted by ω. This is approximated by an average over epochs t. See text for the rest of the notation.

[b] The covariance operators C_ϕ and C_X are approximated in practice by the corresponding matrices $\mathbf{C}_\phi$ and $\mathbf{C}_X$.

Figure captions

Figure 1 : The schematic diagram of a ghost limit cycle. (a) A perspective sketch of the limit cycle (heavy solid) in a three-dimensional Euclidean space. Light-solid lines indicate two distinct trajectories that approach the limit cycle in a direction in which it is stable and leave its neighborhood in an unstable direction (b) Sketch of the projection of four representative trajectories (light solid) onto a Poincaré section (i.e., a plane intersecting transversally the limit cycle) in a neighborhood of the limit cycle. The figure's mutually perpendicular directions (heavy solid) of stability (inward pointing arrows) and instability (outward pointing arrows) can form in general an arbitrary angle with each other. The shape of such a limit cycle need not be elliptic, as shown in the figure, except near the Hopf bifurcation point where it arises. (Reproduced from *Ghil and Yiou* [1996] with the permission of Springer–Verlag.)

Figure 2 : Variations of the Southern Oscillation Index (SOI) between January 1942 and June 1999. Time on the abscissa in calendar years and SOI on the ordinate normalized by its standard deviation.

Figure 3 : Singular spectrum of the SOI time series; the eigenvalues are plotted as usual in decreasing order. The embedding dimension for Figs. 3 to 7 is $M = 60$, which represents a time window of 5 years. The error bars are based on the default option of the SSA-MTM Toolkit (http://www.atmos.ucla.edu/tcd/ssa/). This option uses the *ad hoc* estimate of variance σ_k of an eigenvalue λ_k given by

$$\sigma_k = (2/\hat{N})\lambda_k.$$

The key element of this estimate is the number $\hat{N}$ of independent *d-o-f* in the time series. *Vautard and Ghil* [1989], who first proposed this estimate in SSA, used $\hat{N} = N$, which is very optimistic, while *Ghil and Mo* [1991a] used $\hat{N} = N/M$, which is too conservative. The Toolkit's current Version 4.0, following *Unal and Ghil* [1995], uses $\hat{N} = N/\kappa\tau$, where τ is the time series' estimated decorrelation time [see Eqs. (2.11) and (2.12)] and κ is an empirical

constant, $\kappa = 1.5$.

Figure 4 : First five EOFs of the SOI time series. The leading four EOFs are grouped into two pairs, $(1, 2)$ and $(3, 4)$, in panels (a) and (b).

Figure 5 : First five PCs of the SOI time series. Note phase quadrature in panels (a) and (b). The maximal cross-correlation in panel (a) is reached when PC-2 leads PC-1 by 10 months and equals 0.89. In panel (b) the maximum cross-correlation equals 0.71 and is obtained when PC-4 leads PC-3 by 9 months.

Figure 6 : Partial reconstructions of the SOI time series. (a) Reconstruction based on EOFs 1–4 (heavy solid). The raw SOI series is shown as the light solid curve. El Niños (warm events) are shown as upward-pointing arrows on the lower abscissa, while La Niñas (cold events) are shown as downward-pointing arrows on the upper abscissa; the arrows are aligned with the January closest to the event maximum. (b) The trend of reconstructed component (RC) number 5 (RC-5); see text for details.

Figure 7 : Monte Carlo singular spectrum of the SOI time series. The diamonds indicate projections of the noise eigenvectors onto the data correlation matrix, cf. Eq. (2.13); the lower and upper ticks on the error bars indicate the 5th and 95th noise percentiles. For each EOF, a characteristic frequency was estimated by maximizing its correlation with a sinusoid; therefore, the frequency at which each diamond is plotted is just an average estimate. The frequencies in this figure and subsequent ones are given in cycles per month.

Figure 8 : Instantaneous SOI frequencies, based on the two leading EOFs of multi-scale SSA. These frequencies were computed independently for the local EOFs 1 (solid) and 2 (dashed). The results agree well with each other and demonstrate a sharp jump in frequency between 1961 and 1962 (Reproduced from *Yiou et al.* [2000] with the permission of Elsevier Science B.V.).

Figure 9 : Classical spectral estimates of the SOI time series, from a raw correlogram (light solid) and lag-window estimate (heavy solid), with a Bartlett window of width $m = 40$, are

shown.

Figure 10 : Blackman-Tukey correlogram estimate of the SSA-filtered SOI. The embedding dimension for SSA is $M = 60$ and reconstructed components (RCs) 1–4 and 10–11 were chosen as most significant. A Bartlett window of width $m = 30$ was used for the spectral estimate of the prefiltered SOI. The error bars are based on a chi-square test and reach from the 2.5% to the 97.5% quantile.

Figure 11 : MEM spectral estimates of the SOI time series. The autocorrelation orders are $M' = 10, 20, 40$ (light solid, heavy solid, and dashed curves). The Akaike information-content (AIC) criterion [*Haykin and Kesler*, 1983] predicts an order of $M' \approx 10$, which is obviously too low.

Figure 12 : MEM analysis of the SOI time series after SSA prefiltering. The auto-regression order used is $M' = 20$. The SSA window width is $M = 60$, and RCs 1–4 and 10–11 were chosen, as in Fig. 10.

Figure 13 : The first $K = 3$ eigentapers for the case $p = 2$, computed for the length $N = 690$ of the SOI time series in Fig. 2. Note that the 3rd eigentaper or discrete prolate spheroidal sequence (DPSS) is the least leakage-resistant since its values near the boundaries $t = 0$ and $t = N\Delta t$ of the sampling interval are no longer negligible, like those of the first two DPSSs. The eigentapers shown have been normalized to unit power.

Figure 14 : Adaptively weighted MTM spectrum of the SOI time series. The estimated red-noise background and associated 90%, 95%, and 99% significance levels are shown by four curves (light solid), in this order, from the lowest to the highest curve in the figure. Three signals with interannual and lower frequencies — the trend and ENSO's low-frequency and quasi-biennial signals — are significant at the 99% level. Several higher-frequency peaks appear also to be significant; please see text for a more detailed assessment of their significance. The bandwidth parameter is $p = 2$ and $K = 3$ tapers were used.

Figure 15 : Comparison between MTM spectra obtained by harmonic-analysis test and adap-

tively weighted estimation against red-noise background. (a) Harmonic-signal (variance-ratio) F-test with median, 90%, 95%, and 99% significance levels indicated; (b) reshaped vs. unreshaped adaptively weighted MTM spectrum based on $p = 2$ and $K = 3$, and a 90% F-test significance criterion for reshaping. The dashed line in panel (b) represents the discrepancy between reshaped spectrum (solid) and the unreshaped one; it thus provides a measure of the portion of the MTM spectrum associated with harmonic features. Significance levels for the reshaped spectrum are as in Figure 14.

Figure 16 : Multi-taper reconstruction of the significant interannual components corresponding to (a) the trend, (b) low-frequency band, and (c) quasi-biennial band of ENSO. A minimum-misfit boundary condition was employed in the reconstructions, as described in the text. The parameters $p = 2$ and $K = 3$ were used as above.

Figure 17 : Sum of the three MTM reconstructed components (heavy solid), along with the raw SOI series (light solid).

Figure 18 : SST anomalies averaged over the Niño-3 region. Light solid line: GISST data; heavy curves are RCs 2-3 for $M = 60$ (solid) and $N' = 60$ (short dashes), as well as RCs 3-4 for $M = N' = 270$ (long dashes).

Figure 19 : Phase-composites for the low-frequency ENSO cycle, as given by RCs 2-3 with $N' = 60$. The cycle is divided into eight segments of $45°$ each, using the first spatial PC as an index. Contour interval is $0.1°$C.

Figure 20 : Monte Carlo significance test for oscillations in the GISST data. The first 10 spatial PCs are projected onto AR(1) basis vectors with $M = 270$. Error bars show the 95% confidence interval constructed from an AR(1) surrogate-data ensemble of 1000. In comparing with Fig. 7, please note that the distribution of the points along the abscissa here is uniform, since the attribution of the dominant frequency for AR(1) basis vectors is simple [*Vautard and Ghil*, 1989; *Allen*, 1992].

Figure 21 : Comparison between PCA, SSA, and EEOFs–M-SSA. The axes are spatial coordi-

nate (or spatial PC label) x, time t, and lag s; discrete values of these variables are labeled by l, n, and j, respectively (courtesy of K.-C. Mo).