SSA-MTM Toolkit for Spectral Analysis

Distributed by

SSA-MTM Group, Department of Atmospheric Sciences, University of California, Los Angeles

---

Disclaimer

Reference to the Toolkit

TABLE OF CONTENTS

Disclaimer

The developers of the SSA-MTM Toolkit are researchers attempting to make some useful time-series analysis methods more accessible to the scientific community. Although we use the tools ourselves and have made every effort to ensure their accuracy, we can not make any guarantees. We provide the Toolkit to you for free, but it is copyrighted. In return, you--the user--assume full responsibility for use of the software. The SSA-MTM Toolkit and this Guide come without

Sciences; U.S. Geological Survey; Commissariat à l'Énergie Atomique; and the various individuals involved in development and maintenance of the SSA-MTM Toolkit are not responsible for any damage that may result from correct or incorrect use of this software.

Reference to the Toolkit

If you feel that your research has benefitted from the use of the SSA-MTM Toolkit, you can repay us by citing our articles:

Ghil M., R. M. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou, (2002) (PDF File). "Advanced spectral methods for climatic time series," *Rev. Geophys.*,40(1), pp. 3.1-3.41, 10.1029/2000RG000092.

- **Dettinger, M.D., Ghil, M., Strong, C.M., Weibel, W., and Yiou, P., 1995: Software expedites singular-spectrum analysis of noisy time series, Eos, Trans. American Geophysical Union, v. 76(2), p. 12, 14, 21.**

We also appreciate your references to the original articles that motivated and made the Toolkit possible (see the reference list), and especially to

- **Allen, M.R., and Smith, L.A., 1996: Monte Carlo SSA: detecting oscillations in the presence of coloured noise. *J. Climate*, 9, 3373.**
- **Allen M. R., Robertson A. W. (1996) Distinguishing modulated oscillations from coloured noise in multivariate datasets. *Clim. Dyn. 12*, 775-784**
- **Mann, M.E. and Lees, J.M., 1996: Robust estimation of background noise and signal detection in climatic time series, *Clim. Change*, 33, 409-445.**
- **Vautard, R., Yiou, P., and Ghil, M., 1992: Singular-spectrum analysis: A toolkit for short, noisy chaotic signals, *Physica D*, 58, 95-126.**

URL: http://www.atmos.ucla.edu/tcd/ssa/guide/
Last modified: 4/11/07

SSA-MTM Toolkit for Spectral Analysis

---

1. INTRODUCTION

The Singular Spectrum Analysis - Multitaper Method (SSA-MTM) Toolkit consists of a set of programs that perform detailed spectral analyses and decompositions on input time series (univariate or multivariate). The toolkit contains procedures for:

(a) estimating the spectrum of a time series,

(b) decomposing the time series into trends, oscillatory components, and noise, and

(c) reconstructing the contributions of selected components of the time series.

For univariate time series four methods of spectral-analysis (task a) are provided: Blackman-Tukey correlogram estimation (BT), the Maximum-Entropy Method (MEM), the Multi-Taper Method (MTM), and Singular-spectrum Analysis (SSA). SSA and MTM can then be used for separation of trends, near-periodic and other significant components, and noise (tasks b and c). Both SSA and MTM incorporate sophisticated significance tests against a variety of noise null-hypotheses, including the methods of Allen and Smith (1996) for SSA, and Mann and Lees (1996) for MTM. For multivarate time series Multi-Channel SSA method is provided.
The basic philosophy of the Toolkit is that only the simultaneous and flexible application of more than one spectral estimation method can provide truely reliable information on a given time series, when the signal-to-noise ratio is low.
This document briefly describes some of the theory behind the various methods, and demonstrates how to use the Toolkit. More details about each method can be found in the following references. Those indicated with an asterix are the central works on which the Toolikit is based:

SSA: Vautard et al. (1992)*, Allen and Smith (1996)* PDF file and test dataset, Vautard and Ghil (1989).

MTM: Mann and Lees (1996)* PDF file, Thomson (1982; 1990a-b), Percival and Walden (1993), and Yiou et al. (1991).

Blackman-Tukey correlogram estimation: Kay (1988).

MEM: Childers (1978), Press et al. (1989), and Penland et al. (1991).

Press et al. (1989, chapters 11 and 12) also provide excellent overviews of spectral methods and eigensystems analysis, while Manly (1986) provides a useful introduction to principal-components analysis.

Section 2 gives a brief introduction to the theory behind each tool. Section 3 demonstrates the Toolkit functions. Sections 4 outlines the specifications of the Toolkit and acknowledges the many software contributions that made it possible.

The Toolkit was developed at, and is distributed by, the University of California, Los Angeles, with contributions from the U.S. Geological Survey in San Diego, California, and the Commissariat à l'Énergie Atomique in Gif-sur-Yvette, France.

Next Section | Table of Contents | References

This file last modified: 9/26/03

2. THEORY

a. Blackman-Tukey Correlogram Analysis

The correlogram constructs an estimate of the power spectrum using a windowed fast Fourier transforms (FFT) of the autocorrelation function of the time series. It was developed by Blackman and Tukey (1958) and is based on the Wiener-Khinchin theorem, which states that if the Fourier transform of a series g(t) is G(w), and if the autocorrelation function of the series is R, then the Fourier transform of R is $|G(w)|^2$ or the power spectrum of g (e.g., Press et al., 1989). The resulting power-spectrum estimate is called a correlogram. An alternative that is not included in the Toolkit is direct or windowed FFT of the time series itself, called a periodogram.

Both periodograms and correlograms are usually performed on weighted versions of the time series or autocorrelation functions in order to reduce power leakage (artificially high power estimates at frequencies away from the true peak frequencies). Press et al. (1989, pp. 423-424) note that "when we select a run of N sampled points for periodogram spectral estimation, we are in effect multiplying an infinite run of ... data ... by a window function in time, one which is zero except during the total sampling time [NDt], and is unity during that time." The sharp edges of this window function contain much power at highest frequencies, which is imparted to the windowed signal and leads to power leakage. A similar argument can be made for correlograms. Weighting the data or correlation function by various tapered shapes (high in center and falling off to sides) is an accepted traditional approach to reducing power leakage. In the Blackman-Tukey approach, the power spectrum P(w) is estimated by

$$P(\omega) = \frac{1}{M}\left|\sum_{j=0}^{M-1} \rho_j w_j e^{i\omega j}\right|^2 \qquad (1)$$

where r_j is the autocorrelation function, M is the maximum lag considered and window length, and w_j is the windowing function. Several window shapes are available in the Toolkit: Bartlett (triangular), Hamming (cosinusoidal), Hanning (slightly different cosinusoidal), and none.

You may find that the various windows of the same widths give similar results. The more important choice is how wide the windows should be. The averaging associated with windowing a series reduces the resolution of the methods, from the frequency intervals of 1/N, to a windowed frequency intervals of about 1/M (e.g., Kay 1988, p. 81). Thus, wider windows yield higher spectral resolution, and vice versa.

However, there is a trade-off between higher resolution and increasing variance of the spectral estimate. At the extreme, a single (M=N) direct application of FFT to an unwindowed time series results in a periodogram with a theoretical standard deviation of the estimates equal to the estimates at each frequency, regardless of the number of observations in the time series (Press et al. 1989, p. 423). Averaging the results from many short data windows throughout the series (or autocorrelation) effectively increases the number of independent samples used in estimation and thereby reduces the estimation variance. Kay (1988, section 4.5) shows that the variance of a power spectrum obtained by a windowed correlogram is 2M/3N of the estimated power at each frequency. Thus a narrower window should be used to smooth the spectrum and reduce the sampling errors on the estimate. In practice, Kay (1988) recommends that windows should be no more than one-fifth to one-tenth the total number of data points (to obtain desired estimate-variance reductions) and not too much smaller (in order to retain the ability to distinguish between powers at neighboring frequencies and to obtain the desired leakage reductions).

Theoretical estimates of variance for Blackman-Tukey power spectra are available (e.g., Kay, 1988) and the Toolkit provides error bars constructed from them. These can either be plotted about the estimates themselves, or as a red-noise uncertainty interval. In the latter case, an AR(1) process is fitted to the data, and the the error bars are centered on the theoretical AR(1) spectrum.

As a "traditional" method, the correlogram is intended to provide a familiar benchmark against which the other more modern methods provided in the Toolkit can be judged.

Previous Section | Next Section | Table of Contents | References

---

This file last modified: January 31, 2000

2. THEORY

b. Maximum-Entropy Spectral Estimates

The toolkit also provides spectral estimation by MEM. Given a stationary time series X, and its first M auto-correlation coefficients ϕ_X , the purpose of MEM is to obtain the spectral density P_X by determining the most random (i.e. with the fewest assumptions) process, with the same auto-correlation coefficients as X. In terms of information theory, this is the notion of *maximal entropy*, hence the name of the method.

The entropy h of a Gaussian process is given by

$$h = \int_{-1/2}^{1/2} \log P_X(\omega) d\omega. \qquad (1)$$

From the Wiener-Khintchin identity, the maximal entropy process and the series X will have the same spectral density. Some algebra (Percival and Walden, 1993) shows that under the constraints of ϕ_X , h is maximized by an autoregressive process Y of size M-1:

$$Y_{n+M} = a_M Y_{n+M-1} + \cdots + a_1 Y_n + b_n, \qquad (2)$$

where b_n is a Gaussian white process with variance a_o. And hence P_X is

$$P_X(\omega) = \frac{a_0}{\left|1 + \sum_{k=1}^{M-1} a_k e^{ik\omega}\right|^2} \qquad (3)$$

In summary, the method boils down to looking for an auto-regressive process that ``mimics'' the original time series. This is why it is a called *parametric method*.

The MEM is very efficient for detecting frequency lines in stationary time series. However, if

this time sereis is not-stationary, misleading results can occur, with little chance of being detected otherwise than by cross-checking with other techniques.

The art of using MEM resides in the appropriate choice of M, i.e. the order of regression of Y. The behavior of the spectral estimate depends on the choice of M: it is clear that the number of poles (or even maxima) of Eq. (3) depends on the order of regression M and the auto-regression coefficients a_k , so that, for a given time series, the number of peaks will increase with M! Therefore, a trade-off between a good resolution (high M) and few spurious peaks (low M) has to be found. A few guides are provided by the default values of the toolkit (i.e. M should not exceed half the length of the time series).

The weaknesses can be remedied partly by (a) determining which peaks survive reductions in M, (b) comparing MEM spectra to those produced by correlogram and MTM which generally should not share spurious peaks with MEM, and (c) using SSA to pre-filter the series and thus to decompose the original series into several components, each of which contains only a few harmonics (so that small M values can be chosen; see Penland et al., 1991). The ease with which these various analyses can be interwoven in the Toolkit was a major motivation for its development.

Previous Section | Next Section | Table of Contents | References

---

This file last modified: February 1, 2000

2. THEORY

c. MultiTaper Method (MTM)

The Multi-Taper method (MTM) of spectral analysis provides a novel means for spectral estimation (Thomson, 1982; Percival and Walden, 1993) and signal reconstruction (e.g., Park, 1992) of a time series which is believed to exhibit a spectrum containing both continuous and singular components. This method has been widely applied to problems in geophysical signal analysis, including analyses of atmospheric and oceanic data (Ghil and Vautard, 1991; Kuo et al, 1990; Mann and Park, 1993;1994;1996b; Lall and Mann, 1995; Mann et al, 1995a; Thomson 1995; Mann and Park, 1996a) paleoclimate data (Berger et al, 1991; Chappellaz et al, 1990; Mann et al, 1995b; Mann and Lees, 1996; Mommersteeg et al, 1995; Park and Maasch, 1993; Thomson, 1990ab; Yiou et al, 1991;1994;1995;1995) geochemical tracer data (Koch and Mann, 1995), and seismological data (Park et al, 1987; Lees, 1995). Time-frequency ``evolutive'' analyses based on moving window adaptations of MTM have also been applied (e.g., Yiou et al, 1991; Birchfield and Ghil, 1993; Mann et al, 1995b; Mann and Park, 1996b).

MTM, like the method of Blackman and Tukey (1958), offers the appeal of being nonparametric, in that it does not prescribe an a priori (e.g., autoregressive) model for the process generating the time series under analysis. MTM attempts to reduce 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 Blackman-Tukey methods. A set of independent estimates of the power spectrum is computed, by pre-multiplying the data by orthogonal tapers which are constructed to minimize the spectral leakage due to the finite length of the data set. The optimal tapers or ``eigentapers'' belong to a family of functions known as discrete prolate spheroidal sequences (DPSS) and defined as the eigenvectors of a suitable Rayleigh-Ritz minimization problem, were extensively studied by Slepian (1978). Averaging over this (small) ensemble of spectra yields a better and more stable estimate -- i.e., one with lower variance -- than do single-taper methods (Thomson, 1990b). The tapers are the discrete set of eigenfunctions which solve the variational problem of minimizing leakage outside of a frequency band of half bandwidth pf_R where $f_R = 1/N\delta t$ is the Rayleigh frequency, and p is an integer. Note that the case p=1 reduces to the trivial Blakman-Tukey case of a single tapered

Discrete Fourier Transform (DFT). 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 Box and Jenkins, 1970; Jenkins and Watts, 1968).

Detailed algorithms for the calculation of the eigentapers are readily available (e.g., Thomson, 1990b; 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 of tapers used K should be less than $2p$-1 in any application of MTM. The choice of the bandwidth $2pf_N$ and number of tapers K thus represents the classical tradeoff between spectral resolution and the stability or ``variance'' properties of the spectral estimate (Thomson, 1982). The trivial case K=1 and p=1 is, again, the single-tapered DFT of Blackman and Tukey. For typical length instrumental climate records, the choice K=3, p=2 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 (see e.g., Mann and Park, 1993). Longer datasets can admit the use of a greater number K of tapers while maintaining a desired frequency resolution, and the optimal choice of K and p is in general most decidedly application specific.

c.1 Spectral Estimation

MTM can provide estimates of both the singular components (ie, the ``lines'') and the continuous component (ie, the background) of the spectrum. Once the tapers $w_k(t)$ are computed for a chosen frequency bandwidth, the total power spectrum P_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 where Y_k is the DFT of Xw_k. The *high-resolution* multitaper spectrum is a simple weighted sum of the K eigenspectra,

$$S(f) = \frac{\sum_{k=1}^{K} \lambda_k |Y_k(f)|^2}{\sum_{k=1}^{K} \lambda_k} \qquad (1)$$

Its resolution in frequency is $\pm pf_N$, which means that ``line'' components will be detected as peaks or bumps of width $2pf_N$. For a white noise or ``locally white noise'' process (see Thomson, 1982) the high-resolution spectrum is chi-squared distributed with $2K$ degrees of freedom.

By adjusting the relative weights on the contributions from each of the K eigenspectra, a more leakage-resistant spectral estimate termed the *adaptively weighted multitaper spectrum* is obtained,

$$S(f) = \frac{\sum_{k=1}^{K} b_k^2(f) \lambda_k |Y_k(f)|^2}{\sum_{k=1}^{K} b_k^2(f) \lambda_k} \qquad (2)$$

where $b_k(f)$ is a weighting function that further guards against broadband leakage for a non-white (``colored'') but locally-white process. The adaptive spectrum estimate has an effective degrees of freedom ν that generally departs only slightly from the nominal value $2K$ of the high-resolution multitaper spectrum.

The purpose of *harmonic analysis* is to determine the line components-ie the spikes in the spectrum corresponding to a periodic or quasi-periodic signal-in terms of their frequency, amplitude, and phase. The Fourier transform of a clean periodic signal of infinite length yields a *Dirac function* at the frequency of the signal, viz., a line (or peak of zero width) with infinite magnitude. A spectral estimate based on the methods discussed in earlier sections, gives indirect information on the amplitude of the signal at a given frequency, through the area under the peak centered at that frequency and whose width is, roughly speaking, inversely proportional to the length N of the time series; this area is nearly constant, since the height of the peak is also proportional to N. Harmonic analysis attempts, instead, 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. Other approaches are described by Macdonald (1989).

Assume the time series $X(t)$ is the sum of a sinusoid of frequency f and amplitude μ, plus a ``noise'' $\xi(t)$ which is the sum of other sinusoids and white noise. One can then write

$$X(t) = \mu e^{i\hat{f}t} + \xi(t). \qquad (3)$$

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

$$\hat{\mu}(\hat{f}) = \frac{\sum_{k=1}^{K} U_k^*(0)\, y_k(\hat{f})}{\sum_{k=1}^{K} |U_k(0)|^2}, \qquad (4)$$

where the asterisk denotes complex conjugation. A statistical confidence interval can be given for the least-squares fit 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,

$$\theta = |\hat{\mu}(\hat{f})|^2 \sum_{k=1}^{K} |U_k(0)|^2 \qquad (5)$$

and

$$\psi = \sum_{k=1}^{K} |y_k(\hat{f}) - \hat{\mu}(\hat{f})U_k(0)|^2, \qquad (6)$$

that are respectively the ``explained'' and ``unexplained'' contributions to the variance.

The random variable

$$F(\hat{f}) = (K-1)\frac{\theta}{\psi} \qquad (7)$$

would obey a Fisher-Snedecor law with 2 and 2K-2 degrees of freedom if the time series $X(t)$ were a pure white-noise realization. One can interpret its numerical value for given data by assuming that $\mu = 0$ -- 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 (see 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 $F(\hat{f})$ does not depend -- to first order -- on the magnitude of $\hat{\mu}(\hat{f})$. This feature is an important advantage of MTM over standard methods where error bars are essentially proportional to the amplitude of a peak (e.g., Jenkins and Wattes, 1968). However, the implicit assumption in the harmonic-analysis appproach is that the time series is produced by a process that consists of a superposition of separate, purely periodic fixed-amplitude components. 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 ``signals'' are represented by lines in the spectrum corresponding to phase-coherent harmonic signals, while the ``noise'' is represented by the continuous component of the spectrum. In geophysical applications, signals are frequently associated with narrowband, but not strictly harmonic variability (c.f. the discussion by Park, 1992) and truly harmonic signals are rarely detected. In such cases, the simple noise null hypothesis and signal assumptions implicit in the traditional MTM approach described above loses much of its utility.

The more subtle nature of geophysical signals and noise has led to a more recent generalization of the conventional MTM approach which combines the harmonic signal detection procedure described above, with a criterion for detecting significant narrowband, ``quasi-oscillatory'' signals which may exhibit phase and amplitude modulation, and intermittent oscillatory behavior, but are nonetheless significant relative to some suitably defined null hypothesis (Mann and Lees, 1996). We favour this approach, whichs provides for the detection and significance estimation of both harmonic and anharmonic, narrowband signals in the MTM spectra of time series, making use of a ``robust'' estimate of the background noise. Information from the harmonic peak test of the traditional MTM procedure is retained, but peaks, whether

indicated as harmonic or anharmonic, are tested for significance relative to the null hypothesis of a globally red (or, trivially, white) noise background estimated empirically from the data. This is particularly important in climate studies, where the intrinsic inertia of the system leads to greater power (and greater liklihood of prominent peaks in the spectrum) at lower frequencies, even in the absence of any signals (e.g., Hasselmann, 1976). To accomodate the red noise background assumption requisite in geophysical applications, an AR(1) noise process is assumed, although more complex noise models can be motiviated (see the discussion in Mann and Lees, 1996).

The power spectrum of the AR(1) process is given by (e.g., Bartlett, 1966),

$$S(f) = S_0 \frac{1-\rho^2}{1-2\rho\cos\pi\frac{f}{f_N}+\rho^2} \qquad (8)$$

for frequency f where S_0 is the average value of the power spectrum, related to the white-noise variance by,

$$S_0 = \sigma^2/\left(1-\rho^2\right) \qquad (9)$$

$f_N = 2/\Delta t$, the Nyquist frequency, is the highest frequency that can be resolved for sampling rate Δt . The characteristic noise decay time scale can be estimated

$$\tau = -\frac{\Delta t}{\log\rho} \qquad (10)$$

For periodicities much larger than τ , the spectrum behavies like a white spectrum.

The approach of Mann and Lees (1996) performs a ``robust'' estimate of the red noise background by minimizing (as a function of ρ) the misfit between an analytical AR(1) red noise spectrum and the adaptively weighted multitaper spectrum convoluted with a median smoother. The median smoothing operation in the fit insures that the estimated noise background is insensitive to outliers (most obviously, peaks associated with signals) that, properly, should not influence the estimation of the global noise background. This guards against the typical problem of inflated estimates of noise variance and noise autocorrelation that arise in a conventional Box-Jenkins approach due to the contamination of noise parameters by non-noise contributions (e.g., a significant trend or oscillatory component of the series). Mann and Lees (1996) motivate the choice of a median smoothing width of $\Delta f = min(f_N/4, 2pf)$ as a compromise between describing the full variation of the background spectrum over the Nyquist interval, and insensitivity to narrow spectral features.

Significance levels for harmonic or narrowband spectral features relative to the estimated noise background can be determined from the appropriate quantiles of the chi-squared distribution, approximating the spectrum as being distributed with 2K degrees of freedom (see Mann and Lees, 1996). A reshaped spectrum is determined in which the contributions from harmonic

signals are removed (see Thomson, 1982), depending on a significance threshold for the F variance-ratio test described above. In this way, noise background, harmonic, and anharmonic narrowband signals are each independently isolated. The harmonic peak detection procedure provides information as to whether the signals are best approximated as harmonic (phase-coherent sinusoidal oscillations) or narrowband (amplitude and phase modulated, perhaps intermittant oscillations). In either case, they must be significant relative to a specified (e.g., red) noise hypothesis to be isolated as significant.

c.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 multitaper decomposition. The signals or ``reconstructed components'' are analagous to those of SSA described in previous sections, except that information from a frequency domain eigendecomposition, rather than a correlation-domain eigendecomposition, is used to reconstruct the detected signal. As in the correlation-domain case, a priori assumptions regarding the domain boundaries must be invoked.

The reconstructed signal corresponding to a peak centered at frequency $f0$ is written

$$\tilde{x}(t) = \Re\{A(t)e^{-i2\pi f_0 t}\} \qquad (11)$$

or, for the discrete case at hand,

$$\tilde{x_n} = \Re\{A_n e^{-i2\pi f_n \delta t}\} \qquad (12)$$

where the envelope function $A(t)$ is determined from a time-domain inversion of the spectral domain information contained in the K complex eigenspectra (see Park 1992; Park and Maasch, 1993; Mann and Park, 1994; 1996b). 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 $f0$. The discrete time sequence describing the complex envelope A_n is determined from a discrete inverse problem using the complex amplitudes of each of the the K eigenspectra and appropriate boundary conditions (see Park, 1992; Park and Maasch, 1993). The three lowest order boundary contraints in this inversion involve minimization of the envelope A_n itself near the boundaries, numerical mimimization of the slope of the envelope near the boundaries, or optimization of the smoothness of the envelope (ie, numerical minimization of the 2nd derivative near the boundaries). Any one of these choices might be favoured if a priori information regarding the characteristics of the signal is available (see the discussion of Park, 1992). The amplitude of the seasonal cycle in surface temperature, for example, is nearly constant in time, and a minimum slope constraint is most appropriate for its reconstruction (see Mann and Park, 1996a). Generally, however, a nearly optimal reconstruction can always be obtained through seeking the weighted linear combination of these 3 contraints that minimizes the mean-square misfit of the reconstructed signal with respect to the raw data series (Mann and

Park, 1996b). We favour this method of time-domain inversion.

Previous Section | Next Section | Table of Contents | References

---

This file last modified: Jan. 31, 2000

2. THEORY

d. Singular-Spectrum Analysis

SSA is a nonparametric method. It tries to overcome the problems of finite sample length and noisiness of sampled time series not by fitting an assumed model to the available series, but by using a data-adaptive basis set, instead of the fixed sine and cosine of the BT method.

Colebrook (1978) applied a form of SSA to biological oceanography and noted the duality between *principal component* analysis (PCA) in the space and time domain. Broomhead and King (1986: BK hereafter) applied the ``method of delays" of dynamical systems theory to estimate the dimension of and reconstruct the Lorenz attractor using singular-value decomposition (SVD) on the trajectory matrix formed by lagged copies of a single series obtained from the system.

Vautard and Ghil (1989: VG hereafter) realized the formal similarity between classical lagged-covariance analysis and the method of delays. They exploited this similarity further by pointing out that pairs of SSA eigenmodes corresponding to nearly equal eigenvalues and associated (temporal) principal components that are nearly in phase quadrature can represent efficiently a nonlinear, anharmonic oscillation. This is due to the fact that a single pair of data-adaptive eigenmodes can capture the basic periodicity of a boxcar or seesaw-shaped oscillation, say, rather than necessitate the many overtones that will appear in methods with fixed basis functions.

There are three basic steps in SSA: i) embedding the sampled time series in a vector space of dimension M; ii) computing the $M \times M$ lag-covariance matrix $\mathbf{C_D}$ of the data; and iii) diagonalizing $\mathbf{C_D}$.

Step (i): The time series $\{x(t) : t = 1, \ldots, N\}$ is embedded into a vector space of dimension M by considering M lagged copies $\{x(t-j) : j = 1, \ldots, M\}$ thereof. The choice of the dimension M is not obvious. Vautard *et al.* (1992) suggest that SSA is typically successful at analyzing periods in the range (M/5, M).

Step (ii): One defines the $M \times M$ lag-covariance matrix estimator $\mathbf{C_D}$. There are two distinct methods used widely to define $\mathbf{C_D}$, due to BK and VG, respectively. In the BK algorithm, a window of length *M* is moved along the time series, producing a sequence of $N' \equiv N - M + 1$ vectors in the embedding space. This sequence is used to obtain the $N' \times M$ trajectory matrix $\mathbf{D}$, where the *i*-th row is the *i*-th view of the time series through the window. In this approach, $\mathbf{C_D}$ is defined by:

$$\mathbf{C_D} = N^{-1}\mathbf{D}\mathbf{D}^{\mathbf{T}}. \qquad (1)$$

In the VG algorithm, $\mathbf{C_D}$ is estimated directly from the data as a Toeplitz matrix with constant diagonals. Both the BG and VG algorithms are implemented in the SSA-MTM Toolkit.

Step (iii): The covariance matrix calculated from the *N* sample points, using the BK or VG algorithm, is then diagonalized and the eigenvalues $\{\lambda_k : 1 \le k \le M\}$ are ranked in decreasing order. The eigenvalue λ_k gives the variance of the time series in the direction specified by the corresponding eigenvector $\mathbf{E}_k$; the square roots of the eigenvalues are called *singular values* and the set $\{\lambda_k^{1/2} : 1 \le k \le M\}$ the *singular spectrum*. These terms give SSA its name; BK showed how to obtain the set $\{(\lambda_k^{1/2}, \mathbf{E}_k) : 1 \le k \le M\}$ by SVD applied to the trajectory matrix $\mathbf{D}$. VG called the $\mathbf{E}_k$'s *empirical orthogonal functions* (EOFs), by analogy with the meteorological term used when applying PCA in the spatial domain.

If we arrange and plot the singular values in decreasing order, one can often distinguish an initial steep slope, representing the signal, and a (more or less) ``flat floor,'' representing the noise level (Vautard and Ghil, 1989) To correctly separate signal from noise, this visual observation of the spectral plot must be confirmed by suitable criteria for statistical significance as discussed later.

Once these three steps have been completed, a number of other interesting results can be obtained. For each EOF $\mathbf{E}_k$ with components $\{E_{k,j} : j = 1, \ldots, M\}$, we can construct the time series of length *N*' given by:

$$A_k(t) = \sum_{k=1}^{M} x(t+j)E_{k,j}, \quad t = 0, \ldots, N - M, \qquad (2)$$

called the *k*-th principal component (PC). It represents the projection of the original time series onto the *k*-th EOF. The sum of the power spectra of the PCs is identical to the power spectrum of the time series *x*(*t*) (Vautard et al., 1992); therefore we can study separately the spectral

contribution of the various components. The PCs, however, have length N', not N, and do not contain phase information.

In order to extract time series of length N, corresponding to a chosen subset of $\mathcal{K}$ eigenelements, the associated PCs are combined to form the partial reconstruction $R_{\mathcal{K}}(t)$ of the original time series $x(t)$. Over the central part of the time series,

$$R_{\mathcal{K}}(t) = \frac{1}{M_t} \sum_{k \in \mathcal{K}} \sum_{j=1}^{M} A_k(t-j) E_{k,j}, \quad M \leq t \leq N', \qquad (3)$$

and $M_t = M$; near the endpoints ($1 \leq t \leq M$ or $N'+1 \leq t \leq N$), the summation in j extends from 1 to t or from (t-N+M) to M and $M_t = min\{t, N-t+1\}$ (Vautard et al., 1992). These series of length N are called *reconstructed components* (RCs). They have the important property of preserving the phase of the time series; therefore $x(t)$ and $R_{\mathcal{K}}(t)$ can be superimposed. When $\mathcal{K}$ contains the components of a single eigenvector $\mathbf{E}_k$, $R_{\mathcal{K}}(t)$ is called the k-th RC. If we denote it by $x_k(t)$, one can show that

$$x(t) = \sum_{k=1}^{M} x_k(t); \qquad (4)$$

that is, the time series can be reconstructed completely by summing its RCs, without losing any information. Allen *et al.* (1992a) noted that, when a linear trend is present, reconstruction based on the BK algorithm has certain advantages near the endpoints of the sample.

The correct partial reconstruction of the signal, *i.e.*, the reconstruction over the correct set $\mathcal{K}^* = \{1, 2, \ldots, S\}$ of EOFs, yields the optimal signal-to-noise (S/N) enhancement with respect to white noise. MEM spectral estimation will work very well on this clean signal (Penland et al., 1991, Keppenne and Ghil, 1992) which, unlike the raw time series, is band limited. As a result, a low-order AR-process fit can yield good spectral resolution, without spurious peaks.

In order to perform a good reconstruction of the signal or of its oscillatory part(s) several methods - either heuristic or based on Monte Carlo ideas - have been proposed for S/N separation or for reliable identification of oscillatory pairs and trends (Vautard and Ghil, 1989; Vautard et al., 1992; Ghil and Mo, 1991a, b).

Monte Carlo SSA

The particular Monte Carlo approach to S/N separation introduced by Allen (1992) has become

known as Monte Carlo SSA (MC-SSA), and detailed description can be found in Allen and Smith (1996; AS hereafter). See also the review papers of Ghil and Yiou (1996), and Ghil and Taricco (1997).

Monte Carlo SSA can be used to establish whether a given timeseries is linearly distinguishable from *any* well-defined process, including the output of a deterministic chaotic system, but we will focus on testing against the linear stochastic processes which are normally considered as ``noise". ``Red noise" is often used to refer to any linear stochastic process in which power declines monotonically with increasing frequency, but we prefer to use the term to refer specifically to a first-order auto-regressive, or AR(1), process u_t whose value at a time t depends on the value at time t-1 only,

$$u_t = \gamma(u_{t-1} - u_0) + \alpha z_t + u_0; \qquad (5)$$

where z_t is a gaussian-distributed white-noise process, for which each value is independent of all previous values, u_0 is the process mean and α and γ are constant coefficients.

When testing against a red noise null-hypothesis, the first step in MC-SSA is to determine the red-noise coefficients α and γ from the time series $X(t)$ using a maximum-likelihood criterion. Estimators are provided by Allen (1992) and AS which are asymptotically unbiased in the limit of large N and close to unbiased for shorter series provided the series length is at least an order of magnitude longer than the timescale of decay of autocorrelation, $-1/\log(\gamma)$. Based on these coefficients, an ensemble of surrogate red-noise data can be generated and, for each realization, a covariance matrix C_R is computed. These covariance matrices are then *projected* onto the eigenvector basis E_D of the original data,

$$\Lambda_R = E_D^t C_R E_D. \qquad (6)$$

Since (6) is not the SVD of that realization, the matrix Λ_R is not necessarily diagonal, but it measures the resemblance of a given surrogate set with the data set. This resemblance can be quantified by computing the statistics of the diagonal elements of Λ_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 generic red-noise simulation. 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 confidence. Otherwise, that particular SSA component of the time series cannot be considered as significantly different from red noise.

Allen (1992) stresses that care must be taken at the parameter-estimation stage because, for the results of a test against AR(1) noise to be interesting, we must ensure we are testing against that specific AR(1) process which maximises the likelihood that we will fail to reject the null-

hypothesis. Only if we can reject *this* (the ``worst case'') red noise process, can we be confident of rejecting all other red noise processes at the same or higher confidence level. A second important point is test multiplicity: comparing M data eigenvalues with M confidence intervals computed from the surrogate ensemble, we *expect* $M/10$ to lie above the 90th percentiles even if the null hypothesis is true. Thus a small number of excursions above a relatively low percentile should be interpreted with caution. AS discuss this problem in detail, and propose a number of possible solutions.

The MC-SSA algorithm described above can be adapted to eliminate *known* periodic components and test the residual against noise. This adaptation provides sharper insight into the dynamics captured by the data, since 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 and Smith (1992) and AS describe this refinement of MC-SSA which consists in restricting the projections to the EOFs that do *not* account for known periodic behavior. This refinement is implemented in the SSA-MTM toolkit, where it is referred to as the "Hybrid Basis" (see the Toolkit Demonstration section below).

Previous Section | Next Section | Table of Contents | References

---

This file last modified: Jan. 31, 2000

2. THEORY

e. Multi-channel Singular-Spectrum Analysis

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 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, using both observed data *Jiang et. al-1995a,Unal and Ghil-1995*, and coupled GCM model simulations [*Robertson et al.*, 1995*a, b*].

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 , one usually chooses $L \leq M$ 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 $\{X_{l,n} : l = 1, \ldots, L; n = 1, \ldots, N\}$ 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 $\mathbf{X}(t_n)$, with $t_n = n\Delta t$, is centered and stationary in the weak sense.

The generalization of SSA to a multivariate time series requires the constructiuon of a ``grand'' block-matrix $\mathbf{T_X}$ for the covariances that has the form

$$\mathbf{T_X} = \begin{pmatrix} \mathbf{T}_{1,1} & \mathbf{T}_{1,2} & . & . & . & \mathbf{T}_{1,L} \\ \mathbf{T}_{2,1} & \mathbf{T}_{2,2} & . & & & . \\ . & . & . & . & & . \\ . & . & . & . & \mathbf{T}_{l,l'} & . \\ . & . & . & . & & . \\ . & & & . & . & \mathbf{T}_{L-1,L} \\ \mathbf{T}_{L,1} & . & . & . & \mathbf{T}_{L,L-1} & \mathbf{T}_{L,L} \end{pmatrix}. \quad (1)$$

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*, *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

$$(T_{l,l'})_{j,j'} = \frac{1}{\tilde{N}} \sum_{n=\max(1,1+j-j')}^{\min(N,N+j-j')} X_{l,n} X_{l',n+j-j'}, \quad (2)$$

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

$$\tilde{N} = \min(N, N+j-j') - \max(1, 1+j-j') + 1 \quad (3)$$

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 upper-right 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 lower-left triangle contain the covariances with l leading l'. Equation (2) ensures that $\mathbf{T}_{l',l} = (\mathbf{T}_{l,l'})^{\mathrm{t}}$ so that $\mathbf{T}_{\mathbf{X}}$ 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 $\mathbf{T}_{\mathbf{X}}$ 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*, 1986*a, 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 $\{X_{l,n} : n = 1, \ldots, N\}, 1 \leq l \leq L$, of $\mathbf{X}$ with M lagged copies of itself,

$$\tilde{\mathbf{X}}_l = \begin{pmatrix} X_{l,1} & . & . & . & X_{l,N'} \\ X_{l,2} & . & . & . & X_{l,N'+1} \\ . & . & . & . & . \\ . & . & . & . & . \\ X_{l,M} & . & . & . & X_{l,N} \end{pmatrix}, \quad 1 \leq l \leq L, \tag{4}$$

and then forming the full augmented *trajectory matrix*:

$$\tilde{\mathbf{X}} = \begin{pmatrix} \tilde{\mathbf{X}}_1 \\ \tilde{\mathbf{X}}_2 \\ \vdots \\ \tilde{\mathbf{X}}_L \end{pmatrix}. \tag{5}$$

Thereafter, one computes the grand covariance matrix $\mathbf{C}_{\mathbf{X}}$ as

$$\mathbf{C_X} = \frac{1}{N'}\tilde{\mathbf{X}}\tilde{\mathbf{X}}^{\mathbf{t}} = \begin{pmatrix} \mathbf{C}_{1,1} & \mathbf{C}_{1,2} & . & . & . & \mathbf{C}_{1,L} \\ . & \mathbf{C}_{2,2} & . & . & . & . \\ . & & . & . & . & . \\ . & & . & . & \mathbf{C}_{l,l'} & . \\ . & & . & . & . & . \\ \mathbf{C}_{L,1} & \mathbf{C}_{L,2} & . & . & . & \mathbf{C}_{L,L} \end{pmatrix}. \tag{6}$$

The blocks of $\mathbf{C_X}$ are given by

$$\mathbf{C}_{l,l'} = \frac{1}{N'}\tilde{\mathbf{X}}_l\tilde{\mathbf{X}}_{l'}^{\mathbf{t}} \tag{7}$$

with entries

$$(C_{l,l'})_{j,j'} = C_{IJ} = \frac{1}{N'}\sum_{n=1}^{N'} X_{l,n+j-1}X_{l',n+j'-1}, \tag{8}$$

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

Each block $\mathbf{C}_{l',l}$ contains, like $\mathbf{T}_{l,l'}$ in Eq. (1), 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. (1) and (2) vs. Eqs. (6)-(8). 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 windows 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 $\mathbf{C_X}$ or $\mathbf{T_X}$ 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 $\mathbf{C_X}$ equal those obtained from $\mathbf{T_X}$ 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 $E^k_{l,m}$. The associated space-time PCs $\mathbf{A}^k$ are single-channel time series that are computed by projecting $\tilde{\mathbf{X}}$ onto the EOFs:

$$A^k_n = \sum_{m=1}^{M} \sum_{l=1}^{L} X_{l,n+m} E^k_{l,m}, \tag{9}$$

where *n* varies from 1 to *N*'.

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

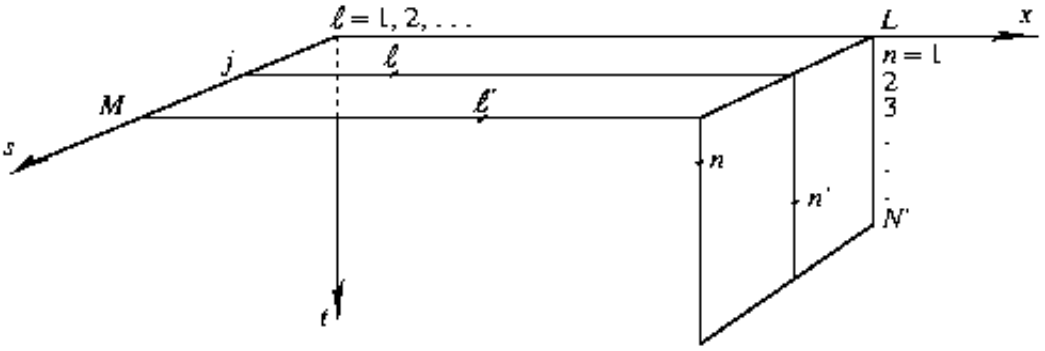
$$R^k_{l,n} = \frac{1}{M_n} \sum_{m=1}^{M} A^k_{n-m} E^k_{l,m}. \tag{10}$$

The normalization factor M_n 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.

Both SSA (*L*=1) and conventional PCA (*M*=1) analysis are special cases of EEOF or M-SSA analysis. Following *Allen and Robertson* [1996], both algorithms can be understood in practice

in terms of a ``window."

These windows are illustrated schematically in the Fig.1



The axes are spatial coordinate (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).

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 Linput channels and identifying high-variance temporal patterns, i.e. the PCs, in the corresponding L views. In Fig.1, 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 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.1. 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. (6)-(8).

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 1

$$C_{IJ}^{(R)} = \frac{1}{L}\sum_{l=1}^{L}\left[\sum_{j=1}^{M} X_{l,I+j-1}X_{l,J+j'-1}\right]. \tag{11}$$

1

This matrix, unlike the $LM \times LM$ matrix of Eq. (6), is given by 1

$$\mathbf{C}^{(R)} = \frac{1}{N'}\tilde{\mathbf{X}}^{\mathrm{t}}\tilde{\mathbf{X}}. \tag{12}$$

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

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. (11) allows the univariate SSA Monte Carlo test of *Allen-1992* 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. *N*' can always be chosen sufficiently small, so that the complementary window *N*' used in Eq. (11) 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. 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. (11).

Previous Section | Next Section | Table of Contents | References

---