EOFs of One-Dimensional Cyclostationary Time Series: Computations, Examples, and Stochastic Modeling

KWANG-Y. KIM, GERALD R. NORTH, AND JIANPING HUANG

Climate System Research Program, Texas A&M University, College Station, Texas

(Manuscript received 11 July 1995, in final form 12 October 1995)

ABSTRACT

Many climatic time series seem to be a mixture of unpredictable fluctuations and changes that occur at a known frequency, as in the case of the annual cycle. Such a time series is called a cyclostationary process. The lagged covariance statistics of a cyclostationary process are periodic in time with the frequency of the nested undulations, and the eigenfunctions are no longer Fourier functions. In this study, examination is made of the properties of cyclostationary empirical orthogonal functions (CSEOFs) and a computational algorithm is developed based on Bloch's theorem for the one-dimensional case. Simple examples are discussed to test the algorithm and clarify the nature and interpretation of CSEOFs. Finally, a stochastic model has been constructed, which reasonably reproduces the cyclostationary statistics of a 100-yr series of the globally averaged, observed surface air temperature field. The simulated CSEOFs and the associated eigenvalues compare fairly with those of the observational data.

1. Introduction

Important issues in the study of climate are the estimation, detection, and prediction of forced signals against the background noise (e.g., Hasselmann and Barnett 1981; Epstein 1982; Barnett 1986, 1991; Hasselmann 1988; von Storch and Zwiers 1988; Zwiers and von Storch 1989; Santer et al. 1991, 1993; Shen et al. 1994; North and Kim 1995; North et al. 1995). We call them collectively the linear estimation problems. Mathematical and statistical prescriptions of the linear estimation problems are rather diverse. One of the most popular quantities derived from the temporal and spatial structure of the covariance kernel of background noise is the frequency-dependent empirical orthogonal functions (EOFs) (e.g., Wallace and Dickinson 1972; Kim and North 1993). A similar concept is the principal oscillation pattern (Hasselmann 1988; von Storch et al. 1988, 1995). These EOFs serve as a natural basis set for designing optimal filters (also called fingerprints). Fingerprint approaches have been adopted by many authors including Barnett (1986, 1991), Barnett and Schlesinger (1987), Barnett et al. (1991), and Santer et al. (1991, 1993, 1994). By projecting a signal onto this frequency-dependent orthogonal basis set, we can calculate the individual statistical mode contribution to the squared signal-to-noise index. The formalism is then used to design a suitable filter for a particular signal (e.g., Hasselmann 1993). We have applied this concept with success in the studies of estimation and sampling error (Shen et al. 1994) and of detection and prediction (North et al. 1995; North and Kim 1995).

Fourier functions are a natural basis set for time series variables that are statistically stationary in time. The term "stationarity" means that various statistics of a variable are independent of time. Technically, the lagged covariance statistics are functions only of lag [see Eq. (1)]. A spectral analysis with the stationary assumption allows the investigator to apportion the variance into distinct frequency bands and examine the contributions from different uncorrelated contributions. There are no cross-correlation terms to contend with, which allows the total variance to be decomposed into a sum of variances. This approach facilitates the construction of optimal filters, which can be used in prediction and detection algorithms. Unfortunately, the assumption of stationarity in time does not hold for many important climate variables for which we would like to formulate optimal estimation algorithms.

Instead of being stationary, however, most climate and weather variables are cyclostationary. The term "cyclostationarity" is used here in the sense that many climatic processes and hence their variability are characterized by multiple, distinct temporal scales. The earth's surface temperature, for example, exhibits not only interannual variability but higher frequency fluctuations that strongly depend on the time of the day and of the year. Not only is the variance a function of season, but so are the spatial autocorrelation lengths (Mad-

Corresponding author address: Dr. Kwang-Y. Kim, Climate System Research Program, College of Geosciences and Maritime Studies, Texas A&M University, College Station, TX 77843-3150.
E-mail: kykim@csrp.tamu.edu

den 1977; Jones and Briffa 1992). There are distinct internal cycles, such as annual and diurnal cycles, that characterize many climatological processes and their associated variability. Various statistics of a cyclostationary process, therefore, depend upon a particular phase of the internal cycle. The lagged covariance statistics of a cyclostationary process, then, are functions of both lag and time. In a strict and technical sense in this study, the lagged covariance statistics of a cyclostationary time series are further restricted to be periodic in time [see Eqs. (1)–(4)]. Each realization, of course, is not periodic in time.

The multiple temporal scales of the meteorological variables mean that the Fourier basis is not a proper basis for a statistically useful decomposition (different frequency components will be correlated, making a straightforward analysis of variance impossible). The nested fluctuations often contain significant information for parameter estimation, signal detection, and linear prediction. To take full advantage of this, we must abandon the stationarity assumption and account for the dependency of the statistics on the phase of the underlying cycle [see Eq. (4)]. Early investigators recognized the importance of this phase dependency of detectability and predictability (e.g., Gardner and Franks 1975; Parzen and Pegano 1979; Hasselmann and Barnett 1981).

The remedy is to look to the eigenfunctions of the lagged covariance function for a particular process involved [see Eq. (7)]. These eigenfunctions are called cyclostationary empirical orthogonal functions (CSEOFs). Unfortunately, the computation of CSEOFs is usually very difficult because of the size of the covariance function involved. We intend to exploit the fact that the eigenfunctions of a cyclostationary time series may be factored into a periodic part whose period is the cycle length and a complex sinusoid of all frequencies (Bloch's Theorem; Smith 1969). It, therefore, takes two indices to characterize a CSEOF [see Eqs. (5)–(7)]. In some limiting cases, the two indices may be interpreted as the interannual frequencies and the harmonics of the seasonal cycle. This factorization tremendously stretches the feasibility of generalizing CSEOFs to spatiotemporal EOFs for cyclostationary time series.

In this study we examine and elaborate on a few examples of one-dimensional cyclostationary processes, including one example that degenerates to a stationary process, to better understand the algorithm and the nature of CSEOFs. We illustrate as much as possible the detailed aspects of the exercises. In the next section we discuss the method of computing CSEOFs based on Bloch's Theorem. In the section of examples we examine several simple cases of cyclostationary processes in which an analytic form of the covariance matrix and its CSEOFs are possible and compare these CSEOFs with numerical solutions to verify the algorithm. Such solutions may also help us understand and develop intuition about CSEOFs. Finally, we will generate a synthetic time series that mimics the statistics of the earth's global average surface temperature field, which is followed by a summary and conclusions.

2. Methods

The covariance kernel of the surface temperature, $T(t)$, is defined as

$$\mathcal{K}(t, t') = K(t, \tau) = \langle T(t)T(t')\rangle, \tag{1}$$

or in a slightly different form,

$$\tilde{K}(t, \tau) = \mathcal{K}(t - \tau/2, t + \tau/2), \tag{2}$$

where $\tau = t' - t$. The covariance kernel of the stationary time series is characterized as a function only of τ, that is, it is independent of t. The covariance kernel of the cyclostationary time series is such that

$$\mathcal{K}(t + d, t' + d) = \mathcal{K}(t, t'), \tag{3}$$

or equivalently,

$$\tilde{K}(t + d, \tau) = \tilde{K}(t, \tau), \tag{4}$$

where d is the period of the cyclic process. In addition to this periodicity property, $K(t, \tau)$ is assumed to be square integrable on the Euclidean space and vanish as τ approaches infinity. Most pioneering work and a significant amount of research on cyclostationarity has been conducted and published in the fields of electrical engineering and econometrics. Gardner (1994) has detailed accounts of this work.

Figure 1 illustrates the structure of the cyclostationary covariance kernel. The covariance kernel is symmetric with respect to t and t'. (The same information emanates in the direction of $\pm\tau$ in each segment delimited by black circles.) This is a strong constraint on the possible modal interactions in the frequency domain. As a result of this constraint the covariance matrix in the frequency domain, namely a covariance matrix in terms of Fourier basis functions, becomes block diagonal and each block matrix represents an outer mode number n and is solved for separately as will be shown below. We elaborate this finding in the form of Bloch's theorem.

We define the CSEOFs

$$\psi_{nm}(t) = e^{2\pi int/T}U_{nm}(t), \tag{5}$$

$$U_{nm}(t) = \sum_l u_{nml}e^{2\pi ilt/d} \tag{6}$$

as functions satisfying

$$\lim_{T\to\infty}\int_{-T}^{T}\mathcal{K}(t, t')\psi_{nm}(t')dt' = \lambda_{nm}\psi_{nm}(t), \tag{7}$$

where T is the record length and d is the period of nested fluctuations. The limit sign here indicates that the exactness of the derivation of the CSEOFs assumes

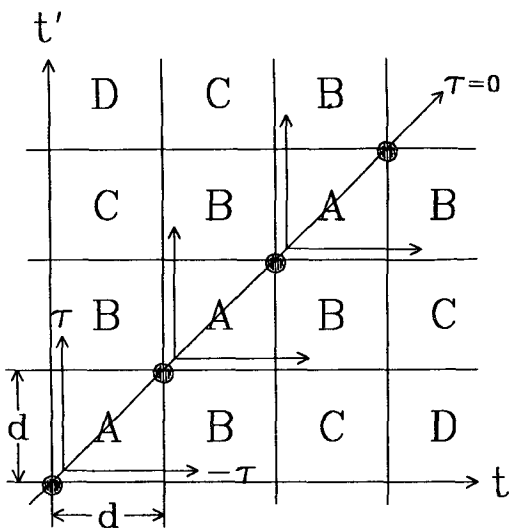

FIG. 1. Schematic diagram showing the structure of a cyclostationary covariance kernel. The covariance kernel is symmetric with respect to t and t'. The same information repeats with periodicity d along the line $\tau = 0$. The same information also emanates along the direction of $\pm\tau$ in each segment delimited by black circles according to (3) and (4). This allows one to carry out the computation in the direction of τ for only one of the segments.

an infinitely long record. In practice, of course, it should suffice to consider a "fairly long" record as in the analysis of stationary time series. The limit sign should be understood accordingly.

This particular form of CSEOFs is suggested by Bloch's theorem, which is well known in the band theory of solids (Smith 1969). [The theorem states that the eigenfunctions of the wave equation for a periodic potential are of the form of the product of a plane wave $\exp(i\mathbf{k}\cdot\mathbf{r})$ times a function $U_{\mathbf{k}}(\mathbf{r})$ with the periodicity of the crystal lattice. The subscript $\mathbf{k}$ indicates that the function $U_k(r)$ depends on the wave vector $\mathbf{k}$. Equation (5) is a similar statement in the time domain.] The function $U_{nm}(t)$ is called the Bloch function and is cyclic with period d. Note that there are two indices n and m attached to CSEOFs. We call them "outer" and "inner" mode numbers, respectively. The n pertains to the frequency of outer modulation and the m identifies a mode of nested fluctuations as implied in (5) and (6). The limits of summation with respect to l in (6) are given in (16) and (17). The upper limit determines the truncation of the representation of the Bloch's function and can be chosen smaller than that specified in (16).

Since the covariance kernel $K(t, \tau)$ is cyclic in t with periodicity d, it can be written as

$$K(t, \tau) = \sum_l \hat{K}_l(\tau) e^{2\pi i l t/d} e^{\pi i l \tau/d}, \tag{8}$$

or equivalently,

$$\tilde{K}(t, \tau) = \sum_l \hat{K}_l(\tau) e^{2\pi i l t/d}. \tag{9}$$

Since $\tilde{K}(t, \tau)$ is real, $\hat{K}_{-l}(\tau)$ is the complex conjugate of $\hat{K}_l(\tau)$. Note also that the factor $\exp(\pi i l \tau/d)$ is introduced to make $\hat{K}_l(\tau)$ symmetric in τ as will be apparent in the following. Now, the left-hand side of (7) is rewritten as

$$\lim_{T\to\infty} \int_{-T}^{T} \sum_{l'} \hat{K}_{l'}(\tau) e^{2\pi i l' t/d} e^{\pi i l' \tau/d} \cdot e^{2\pi i n(t+\tau)/T} \times \sum_{l''} u_{nml''} e^{2\pi i l''(t+\tau)/d} d\tau = \sum_{l'} \sum_{l''} \lim_{T\to\infty} \int_{-T}^{T} \hat{K}_{l'}(\tau) \times e^{2\pi i (n/T + l''/d + l'/2d)\tau} d\tau \, u_{nml''} e^{2\pi i (n/T + l''/d + l'/2d)t}, \tag{10}$$

whereas the right-hand side of (7) is

$$\lambda_{nm} \sum_l u_{nml} e^{2\pi i l t/d} e^{2\pi i n t/T}. \tag{11}$$

Since (10) and (11) should be equal for all t, it follows that

$$\sum_{l'} \sum_{l''} \lim_{T\to\infty} \int_{-T}^{T} \hat{K}_{l'}(\tau) e^{2\pi i (n/T + l''/d + l'/2d)\tau} d\tau \, u_{nml''} \delta_{l'+l'',l} = \lambda_{nm} u_{nml}, \tag{12}$$

or in a compact form,

$$\sum_{l'} C^{(n)}_{ll'} u_{nml'} = \lambda_{nm} u_{nml}, \quad \text{for any } l \text{ and } m, \tag{13}$$

where

$$C^{(n)}_{ll'} = S_{l-l'}(n/T + (l + l')/2d), \tag{14}$$

and

$$S_l(f) = \lim_{T\to\infty} \int_{-T}^{T} \hat{K}_l(\tau) e^{2\pi i f \tau} d\tau \tag{15}$$

is the cyclic spectrum.

Equation (13) constitutes an eigenvalue problem with the eigenvector $\{u_{nml}\}$ and the eigenvalue λ_{nm} for the complex matrix $\mathbf{C}^{(n)}$. Although $\mathbf{C}^{(n)}$ consists of the spectral density functions (Fourier transformation of the covariance function), we will simply call it a covariance matrix. As was mentioned earlier, the covariance matrix is block diagonal and (13) represents one of the block matrices. Note that the covariance matrix (14) is Hermitian and is symmetric with respect to the two indices l and l'. This symmetry is achieved by introducing a factor $\exp(\pi i l \tau/d)$ in defining the Fourier expansion of the covariance matrix in (7). Since the covariance matrix is Hermitian, this complex eigenvalue problem can be converted into a real eigenvalue problem. Then the problem is solved using various orthogonalization transformation techniques (Press et al. 1988).

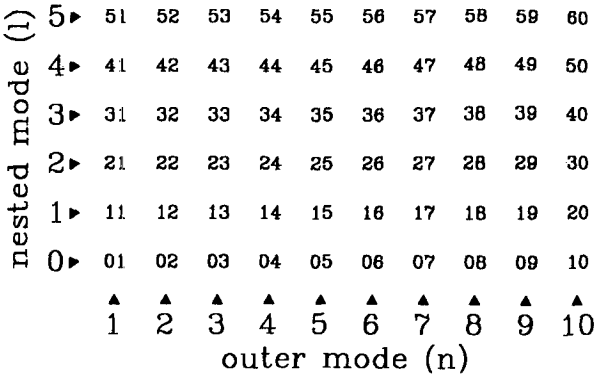

FIG. 2. Illustration of the global mode numbers in terms of the outer mode number (n) and the mode number of the nested fluctuations (l) according to (17). Each number in small print represents a mode number of a discrete Fourier analysis. Fourier functions along two different columns (different n's) are uncorrelated but are generally correlated along the same column. Therefore, the covariance matrix is block diagonal and is partitioned into 10 smaller covariance matrices [Eq. (10)] each of which represents an outer mode number n.

3. Properties of CSEOFs

a. Uniqueness of eigenfunctions

Equation (13) depicts an eigenvalue problem for a cyclic spectrum matrix, each element of which is a function of frequency. For each covariance matrix, which is a subset of the whole covariance matrix according to the factorization theorem presented by Smith (1969), one can compute the eigenfunctions that are orthogonal to each other. However, this orthogonality within a subset does not guarantee the uniqueness of CSEOFs across different outer mode numbers. Therefore, we need to constrain the outer and inner mode numbers to ensure the uniqueness of CSEOFs.

As indicated in (14), the covariance matrix is dependent upon the frequencies of the outer and nested fluctuations. To ensure the uniqueness of CSEOFs, we define n and l such that the frequency band of the covariance matrix does not overlap for different n values. Such n and l are found to be

$$n = 1, \cdots, T/d, \quad l = 0, \cdots, d/2 - 1, \quad (16)$$

where d is the period of the nested fluctuations and T is the total number of sampling points, which is a multiple of d. Note that the range of the discrete frequencies of the eigenfunctions defined by (5) and (6) for the above choice of n and l is

$$n/T + l/d = [n + l*(T/d)]/T = k/T,$$
$$k = 1, \cdots, T/2. \quad (17)$$

An example for $T = 120$ (say, 120 months) and $d = 12$ (12 months) is illustrated in Fig. 2.

b. Orthogonality

From (5) and (13) it follows that Bloch functions are orthogonal to each other:

$$\lim_{T\to\infty} \int_{-T}^{T} U_{nm}(t) U^*_{nm'}(t)dt = \sum_l \sum_{l'} u_{nml} u^*_{nm'l'} \times \lim_{T\to\infty} \int_{-T}^{T} e^{2\pi i(l-l')t/d} dt = \sum_l u_{nml} u^*_{nm'l} = \delta_{m,m'}. \quad (18)$$

We then prove that the complete set of CSEOFs also satisfies the orthogonality property:

$$\lim_{T\to\infty} \int_{-T}^{T} \psi_{nm}(t) \psi^*_{n'm'}(t)dt = \sum_l \sum_{l'} u_{nml} u^*_{n'm'l'} \times \lim_{T\to\infty} \int_{-T}^{T} e^{2\pi i(n-n')t/T} e^{2\pi i(l-l')t/d} dt = \sum_l \sum_{l'} u_{nml} u^*_{n'm'l'} \delta_{n,n'} \delta_{l,l'} = \delta_{n,n'} \delta_{m,m'}. \quad (19)$$

Thus, CSEOFs are orthogonal with respect to both the outer and nested mode numbers n and m.

c. Completeness

Complemented by a constant value for $n, l = 0$, (16) defines a set of complete eigenfunctions (see also Fig. 2). Namely,

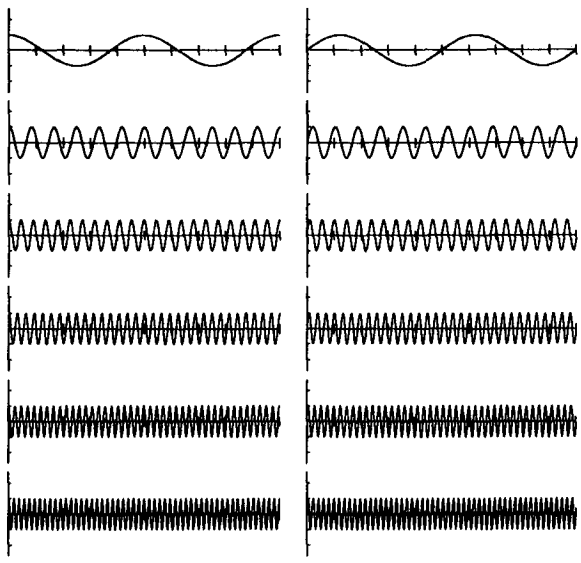

FIG. 3. Real (left) and imaginary (right) parts of CSEOFs of a stationary process with a length of 10 years. The slope of the covariance spectrum (α) is 0.1. The outer mode number is $n = 2$ and the inner mode number is $m = 1, \cdots, 6$ (from top to bottom). With the assumed period of $d = 12$ mo for the nested fluctuations, each CSEOF is identified with the cosine and sine functions of frequency ($n/T + m/d$), where $T = 120$ mo is the total record length.

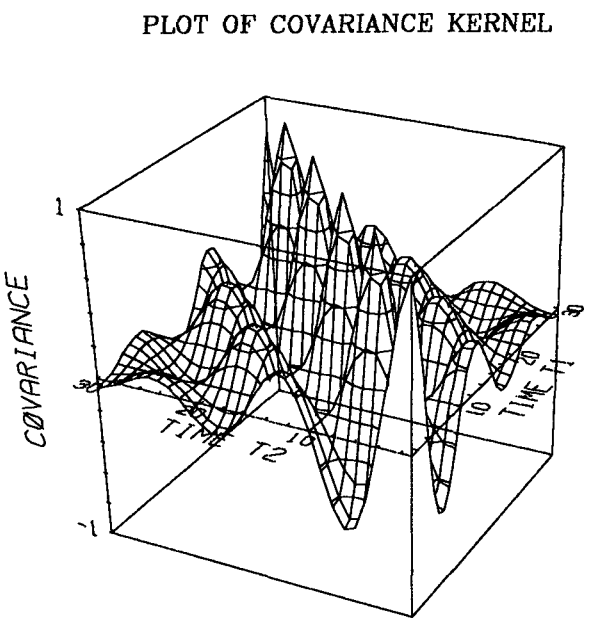

FIG. 4. Plot of (normalized) covariance kernel, $\mathcal{K}(t, t')$, of a periodically modulated stationary process. The period of modulation is 12 mo and the slope of the covariance spectrum (α) is 0.1. T1 and T2 represent two different times at which covariance is computed.

$$\sum_{n,m} \psi_{nm}(t)\psi^*_{nm}(t') = \delta(t - t'), \tag{20}$$

when summed over all n and m. Therefore, it can be shown from (7) that

$$\mathcal{K}(t, t') = \sum_{n,m} \lambda_{nm}\psi_{nm}(t)\psi^*_{nm}(t'). \tag{21}$$

This equation convincingly shows why Bloch's factorization theorem works. Along $t = t'$, $\mathcal{K}(t, t')$ should be a periodic function with period d (see Fig. 1). Thus, the product of a CSEOF with the complex conjugate of itself on the right-hand side of (21) should not produce any harmonics with a period longer than d. This is possible only when the CSEOFs are defined as in (5) and (6). By substituting the definition of CSEOFs in (21) one can interpret the covariance matrix (13) as

$$C^{(n)}_{ll'} = \iint \mathcal{K}(t, t')e^{2\pi i(n/T+l'/d)t'} e^{-2\pi i(n/T+l/d)t} dt'dt. \tag{22}$$

4. Examples

Simple examples below are used to validate the developed computational algorithm and to understand and develop intuition about CSEOFs. For validation we compare solutions from the analytical and numerical approaches. We first assume a cyclostationary time series for which an analytic form of the covariance kernel $\tilde{K}(t, \tau)$ is known. We then proceed to compute $\hat{K}_l(\tau)$ in (8) using a discrete Fourier transformation technique or other equivalent methods. We calculate the analytic form of $\hat{K}_l(\tau)$, too. Then, the cyclic spectrum $S_l(f)$ is given by (15) and the computation of (15) is carried out using a Fourier transformation technique. We also calculate the analytical form of the cyclic spectrum. Note that this particular process is somewhat prone to computational error and the discrete Fourier transformation sometimes yields significant error. Therefore, caution should be exercised in computing (15). Finally, a covariance matrix is set up according to (13) and (14). In the examples below both approaches yield essentially the same CSEOFs. Therefore, only one set of the solutions is shown.

a. Stationary case

We first examine a degenerate case of a stationary time series. Let us consider an autoregressive model of order one (hereafter, AR-1) as our example. Then, the lagged covariance kernel is independent of time and is given by

$$K(t, \tau) \propto \exp(-\alpha|\tau|), \tag{23}$$

where α is the regression coefficient of the model. We know that CSEOFs of a stationary time series of infinite length should collapse into Fourier functions. The eigenvalues of the covariance kernel (4) are given by $\alpha/[\alpha^2 + (2\pi f)^2]$.

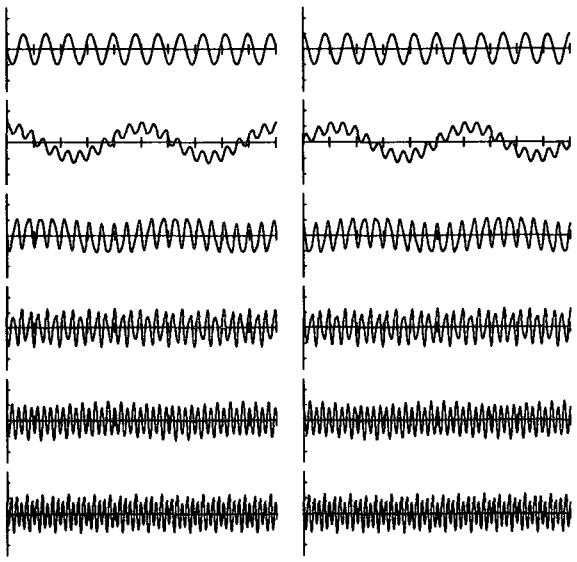

FIG. 5. Real (left) and imaginary (right) parts of CSEOFs of a periodically modulated stationary process with a modulation period of 12 mo. The length of record is 10 years. The slope of the covariance spectrum (α) is 0.1. The outer mode number is $n = 2$ and the inner mode number is $m = 1, \cdots, 6$ (from top to bottom).

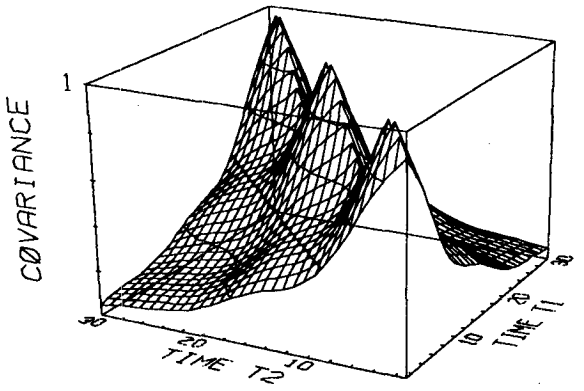

FIG. 6. Plot of (normalized) covariance kernel of an order-1 autoregressive model with a periodic coefficient. The AR coefficient is $0.1 - 0.05\cos(2\pi f_0 t)$, where the frequency f_0 is 1/12 month.

As shown in Fig. 3, the real and imaginary parts of CSEOFs are identified with the cosine and sine functions as expected. The Bloch functions are also pure sinusoids. As shown in (5) and (6) the frequency of each function is $n/T + m/d$, $m = 1, \cdots, 6$, where n and m are the inner and outer mode numbers, T is the total record length, and d is the period of the assumed nested fluctuations. The eigenvalues were also computed correctly. The estimated eigenvalues are very close to the theoretical values above.

Since covariance is independent of t this case is equivalent to assuming $d = 1$. Therefore, each segment in Fig. 1 collapses into a line that represents the lagged covariance of a stationary time series. In line with the earlier argument, the full covariance matrix in the frequency domain is diagonal and each basis function, Fourier function, itself is an eigenfunction. This is also consistent with the computational algorithm. Since the covariance kernel is invariant in t, $\hat{K}_l(\tau)$ in (9) is zero for nonzero l. Therefore, a covariance matrix is diagonal from (14) and (15).

b. *Periodically modulated autoregressive model*

Let us consider a time series

$$T(t) = R(t)\cos(2\pi f_0 t), \tag{24}$$

where $R(t)$ is an AR-1 model such that

$$\langle R(t)R(t+\tau)\rangle = \exp(-\alpha|\tau|), \tag{25}$$

and f_0 is the modulation frequency of the fluctuations. Then, the lagged covariance kernel is given by

$$K(t,\tau) = \exp(-\alpha|\tau|) \times \cos(2\pi f_0 t)\cos[2\pi f_0(t+\tau)]. \tag{26}$$

Figure 4 is a plot of the lagged covariance kernel, $\mathcal{K}(t, t')$. We see the periodic nature of the covariance kernel in accordance with (3).

From (8) and (15) it follows that

$$S_l(f) = \begin{cases} [S_R(f+f_0) + S_R(f-f_0)]/4, & l/d = 0; \\ S_R(f)/4, & l/d = \pm 2f_0, \end{cases} \tag{27}$$

where

$$S_R(f) = \frac{2\alpha}{\alpha^2 + (2\pi f)^2} \tag{28}$$

is the spectrum of an AR-1 model. In view of (14) and (27), all the off-diagonal elements of the covariance matrix are zero except for those elements that are two columns apart from the diagonal. The position of these off-diagonal elements are related to the modulation frequency according to (27).

Figure 5 shows the real and imaginary parts of the CSEOFs based on the lagged covariance kernel (26). These CSEOFs compare accurately with those based on (27). They are not pure sinusoids but are perturbed, which is due to nonzero off-diagonal elements of the covariance kernel. Note also that the position of the first two modes are switched com-

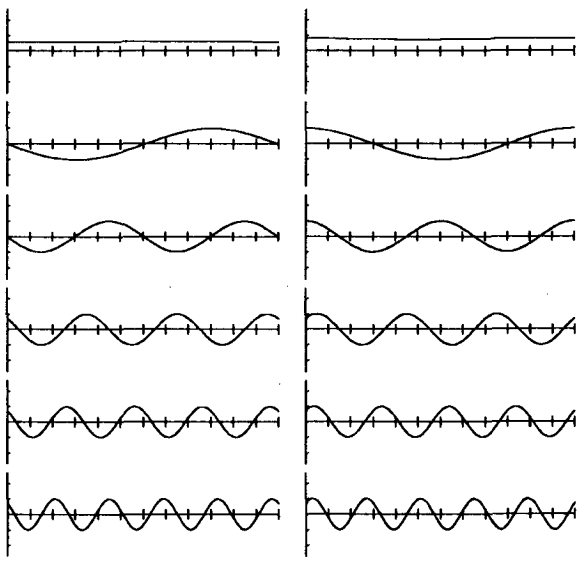

FIG. 7. Real (left) and imaginary (right) parts of Bloch functions of an order-1 autoregressive model with a periodic coefficient. The AR coefficient is $0.1 - 0.05\cos(2\pi f_0 t)$, where the frequency f_0 is 1/12 mo. The length of record is 10 yr. The outer mode number is $n = 2$ and the inner mode number is $m = 1, \cdots, 6$ (from top to bottom).

pared with those in Fig. 3. This is due to strong fluctuation of the covariance at the frequency f_0 as shown in Fig. 4.

c. Autoregressive model with periodic coefficient

Let us consider a model

$$\frac{dT(t)}{dt} + [A + B\cos(f_0 t)]T(t) = w(t), \quad (29)$$

where $w(t)$ is a zero-mean white noise with standard deviation σ_N. The solution to (29) is derived as

$$T(t) = e^{-[At+B'\sin(f_0 t)]} \times \left[T_0 + \int^t w(p) e^{[Ap+B'\sin(f_0 p)]} dp\right], \quad (30)$$

where $B' = B/f_0$ and T_0 is $T(t)$ at a reference time. The process becomes cyclostationary when $t \gg 1/A$ after which the initial condition T_0 is forgotten.

The covariance kernel is given by

$$\langle T(t)T(t')\rangle = e^{-[At+B'\sin(f_0 t)]} e^{-[At'+B'\sin(f_0 t')]} \times \int^t \int^{t'} \langle w(p)w(q)\rangle e^{[Ap+B'\sin(f_0 p)]} \times e^{[Aq+B'\sin(f_0 q)]} dpdq = e^{-[At+B'\sin(f_0 t)]} \times e^{-[At'+B'\sin(f_0 t')]} \int^{t^*} \sigma_N^2 e^{2[Ap+B'\sin(f_0 p)]} dp, \quad (31)$$

where $t^* = \min(t, t')$. To carry out the computation further, we use

$$e^{x\sin\theta} = \sum_{n=-\infty}^{\infty} J_n(2B'i) e^{-inf_0 t} = I_0(x) + 2\sum_{k=0}^{\infty} (-1)^k I_{2k+1}(x)\sin[(2k+1)\theta] + 2\sum_{k=0}^{\infty} I_{2k}(x)\cos(2k\theta), \quad (32)$$

where $J_n(x)$, $I_n(x)$ are the Bessel function of order n and the modified Bessel function of order n, respectively. Then, (31) is rewritten as

$$\mathcal{K}(t,t') = \sigma_N^2 e^{-[At+B'\sin(f_0 t)]} e^{-[At'+B'\sin(f_0 t')]} \int^{t^*} \sum_{n=-\infty}^{\infty} J_n(2B'i) e^{-inf_0 p} dp$$

$$= \sigma_N^2 e^{-A|\tau|} e^{-B'\sin(f_0 t)} e^{-B'\sin(f_0 t')} \sum_{n=-\infty}^{\infty} J_n(2B'i) \frac{e^{-inf_0 t^*}}{2A - inf_0} = \sigma_N^2 e^{-A|\tau|} e^{-B'\sin(f_0 t)} e^{-B'\sin(f_0 t')} \left\{ \frac{I_0(2B')}{2A} + 2\sum_{k=0}^{\infty} (-1)^k I_{2k+1}(2B') \frac{2A\sin[(2k+1)f_0 t^*] - (2k+1)f_0\cos[(2k+1)f_0 t^*]}{(2A)^2 + ((2k+1)f_0)^2} + 2\sum_{k=0}^{\infty} (-1)^k I_{2k}(2B') \frac{2A\sin[2kf_0 t^*] - 2kf_0\cos[2kf_0 t^*]}{(2A)^2 + (2kf_0)^2} \right\}. \quad (33)$$

Figure 6 shows the covariance kernel. After some manipulations, we find the cyclic spectrum in the form

$$S_l(f) = \sigma_N^2 \sum J_k(B'i) \times \sum J_m(B'i) \sum J_n(2B'i)\delta[(k+m-n)f_0 - 2\pi l/d] \times [S_A(mf_0 - \pi l/d + 2\pi f) + S_A(mf_0 - \pi l/d + 2\pi f)]/(2A - inf_0), \quad (34)$$

where

$$S_A(f) = \frac{A + if}{A^2 + f^2}. \quad (35)$$

Figure 7 shows the Bloch functions based on (31). They are essentially the same as the analytic counterparts based on (34). These Bloch functions (and also the CSEOFs) are very close to pure sinusoids. The eigenvalues are also close to those of a stationary case. This behavior is explained by the almost exponential decay of the covariance kernel with respect to lag, which is a characteristic structure of a stationary covariance kernel (Fig. 6). Contribution from each set of inner modes to the variance along $t = t'$ is small. All the modes may be required to fully explain the nested fluctuations along $t = t'$.

5. Numerical model

We have examined the cases that have exact solutions for the cyclic spectrum for the purpose of testing each component of the CSEOF algorithm. Let us finally turn to a real problem and compute CSEOFs of the observational data and those of a stochastic model that mimics the observed surface temperature anomaly field. The observation is the 100-yr (1890–1989) U.K.

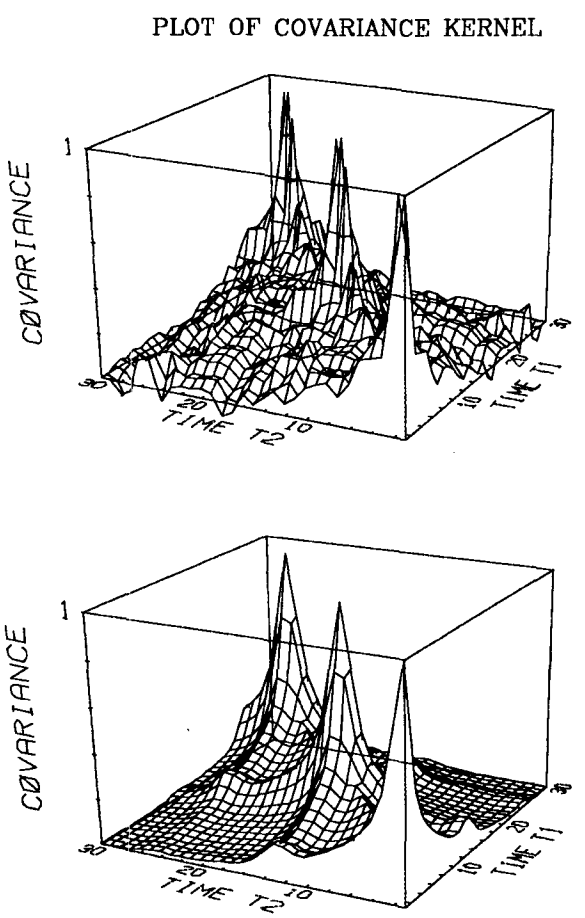

FIG. 8. Normalized lagged covariance kernel, $\mathcal{K}(t, t')$, of (a) the observational data and (b) a stochastic model. Here, T1 and T2 represent two different times. The parameters used in the stochastic model are $A = 1$, $B = 3$, $C = 1$, and $\alpha = 0.15$.

dataset archived at the National Center for Atmospheric Research. The dataset is the monthly average surface air temperature anomaly field averaged over a 5° × 5° lat–long box. There are a number of missing observations. Processing of the observational data is the same as in Kim and North (1991, 1992), and detailed discussion is omitted here.

The observed mean and standard deviation of the earth's surface temperature field have seasonal dependency; namely,

$$\mu(t) = \mu(t + d), \quad \sigma(t) = \sigma(t + d). \tag{36}$$

To remove the seasonal dependency of the mean and standard deviation, it is sensible to construct the fluctuating function (Parzen and Pegano 1979)

$$R(t) = \frac{T(t) - \mu(t)}{\sigma(t)}. \tag{37}$$

Then, $T(t)$ may be modeled by the function $R(t)$ and the known seasonal mean and standard deviation:

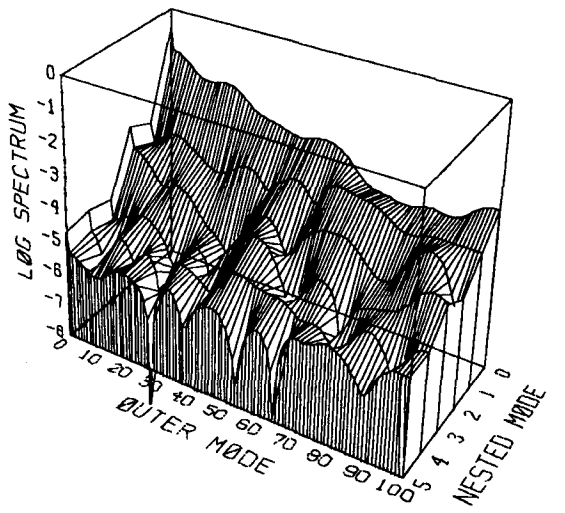

FIG. 9. Cyclic spectrum, $S_l(f)$, on a logarithmic scale of the observational data smoothed by using a Parzen covariance averaging kernel with a lag of 120 [see Eq. (15)]. The frequency, f, is discretized into $\{n/T | n = 0, \cdots, 100\}$, where n is the outer mode number and $T = 1200$ is the total number of sampling points. The nested mode l is the mode number of the nested fluctuations.

$$T(t) = \sigma(t)R(t) + \mu(t). \tag{38}$$

The fluctuating function may not necessarily be a stationary process. The $R(t)$ from the observational data, however, turns out to be almost stationary and is well modeled by a low-order AR process.

Thus, we assume a stochastic model in the form

$$T(t) = R(t)*\{1 + A \exp[-B*(1 - C \cos(f_0 t))]\}. \tag{39}$$

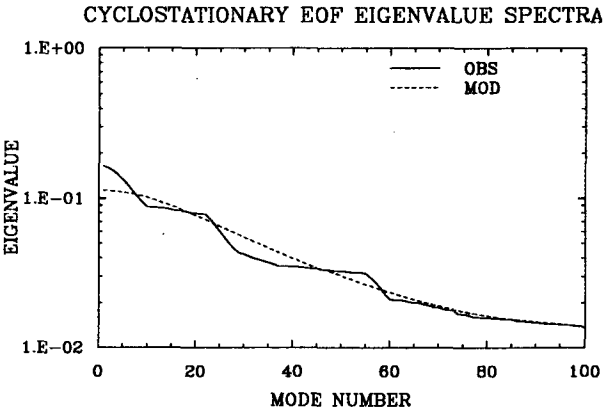

FIG. 10. Cyclostationary EOF eigenvalue spectra of observation and model. The model parameters are $A = 1$, $B = 3$, $C = 1$, and $\alpha = 0.15$.

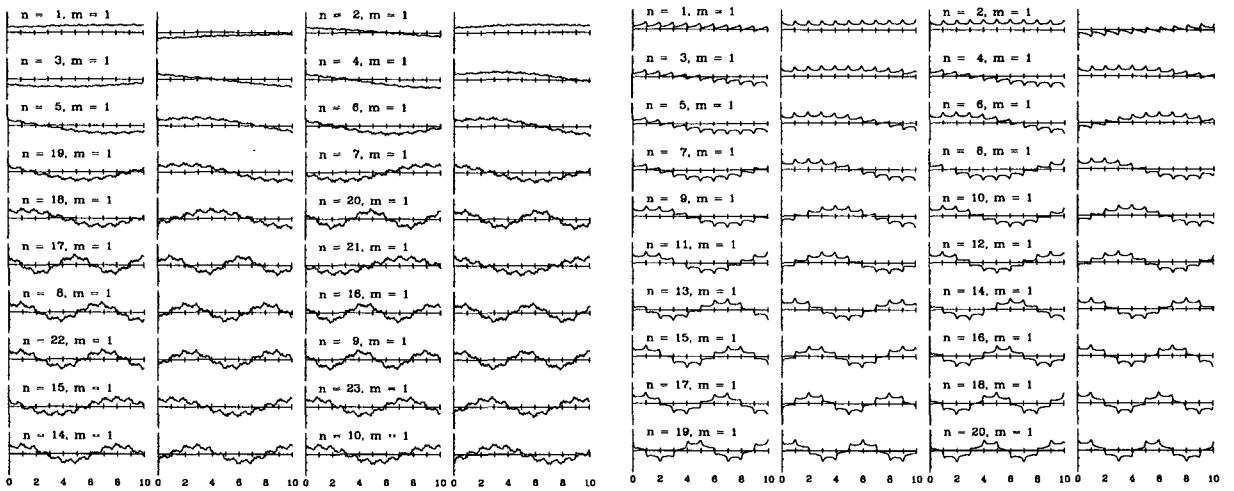

FIG. 11. The first 10 yr of the real and imaginary parts (left and right of each set, respectively) of the first 20 CSEOFs: (a) observation and (b) model. The model parameters are $A = 1$, $B = 3$, $C = 1$, and $\alpha = 0.15$. The first mode is at top left and the last mode is at bottom right. The inner mode number is 1 for all the CSEOFs shown. The respective outer mode numbers are 1, 2, 3, 4, 5, 6, 7, 8, 9, 18, 19, 10, 17, 20, 16, 21, 11, 15, 14, and 12 for the observation. For the model, outer mode numbers are sequential from 1 to 20.

Here $R(t)$ is an AR-1 model with the lagged covariance kernel satisfying (25) and the factor multiplying $R(t)$ is a modulation function. The f_0 is the modulation frequency, and A, B, and C are parameters regarding the shape of the modulation function. Of course, the modulating function mimics the seasonal standard deviation of the observational data as in (38). The model parameters may repeatedly be adjusted until the shape of the modulating function becomes reasonably similar to the seasonal standard deviation of the observational data.

Figure 8 shows the covariance kernels of the observational data and model. The covariance kernel, $K(t, \tau)$, of the observational data was computed from (1). The periodicity of the covariance kernel in t was enforced by averaging the covariance kernel over the same months. As τ becomes larger there are not many ensembles available for the average. The observational data was also smoothed using a Parzen covariance averaging kernel with a lag of 120 (Newton 1988). A visual inspection indicates that the modeled covariance kernel is in fair agreement with that of observation. Variance along the diagonal direction of the covariance kernel clearly shows stronger variability in the winter. Variance is very small for the rest of the year. Perhaps the most serious discordance is the model's inability to mimic the peak splitting (see Fig. 8a) of the observational data.

A frequency-domain counterpart of covariance is the cyclic spectrum as defined in (15). The cyclic spectrum is a useful generalization of the power spectrum in the stationary time series analysis. Figure 9 shows the cyclic spectrum, $S_l(f)$, on a logarithmic scale of the observational data smoothed by using a Parzen covariance averaging kernel with a lag of 120. The frequency, f, is discretized into $\{n/T | n = 0, \cdots, 100\}$, where n is the outer mode number and $T = 1200$ is the total number of sampling points. The nested mode represents l, which is the mode number of the nested fluctuations. For $l = 0$, the cyclic spectrum represents the smoothed spectrum of the annual average observations, which is typically a red spectrum. The cyclic spectrum also exhibits a small but nonnegligible annual ($l = 1$) and seasonal ($l = 2$) excursion of interannual fluctuations. As l increases, the spectrum becomes essentially flat with no preferred long-term temporal scales associated with the bimonthly variation. The variance, of course, is not partitioned in the cyclic spectrum and hence the CSEOFs are needed.

Figure 10 depicts the first 100 eigenvalues of the ordered CSEOFs from the observation and model. The percent variance explained by the first 100 CSEOFs is 86% for the observational data and 84% for the model. The absolute magnitude of the eigenvalue, of course, is irrelevant because it can be tuned. The model seems to capture the general trend of the eigenvalue spectrum rather well.

Figure 11 shows the first 20 CSEOFs of the observation and model. As shown in Table 1, 19 of the first 20 CSEOFs of the observational data (Fig. 11a) are identified with one of the first 20 CSEOFs of the model (Fig. 11b) with very high pattern correlation (>0.97). There is no correlation among the modes with different outer mode numbers because of the orthogonality property of CSEOFs. Thus, pattern correlation measures the similarity between two different datasets of Bloch functions with the same outer mode number. The largest decades of CSEOFs of both the observation and model are similar in disposition to EOFs of a stationary process. They are characterized by peaks in the winter season, which otherwise are cosine and sine functions with

TABLE 1. Pattern correlation (mode number in parentheses) between the CSEOFs of observational data and the best model estimates.

Obs mode	Best modeled	Obs mode	Best modeled	Obs mode	Best modeled	Obs mode	Best modeled
0	0.97 (0)	1	0.97 (1)	2	0.98 (2)	3	0.98 (3)
4	0.98 (4)	5	0.98 (5)	6	0.99 (6)	7	0.99 (7)
8	0.98 (8)	9	0.98 (17)	10	0.98 (18)	11	0.98 (9)
12	0.98 (16)	13	0.98 (19)	14	0.98 (15)	15	0.98 (20)
16	0.98 (10)	17	0.98 (14)	18	0.98 (13)	19	0.98 (11)

the frequency of outer mode numbers. These peaks represent higher variability during the winter of the surface temperature field. In addition to the winter peaks, CSEOFs of observation show small wiggles throughout the year. Such wiggles are due to the split peaks of the covariance kernel (Fig. 8a) and cannot be simulated in the present model.

6. Summary and concluding remarks

This study developed a computational algorithm for the one-dimensional CSEOFs and examined the properties of them. Simple one-dimensional examples have been solved and analyzed to gain insight into computing CSEOFs and interpreting them. The lagged covariance statistics of a cyclostationary process are periodic with the frequency of the nested undulations. This fact is made use of in Bloch's theorem and facilitates the computation of CSEOFs tremendously. Finally, a stochastic model of a one-dimensional cyclostationary process was developed, which mimics the statistics of the observed global-average surface air temperatures.

The greatest significance of Bloch's theorem is in the decoupling of the nested fluctuations (such as the annual cycle) from the long-term stationary component of fluctuations. This decoupling allows us to write covariance kernel and eigenfunctions as in (9) and (5). A consequence of decoupling is a separate eigenvalue problem for each outer mode n as in (13). This greatly facilitates the computation of CSEOFs. It is also important to note that the covariance kernel is decomposed as in (22).

The developed computational algorithm was validated using simple examples for which analytic solutions are possible. The primary strategy in this validation procedure was to compare the eigenfunctions and eigenvalues based upon the analytically derived eigenvalue problem and those from a numerical calculation. This includes the trivial case of a stationary process. Test results show that the developed computer codes produce CSEOFs and the associated eigenvalues accurately.

We constructed a stochastic model that reasonably (to a first approximation) reproduces a lagged covariance kernel. The largest CSEOFs of the monthly global-average surface temperature field were faithfully reproduced by the model. Specifically, 19 of the first 20 CSEOFs of observational data were identified with one of the first 20 CSEOFs of the model with high pattern correlation (>0.97). Many of the first 100 CSEOFs are essentially the cosine and sine functions with the frequency of outer mode numbers and are characterized by peaks in the winter. These functions are an analogy to EOFs of a stationary process. The peaks in the winter season signify the stronger variability during winter months. The eigenvalue spectra of the first 100 CSEOFs, which explain, respectively, 86% and 84% of total variance, were consistent between observation and model.

While cyclostationary processes are commonly encountered in studies of the earth's climate, scientists have been reluctant to use CSEOFs mainly because they are cumbersome and difficult to compute. The difficulty of interpretation of CSEOFs may also be a big disincentive. This study shows that CSEOFs can be computed easily by utilizing the factorization theorem by Smith (1969). Two key issues we did not address in this study are (1) the sensitivity of CSEOFs and (2) a two-dimensional generalization of CSEOFs as in Kim and North (1993). These issues should be examined before CSEOFs are applied to the linear estimation problems in climate studies. We hope this study represents an essential stepping stone for future studies.

Acknowledgments. We greatly benefited from the valuable comments of the anonymous reviewers. We are grateful to Mr. Neil Smith, who carefully proofread the manuscript. We also thank the National Science Foundation for its support of this work via Grant ATM-9423335 to Texas A&M University.

REFERENCES

Barnett, T., 1986: Detection of changes in the global tropospheric temperature field induced by greenhouse gases. *J. Geophys. Res.*, **91**, 6659–6667.

——, 1991: An attempt to detect the greenhouse-gas signal in a transient GCM simulation. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. E. Schlesinger, Ed., Elsevier, 559–568.

——, and M. Schlesinger, 1987: Detecting changes in global climate induced by greenhouse gases. *J. Geophys. Res.*, **92**, 14 472–14 780.

——, ——, and X. Jiang, 1991: On greenhouse gas signal detection strategies. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. E. Schlesinger, Ed., Elsevier, 537–558.

Epstein, E. S., 1982: Detecting climate change. *J. Appl. Meteor.*, **21,** 1172–1182.

Gardner, W. A., 1994: *Cyclostationarity in Communications and Signal Processing.* IEEE Press, 504 pp.

——, and L. E. Franks, 1975: Characterization of cyclostationary random signal processes. *IEEE Trans. Inform. Theory,* **21,** 4–14.

Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.,* **93,** 11 015–11 021.

——, 1993: Optimal fingerprints for the detection of time-dependent climate change. *J. Climate,* **6,** 1957–1971.

——, and T. P. Barnett, 1981: Techniques of linear prediction for systems with periodic statistics. *J. Atmos. Sci.,* **38,** 2275–2283.

Jones, P. D., and K. R. Briffa, 1992: Global surface air temperature variations during the twentieth century. Part 1: Spatial, temporal and seasonal details. *Holocene,* **2,** 165–179.

Kim, K.-Y., and G. R. North, 1991: Surface temperature fluctuations in a stochastic climate model. *J. Geophys. Res.,* **96,** 18 573–18 580.

——, and ——, 1992: Seasonal cycle and second-moment statistics of a simple coupled climate system. *J. Geophys. Res.,* **97,** 20 437–20 448.

——, and ——, 1993: EOF analysis of surface temperature field in a stochastic climate model. *J. Climate,* **6,** 1681–1690.

Madden, R. A., 1977: Estimates of the autocorrelations and spectra of seasonal mean temperatures over North America. *Mon. Wea. Rev.,* **105,** 9–18.

Newton, H. J., 1988: *TIMESLAB: A Time Series Analysis Laboratory.* Wadsworth and Brooks, 623 pp.

North, G. R., and K.-Y. Kim, 1995: Detection of forced climate signals. Part II: Numerical simulations. *J. Climate,* **8,** 409–417.

——, ——, S. S. P. Shen, and J. W. Hardin, 1995: Detection of forced climate signals. Part I: Theory. *J. Climate,* **8,** 401–408.

Parzen, E., and M. Pagano, 1979: An approach to modeling seasonally stationary time series. *J. Econometrics,* **9,** 137–153.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1988: *Numerical Recipes.* 3d ed. Cambridge University Press, 818 pp.

Santer, B. D., T. M. L. Wigley, P. D. Jones, and M. E. Schlesinger, 1991: Multivariate methods for the detection of greenhouse-gas-induced climate change. *Greenhouse-Gas-Induced Climatic Change: A Critical Appriaisal of Simulations and Observations,* M. Schlesinger, Ed., Elsevier, 511–536.

——, ——, and ——, 1993: Correlation methods in fingerprint detection studies. *Climate Dyn.,* **8,** 265–276.

——, W. Brüggemann, U. Cubasch, K. Hasselmann, H. Höck, E. Maier-Reimer, and U. Mikolajewicz, 1994: Signal-to-noise analysis of time-dependent greenhouse warming experiments. *Climate Dyn.,* **9,** 267–285.

Shen, S. S., G. R. North, and K.-Y. Kim, 1994: Spectral approach to optimal estimation of the global average temperature. *J. Climate,* **7,** 1999–2007.

Smith, R. A., 1969: *Wave Mechanics of Crystalline Solids.* Chapman and Hall, 553 pp.

von Storch, H., and F. W. Zwiers, 1988: Recurrence analysis of climate sensitivity experiments. *J. Climate,* **1,** 157–171.

——, T. Bruns, I. Fischer-Bruns, and K. Hasselmann, 1988: Principal oscillation pattern analysis of the 30 to 60 day oscillation in general climate model equatorial troposphere. *J. Geophys. Res.,* **93,** 11 022–11 036.

——, G. Bürger, R. Schnur, and J.-S. von Storch, 1995: Principal oscillation patterns: A review. *J. Climate,* **8,** 377–400.

Wallace, J. M., and R. Dickinson, 1972: Empirical orthogonal representation of time series in the frequency domain. Part I: Theoretical considerations. *J. Appl. Meteor.,* **11,** 887–892.

Zwiers, F. W., and H. von Storch, 1989: Multivariate recurrence analysis. *J. Climate,* **2,** 1538–1553.