Optimal Fingerprints for the Detection of Time-dependent Climate Change

K. HASSELMANN

Max-Planck-Institut für Meteorologie, Hamburg, Germany

(Manuscript received 24 August 1992, in final form 17 March 1993)

ABSTRACT

An optimal linear filter (fingerprint) is derived for the detection of a given time-dependent, multivariate climate change signal in the presence of natural climate variability noise. Application of the fingerprint to the observed (or model simulated) climate data yields a climate change detection variable (detector) with maximal signal-to-noise ratio. The optimal fingerprint is given by the product of the assumed signal pattern and the inverse of the climate variability covariance matrix. The data can consist of any, not necessarily dynamically complete, climate dataset for which estimates of the natural variability covariance matrix exist. The single-pattern analysis readily generalizes to the multipattern case of a climate change signal lying in a prescribed (in practice relatively low dimensional) signal pattern space: the single-pattern result is simply applied separately to each individual base pattern spanning the signal pattern space. Multipattern detection methods can be applied either to test the statistical significance of individual components of a predicted multicomponent climate change response, using separate single-pattern detection tests, or to determine the statistical significance of the complete signal, using a multivariate test. Both detection modes make use of the same set of detectors. The difference in direction of the assumed signal pattern and computed optimal fingerprint vector allows alternative interpretations of the estimated signal associated with the set of optimal detectors. The present analysis yields an estimated signal lying in the assumed signal space, whereas an earlier analysis of the time-independent detection problem by Hasselmann yielded an estimated signal in the computed fingerprint space. The different interpretations can be explained by different choices of the metric used to relate the signal space to the fingerprint space (inverse covariance matrix versus standard Euclidean metric, respectively). Two simple natural variability models are considered: a space–time separability model, and an expansion in terms of POPs (principal oscillation patterns). For each model the application of the optimal fingerprint method is illustrated by an example.

1. Introduction

The general public concern and ongoing scientific debate on the anticipated global warming due to increasing greenhouse gas concentrations and on the impact of other activities of man on the earth's climate has generated a strong demand for the development of improved techniques for the early detection of the predicted climate change signal. A clear identification of the anthropogenic signal in climate observations would reduce the present scientific uncertainties regarding the magnitude and form of the anticipated climate change and would provide a more reliable quantitative basis for the development of rational political abatement and adaptation strategies.

At the core of the detection problem is the development of a suitable strategy for distinguishing between the anticipated externally generated time-dependent climate change signal and the natural internal variability of the climate system. The problem can be divided into three parts: 1) identification of the climate change signal that one wishes to detect, 2) determination of the relevant statistical properties of the natural climate variability background, and 3) development of an optimal detection method. This paper addresses the third problem. However, the question of the representation (but not the estimation) of the second-moment statistics of the natural climate variability noise needed for optimal detection, which relates to the second problem, will also be considered briefly.

It will be assumed throughout that the first problem has been resolved and the general structure of the space- and time-dependent climate change signal that one is seeking to detect has been determined, for example, from model simulations. It is well known that in attempting to detect signals in noisy multivariate data, the number of degrees of freedom of the signal must be severely curtailed in order to extract statistically significant results. Specifically, it will be assumed that the signal is defined to within the unknown coefficients of a relatively small set of prescribed time-varying response patterns (cf. Cubasch et al. 1992).

This should not gloss over the difficulties, however, of defining the anticipated climate change signal. This is generally a nontrivial problem requiring an intercomparison and detailed assessment of different climate change simulations with different models. The problem

Corresponding author address: Dr. Klaus Hasselmann, Max-Planck-Institut für Meteorologie, Bundestrasse 55, D-2000 Hamburg 13, Germany.

is compounded by the fact that realistic time-dependent climate change simulations can be carried out only with coupled ocean–atmosphere general circulation models (CGCMs), which generate their own natural climate variability (Washington and Meehl 1989; Stouffer et al. 1989; Manabe et al. 1991; Cubasch et al. 1992; Santer et al. 1993; see also the stochastically forced ocean experiment of Mikolajewicz and Maier-Reimer 1990). Therefore, the signal detection problem already arises in the attempt to define the climate response signal in model simulations. This calls into question the basic premise of the separability of the signal definition and signal detection problems, that is, the assumption that the signal to be detected in an observed or model-simulated dataset can be defined a priori in an independent model experiment. This will nevertheless be assumed in the following as a conceptual starting point.

In the same spirit, the estimation of the statistical properties of the natural climate variability required for optimal detection, which is at the core of the second problem, will also be regarded as resolved. This is again a nontrivial assumption, but is necessary in order to pose a well-defined detection problem. In practice, much of the ongoing debate on whether an anthropogenic climate signal can already be detected today revolves around uncertainties over the structure and magnitude of the natural variability of the real climate. Nevertheless, in order to develop a consistent conceptual framework it will be assumed, again as a starting point, that estimates of the required natural climate variability statistics are available. All fields are assumed to be statistically stationary (indeed, if this is not the case it is generally not possible to construct a statistical ensemble of realizations, which is a prerequisite for the development of a meaningful statistical theory of signal detection).

The assumption of known statistics is not as severe a restriction, however, as it may appear at first sight. The detection strategy will be developed in the following for an "observed climate state vector," which refers to any set of climate variables or indices for which adequate statistical data are available. The observed climate state need not be dynamically complete, as required, for example, for a climate model. Thus the "observed climate state" can consist of time series of varying length and for different climate variables, measured at inhomogeneously distributed stations. Time-dependent signal detection is concerned only with the kinematics (the description of the time-dependent evolution), not the dynamics (the causes of the evolution) of climate change. A dynamically incomplete representation of the climate state therefore has no consequences for the method of detection (apart, of course, from the unavoidable loss of detection power associated with a loss of information).

After an optimal detection strategy has been developed under these two working assumptions, the general question of the interdependence of the three elements of the overall detection problem can be revisited in further iterations. The discussion of some possible approaches to modeling the relevant second-moment properties of the natural climate variability presented at the end of this paper represents a first step in this direction.

The detection problem, in the separable form discussed here, is often viewed as the task of identifying the most sensitive climate index, from a large set of potentially available indices, for which the anticipated anthropogenic climate signal can be most readily distinguished from the natural climate noise. Global or regional mean surface temperature, vertical temperature differences, sea ice extent, sea level change, and integrated deep ocean temperatures are examples of indices that have been discussed in this context (e.g., Wigley and Jones 1981; Barnett 1986; Barnett and Schlesinger 1987; Karoly 1987, 1989; Munk and Forbes 1989; Mikolajewicz et al. 1993). A more systematic approach, however, is the fingerprint method. Here, all climate variables are regarded as containing potentially useful information on climate change, and the task is to extract from the full set of available observed climate variables an optimal net climate change detection index (cf. Madden and Ramanathan 1980; MacCracken and Moses 1982). This approach will be pursued in this paper. While it is theoretically conceivable that a single climate variable will turn out to be the most effective detector, in the general case all variables will carry some signal information, although with varying levels of noise contamination. The optimal detector, defined as the variable that has a maximal signal-to-noise ratio, will consist of a weighted linear combination (the "fingerprint") of all variables.

Most investigations in the past have considered only partial aspects of the full multivariate, space–time-dependent problem. The detection of a time-independent equilibrium change of a multivariate climate system in response to a constant external forcing has been investigated by Hasselmann (1979, referred to in the following as H), Hannoschöck and Frankignoul (1985), Hense (1986), Hense et al. (1990), Bell (1982, 1986), and others. The complementary problem of detecting a time-dependent climate change signal, while disregarding the spatial and multivariable nature of the climate signal, has been studied by Bloomfield and Nychka (1992). These authors did not address, however, the question of the optimization of the detection variable. Various other facets of the signal detection problem have been investigated or reviewed by Preisendorfer (1988), Wigley and Barnett (1990), Barnett (1991), Barnett et al. (1991), Solow (1991), and other authors. Recently, however, North et al. (1993) and North and Kim (1993) have independently investigated the full space–time-dependent climate change detection problem, also using pattern analysis techniques rather similar to the approach pursued here.

In comparing the theoretical signal predictions with

observed data, many studies apply some form of filtering or pattern correlation technique that in effect projects the observed data onto the predicted signal pattern. This essentially corresponds to the usual least-squares fitting of a signal pattern to data such that the variance of the residual is minimized. While this removes much of the irrelevant noise in the data, it does not represent the optimal signal detection solution. This should maximize the signal-to-noise ratio rather than the explained variance.

In this paper an optimal space–time-dependent filter (fingerprint) is derived that maximizes the signal-to-noise ratio for the associated detector for any prescribed multivariable, space–time-dependent climate signal. The single-pattern solution is then generalized to the multipattern case to determine a set of p optimal fingerprints and associated detectors for any climate change signal lying in a prescribed space spanned by p given climate change patterns. To determine the statistical significance of the estimated climate signal it must be assumed that the natural climate variability is Gaussian (or is otherwise known). However, the definition of the optimal fingerprints as such is not dependent on the Gaussian hypothesis.

The detection problem is defined in section 2. Assuming that the space–time-dependent structure of the externally generated climate signal is prescribed to within an unknown amplitude factor, the optimal space–time-integrated detector is derived in section 3 as a linear combination (the optimal "fingerprint") of the complete set of time-dependent climate variables. The result is generalized in section 4 to the case of a signal defined only as an unknown linear combination of a finite set of prescribed time-dependent patterns. The analysis follows the basic ideas of H, extended in a straightforward manner to include the time dimension, but is simplified by the introduction of the "fingerprint" terminology. The problem of modeling the complex space–time- (or space–frequency) dependent second moments of the natural climate variability required for optimal detection is discussed in section 5. Two simplifications are considered and illustrated by examples: the assumption of space–time separability, and the introduction of approximate POP (principal oscillation pattern) representations. The results are summarized and some open questions mentioned in the concluding section 6.

2. The detection problem for a multicomponent space–time-dependent signal

Consider the evolution of an "observed climate state" $\varphi_{v,\mathbf{x}}(t) = \phi_i(t)$ in response to some time-dependent external forcing over a time interval $0 \le t \le T$. The "observed climate state" can be represented as a discretized composite vector $\phi = (\phi_i)$ whose indices $i = (v, \mathbf{x})$ run through the climate variables v (temperature, pressure, moisture, . . .) and the discrete spatial

coordinates $\mathbf{x}$, which can refer, for example, to a set of observing stations or a model grid. The climate trajectory $\phi(t)$ can represent either a set of variables of the real climate system or the simulated response in some numerical model experiment. As has been pointed out, it is not necessary that $\phi(t)$ provide a dynamically complete description of the climate system. The "observed state vector" can (and, indeed, must) be limited to variables for which sufficient observational information is available to adequately define the statistics of the ensemble of trajectories $\tilde{\phi}(t)$ characterizing the natural variability of the system. The detection problem is then to decide whether the climate time history $\phi(t)$ generated by the external forcing can be distinguished, at some given level of significance, from the statistical ensemble of natural variability trajectories $\tilde{\phi}(t)$.

For the time-independent problem, an optimal signal detection strategy has been presented in H. In the following, the analysis of H is extended to the general multivariate, time-dependent case and recast in a simpler "fingerprint" terminology.

Formally, the approach of H can be immediately generalized to the time-dependent case by simply discretizing the time variable and incorporating the time index t in a combined index $a = (i, t)$ running from 1 to n. The extended climate vector $\pmb{\psi} = (\psi_a)$ summarizes then the complete climate trajectory: ψ_a = set of all $\phi_i(t)$ in the discretized time interval $t = 1, 2, \cdots T$ (the time-discretization interval is set to unity). The introduction of a compact variable space–time index $a = (v, \mathbf{x}, t)$ not only simplifies the notation, but also focuses on the essentially very simple linear algebra geometry of the detection problem.

An important requirement for a successful detection strategy is the reduction of the number of degrees of freedom of the signal. This is achieved in H by considering a signal that is defined a priori only in a relatively low-dimensional subspace of the full climate system. Attempts to test whether the full climate response vector ψ can be distinguished from an element $\tilde{\psi}$ of the natural variability ensemble without a reduction in dimensionality of the signal will generally fail for the following reasons (cf. H; Barnett and Hasselmann 1979).

The climate response to given external forcing can be represented generally (ignoring nonlinear interactions between the natural climate variability and the externally forced response) as a superposition

$$\psi = \psi^s + \tilde{\psi} \tag{1}$$

of the forced deterministic climate signal ψ^s and a particular realization $\tilde{\psi}$ of the natural variability ensemble. Let $\mathbf{v}$ be a unit-length vector denoting the direction of the signal, $\psi^s = |\psi^s|\mathbf{v}$. Assume now that the signal is statistically significant in the sense that the signal amplitude $|\psi^s|$ is large compared with the noise component in the $\mathbf{v}$ direction,

$$|\psi^s|^2 \gg \langle (\mathbf{v}^T\tilde{\psi})^2 \rangle, \qquad (2)$$

where the cornered parentheses denote ensemble means and matrix notation is used, a superscript T denoting the transpose. The mean of the climate noise is taken to be zero, $\langle \tilde{\psi} \rangle = 0$. If the direction $\mathbf{v}$ is known a priori, it is then possible to test for the statistical significance of the net-square response $(\mathbf{v}^T\psi)^2$ in the $\mathbf{v}$ direction, and the outcome under the inequality (2) would be positive.

If, on the other hand, the signal pattern is unknown, one can test only the statistical significance of the complete n-dimensional response vector ψ. This requires considering the magnitude and orientation of the response vector in relation to the joint probability density of the ensemble of vectors $\tilde{\psi}$ in the full n-dimensional climate trajectory space. If the probability density decreases monotonically with distance from the origin, as is normally the case, each $(n-1)$-dimensional hypersurface of constant probability density will divide the n-dimensional space into an internal closed region of some probability measure P around the origin and an external open region of probability measure $(1 - P)$, in which the probability density is everywhere smaller than in the internal region. The response vector ψ is then normally termed statistically significant at the significance level P if the end point of the response vector, drawn from the origin, lies in the external region. Without entering here into the details of the analysis, it is qualitatively apparent that the larger the number of irrelevant noise dimensions, the smaller the relative contribution of the signal to the total magnitude of the response vector, and the more difficult it will be to detect the signal in the full n-dimensional space—even when the signal component is significantly larger than the noise component in the one (unknown) direction $\mathbf{v}$.

Fortunately, this second "needle in a haystack" situation will normally not apply in practice: the direction of the hypothesized climate signal can be assumed to be known from model simulations, or at least to lie within a known subspace of relatively small dimension.

The case that the climate change signal is known exactly applies for the simplest yes–no question of climate change detection: one wishes to determine whether a specific time-dependent global climate signal that has been predicted by a model can be detected in the data (or at what future time it should become detectable in the data). The more general case that the response signal is assumed to lie only within some prescribed low-dimensional subspace of the full climate trajectory space arises if the predicted climate change signal is only imperfectly known (for example, because different models have predicted different climate change patterns) or if one wishes to distinguish between different climate change signals produced by different anthropogenic or natural external forcing mechanisms.

A multipattern analysis is necessary also if one wishes to test the statistical significance not only of the complete global climate change signal, but also of particular subcomponents of the global signal. It can be anticipated, for example, that the most effective single global climate change detector will be based primarily on the large-scale features of the climate fields. But for policymakers, the regional climate changes (which at present are not predicted very reliably by climate models) will presumably be of greater concern than globally integrated quantities. They will therefore wish to know not only whether the global climate change predicted by models has been detected, but also whether the model predictions of climate change on the regional scale can be confirmed by observations.

The analysis of this paper can be extended to the case of statistically rather than deterministically prescribed climate signals. This could be appropriate, for example, if the predicted signal is inferred from an ensemble of different model simulations with different levels of credibility, or from a mixture of model simulations and general theoretical considerations (suggesting, for example, a land–sea contrast signal). However, this Bayesian generalization will not be pursued further here.

3. The optimal fingerprint for a prescribed signal pattern

If the pattern $\mathbf{g}$ of the space–time signal trajectory ψ^s is known to within an unknown constant amplitude factor c,

$$\psi^s = c\mathbf{g}, \qquad (3)$$

the signal detection problem reduces (if one limits oneself to linear techniques) to the task of deriving an optimal detection variable or "detector,"

$$d = (\mathbf{f}^T\psi), \qquad (4)$$

computed from ψ by application of a linear "filter function" or "fingerprint" $\mathbf{f} = (f_a)$, for which the square signal-to-noise ratio

$$R^2 = \frac{(d^s)^2}{\langle \tilde{d}^2 \rangle}, \qquad (5)$$

is maximized. Here

$$d^s = (\mathbf{f}^T\psi^s) \qquad (6)$$

and

$$\tilde{d} = (\mathbf{f}^T\tilde{\psi}) \qquad (7)$$

represent, respectively, the signal and climate-variability noise components of the net detector $d = d^s + \tilde{d}$.

The fingerprint vector $\mathbf{f}$ is determined through the condition $R^2 = \text{max}$ only to within a factor. Similarly, the signal pattern $\mathbf{g}$ need be defined only to within an arbitrary factor since the amplitude of the signal will

be estimated by the detection procedure. Although **f** and **g** could therefore be normalized to unit vectors, it is notationally more convenient to leave the vectors unnormalized at this point.

The maximization of R^2 with respect to the arbitrarily normalized vector **f** is equivalent to the minimization of $\langle d^2 \rangle$ under the side condition $(d^s)^2$ = const. This yields as determining equation for the fingerprint **f**,

$$\mathbf{Cf} + \lambda(\mathbf{f}^T\mathbf{g})\mathbf{g} = 0, \qquad (8)$$

where $\mathbf{C} = C_{ab} = \langle \tilde{\psi}_a \tilde{\psi}_b \rangle$ is the covariance matrix of the natural climate variability and λ is a Lagrange multiplier whose value depends on the normalization chosen for **f**. The solution is

$$\mathbf{f} = \mathbf{C}^{-1}\mathbf{g}, \qquad (9)$$

with $\lambda = -(\mathbf{g}^T\mathbf{C}^{-1}\mathbf{g})^{-1}$.

The optimal fingerprint direction is in general not parallel to the assumed signal direction, as may perhaps have been expected intuitively (and is often assumed in detection studies). This is best understood by transforming to statistically orthogonal coordinates (denoted in the following by primes),

$$\psi_a = \sum_b \psi'_b e_{ba}, \qquad (10)$$

which are defined with respect to an orthonormal base $e_{ba} = e_b$ consisting of the eigenvectors (empirical orthogonal functions, EOFs) e_b of **C**,

$$\mathbf{Ce}_b = \sigma_b^2 \mathbf{e}_b, \qquad (11)$$

with

$$(\mathbf{e}_a^T\mathbf{e}_b) = \delta_{ab}. \qquad (12)$$

In EOF coordinates, the covariance matrix of the natural climate variability takes the diagonal form

$$\langle \tilde{\psi}'_a \tilde{\psi}'_b \rangle = \sigma_a^2 \delta_{ab}, \qquad (13)$$

where σ_a^2 is the variance associated with the EOF e_a. Equation (9) thus becomes

$$f'_a = g'_a \sigma_a^{-2}. \qquad (14)$$

The multiplication of the signal with the inverse of the covariance matrix is seen to weight the fingerprint components f'_a in the EOF frame relative to the signal components g'_a by the inverse σ_a^{-2} of the EOF variances, thereby slewing the fingerprint vector away from the EOF directions with high noise levels toward the low-noise directions. [In practice, the EOF spectrum, if estimated from data, should be truncated after a finite number of terms, since the higher-index eigenvalues tend to be underestimated, leading to a spurious amplification of the higher-index fingerprint components (cf. v. Storch and Hannoschöck 1985; Preisendorfer 1988).]

For the special case of a single time-dependent variable, the result (13) is well known from classical signal processing theory (cf. Wainstein and Zubakov 1962). The EOFs for a statistically stationary time series are simply the harmonic functions of the Fourier series representation, so that Eq. (14) reproduces in this case the basic theorem that the optimal signal detection filter for a stationary time series is given by the Fourier transform of the signal divided by the noise variance spectrum.

Implicit in the definition of the optimal detector as the linear variable that maximizes the square signal-to-noise ratio is the assumption that the statistical significance of the detector increases monotonically with the signal-to-noise ratio. For most climate variability distributions, this will be the case. It has also been assumed that the only source of statistical noise in the detector is the natural climate variability. In practice, data errors will also contribute to the detector noise. However, these are generally small compared with the climate variability, and, to avoid complicating the analysis, will be ignored.

The statistical significance of the optimally detected signal d can be computed from the probability distribution of the noise variable $\tilde{d}$ for the null hypothesis that there is no signal. For this purpose it is normally assumed that the natural climate variability is Gaussian, so that all distributions may be derived from the covariance matrix **C**.

In this case, if **C** is known exactly (as opposed to being estimated from a finite dataset), $\tilde{d}$ is also Gaussian with variance

$$\langle \tilde{d}^2 \rangle = \sum_a \sigma_a^{-2} g'^2_a. \qquad (15)$$

If the covariance matrix is estimated from a finite dataset, $\tilde{d}$ has statistics generally similar to a Student's t variable (Morrison 1990). However, $\tilde{d}$ differs from a Student's distribution in that the direction for which the variance of $\tilde{d}$ is estimated is not prescribed a priori, but is modified relative to the prescribed signal direction by multiplication with the inverse of the estimated covariance matrix. The resultant variable does not correspond to a standard tabulated statistical variable, and its distribution must therefore be estimated by approximate analytical techniques or Monte Carlo simulations.

In principle, the significance level can also, of course, be computed for an arbitrary non-Gaussian, but known, statistical distribution. However, it will normally be difficult in the general time-dependent case to obtain reliable direct estimates of statistical distribution using, for example, permutation methods. This requires creating an ensemble of realizations, for which time series are needed that are significantly longer than the analysis time interval T (the same problem arises also in the estimation of the covariance matrix **C**). The only recourse in this case may be to augment the observational or model-simulated data with still longer model simulations of the natural climate variability.

The analysis so far has addressed only the problem of *detecting* a signal with known direction g. How can the optimal detector d and associated fingerprint f be translated now into an *estimate* of the signal? This requires defining the direction and magnitude of the estimated signal as a function of g, f, and d. The answer is not unique and depends on how the fact that f and g are not parallel is interpreted.

In H, the optimal detection problem was formulated as the task of finding an optimal unit-length signal detection vector b, given the signal direction g, which maximizes the signal-to-noise ratio for the coefficient c_1 of the estimated signal

$$\psi^e = c_1 b. \tag{16}$$

The coefficient c_1 itself was determined in the standard manner by minimizing the mean square of the residual ψ^r of the net response

$$\psi = \psi^e + \psi^r \tag{17}$$

which yielded

$$c_1 = \langle b^T \psi \rangle. \tag{18}$$

The optimal detection direction was found to lie in the direction of the optimal fingerprint, as given by Eq. (9),

$$b = \frac{f}{|f|}, \tag{19}$$

so that the best signal estimate was given by

$$\psi^e = \frac{\langle f^T \psi \rangle}{|f|^2 f}. \tag{20}$$

Alternatively, one can adopt the view that since the direction of the signal is assumed to be g, the estimated signal should also be taken to lie in this direction,

$$\psi^e = c_2 g. \tag{21}$$

The coefficient c_2 should then be determined by the condition that in the absence of noise one should recover the true signal (or, equivalently, that in the presence of noise the mean-square deviation from the true signal should be minimized). This yields

$$c_2 = d(g^T C^{-1} g)^{-1}. \tag{22}$$

Thus the estimated signal is given in this case by

$$\psi^e = \langle f^T \psi \rangle (g^T C^{-1} g)^{-1} g. \tag{23}$$

This interpretation will be adopted in the following discussion.

One can argue either viewpoint. The suppression in the fingerprint pattern of the signal components associated with high noise levels can be interpreted with H to imply that one has actually tested the climate change pattern only in the low-noise fingerprint direction, and that the existence of climate change signal components in the suppressed noise-contaminated di-

rections cannot be supported by the data. On the other hand, one can adopt the present view that the climate change signal, if it exists, is specified a priori as a complete vector, and one is at liberty to test any linear projection of the signal onto some chosen direction as evidence of the existence of the complete signal.

4. The multipattern problem

The single-pattern analysis can be readily generalized to the case in which the direction of the signal vector is no longer prescribed but is postulated to lie only in a space spanned by p given guess vectors g_ν, $\nu = 1$, $\cdots p$,

$$\psi^s = \sum_{\nu=1}^{p} c_\nu g_\nu. \tag{24}$$

The set of patterns is nonhierarchical and the base used to represent the prescribed signal space can be chosen arbitrarily. The guess vectors will not be orthonormalized, however, in order to preserve their original physical meaning. The guess vectors may represent, for example, different possible time-dependent climate change patterns induced by a CO_2 increase or enhanced aerosol concentrations, or the climate change associated with variations in the solar constant, or some regional climate change pattern. In general, these space–time patterns will not be orthogonal.

In most applications, the coefficients c_ν are estimated from the net response ψ, Eq. (1), by minimizing the mean-square residual ψ, yielding the standard least-squares solution

$$c_\nu = \sum_\mu H_{\nu\mu}^{-1} (g_\mu^T \psi), \tag{25}$$

where

$$H_{\nu\mu} = (g_\nu^T g_\mu). \tag{26}$$

However, in the present application the goal is not to maximize the explained variance but to maximize the signal-to-noise ratio for the estimated signal. This yields a different solution. The following derivation is based on H, but is simplified through the application of the single-pattern fingerprint concept.

As before, the problem is solved in two steps. First, a set of p detectors d_ν is derived for which the p-dimensional statistical significance (assuming a Gaussian distribution) is maximized. It will be shown that these are just the single-pattern detectors of the individual patterns g_ν,

$$d_\nu = (f_\nu \psi), \tag{27}$$

where

$$f_\nu = C^{-1} g_\nu. \tag{28}$$

In a second step, the coefficients d_ν are then assigned to p base vectors b_ν to construct an estimate of the signal.

Consider first the case that the signal direction is known. Thus, although the signal can in principle lie anywhere in the signal pattern space g_ν, the coefficients c_ν are in fact given. The detection problem reduces then to the previous single signal pattern detection problem, the only difference being that the signal is now represented in the form (24) with respect to the vector base g_ν. From Eqs. (4) and (9) it follows then that the optimal fingerprint and associated optimal detector are given by, respectively,

$$\mathbf{f} = \sum_\nu c_\nu \mathbf{f}_\nu \qquad (29)$$

and

$$d = (\mathbf{f}^T \psi) = \sum_\nu c_\nu d_\nu, \qquad (30)$$

where $\mathbf{f}_\nu$ and d_ν are given by Eqs. (28) and (27), respectively. Thus the set of fingerprints $\mathbf{f}_\nu$ and detectors d_ν represent a straightforward generalization of the single-pattern solution to the multipattern case in a first simple sense that they yield the optimal fingerprint, Eq. (29), and maximal signal-to-noise detector, Eq. (30), for any signal, Eq. (24), with a prescribed direction in the space spanned by the set of patterns g_ν.

This result is relevant if one wishes to test the statistical significance of individual a priori defined components of a multipattern signal. For example, one may wish to decide whether the global mean temperature trend or the change in precipitation over a limited specified region are statistically significant. Although this is important for many applications, the more general situation in multipattern signal detection is that the direction of the signal is unknown, except that it is assumed to lie in the p-dimensional signal pattern space spanned by the set of p prescribed guess patterns g_ν. The task is then to decide whether the climate change signal found within this space is statistically significant, without specifying the direction of the signal a priori.

One needs in this case to find a set of p detectors d_ν, given by the general form (27) with suitably defined, but as yet unknown, fingerprints f_ν, such that the relevant p-dimensional statistical significance measure of the vector $\underline{d} = (d_\nu)$ is maximized for any signal lying in the p-dimensional guess pattern space g_ν (underlined symbols denote vectors or matrices in the p-dimensional space with component indices $\mu, \nu \ldots$, in contrast to the vectors considered hitherto in the n-dimensional observed climate-trajectory space with component indices $a, b, \ldots$). It will be shown that the solution to this problem is again given by the fingerprints f_ν defined by Eq. (28).

The standard measure of the statistical significance of the p-dimensional vector $\underline{d} = (d_\nu)$ in the presence of a Gaussian background noise field is the ρ^2 statistic

$$\rho^2(\underline{d}) = \underline{d}^T \underline{\underline{D}}^{-1} \underline{d}, \qquad (31)$$

where

$$\underline{\underline{D}} = D_{\mu\nu} = \langle \tilde{d}_\mu \tilde{d}_\nu \rangle = \mathbf{g}_\mu \mathbf{C}^{-1} \mathbf{g}_\nu \qquad (32)$$

and

$$\tilde{d}_\nu = (\mathbf{f}_\nu^T \tilde{\psi}) \qquad (33)$$

represent the covariance matrix and individual components, respectively, of the detector vector $\underline{\tilde{d}}$ associated with the natural variability noise in the absence of a signal.

If the climate variability, and thus the probability distribution, of $\tilde{d}_\nu$ is indeed Gaussian, the $(p-1)$-dimensional hypersurfaces $\rho^2(\underline{d}) = $ constant represent surfaces of constant probability density. Thus $\rho^2(\underline{d})$ provides a measure of the statistical significance of the coefficient vector, as discussed above. If the joint probability distribution is non-Gaussian, a rigorous computation of the statistical significance will generally be fraught with still greater sampling uncertainties than in the one-dimensional case, so that ρ^2 will remain a useful statistic also in this case, providing at least a lowest-order (second moment) estimate of the statistical significance.

For a Gaussian distribution and known covariance matrix $\mathbf{C}$ (and therefore $\underline{\underline{D}}$), $\rho^2(\tilde{d})$ is a χ^2 variable with p degrees of freedom. If $\mathbf{C}$ is estimated from data, $\rho^2(\tilde{d})$ is a Hotelling-type variable (cf. Morrison 1990). However, in analogy with the single-pattern case, the ρ^2 distribution is not strictly Hotelling, as the set of estimated coefficients $\tilde{d}_\nu$ is modified by the multiplication with the inverse $\mathbf{C}^{-1}$ of the estimated covariance matrix. Therefore, the ρ^2 distribution must also be estimated in this case by approximate analytical or Monte Carlo techniques.

The problem of the optimal detection of an arbitrary unknown signal lying in the space spanned by the set of p guess signals g_ν can be defined now as the problem of determining a set of fingerprints $\mathbf{f}_\nu$ that maximizes ρ^2 for any signal lying in the prescribed signal space. The condition $\rho^2 = $ max clearly defines only the space spanned by the fingerprints, not the fingerprints themselves, since the scalar form ρ^2 is invariant with respect to any linear transformation to a new signal pattern base. This property will be used now to require that the optimal set of fingerprints forms an orthonormal base. Similarly, the set of signal patterns will also be redefined to represent an orthonormal base. (It was stated at the beginning of this section that orthonormality would not be presumed for the signal base g_ν. The resultant fingerprint solutions (28) were then similarly nonorthonormal. Orthonormality is invoked here only as an interim convenience and will be dropped again in the final result.)

Finally, the climate state space will be transformed,

$$\psi \rightarrow \psi'' = \mathbf{A}\psi, \qquad (34)$$

to a new coordinate system in which the covariance matrix **C** becomes the unit matrix,

$$\mathbf{C}'' = \mathbf{ACA}^\mathrm{T} = \mathbf{I}. \tag{35}$$

The transformation **A** can be obtained, for example, by first transforming to EOF coordinates ψ' and then normalizing the EOF coordinates to unit variance, $\psi_a'' = \psi_a'/\sigma_a$.

In order that the signal and detectors c_ν, d_ν remain invariant under this transformation, the signal patterns must transform in the same way as the climate state vector,

$$\mathbf{g}_\nu'' = \mathbf{A}\mathbf{g}_\nu, \tag{36}$$

while the fingerprints transform as adjoint vectors,

$$\mathbf{f}_\nu'' = (\mathbf{A}^\mathrm{T})^{-1}\mathbf{f}_\nu. \tag{37}$$

After these transformations, the statistic ρ^2 reduces to the Euclidean form

$$\rho^2 = \sum_\nu (d_\nu)^2 = \sum_\nu (\mathbf{f}_\nu''^\mathrm{T}\psi)^2. \tag{38}$$

It follows immediately that the optimal set of fingerprints that maximizes ρ^2 for any signal ψ^s lying in the space spanned by the set of signal patterns $\mathbf{g}_\nu''$ is given by

$$\mathbf{f}_\nu'' = \mathbf{g}_\nu'' \tag{39}$$

(or any equivalent rotated orthonormal base that spans the same space as the base $\mathbf{g}_\nu''$). For this solution, $\rho^2 = |\psi^s|^2$, while for any other fingerprint space, part of the signal will be lost in the projection onto the fingerprint space, so that $\rho^2 \leqslant |\psi^s|^2$.

Transforming back to the original coordinates, Eq. (39) becomes, applying Eqs. (36) and (37),

$$\mathbf{f}_\nu = \mathbf{A}^\mathrm{T}A\mathbf{g}_\nu \tag{40}$$

or, invoking Eq. (35),

$$\mathbf{f}_\nu = \mathbf{C}^{-1}\mathbf{g}_\nu. \tag{41}[=(28)]$$

The orthonormality conditions imposed on the guess patterns $\mathbf{g}_\nu''$ and fingerprints $\mathbf{f}_\nu''$ in the transformed climate state space transform into similar conditions in the original climate state coordinate system, but with the scalar products defined now with respect to the non-Euclidean metrics $\mathbf{C}^{-1}$ and $\mathbf{C}$, respectively. However, these orthonormality conditions can now be dropped, since only the signal and fingerprint spaces as such are of interest, the choice of base for either space being arbitrary. The signal base vectors can therefore be identified with the original guess patterns, without normalization and orthogonality restrictions, and the fingerprints can be similarly defined, without orthonormality considerations, by the relation (41). Thus the optimal p-dimensional detector vector $\underline{d}$ is identical to the detector vector found previously for the prescribed-pattern case.

It can be shown for the optimal fingerprint solution [using Eqs. (4)–(7) and (27)–(33)] that the multivariate significance measure ρ^2 for any signal $\psi^s = \mathbf{g}$ lying in the signal space $\mathbf{g}_\nu$ is identical to the single-pattern square signal-to-noise ratio R^2:

$$\rho^2(\underline{d}^s) = \underline{d}^\mathrm{T}\underline{D}^{-1}\underline{d} = \mathbf{g}^\mathrm{T}\mathbf{C}^{-1}\mathbf{g} = \frac{(d^s)^2}{\langle \tilde{d}^2 \rangle} = R^2. \tag{42}$$

The result (42) holds only for the optimal fingerprints and for the signal ψ^s itself, not for the net response ψ consisting of the signal plus noise. (For the case that ψ contains no signal, for example, $\psi = \tilde{\psi}$, one finds $\langle \rho^2(\underline{\tilde{d}}) \rangle = p$, while $R^2 = 1$.)

There remains now the second step of attributing an estimated signal ψ^e to the detection vector $\underline{d}$. Adopting again the view that the estimated signal should reproduce the true signal when the noise contamination of the observed response ψ is negligible (or, equivalently, should exhibit the smallest mean-square deviation from the true signal in the presence of noise) one finds, in analogy with the single-pattern case,

$$\psi^e = \sum_\nu d_\nu \mathbf{b}_\nu, \tag{43}$$

where the base vectors $\mathbf{b}_\nu$ of the estimated signal are given by

$$\mathbf{b}_\nu = \sum_\mu (\underline{G}^{-1})_{\nu\mu}\mathbf{g}_\mu \tag{44}$$

with

$$\underline{G}_{\mu\nu} = (\mathbf{g}_\mu^\mathrm{T}\mathbf{C}^{-1}\mathbf{g}_\nu). \tag{45}$$

Thus the estimated signal associated with the set of maximally significant detectors d_ν is given by

$$\psi^e = \sum_{\mu,\nu} (\mathbf{g}_\mu^\mathrm{T}\mathbf{C}^{-1}\psi)(\underline{G}^{-1})_{\mu\nu}\mathbf{g}_\nu. \tag{46}$$

In H, an alternative derivation of the maximally significant signal estimate was given, which yielded the same set of optimal detectors d_ν, but a different set of base vectors $\hat{\mathbf{b}}_\nu$, and thus a different optimally estimated signal pattern. Starting from the general form (43) (with $\mathbf{b}_\nu$ replaced by $\hat{\mathbf{b}}_\nu$), the detectors d_ν were first determined by a least-squares fit to each realization ψ for a fixed base $\hat{\mathbf{b}}_\nu$. Subsequently, the base was determined by applying a maximal-significance condition to the set of detectors. This yielded (as already mentioned for the single-pattern case) the solution $\hat{\mathbf{b}}_\nu = \mathbf{f}_\nu = \mathbf{C}^{-1}\mathbf{g}_\nu$. In contrast to the base $\mathbf{b}_\nu$ found in the present analysis [Eq. (44)], which defines the same space as the signal space $\mathbf{g}_\nu$, the space spanned by the fingerprint base $\hat{\mathbf{b}}_\nu$ is generally different from the signal space.

The present result (46) can also be derived as a least-squares error solution using an alternative definition of the square error based on the "significance metric" $\mathbf{C}^{-1}$ rather than the usual Euclidean metric **I**. This corresponds to finding an estimated signal in the prescribed

signal space for which the probability is maximized that the residual error represents a realization of the natural variability ensemble.

Representing the observed climate response as the superposition

$$\psi = \psi^e + \psi^r \qquad (47)$$

of an estimated signal

$$\psi^e = \sum_\nu \hat{c}_\nu \mathbf{g}_\nu \qquad (48)$$

in the signal space $\mathbf{g}_\nu$ and a residual error ψ^r, minimization of the square error norm

$$\|\psi^r\|^2 = (\psi^{rT}\mathbf{C}^{-1}\psi^r), \qquad (49)$$

one obtains in this case the solution

$$\hat{c}_\nu = \sum_\mu (\underline{G}^{-1})_{\nu\mu}(\mathbf{g}_\mu^T\mathbf{C}^{-1}\psi). \qquad (50)$$

Equations (48) and (50) yield the estimated signal (46) derived previously but expressed now in terms of the original set of base functions $\mathbf{g}^\nu$ instead of the transformed base $\mathbf{b}^\nu$.

It should be emphasized again that the statistical significance of the detectors $d_\nu = (\mathbf{f}_\nu^T\psi)$ or $\hat{c}_\nu$ is maximized and is identical for all three derivations, independent of the definition of the associated estimated signal.

In practice, multipattern signal detection will normally be carried out in a hierachical mode: first, a single pattern is tested; if this is detected as significant, a second pattern is added, and so forth until the ρ^2 statistic is no longer accepted as statistically significant. The success of the method depends critically, as always in signal detection problems, on the realistic choice of the prescribed signal patterns $\mathbf{g}_\nu$.

5. Natural variability models

The EOFs e_{ab} characterizing the space–time covariance matrix $\mathbf{C}_{ab}$ represent familiar functions. For a fixed subindex i (suppressed in the following relations), the vector $\psi_{(i,t)} = \psi_t$ can first be decomposed into EOFs e_{ft} with respect to the time index. As pointed out in section 3, for a statistically stationary process, the time-domain EOFs are simply the harmonic functions of the Fourier representation (this follows from the statistical orthogonality of the Fourier components, although these are not normally ordered with respect to variance, as in other applications).

Diagonalization of the covariance spectra with respect to the remaining index i characterizing the climate variables and spatial coordinates (referred to for brevity in the following as the "spatial" index) then yields the standard complex EOF representation of the covariance spectrum of a multivariate process (Wallace and Dickinson 1972; Barnett 1983).

The complete space–time set of EOFs e_{ab} is generally very large: it consists of a different set of complex EOFs

with respect to the spatial index i for each frequency band f of the spectrum. In practice, it will be difficult both to estimate such a large set of functions from a finite dataset and to work with the complete representation in numerical computations. One will therefore need to resort to some form of approximation or simplified model based on a reduced set of functions. Two such models are considered in the following. The first assumes statistical separability of the time and spatial dimensions, while the second uses a reduced POP representation (Hasselmann 1988; v. Storch et al. 1988).

a. Space–time separability

Separability of the space and time coordinates i and t, respectively, is defined here as the property that the eigenvector e_{ba} can be factored into a spatial eigenvector e_{ji}^s and a temporal eigenvector e_{ft}^t,

$$e_{ba} = e_{ji}^s e_{ft}^t, \qquad (51)$$

where $a = (i, t)$, $b = (j, f)$, and the upper indices s and t will be used generally to distinguish between spatial and temporal quantities, respectively.

This property holds if the time-dependent coefficients of the standard spatial EOFs are completely uncorrelated, that is, if the coefficients are uncorrelated not only for the same time arguments, as implicit in the definition of the spatial EOFs, but also for nonzero time lags.

Under these conditions the spatial eigenvectors satisfy the eigenvalue equations

$$\sum_j C_{(i,t)(j,u)}e_{kj}^s = (\sigma_{tu,k}^s)^2 e_{ki}^s \qquad (52)$$

in which the time indices t, u appear as parameters. Only the eigenvalues $(\sigma_{tu,k}^s)^2$ depend on these parameters, not the eigenvectors themselves.

The temporal eigenvectors satisfy the eigenvalue equation

$$\sum_u (\sigma_{tu,k}^s)^2 e_{fu}^t = (\sigma_{(k,f)})^2 e_{ft}^t, \qquad (53)$$

which then yields for the full system the eigenvalue equation

$$\sum_{j,u} C_{(i,t)(j,u)}e_{kj}^s e_{fu}^t = (\sigma_{(k,f)})^2 e_{ki}^s e_{ft}^t. \qquad (54)$$

The eigenvalue $(\sigma_{(k,f)})^2$ represents the variance of the f'th spectral band of the autovariance spectrum of the k'th spatial EOF. Thus, for a separable system, the full space–time climate covariance matrix C_{ab} is completely characterized by the spatial EOFs e_{ki} and their autovariance spectra $\sigma_{(k,f)}^2$.

Transforming to EOF coordinates $(g_{(k,f)}^\nu)'$ (denoted by primes, and writing the signal pattern index ν now as a superscript to relieve index congestion) the expressions for the guess signal patterns take the form

$$g_{(i,t)}^\nu = \sum_{k,f} (g_{(k,f)}^\nu)' e_{ki}^s e_{ft}^t, \qquad (55)$$

which yields for the fingerprints

$$f^\nu_{(i,t)} = \sum_{k,f} (g^\nu_{(k,f)})' \sigma^{-2}_{(k,f)} e^s_{ki} e^t_{ft}. \qquad (56)$$

The basic advantage of a separable system is that the spatial EOFs can first be determined in the standard manner without filtering in the time domain. The analysis in the time frequency domain can then be carried out for each EOF separately as a second step.

Example 1

To illustrate, consider a single pattern signal consisting of a superposition of a number of spatial EOFs e^s_{ki},

$$g_{(i,t)} = \sum_k \gamma_{kt} e^s_{ki}, \qquad (57)$$

where each of the time-dependent spatial EOF coefficients γ_{kt} can be represented as a linear trend,

$$\gamma_{kt} = m_k(t - 1), \qquad (58)$$

with $m_k = $ const.

Taking the Fourier transform of γ_{kt}, one obtains for the coefficient of the signal in EOF coordinates,

$$g'_{(k,f)} = m_k \begin{cases} i \sqrt{T} \dfrac{\exp(-\pi i f/T)}{2 \sin(\pi f/T)} & \text{for } f \neq 0 \\[2ex] T^{1/2} \dfrac{(T-1)}{2} & \text{for } f = 0 \end{cases} \qquad (59)$$

The fingerprint can then be obtained from Eq. (56).

Figure 1 shows the spectra and typical fingerprints $f^\nu_{(i,t)}$ computed for red power-law spectra, $\sigma^2_{(k,f)} = \sigma^2_k(f + T^{-1})^{-q}$ for various values of $q > 0$ (the frequency is offset by one discrete frequency unit T^{-1} to avoid the singularity at $f = 0$). The enhancement of the square signal-to-noise ratio R^2 computed with the optimal fingerprint relative to the reference nonoptimized case in which the fingerprint is simply set equal to the signal pattern is shown in Table 1. The computations were carried out for $T = 100$ discrete time steps.

The curves demonstrate that the optimal detection fingerprint generally differs significantly from the nonoptimized fingerprint, which is given by a straight line. The nonoptimized solution reduces in the present case to the estimation of a linearly increasing signal by the standard method of constructing a regression line through the data. However, this represents the optimal analysis method only for a white-noise natural variability spectrum. For red spectra, the optimal weighting is distributed more toward the end points of the time series. For power-law spectra steeper than f^{-1}, the optimal detection strategy is to use the end points of the time series only. The enhancement of the statistical significance through optimization of the fingerprint can be quite large for steep spectra (cf. Table 1).
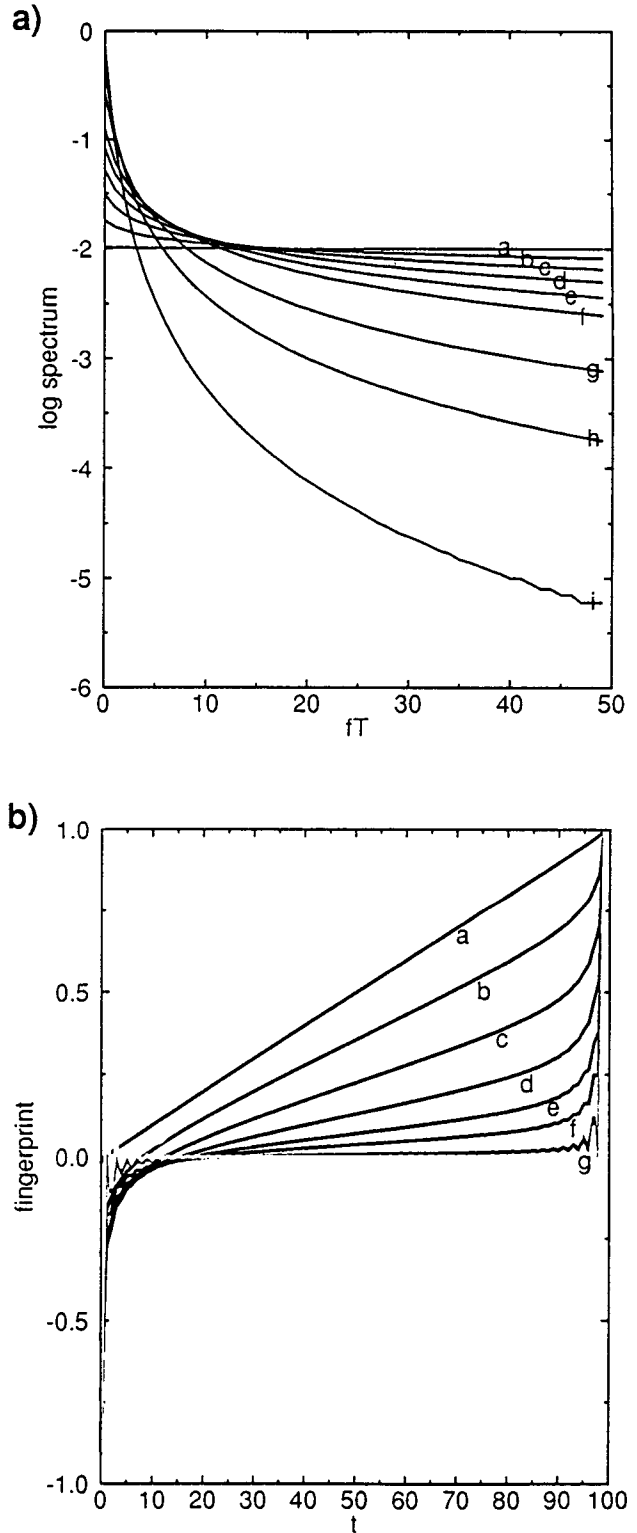

FIG. 1. (a) Noise spectra σ^2_f and (b) optimal fingerprints for the detection of a linear time-dependent signal in presence of power-law noise. Definitions of cases a–i and values of the associated signal-to-noise enhancement factors are given in Table 1. The length of the time series is 100 ($0 \leq t \leq 99$; $-50 \leq fT \leq 49$; only the positive-frequency branches of the spectra are shown.)

TABLE 1. Signal-to-noise enhancement factors $E = [R^2 \text{ (optimal fingerprint)}/R^2 \text{ (signal pattern)}]$ for linear signal and various power law spectra.

Case (cf. Fig. 1)	a	b	c	d	e	f	g	h	i
Spectral power q	0	0.2	0.4	0.6	0.8	1.0	1.5	2	3
Enhancement factor E	1	1.01	1.05	1.14	1.27	1.5	3.1	10.4	>100

b. POP representation

An alternative method of reducing the complexity of the full covariance matrix C_{ab} without introducing the rather stringent assumption of space–time separability is to represent the natural climate variability as a superposition of a finite number of principal oscillation patterns (POPs, cf. Hasselmann 1988; v. Storch et al. 1988). The basic idea of the POP method is to combine an EOF-type pattern expansion in the spatial domain with an ARMA-type dynamical modeling approach in the time domain. In the original papers of Hasselmann and v. Storch et al. [and in a number of subsequent applications, cf. Xu (1992, 1993), Latif et al. (1993a, 1993b)], the POP method was regarded primarily as a technique for constructing simple dynamical models, usually for forecasting or diagnostic purposes. However, the POP method is equally useful as an approximate multivariate spectral compression technique.

The POP method approximates the natural variability $\tilde{\phi}_a = \tilde{\phi}_i(t)$ (returning here to the usual decomposed index notation) as a superposition of a number of damped oscillations α,

$$\phi_i^\alpha(t) = p_i^\alpha a^\alpha(t) + \text{complex conjugate}, \quad (60)$$

where p_i^α is a constant complex pattern,

$$p_i^\alpha = p_i^{\alpha(1)} + ip_i^{\alpha(2)} \quad (61)$$

and the complex amplitude

$$a^\alpha(t) = a^{\alpha(1)} + ia^{\alpha(2)} \quad (62)$$

satisfies the damped oscillator equation

$$(d/dt - 2\pi i f^\alpha + \lambda^\alpha)a^\alpha = n^\alpha(t) \quad (63)$$

with eigenfrequency f^α (>0) and damping factor λ^α (>0); $n^\alpha(t)$ is a complex white-noise forcing function,

$$n^\alpha(t) = n^{\alpha(1)} + in^{\alpha(2)}. \quad (64)$$

Decomposed into spectral components $\phi_i^\alpha(f)$ (to avoid notational proliferation the same symbols will be used for time- and frequency-domain functions) the solution of (63) is given by

$$a^\alpha(f) = T^\alpha(f)n^\alpha(f), \quad (65)$$

where

$$T^\alpha(f) = \{2\pi i(f - f^\alpha) + \lambda^\alpha\}^{-1} \quad (66)$$

is the POP transfer function, yielding for the POP oscillation (60)

$$\phi_i^\alpha(t) = \sum_f T^\alpha(f)n^\alpha(f)P_i^\alpha \exp(2\pi ift)$$

$$+ \text{complex conjugate} = \sum_f \{T^\alpha(f)n^\alpha(f)P_i^\alpha$$

$$+ T^\alpha(-f)^*n^\alpha(-f)^*P_i^{\alpha*}\} \exp(2\pi ift) \quad (67)$$

(note that $T^\alpha(f) \neq T^\alpha(-f)^*$, $n^\alpha(f) \neq n^\alpha(-f)^*$ since both the POPs and the noise forcing are complex).

While the free POP solution consists of a single damped oscillation rotating clockwise in the complex plane, characterized by the pattern sequence $p_i^{\alpha(1)} \rightarrow -p_i^{\alpha(2)} \rightarrow -p_i^{\alpha(1)} \rightarrow p_i^{\alpha(2)} \rightarrow p_i^{\alpha(1)}$, the forced solution consists generally of a superposition of two oscillations rotating in opposite directions. These arise from the positive and negative frequencies in the spectral representation of $n(t)$, which force both the basic POP pair and the complex conjugate pattern pair.

Assuming that the forcing components n^α, n^β for different POPs α and β are uncorrelated, the complex cross-spectrum of the process $\tilde{\Phi}_i$ is then given by

$$F_{ij}(f)$$

$$= \langle \{\tilde{\phi}_i(f)\}^*\tilde{\phi}_j(f)\rangle = \sum_\alpha \{N^\alpha[|T^\alpha(f)|^2(p_i^\alpha)^*p_j^\alpha$$

$$+ |T^\alpha(-f)|^2 p_i^\alpha(p_j^\alpha)^*]$$

$$+ 2 \text{Re}[M^\alpha T^\alpha(f)T^\alpha(-f)p_i^\alpha p_j^\alpha]\}, \quad (68)$$

where

$$N^\alpha = \langle n^\alpha(f)^*n^\alpha(f)\rangle = \text{const}, \quad (69)$$

$$M^\alpha = \langle n^\alpha(f)n^\alpha(-f)\rangle = \text{const}. \quad (70)$$

The general expression (68) can be simplified by assuming that the excitation of the conjugate POP pair, which occurs at negative frequencies far removed from the resonant POP eigenfrequencies, is negligible. In this case the second and third terms in (68) are small for positive f (and similarly the first and third terms for negative f), so that for $f > 0$, (68) reduces to

$$F_{ij}(f) = \langle \{\tilde{\phi}_i(f)\}^*\tilde{\phi}_j(f)\rangle$$

$$= \sum_\alpha N^\alpha |T^\alpha(f)|^2(p_i^\alpha)^*p_j^\alpha. \quad (71)$$

The corresponding expression for $f < 0$ follows from (71) and the symmetry relation $F_{ij}(f) = F_{ij}(-f)^*$. (In applications, such as the numerical example considered below, it will generally be more convenient to drop only the third term in (68), as this avoids a dis-
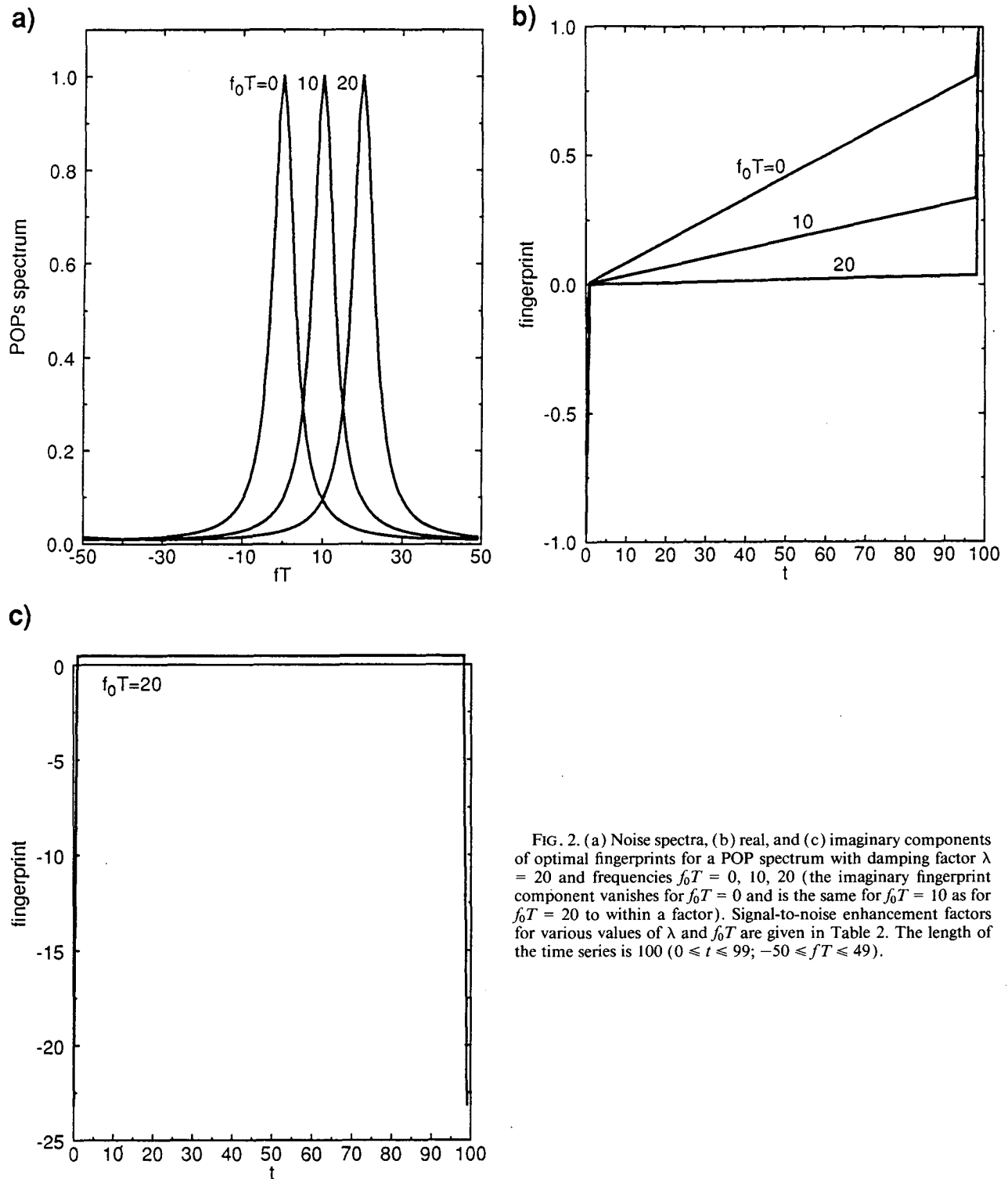
FIG. 2. (a) Noise spectra, (b) real, and (c) imaginary components of optimal fingerprints for a POP spectrum with damping factor λ = 20 and frequencies $f_0 T$ = 0, 10, 20 (the imaginary fingerprint component vanishes for $f_0 T$ = 0 and is the same for $f_0 T$ = 10 as for $f_0 T$ = 20 to within a factor). Signal-to-noise enhancement factors for various values of λ and $f_0 T$ are given in Table 2. The length of the time series is 100 ($0 \leqslant t \leqslant 99$; $-50 \leqslant fT \leqslant 49$).

continuous change in the expression for the covariance spectrum which otherwise occurs at zero frequency.)

The simplified POP representation of the cross-spectrum is seen to have the same form as the standard complex EOF representation

$$F_{ij}(f) = \sum_{\nu} F^{\nu}(f)(e_i^{\nu})^* e_j^{\nu} \qquad (72)$$

of the cross-spectrum in terms of complex EOFs e^{ν}. The variance spectrum $F^{\nu}(f) = \langle |c^{\nu}(f)|^2 \rangle$ of the coef-

ficients $c^r(f)$ of the complex EOF expansion correspond to the spectrum $N^\alpha |T^\alpha(f)|^2$ of the POP representation while the complex EOF e^r_i itself corresponds to the principal oscillation pattern p^α_i. However, in contrast to the standard representation, the cross-spectrum is now no longer decomposed into a different set of EOFs for each spectral band, but into a single set of complex POP patterns applicable for the entire spectrum.

The contribution of individual POPs to different spectral bands is determined by the weighting factor $|T^\alpha(f)|^2$ in (71). As pointed out, individual POPs will contribute mainly to spectral bands in the neighborhood of the POP eigenfrequencies f^α, the effective spectral bandwidth being proportional to the damping factor λ^α.

In contrast to the true EOFs e^r_i, the complex POPs do not represent the eigenvectors of a Hermitian matrix and will therefore generally not be (spatially) orthogonal. They can be readily orthogonalized, however, through a suitable (frequency dependent) complex rotation $p^\alpha_i \to p^{\alpha'}_i = \sum_\beta S^{\alpha\beta} p^\beta_i$, where $S^{\alpha\beta}$ is a unitary matrix: $\sum_\gamma (S^{\alpha\gamma})^* S^{\beta\gamma} = \delta^{\alpha\beta}$. The transformation preserves the essential statistical orthogonality of the POP coefficients.

The determination of the approximate form (72) is normally carried out in the time domain by fitting a first-order vector Markov process to the data time series (cf. v. Storch et al. 1988). However, a direct fit of the model covariance spectrum to the observed covariance spectrum in the frequency domain, using methods applied by Frankignoul and Hasselmann (1977), Lemke et al. (1980), Herterich and Hasselmann (1987), and Dobrovolski (1992) in similar problems of stochastic model fitting, should also be feasible in the POP model case. Model fitting in the spectral domain generally has the advantage of providing better quantitative estimates of error bounds.

Example 2

To illustrate the impact of optimal filtering for a POP noise spectrum, consider the case of a signal lying in the pattern space spanned by the POPs. Since the different POP pairs are assumed to be statistically orthogonal, one can consider each POP signal component

$$g_i(t) = \gamma(t)p_i + \text{complex conjugate} \qquad (73)$$

separately [$\gamma(t)$ denotes a complex time-dependent coefficient, and the POP index α has been dropped].

The optimal fingerprint

$$f_i(t) = \phi(t)p_i + \text{complex conjugate} \qquad (74)$$

can be represented in this case in closed form. In the Fourier domain, the complex fingerprint coefficient is given by [Eqs. (14), (71)]

$$\phi(f) = \gamma(f) \frac{\{[2\pi(f - f_0)]^2 + \lambda^2\}}{N}, \qquad (75)$$

which yields in the time domain

$$\phi(t) = \text{const}\{[(2\pi f_0)^2 + \lambda^2]\gamma(t)$$
$$- d^2\gamma(t)/dt^2 + i4\pi f_0 d\gamma(t)/dt\}. \qquad (76)$$

The fingerprint is seen to have the right structure to reduce the POP noise contributions. For a modulation factor $\gamma(t) \sim \exp(2\pi i f_0 t - \lambda t)$ corresponding to a pure POP oscillation, Eq. (76) yields $\phi(t) = 0$, that is, pure POP oscillations are rejected by the fingerprint. (This result appears paradoxical, since the fingerprint is defined in the Fourier domain as the product of the signal and the inverse noise spectrum, which cannot vanish identically. The explanation is that a pure POP oscillation is not a permitted signal form, since it becomes infinite for $t \to -\infty$, so that its Fourier transform does not exist. A signal that is zero for $t \leq 0$ and represents a POP oscillation only for $t > 0$ is not completely removed by the fingerprint.)

Figure 2 shows the POP spectra $|T^\alpha(f)|^2$ and fingerprint modulation factors $\phi(t)$ for the case of a linear signal modulation factor $\gamma(t)$ [Eq. (58)], taking again $T = 100$ time increments, for the frequency values $f_0 = 0$, 10, and 20 and the damping factor $\lambda = 20$ (measured in frequency increment units T^{-1}). For a linear signal (which appears as a periodic sawtooth in the Fourier sum representation) the first derivative in Eq. (76) consists of the sum of a constant term and a negative δ-function (negative spike) at each end of the time interval. The second derivative is given by the derivative of a δ function, which in the discrete representation takes the form of a positive and negative spike at the beginning and end, respectively, of the time interval.

The real part of the fingerprint modulation factor consists then of the linear signal modulation factor itself plus a term representing the difference in the response at the end points of the time interval, the latter term gaining more weight the more the spectrum differs from a uniform white spectrum. This is qualitatively similar to the result found for red power-law spectra in the previous example of a space–time separability model.

The imaginary component of the fingerprint modulation factor consists of two equal negative contributions from the end points of the time interval and an equally weighted constant positive contribution acting over the full time interval.

The enhancement of the square signal-to-noise ratio R^2 achieved using the optimal fingerprint solution relative to the reference case without optimization is shown in Table 2. As in the previous example, the largest enhancement is achieved for noise spectra containing large variance contributions at low frequencies.

6. Summary

The signal pattern detection method developed by H for the time-independent problem can be readily extended to the time-dependent case. The introduction

TABLE 2. Signal-to-noise enhancement factors $E = [R^2$ (optimal fingerprint)$/R^2$ (signal pattern)] for linear signal and *POP* variance spectra for various eigenfrequencies f_0 and damping factors λ.

λ	f_0T		
	0	10	20
10	3.5	1.21	1.16
20	1.63	1.12	1.08
40	1.14	1.06	1.04

of the fingerprint concept leads to a significant simplification of the theory both conceptually and analytically. For a given space–time-dependent signal **g**, the optimal filter **f** for the detection of the signal **g** in the presence of natural climate variability noise is given by $\mathbf{f} = \mathbf{C}^{-1}\mathbf{g}$, where **C** is the space–time covariance matrix of the natural variability. Application of the optimal fingerprint to an observed climate trajectory ψ yields a detector $d = (\mathbf{f}^T\psi)$ with maximal signal-to-noise ratio.

This result generalizes immediately to the multi-pattern case: the set of fingerprints $\mathbf{f}_\nu = \mathbf{C}^{-1}\mathbf{g}_\nu$ associated with a set of p guess patterns $\mathbf{g}_\nu$ yields a set of detectors $d_\nu = (\mathbf{f}_\nu^T\psi^s)$ for which the relevant p-dimensional statistical-significance statistic $\rho^2 = \sum_{\nu,\mu} d_\nu D_{\nu\mu}^{-1} d_\mu$ is maximized for any signal ψ^s lying in the space spanned by the set of guess patterns $\mathbf{g}_\nu$. Here, $D_{\nu\mu} = (\mathbf{g}_\nu^T\mathbf{C}^{-1}\mathbf{g}_\mu)$ represents the covariance matrix of the detector noise $\tilde{d}_\nu = (\mathbf{f}_\nu^T\tilde{\psi})$ induced by the climate variability $\tilde{\psi}$.

The fingerprint emphasizes those components of the signal that are least contaminated by noise. The direction of the fingerprint vector therefore normally differs from the direction of the associated signal. This permits alternative interpretations of the estimated signal inferred from the set of detectors. In H, the detection and estimation problems were regarded as coupled parts of a single problem, and the estimated signal was defined to lie in the space spanned by the set of fingerprint vectors. In the present analysis, the detection problem was solved first and the attribution of an estimated signal to the set of detectors was addressed subsequently as an independent problem. From this viewpoint it appears more consistent to regard the estimated signal as lying in the space spanned by the prescribed signal patterns. Arguments can be given for either interpretation. The statistical significance of the estimated signal is determined in both cases by identical sets of detectors and is thus independent of the interpretation.

The detection technique can be applied to any set of observed or model-simulated data for which the second moments can be adequately estimated, independent of the completeness of the dataset with regard to the dynamical description of the climate system.

Two practical difficulties are encountered in applying the technique. First, a complete description of the space–time-dependent covariance matrix of the natural climate variability noise involves large quantities of information that cannot normally be effectively handled—and also cannot be inferred from a finite amount of observed or simulated data. Thus some form of simplified statistical model must be introduced. An expansion in POPs (principal oscillation patterns) provides an effective general reduction technique. In some cases a still simpler space–time separability model may be applicable. Examples given for both types of model demonstrate that the optimal fingerprints can deviate significantly from the original signal pattern and yield considerably enhanced signal-to-noise ratios compared with a straightforward projection onto the signal pattern. The strongest enhancement is obtained for red spectra with high-variance contributions at very low frequencies.

Secondly, the optimally estimated detectors d_ν have a known Gaussian distribution only if the natural climate variability is Gaussian with known (rather than estimated) covariance matrix. In the normal case that the climate variability is non-Gaussian or the covariance matrix is estimated from a limited dataset, the statistical significance of the computed detectors must be estimated by Monte Carlo simulations or other approximate techniques.

Not addressed in the present paper were problems of data errors (uncertainties were associated solely with the natural climate variability) or questions related to the definition of the signal pattern, including extensions of the theory to allow for a priori probabilities of the anticipated signal distribution within a prescribed signal space. The author intends to pursue these questions later in applications of the theory.

REFERENCES

Barnett, T. P., 1983: Interaction of the monsoon and Pacific trade wind system at interannual time scales. Part I: The equatorial zone. *Mon. Wea. Rev.,* **111,** 756–773.

——, 1986: Detection of changes in 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., 559–568.

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

——, and M. E. Schlesinger, 1987: Detecting changes in global climate induced by greenhouse gases. *J. Geophys. Res.,* **92,** 14 772–14 780.

——, ——, 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., 537–558.

Bell, T. L., 1982: Optimal weighting of data to detect climatic change: Application to the carbon dioxide problem. *J. Geophys. Res.,* **87,** 11 161–11 170.

——, 1986: Theory of optimal weighting of data to detect climatic change. *J. Atmos. Sci.,* **43,** 1694–1710.

Bloomfield, P., and D. Nychka, 1992: Climate spectra and detecting climate change. *Clim. Change,* **21,** 275–287.

Cubasch, U., K. Hasselmann, H. Höck, E. Maier-Reimer, U. Mikolajewicz, B. D. Santer, and R. Sausen, 1992: Time dependent greenhouse warming computations with a coupled ocean-atmosphere model. *Climate Dyn.,* **8,** 55–69.

Dobrovolski, S. G., 1992: Global climate changes in water and heat transfer accumulation processes. *Dev. Atmos. Sci.,* **21,** 265 pp.

Frankignoul, C., and K. Hasselmann, 1977: Stochastic climate models. Part 2: Application to sea surface temperature anomalies and thermocline variability. *Tellus,* **29,** 289–305.

Hannoschöck, G., and C. Frankignoul, 1985: Multivariate statistical analysis of sea surface temperature anomaly experiments with the GISS general circulation model. *J. Atmos. Sci.,* **42,** 1430–1450.

Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies. *Meteorology of Tropical Oceans,* Royal Meteorology Society, 251–259.

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

Hense, A., 1986: Multivariate statistical investigations of the Northern Hemisphere circulation during the El Niño event 1982/83. *Tellus,* **38A,** 189–204.

——, R. Glowienka-Hense, H. v. Storch, and U. Stähler, 1990: Northern Hemisphere atmospheric response to changes of Atlantic Ocean SST on decadal time scales: A GCM experiment. *Climate Dyn.,* **4,** 157–174.

Herterich, K., and K. Hasselmann, 1987: Extraction of mixed layer advection velocities, diffusion coefficients, feedback factors and atmospheric forcing parameters from the statistical analysis of North Pacific SST anomaly fields. *J. Phys. Oceanogr.,* **17,** 2145–2156.

Karoly, D. J., 1987: Southern Hemisphere temperature trends: A possible greenhouse gas effect? *Geophys. Res. Lett.,* **14,** 1139–1141.

——, 1989: Northern Hemisphere temperature trends: A possible greenhouse gas effect? *Geophys. Res. Lett.,* **16,** 465–468.

Latif, M., A. Sterl, E. Maier-Reimer, and M. M. Junge, 1993a: Climate variability in a coupled general circulation model. Part I: The tropical Pacific. *J. Climate,* **6,** 5–21.

——, ——, ——, and ——, 1993b: Structure and predictability of the El Niño/Southern Oscillation phenomenon. *J. Climate,* **6,** 700–708.

Lemke, P., E. W. Trinkl, and K. Hasselmann, 1980: Stochastic dynamic analysis of polar sea ice variability. *J. Phys. Oceanogr.,* **10,** 2100–2120.

MacCracken, M. C., and H. Moses, 1982: The first detection of carbon dioxide effects: Workshop Summary, Harpers Ferry, West Virginia. *Bull. Amer. Meteor. Soc.,* **63,** 1164–1178.

Madden, R. A., and V. Ramanathan, 1980: Detecting climate change due to increasing carbon dioxide. *Science,* **209,** 763–768.

Manabe, S., R. J. Stouffer, M. J. Spelman, and K. Bryan, 1991: Transient responses of a coupled ocean–atmosphere model to gradual changes of atmospheric CO_2. Part I: Annual mean response. *J. Climate,* **4,** 785–818.

Mikolajewicz, U., and E. Maier-Reimer, 1990: Internal secular variability in an ocean general circulation model. *Climate Dyn.,* **4,** 145–156.

——, ——, and T. P. Barnett, 1993: Acoustic detection of greenhouse-induced climate changes in the presence of slow fluctuations of the thermo haline circulation. *J. Phys. Oceanogr.,* **23,** 1099–1109.

Morrison, D. F., 1990: *Multivariate Statistical Methods.* 3d ed. McGraw Hill, 495 pp.

Munk, W. H., and A. M. G. Forbes, 1989: Global ocean warming: An acoustic measure? *J. Phys. Oceanogr.,* **19,** 1765–1778.

North, G. R., K.-Y. Kim, 1993: Detection of forced climate signals. Part II: Simulation results. *J. Climate,* submitted.

——, ——, S. P. Shen, and J. W. Hardin, 1993: Detection of forced climate signals. Part I: Filter theory. *J. Climate,* submitted.

Preisendorfer, R. W., 1988: Principal component analysis in meteorology and oceanography. *Dev. Atmos. Sci.,* **17,** Elsevier, 1–425.

Santer, B. J., W. Brüggemann, U. Cubasch, K. Hasselmann, H. Höck, E. Maier-Reimer, and U. Mikolajewicz, 1993: Signal-to-noise analysis of time-dependent greenhouse warming experiments. Part 1: Transient pattern analysis; Part 2: Standard error estimates. *Climate Dyn.,* submitted.

Solow, A., 1991: On the statistical comparison of climate model output and climate data. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations.* M. E. Schlesinger, Ed., 505–510.

Storch, H. v., and G. Hannoschöck, 1985: Statistical aspects of estimated principal vectors ("EOFs") based on small sample sizes. *J. Climate Appl. Meteor.,* **24,** 716–724.

——, 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.

Stouffer, R. J., S. Manabe, and K. Bryan, 1989: Interhemispheric asymmetry in climate response to a gradual increase of atmospheric CO_2. *Nature,* **342,** 660–662.

Wainstein, L. A., and V. D. Zubakov, 1962: *Extraction of signals from noise.* R. A. Silverman, Trans., Prentice-Hall, 362 pp.

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

Washington, W. M., and G. A. Meehl, 1989: Climate sensitivity due to increased CO_2: Experiments with a coupled atmosphere and ocean general circulation model. *Climate Dyn.,* **4,** 1–38.

Wigley, T. M. L., and P. D. Jones, 1981: Detecting CO_2-induced climatic change. *Nature,* **292,** 205–208.

——, and T. P. Barnett, 1990: Detection of the greenhouse effect in the observations. *Climate Change. The IPCC Scientific Assessment,* J. T. Houghton, G. J. Jenkins, and J. J. Ephraums, Eds., Cambridge University Press, 239–256.

Xu, J. S., 1992: On the relationship between the stratospheric quasi-biennial oscillation and the tropospheric Southern Oscillation. *J. Atmos. Sci.,* **49,** 725–734.

——, 1993: The joint modes of the coupled atmosphere–ocean system observed from 1967–1986. *J. Climate,* **6,** 816–838.